Enhancing museum experiences: Using immersive environments to evaluate soundscape preferences

Milena J. Bem, Samuel R.V. Chabot, Vic Brooks, and Jonas Braasch

Revision Report for the first revision, Oct. 7, 2024

Before we go into the details of our revision report, we would like to thank you, Tin Oberman, and the reviewers for your substantial effort in helping us improve this paper. The literature on museum acoustics is relatively scarce, making this special issue an ideal platform to address the gap in publications on museum soundscape. The comments were very productive and instrumental in enhancing our work. We divided our response into two parts. First, we addressed the two reviewers' comments individually. Second, we focus on the introduction, where we have made significant changes to incorporate references to literature on comfort within museums, museum fatigue, and the role of sound maskers in museums, as requested by reviewer 2.

Reviewer 1

"The paper presents relevant research on soundscapes in indoor museum environments, with interesting findings. The comments and remarks are listed below."

Reviewer 1 Comment: "Abstract, line 13: please use the full name of "MASS MoCA"."

Response:

Added: "Massachusetts Museum of Contemporary Art", on page 2, line 11.

Reviewer 1 Comment (Line 182): "Introduction: I would finish the Introduction section at line 95 and move the content from line 96 to line 111 to the Methods section, as it contains very specific information on the methodology and research questions to be answered."

Response: Lines 96 to 111 were moved to the "Methods" Section on page 7, lines 125-139:

"Utilizing a method developed for subjective assessment of soundscape preference in immersive environments (Bem, 2023), 23 participants were immersed in an audiovisual experiment using the Panorama Screen System at Rensselaer Polytechnic Institute. Six museum rooms combined with four soundscapes were randomly presented to participants. After each scenario, test participants were asked to answer a questionnaire using a 7-point Likert Scale. They used a wireless controller as a user interface while virtually immersed, allowing for a nuanced evaluation of each factor. Statistical tests, including Kruskal-Wallis and Holm Bonferroni, were then employed to elucidate the role of soundscapes in enhancing the museum experience. Our research seeks to address the following questions through a comparative analysis of four different soundscapes:

1. Are sounds congruent with the museum environment more effective than conventional masking sounds in reducing background noise distractions, thereby enhancing the experience and engagement with the artwork and creating a more immersive museum experience?

2. Are there notable differences in preferences among the soundscapes?

3. Do soundscape preferences vary across exhibitions?”

Reviewer 1 Comment: “Lines 127-129: Hong did not examine complex systems such as 3D HOA which would surely show improved performance over the 2D FOA, but I guess their goal was to examine the systems that are within reach of the "average" user.”

Response: You are correct. To address your comment, we replaced the following text on page 9, line 154.

Replaced:

“Specifically, the FOA 2-dimensional (2D) octagonal array offers the best overall listening experience and fidelity for evaluating soundscapes in virtual reality (Hong et al., 2019).”

Replaced with (on page 9, line 154):

“For example, Hong et al. examined various systems and highlighted the FOA 2-dimensional (2D) octagonal array for its effective performance and overall listening experience in virtual reality (Hong et al., 2019). While more complex systems, such as 3D Higher-Order Ambisonics (HOA), offer superior spatial accuracy and fidelity, the 2D FOA remains more accessible.”

Reviewer 1 Comment: “Lines 133-139: The term "emphasis" in this sentence would function better if "equal" was left out.”

Response: Deleted “equal” on page 9, line 165.

Reviewer 1 Comment: “Line 138: Does this mean that the loudspeakers are hidden behind the projection screen?”

Response: You are correct. To address your comment, we added the following text:

Added (on page 10, line 169): “Hidden behind the screen are”

Reviewer 1 Comment: “Lines 139-140: "congruent audiovisual content with balanced fidelity". I am not sure how to interpret this. From an engineer's point of view, it sounds very artistic.”

Response: Thank you for your comment. We agree that the original phrasing could be misinterpreted. To clarify the meaning, we have revised the sentence to be more objective and technically precise.

Replaced:

“Together, these elements deliver congruent audiovisual content with balanced fidelity”

Replaced with (on page 10, line 171):

“Together, these elements deliver synchronized audiovisual content, ensuring that the spatial audio rendering is accurately aligned with the visual projection, resulting in a cohesive experience where both sound and visuals are precisely timed.”

Reviewer 1 Comment: “Lines 162-163: "each participant was exposed to only six rooms". It should be explained here that one of the rooms was used for practice, and the other five were evaluated.”

Response: Thank you for your observation. We have addressed this by adding a more precise explanation to the “Visual Stimulus” section and removing redundant information in the “Experiment Design” section. This improves clarity and centralizes the relevant content in one place. Following are all the changes that we made to address this comment:

Replaced:

“However, each participant was exposed to only six rooms, with a randomized presentation methodology implemented to ensure diversity in the rooms presented to each participant.”

Replaced with (page 11, line 201):

“A total of fifteen rooms were documented for this study. However, each participant was exposed to only six rooms: one room was used for practice, and the remaining five were evaluated. The purpose of the practice tests was to evaluate their proficiency rather than to assess the task’s learning curve. Furthermore, the rooms and the sound scenarios were presented in a random order for each participant, ensuring that no two participants experienced the same sequence of rooms, which helped minimize order effects and bias in their responses.”

Deleted (page 19, line 355):

“The experiment ended after each participant performed this cycle six times, which means that each participant was randomly audiovisually exposed to six rooms. However, the first room was designated as a test round to ensure precision and accuracy in their assessments. The duration of practice tests varied, lasting between 5 and 8 minutes, a timeframe determined by the participants. The purpose of these practice tests was to evaluate their proficiency rather than to assess the task’s learning curve. The order of sound stimuli was also randomized.”

Reviewer 1 Comment: “Line 163: “randomized presentation methodology". A short description would be appreciated.”

Response: I added a description to make that more straightforward. Please see my previous comment.

Reviewer 1 Comment: "Lines 167-168 the sounds to be played are divided into onsite recordings and artificial soundscapes."

Response: You are correct. To address your comment, we replaced the following text:

Replaced:

"The presented sound was a mix of onsite sound recordings at the museums and artificially generated soundscapes."

Replaced with (page 13, line 212):

"The auditory elements comprise three-minute stimulus sections, with the sounds divided into two categories: onsite sound recordings from the museums and artificial soundscapes."

Reviewer 1 Comment: "Line 182: The A-weighted equivalent continuous sound level of 41.1 dBA is rather high. I am wondering how the room meets the RC-15 criteria."

Response: Thank you for pointing this out. Upon reviewing the data, we realized there was a misunderstanding in reporting the sound level. We measured the background sound level in different positions. The most critical position for this experiment is at the center of the screen at the average height of the human ear (we used 1.5 meters), where the participants are located most of the time. The correct A-weighted equivalent continuous sound level measured at this position when the projectors were on was 31.1 dBA. The first reported level of 41.1 dBA was measured above the screen, on the catwalk, one meter from the projectors, which are mounted high up. The inverse square law applies here.

The RC-15 criteria were met when the projectors were off. However, since the projectors were running continuously during the experiment, the room did not meet the RC-15 criteria under actual experimental conditions, as 31.1 dBA is slightly above RC-15. We previously mentioned that the empty Studio 2 met the RC-15 standard:

"Furthermore, the empty Studio 2, which hosts the panorama screen, features a low-noise floor that meets the RC-15 standard criteria (Braasch et al., 2010)."

This phrasing needed to be clarified. By "empty," we would like to say that without the panoramic screen system installed, the room meets the RC-15, as the panoramic screen is not a permanent fixture in Studio 2. The screen system is set up for only a few months each year, after which Studio 2 reverts to a performance space. When the projectors were installed, the sound levels increased. To clarify this, we have replaced the original statement to make it more straightforward and to avoid confusion regarding the room's noise levels during the experiment.

Replaced:

“Furthermore, the empty Studio 2, which hosts the panorama screen, features a low-noise floor that meets the RC-15 standard criteria (Braasch et al., 2010).”

Replaced with (on page 10, line 175):

“Studio 2, when empty, meets the RC-15 standard criteria for low-noise floors (Braasch et al., 2010). However, during the experiment, the installation of the projectors and the panoramic screen system raised the sound levels of the RC-15-rated studio to 31.1 dBA. This equivalent sound level was measured at the center of the screen, at a height of 1.5 meters, using an Extech 407740 sound level meter.”

Reviewer 1 Comment: “Line 181: “In this scenario, no artificial sound was added." I assume that "Silence" means that there was no playback of any kind.”

Response: You are correct. To address your comment, we replaced the following text:

Replaced:

“In this scenario, no artificial sound was added. In this condition, the broadband fan noise of the projectors was barely audible (LAeq = 41.1 dBA).”

Replaced with (on page 13, line 228):

“In this scenario, no sound was played back, and no artificial sound was added. The only audible sound was the broadband fan noise of the projectors, which was barely perceptible (LAeq = 31.1 dBA).”

Reviewer 1 Comment: “Line 195: What was the range of playback levels of the recordings in the "Original scenario"? Was the projector audible during playback?”

Response: Thank you for highlighting this critical detail, which we had omitted. We have now added the relevant information to the manuscript. The projectors were not audible because the playback levels ranged from 39 dBA to 50 dBA, so they were at least 8 dBA louder than the A-weighted equivalent continuous sound level measured at the center of the screen when the projectors were on (31.1 dBA).

Added (on page 14, line 248): “For the experiment, the authors selected a representative two-minute sound excerpt for each location. The playback levels ranged from 39 dBA to 50 dBA. They were calibrated, using an Extech 407740 sound level meter, specifically for these two-minute excerpts to avoid exceptional sound events that could influence the LAeq value.”

Deleted (on page 15, line 259): “For the experiment, the authors selected a representative two-minute sound excerpt for each location”.

Reviewer 1 Comment: “Line 200: typo "HVAC""

Replaced: “HVCA”

Replaced with (on page 13, line 217): “HVAC”

Reviewer 1 Comment: “Line 202: "the microphone was positioned 1.5 meters above the ground". Already mentioned in line 188.”

Deleted (on page 14, line 237):

“The ambisonics microphone was placed on a tripod at a height of 1.5 m from the ground.”

Reviewer 1 Comment: “Line 205: “representative two-minute sound excerpt for each location". It would be better if the playback levels had been calibrated for the exact two-minute excerpts that were used. The one minute of audio that was not used as sound stimuli could contain "exceptional" sound events that could influence the value of the LaEQ.”

Response: That is precisely what we did. However, we should have detailed that in the paper. Thank you for pointing that out.

Added (on page 14, line 248):

“For the experiment, the authors selected a representative two-minute sound excerpt for each location. The playback levels ranged from 39 dBA to 50 dBA and were calibrated specifically for these two-minute excerpts to avoid ‘exceptional’ sound events that could influence the LAeq value.”

Reviewer 1 Comment: “Lines 206-213: The provided description offers a lot a possibilities when it comes to masker design. Please describe the exact features of the masker used in this case.”

Response: Thank you for pointing this out. We acknowledge that the previous version of the manuscript may have caused some confusion. We focused more on describing the general industry-standard guidelines for spectrum contour and overall sound levels for conventional sound maskers rather than detailing the specific design choices we made in this study. In response to your feedback, we have revised the section to clarify the exact features of the masker used in the "Conventional" scenario. Specifically, we employed a neutral masking spectrum contour with a consistent 5 dB per octave decrease to simplify the design process and minimize potential issues such as rumble or hiss often associated with masking systems.

Added (on page 15, line 267):

“For this research, a standard “neutral” masking spectrum contour was used, with a consistent decrease of 5 dB per octave band. This approach was chosen to simplify the design process and avoid common complaints about rumble or hiss often associated with masking systems.”

Reviewer 1 Comment: “Line 210: the masker level is in the "range between 40 and 45 dBA". This is very similar to the noise level of the projectors. Can the masker be mistaken for projector noise or vice versa?”

Response: As mentioned above, the background level measured in the experiment room was 31.1 dBA, which ensured that the projector was not audible during playback. However, we understand that the previous explanation, like the earlier comment, focused too much on describing industry-standard suggestions rather than the specifics of our study. To address this, we have added a paragraph specifying that the overall sound level of the masker was set at 5 dBA above the ambient noise level recorded in the "Original" scenario rather than adhering strictly to the 40-45 dBA range suggested by industry standards. This adjustment helped ensure the masker sound was distinct from other noises, including the projector.

Added (on page 15, line 272):

“However, to ensure consistency across all scenarios and achieve a proper blend with the existing environmental noise, the masking sound level was set at 5 dBA above the ambient noise level recorded in the “Original” scenario. For instance, in a room where the overall sound level of the “Original” scenario measured 45 dBA, the sound masker level was adjusted to playback at 50 dBA. Consequently, the playback levels for this scenario ranged from 44 dBA to 55 dBA.”

Reviewer 1 Comment: “Lines 214-231: What are the playback levels of the sounds used in the "congruent" scenario? In addition, at this point, it is difficult to understand what these sounds are, based only on the general description provided in the text, and the supplementary material is not available at this time.”

Response: Thank you for pointing this out. We understand that the previous description lacked specific details about the playback levels and the sounds used in the "Congruent" scenario. To address your concern, we have added examples of the sounds used to improve clarity. Furthermore, like the "Conventional" scenario, we have clarified that the playback levels for the "Congruent" scenario were set at 5 dBA above the ambient noise level for each room, resulting in overall playback levels ranging from 44 dBA to 55 dBA. Also, we will attach the supplementary material in this submission so you can better understand the sounds that we used in the “Congruent” scenario.

Added (on page 16, line 289):

“–for example, Philip Glass’s composition “Glassworks” was used for the Sol LeWitt: A Wall Drawing Retrospective exhibition–”

Added (on page 16, line 291):

“–for example, birds sounds were chosen for the Birds of New York exhibition”

Added (on page 17, line 297):

“Similar to the “Conventional” scenario, the overall playback levels in this scenario were adjusted so that the masking level was 5 dbA above the ambient noise level in each room. Therefore, the playback levels of this scenario also ranged from 44 dBA to 55 dBA, ensuring consistency across different rooms and scenarios.”

Reviewer 1 Comment: “Line 258: "The background sound level __ me". This is not the same as presented in Figure 3, where it is correctly hinted that it is the sound that is distracting. The level is just one of the parameters that describe sound. In addition, why is it referred to as background sound in this question if the term "soundscape" is used in the remaining questions?”

Response: Thank you for bringing this to our attention. You are correct—this was an oversight. The figure was created during the experiment design phase, and we mistakenly left the original version in the manuscript. Initially, we planned to ask about the sound level, but we later agreed that this was not appropriate for the same reasons you mentioned. Instead, we changed the question to focus on whether the soundscape itself contributed to either enhancing focus or causing distraction for participants. Unfortunately, we neglected to update Figure 3 to reflect this change. We have now corrected the figure to align with the updated question, which focuses on the overall soundscape rather than just the sound level.

Updated Figure:

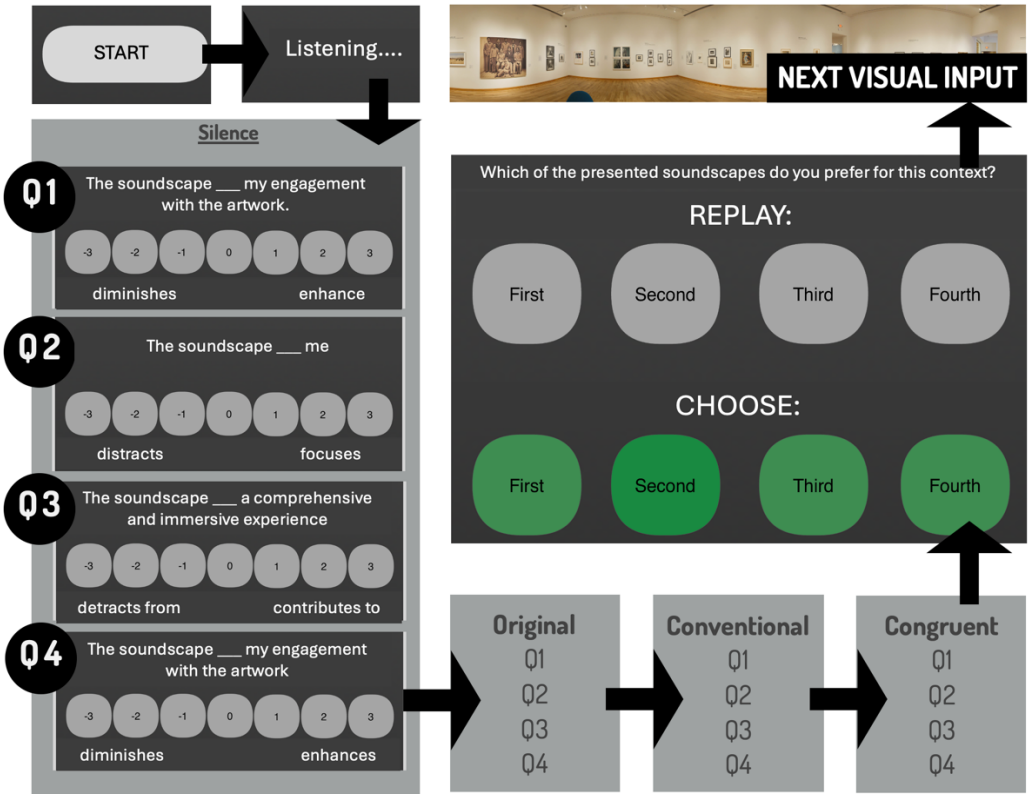

FIG. 3. Scheme showing the experiment’s steps and the GUI’s appearance.

Reviewer 1 Comment: “Lines 286-287: “However, the first room was designated as a test round to ensure precision and accuracy in their assessments." This explanation should be moved to lines 162-163 (see earlier comment).”

Response: Thank you for the suggestion. This issue has already been addressed in a previous response. The explanation about the first room being used as a test round has been moved accordingly.

Reviewer 1 Comment: “Line 295: significance level”

Replaced: “alpha level”

Replaced with (on page 21, line 366): “significance level”

Reviewer 1 Comment: “Line 328: please change the research question to match the GUI in Figure 3. Figure 5 - caption: please change "background level" to "background sound"”

Response:

Replaced (on page 18, line 328): “background sound level” **replaced with:** “soundscape”

Replaced (on page 23, line 402): “background sound level” **replaced with:** “soundscape”

Reviewer 1 Comment: “Lines 393-406: The data shown in part A are also the result of statistical analysis. This subsection should be incorporated into the previous one. Pairwise comparison is already shown above.”

Response: Thank you for your suggestion. You are correct. We have deleted the entire subsection “Pairwise Comparisons” and the table with the Kruskal-Wallis Test results. All the content has now been incorporated into the main Results section to streamline the presentation of the statistical analysis. Below are the specific changes made:

Added (on page 21, line 380): “The Dunn test results and significant differences are indicated at the top of each figure. **“**

Added (on page 22, line 386): “Statistical analysis using the Kruskal-Wallis test revealed significant differences among scenarios ($H = 15.001$, $p = 1.8e{-}3$). Pairwise comparisons showed significant differences between “Congruent” and “Silence” scenarios.”

Added (on page 23, line 404): “The Kruskal-Wallis test showed significant differences among the scenarios ($H = 35.99$, $p = 7.525e{-}8$). Pairwise comparison showed significant differences between “Congruent” and all the other scenarios. Additionally, significant differences were found between “Conventional” and “Silence”, and between “Original” and “Silence”.”

Added (on page 25, line 427): "The Kruskal-Wallis test indicated significant differences among the soundscapes (H = 45.467, p = 7.363e−10). Pairwise comparison showed significant differences between "Congruent" and all the other scenarios."

Added (on page 27, line 446): "The Kruskal-Wallis test showed significant differences among the scenarios (H = 34.816, p = 1.332e−7). Pairwise comparison showed significant differences between "Congruent" and all the other soundscapes."

Deleted (on page 28, line 469): TABLE II. Kruskal-Wallis Test results for each question.

TABLE II. Kruskal-Wallis Test results for each question.

Question	$H - value(3)$	$p - value$
01	15.001	$1.8e^{-3}$
02	35.99	$7.525e^{-8}$
03	45.467	$7.363e^{-10}$
04	34.816	$1.332e^{-7}$

Deleted (on page 28, line 471): "Subsequently, the Dunn tests were conducted to identify significant pairwise differences. The pairwise comparisons are shown at the top of the graphics in Figs. 4 to 7. The results show that the "Congruent" scenario is involved in a pair, with a significant difference from all three other scenarios for all questions. This suggests that there is a clear preference for the "Congruent" scenario compared to the different scenarios in the pair. Regarding the statistical significance between the groups, these significant differences between the "Congruent" and all the other soundscapes imply that the "Congruent" soundscape is perceived as significantly more effective in contributing to a comprehensive and immersive experience compared to the other soundscapes. The statistical analysis revealed significant differences in participants' responses across the analyzed soundscape scenarios. The Kruskal-Wallis tests for all questions were significant, indicating statistically significant differences in participants' perceptions of the different soundscape scenarios. The H and p values are presented in Table II."

Reviewer 1 Comment: "Lines 529-532: I am not sure that the use of the questionnaire defined in ISO 12913-2 would be an improvement. I believe you developed a better investigation tool for this specific purpose."

Response: Thank you for your feedback. We agree with your assessment and have decided to delete the reference to ISO 12913-2. Based on your comment and input from Reviewer 2, we have refocused the future research suggestions to align with topics that are relevant to museum environments, such as museum fatigue.

Deleted (on page 37, line 644):

“Lastly, future investigations in this field should delve into traditional data collection methods outlined in ISO 12913-2, both in laboratory settings and onsite within museum environments. For instance, employing the questionnaire method suggested in ISO 12913-2 (ISO 12913-2, 2018) can yield valuable insights into how individuals perceive acoustic environments, encompassing aspects such as sound source identification, affective quality, and assessment of the surrounding sound environment. Future research can apply the same program developed for administering the questionnaire in this study but incorporating questions aligned with the ISO 12913-2 (ISO 12913-2, 2018). Alternatively, conducting in situ experiments using the soundwalk method offers another avenue for gathering comprehensive data on the acoustic environment or a proposed environment within museums.”

Added (on page 36, line 619)**:**

“Future studies should also explore how soundscapes might alleviate museum fatigue, especially its physical aspects, which result from prolonged walking and standing while engaging with exhibits. Investigating how soundscapes could reduce cognitive overload and sustain visitor attention throughout a visit is crucial for a more comprehensive understanding of how they interact with the dynamic nature of museum environments. Additionally, future work should assess whether the findings from this research hold consistently from the beginning to the end of a museum visit, as cognitive and physical fatigue may influence how soundscapes are perceived at different stages. Such insights could prove invaluable for museum curators, helping them to strategically place certain types of audio content to enhance both engagement and visitor well-being.”

Reviewer 2

“The manuscript reports an experiment on the virtual recreation of museum spaces with evaluations of four soundscape scenarios and assessments related to a sample of assessors. The research question and method are precise, and the results are presented very clearly. The discussions are limited, also due to the relatively scarce literature on the museum theme. The strengths of the article are also its limitations: the context of the article seems to be more related to laboratory technology than to a generalizable problem within museums.”

Reviewer 2 Comment: “In fact, the article lacks some references to the literature on comfort within museums, the fact that visitors can move around during the visit, and the rather well-developed concept in the literature of 'museum fatigue'. The reviewer suggests, therefore, that the authors conduct a careful survey on the indicated themes and contextualize the discussion of the results.”

Response: Thank you for your valuable feedback. In response to your suggestion, we have carefully reviewed the literature on museum comfort and the concept of "museum fatigue." We have expanded the Introduction to better contextualize the importance of these aspects and have referenced key studies that highlight how acoustic environments influence cognitive and physical fatigue in museums. We have also enhanced our “Discussion” section to link the

impact of different soundscapes with visitor engagement and museum fatigue, explicitly addressing how congruent soundscapes might alleviate cognitive load, reducing fatigue over time. We hope these revisions improve the clarity and scope of our findings in relation to visitor comfort and museum fatigue.

Added in the Introduction section (on page 3, line 31)

“In contrast, despite the fundamental role of the acoustic environment in visitor comfort, it has often been overlooked (D’Orazio et al., 2020). Visitor comfort is an essential aspect of museum design, mainly as museums are spaces where individuals usually spend extended periods moving between exhibitions. Prolonged standing, walking, and mental engagement can lead to a phenomenon known as “museum fatigue,” where physical and cognitive exhaustion can diminish visitor engagement and retention (Bitgood, 2009, 2010; Kim et al., 2020), and the environment can exacerbate or alleviate this fatigue..”

Added in the Discussion section (on page 31, line 511):

“Second, the analysis demonstrates that the “Congruent” soundscape significantly enhances focus compared to other soundscapes. The “Original” and “Silence” soundscapes show neutrality, with neither strong focusing nor distracting effects, though they differ significantly from each other. The “Conventional” soundscape shows a negative impact on focus, with a tendency toward distracting effects. This suggests that conventional sound maskers could potentially detract from the visitor’s attention, **(Added: while aligning the auditory environment with the visual content reduces potential distractions. This reduction in distractions can potentially reduce museum fatigue, as prior research emphasized the importance of minimizing disturbances, such as background noise, to prevent museum fatigue and help visitors sustain focus for more extended periods (Bitgood, 2010).”**

Added in the Discussion section (on page 32, line 521):

“Furthermore, the analysis demonstrates that the “Congruent” soundscape significantly enhances the immersive experience compared to other soundscapes. The “Conventional” and “Original” soundscapes show neutrality, with neither strong contributing nor detracting effects, though they differ significantly from the “Congruent” soundscapes. “Silence” shows a varied perception, slightly tending towards positive perceptions, but does not significantly enhance the immersive experience. **(Added: “This highlights the importance of contextually congruent soundscapes in fostering a more immersive museum environment, which may also delay cognitive exhaustion, a critical factor in museum fatigue. As Bitgood (2009) points out, engaging exhibits play a crucial role in minimizing fatigue and boredom.”)**

Reviewer 2 Comment: “The reviewer has some doubts regarding the interpretation of the use of sound maskers, at least in the proposed utilization: sound maskers have been proposed to increase privacy within open-plan offices by decreasing the intelligibility of speech from adjacent workstations.”

Response: Thank you for your thoughtful feedback. To clarify and address this point, we have added a more detailed explanation in the “Introduction” section that distinguishes between the conventional use of sound maskers in office environments and their proposed alternative function in museums. We elaborate on how sound maskers, when applied in a museum setting, serve a different purpose by aligning auditory cues with the exhibition’s thematic and visual elements. Instead of focusing on privacy, as in offices, the aim here is to enhance visitor engagement and immersion by using *thematic sound maskers* that contribute to a cohesive multisensory experience. In the revision, I also provide a theoretical framework for this alternative use of sound maskers, supported by relevant literature. I emphasize that this research proposes a shift from traditional sound masking techniques (low-level broadband noise for privacy) to thematic sound maskers that reduce background distractions while enriching the overall visitor experience. This revision addresses the concerns about interpretation by highlighting that the study aims to explore how sound maskers can be adapted and optimized for immersive environments like museums.

Added in the Introduction section (on page 5, lines from 78 to 113):

“In addition to enhancing comfort and engagement, sounds can be added as a component of the overall soundscape to mask distractions, particularly in settings where focus is critical. In open-plan offices, for instance, sound maskers are commonly employed to increase privacy by reducing the intelligibility of speech between adjacent workstations (DeLoach et al., 2015; Haapakangas and Hongisto, 2008; Lenne et al., 2020). This well-established technique involves intentionally introducing electronic background sound, typically low-level broadband noise, to mask or reduce the perception of unwanted sounds. However, the use of sound maskers in museums—where the visitor’s immersion is essential and privacy is less of a concern—remains unexplored. This study proposes that, in a museum context, sound maskers can serve a different purpose. They can be tuned to align auditory cues with visual and thematic elements of exhibitions to create a more cohesive and immersive multisensory experience. This idea is backed up by previous research, which has shown that human perception is heavily influenced by context, further emphasizing the potential for soundscapes aligned with the exhibition context to enhance the visitor experience (Acun and Yilmazer, 2018; Hong and Jeon, 2015; Jambrošić et al., 2013), and the importance of integrating visual and aural cues to create multisensory experiences in museums (Jelinčić et al., 2022).

Furthermore, given the acoustic challenges and the increasing use of auditory elements in museums, further research is needed to understand visitors’ subjective perceptions of indoor soundscapes. While previous studies have explored some aspects of sound in museums, there remains a limited understanding of how different soundscapes impact the visitor experience. This study seeks to address these gaps by analyzing soundscape perceptions in museums, focusing on four key factors: context, background distraction, engagement, and immersiveness. We hypothesize that soundscapes composed of sounds in congruency with the exhibition context, used as thematic maskers, are more effective than traditional sound maskers (e.g., low-level broadband noise) in:

- Reducing background noise distractions
- Enhancing visitors’ experience and engagement with the artwork

- Creating a comprehensive and immersive museum experience

In this context, thematic sound maskers not only mask unwanted background noise but also reinforce the exhibition's content through congruent auditory cues, creating a more cohesive and immersive experience. Thus, this study explores how sound maskers, beyond their traditional use for privacy, can shape the museum's acoustic environment to deepen visitor engagement and immersion."

Reviewer 2 Comment: "The fact that 'original' and 'conventional' scenarios yield comparable results might indicate that the original soundscape was already sufficiently 'unfocused,' which is predictable given the poor absorption and absence of acoustic screens in museum spaces. Conversely, the use of sound maskers would have been interesting in evaluating the 'privacy' of each visitor's experience. The overall judgment between 'original' and 'conventional' seems to indicate that the visitor tends not to prefer the 'conventional' scenario, but the reason for this is not sufficiently explained. Perhaps the museum experience does not require privacy?"

Response: Thank you for your insightful comments regarding the comparison between the "Original" and "Conventional" soundscapes. Your point about the original soundscape possibly being "unfocused" due to common acoustic challenges found in museum spaces makes total sense. Additionally, your observation about the role of privacy has prompted me to clarify further why the "Conventional" soundscape was less preferred and why privacy may not be as relevant in the museum context.

Added in the Discussion section (on page 33, line 558):

"Finally, the results for "Original" soundscapes might indicate that the on-site recorded soundscape was already sufficiently unfocused, distracting, and diminished engagement. This is predictable and aligns with previous research identifying the poor absorption and absence of acoustic screens in museum environments, which result in common acoustic challenges such as excessive reverberation, poor speech clarity, limited privacy, and increased background noise—all of which can detract from the overall visitor experience (Barbo and Brandao, 2019; Carvalho et al., 2013; Monica et al., 2022; Paxton et al., 2018; Pon et al., 2016; Su Gul, 2021). Additionally, the lower preference for the "Conventional" soundscape may suggest that privacy, typically prioritized in environments like offices, is not a primary concern for museum visitors. Instead, visitors may prioritize immersive and engaging auditory experiences that align with the exhibitions, which the "Conventional" soundscape may not provide as effectively."

Reviewer 2 Comment: "The most important limitation, according to the reviewer, is the fact that the experiment does not consider that the visitor is moving within a space for a certain period of time, which leads to physical and cognitive fatigue. In other word, the scenario is not fixed but it changes dynamically. What role can a certain soundscape play compared to another in limiting the so-called museum fatigue and other factors proposed by scholars? Do the considerations reported by the authors hold in the same way at the beginning and at the end of a museum visit? This aspect could be very helpful for museum curators, indicating where to focus a certain type of audio content. Even though the experiment did not take this aspect into account, it should be mentioned and explained in the limitations of the work."

Response: Thank you for your insightful comments regarding the dynamic nature of museum visits and the role of visitor movement over time. We acknowledge that the controlled laboratory environment in our study did not fully capture the complexities of real-world museum fatigue, which includes both cognitive and physical aspects. In response, we have added the following to the "Future Research and Limitations" section to reflect these considerations and propose future directions:

Added in the Future Research and Limitations section (on page 36, line 608):

“Firstly, while the controlled laboratory setting allowed for precise manipulation and testing of various soundscapes, it is important to recognize the limitations of this approach in fully replicating the dynamic and complex environments of actual museums. Our experiment did not account for visitor movement over time, which plays a key role in museum fatigue—both cognitive and physical. This controlled context may influence the generalizability of our findings. To address these limitations, future research should conduct in-situ experiments in real-world museum settings, where factors such as architecture, environmental variability, and visitor interactions can be more accurately assessed.”

Added in the Future Research and Limitations section (on page 36, line 617):

“Future studies should also explore how soundscapes might alleviate museum fatigue, especially its physical aspects, which result from prolonged walking and standing while engaging with exhibits. Investigating how soundscapes could reduce cognitive overload and sustain visitor attention throughout a visit is crucial for a more comprehensive understanding of how they interact with the dynamic nature of museum environments. Additionally, future work should assess whether the findings from this research hold consistently from the beginning to the end of a museum visit, as cognitive and physical fatigue may influence how soundscapes are perceived at different stages. Such insights could prove invaluable for museum curators, helping them to strategically place certain types of audio content to enhance both engagement and visitor well-being.”

Part 2: Due to the significant changes made to the “Introduction” section in response to Reviewer 2’s comments, we are including the entire revised section in this report. The modifications primarily address the reviewer's suggestions regarding references to literature on museum comfort, the concept of museum fatigue, and the role of sound maskers in museums proposed by this research. These updates enhance the contextual framing of the study, ensuring a more thorough discussion of the relevant themes as requested. By presenting the complete introduction, we aim to demonstrate clearly how these critical revisions have been incorporated. Following is the final version of the Introduction Section with the added paragraphs:

I. INTRODUCTION

Traditional museums are often perceived as silent spaces where artworks are the primary focus. However, contemporary museums increasingly incorporate multisensory experiences, including auditory elements (Wang, 2020). In this shifting context, it is essential to discuss “wanted” sounds and the subjective effects of indoor soundscapes on the museum experience. Despite this, there is a limited exploration of visitors’ subjective perceptions compared to other public venues like concert halls. This gap is particularly notable in museums that use augmented sounds, such as noise masking systems or artificial soundscapes.

Research on indoor environmental quality in museums has predominantly focused on lighting (del Hoyo-Meléndez et al., 2011; Kaya and Afacan, 2018; Kim and Seo, 2012; Wahab and Zuhardi, 2013), energy efficiency (Pavlogeorgatos, 2003; Sharif-Askari and Abu-Hijleh, 2018), and air quality (Proietti et al., 2015). **ADDED (on page 3, line 31) {**In contrast, despite the fundamental role of the acoustic environment in visitor comfort, it has often been overlooked (D’Orazio et al., 2020). Visitor comfort is an essential aspect of museum design, mainly as museums are spaces where individuals usually spend extended periods moving between exhibitions. Prolonged standing, walking, and mental engagement can lead to a phenomenon known as “museum fatigue,” where physical and cognitive exhaustion can diminish visitor engagement and retention (Bitgood, 2009; Kim et al., 2020), and the environment can exacerbate or alleviate this fatigue.**}**

Early studies acknowledged the importance of acoustics (D’Antonio et al., 2003; Reber, 2003), with subsequent research highlighting the unique challenges presented by museum environments, primarily focusing on objective parameters (Barbo and Brandão, 2019; Carvalho et al., 2013; D’Orazio et al., 2020; Mónica et al., 2022; Paxton et al., 2018; Pon et al., 2016; Sür Gül, 2021). These studies have consistently highlighted several challenges, including high reverberation, echoes, poor speech intelligibility, limited privacy, and increased background noise. Such acoustic challenges are predominantly associated with specific architectural features, such as open-plan exhibition halls, coupled rooms, large volumes, and highly reflective surfaces (Barbo and Brandão, 2019; Carvalho et al., 2013).

ADDED (on page 4, line 48) {While such objective parameters provide valuable insights into the physical properties of museum acoustics, they do not fully capture the nuances of visitors’ subjective experiences. However, studies focusing on visitors’ perceptions and subjective experiences remain sparse. For instance, Vi et al. (2017) conducted subjective evaluations to explore the role of auditory senses in multisensory experiences, and Yang and Kang (2020) examined the effects of a sound source on subjective experiences in sequential spaces in the Tate Modern Museum in London. Carvalho et al. (2014) applied questionnaires and interviews for subjective responses of visitors combined with objective parameter measurements to find the IAQM (Index of Acoustic Quality in Museums).**}**

ADDED (on page 4, line 57) {Furthermore, recent literature has begun to emphasize the potential of soundscapes to enrich museum experiences. For example, Acun et al. (2018) demonstrated how sound, under the right conditions, can enhance the museum environment, making even the noisiest places be perceived as expected if the soundscape aligns with the overall context. Building upon this premise, Orhan and Yilmazer (2021) illustrated how the

strategic integration of sound can enhance visitor engagement with objects and deepen immersion within historical exhibitions. These findings underscore the impact of sound as a design element in museum environments, suggesting that a well-designed soundscape could contribute not only to visitor comfort but also to emotional and cognitive engagement.**}**

The ISO 12913-1 standard defines a soundscape as the "acoustic environment as perceived and understood by a person or people in context" (ISO 12913-1, 2014). While traditional soundscape studies have focused on outdoor environments, there is a growing recognition of the importance of studying indoor soundscapes for designing healthier and more comfortable indoor spaces (Visentin et al., 2023). Recent literature discusses soundscape preferences in various indoor environments, including residential buildings (Torresin et al., 2019), educational settings (Visentin et al., 2023), and open-plan offices (Jo and Jeon, 2022). **ADDED (on page 5, line 72) {**Studies have shown that soundscapes featuring natural sounds and music are often perceived as more pleasant and can creating an ideal acoustic environment in indoor settings (Mohamed and Dokmeci Yorukoglu, 2020; Torresin et al., 2021, 2020, 2022; Visentin et al., 2023). These findings are particularly relevant in museum contexts, where the auditory environment can affect how visitors experience and interact with exhibits.**}**

ADDED (on page 5, line 78) {In addition to enhancing comfort and engagement, sounds can be added as a component of the overall soundscape to mask distractions, particularly in settings where concentration is critical. In open-plan offices, for instance, sound maskers are commonly employed to increase privacy by reducing the intelligibility of speech between adjacent workstations (DeLoach et al., 2015; Haapakangas and Hongisto, 2008; Lenne et al., 2020). This well-established technique involves intentionally introducing electronic background sound, typically low-level broadband noise, to mask or reduce the perception of unwanted sounds. However, the use of sound maskers in museums—where the acoustic environment is more complex and privacy is less of a concern—remains unexplored. In a museum context, sound maskers could serve a different purpose. They could align auditory cues with visual and thematic elements of exhibitions to create a more cohesive and immersive multisensory experience. This idea is backed up by previous research, which has shown that human perception is heavily influenced by context, further emphasizing the potential for soundscapes aligned with the exhibition context to enhance the visitor experience (Acun and Yilmazer, 2018; Hong and Jeon, 2015; Jambrošić et al., 2013), and the integration of visual and aural cues to create multisensory experiences in museums (Jelinčić et al., 2022).**}**

ADDED (on page 6, line 95) {Furthermore, given the acoustic challenges and the increasing use of auditory elements in museums, further research is needed to understand visitors' subjective perceptions of indoor soundscapes. While previous studies have explored some aspects of sound in museums, there remains a limited understanding of how different soundscapes impact the visitor experience. This study seeks to address that gap by analyzing soundscape perceptions in museums, focusing on four key factors: context, background distraction, engagement, and immersiveness. We hypothesize that soundscapes composed of sounds in congruency with the exhibition context, used as thematic maskers, are more effective than traditional sound maskers (e.g., low-level broadband noise) in:

- Reducing background noise distractions
- Enhancing visitors' experience and engagement with the artwork
- Creating a comprehensive and immersive museum experience

In this context, thematic sound maskers not only mask unwanted background noise but also reinforce the exhibition's content through congruent auditory cues, creating a more cohesive and immersive experience. Thus, this study explores how sound maskers, beyond their traditional use for privacy, can shape the museum's acoustic environment to deepen visitor engagement and immersion.}