ACOUSTICS IN THE BUILT ENVIRONMENT: A CHALLENGE FOR IMPROVING THE QUALITY OF LIFE

EDITED BY: Arianna Astolfi, Giuseppina Emma Puglisi, Nicola Prodi, Jian Kang, Louena Shtrepi and Chiara Visentin
PUBLISHED IN: Frontiers in Built Environment

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-273-6
DOI 10.3389/978-2-88976-273-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ACOUSTICS IN THE BUILT ENVIRONMENT: A CHALLENGE FOR IMPROVING THE QUALITY OF LIFE

Topic Editors:
Arianna Astolfi, Politecnico di Torino, Italy
Giuseppina Emma Puglisi, Politecnico di Torino, Italy
Nicola Prodi, University of Ferrara, Italy
Jian Kang, University College London, United Kingdom
Louena Shtrepi, Politecnico di Torino, Italy
Chiara Visentin, University of Ferrara, Italy

Citation: Astolfi, A., Puglisi, G. E., Prodi, N., Kang, J., Shtrepi, L., Visentin, C., eds. (2022). Acoustics in the Built Environment: A Challenge for Improving the Quality of Life. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-273-6

Table of Contents

frontiers | Frontiers in Built Environment

EDITORIAL
published: 05 April 2022
doi: 10.3389/fbuil.2022.887324

Editorial: Acoustics in the Built Environment: A Challenge for Improving the Quality of Life

Arianna Astolfi[1]*, *Giuseppina Emma Puglisi*[1], *Nicola Prodi*[2], *Jian Kang*[3], *Louena Shtrepi*[1] and *Chiara Visentin*[2]

[1]Department of Energy, Politecnico di Torino, Torino, Italy, [2]Department of Engineering, University of Ferrara, Ferrara, Italy, [3]UCL Institute for Environmental Design and Engineering, The Bartlett, University College London, London, United Kingdom

Keywords: soundscape, well-being, acoustic quality, COVID 19, sound environment

Editorial on the Research Topic

Acoustics in the Built Environment: A Challenge for Improving the Quality of Life

The acoustics of the environment in which we live influences our health, comfort, performance, and well-being. In this Research Topic, it is showcased that quality of life can be changed and improved by optimizing the acoustics of the built environment. This is done by addressing the complexity of the interactions between the occupants and the sonic environment. A variety of indoor and outdoor settings, in which either communication or perception are targeted, have been taken into consideration by the contributions in this special issue. In particular, the collected papers focused on four critical aspects: effects of noise in learning environments, communication in noise and reverberation, soundscape optimization for outdoor and indoor applications, and development of perception-based criteria in acoustical design.

OPEN ACCESS

Edited by:
Hasim Altan,
Arkin University of Creative Arts and Design (ARUCAD), Cyprus

Reviewed by:
Michael Vorländer,
RWTH Aachen University, Germany
Woon-Seng Gan,
Nanyang Technological University, Singapore

***Correspondence:**
Arianna Astolfi
arianna.astolfi@polito.it

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *01 March 2022*
Accepted: *14 March 2022*
Published: *05 April 2022*

Citation:
Astolfi A, Puglisi GE, Prodi N, Kang J, Shtrepi L and Visentin C (2022) Editorial: Acoustics in the Built Environment: A Challenge for Improving the Quality of Life. Front. Built Environ. 8:887324. doi: 10.3389/fbuil.2022.887324

EFFECTS OF SOUNDSCAPE IN LEARNING ENVIRONMENTS

The sonic environment of learning spaces impacts students' perception and learning. Young children are especially vulnerable to the effects of background noise and reverberation due to their still immature cognitive and linguistic skills. Furthermore, task performance and effort of older students in complex academic tasks might be negatively affected by unfavorable classroom acoustics.

Loh et al. investigated the sonic environment of classrooms and playrooms with a campaign of measurements in both occupied and unoccupied conditions. Importantly, the authors point out the necessity to use appropriate transducers in measurements with children (i.e., children head-and-torso simulator) and to consider complementing traditional acoustic parameters with psychoacoustic ones. The link between objective measurements and children's perceptual evaluation was explored by Persson Waye and Karlberg for preschool-age children. It was found that, despite small changes in background noise levels and reverberation, the improvement in the acoustic conditions was clearly perceived by the children. In particular, favorable acoustic conditions allowed for a significant reduction in the perception of sounds along with a reduction in children's reaction to them. Brill and Wang surveyed the acoustic conditions of occupied classrooms in primary and secondary schools and analyzed their relationship with standardized achievement test results in the math and reading areas. Daily non-speech levels were found to be negatively correlated with math test scores whereas reading achievement was not correlated with any of the acoustic parameters included in the survey.

Regarding older students, Braat-Eggen et al. investigated how changes in the sound environment (i.e., reverberation time or number of talkers) affect complex cognitive tasks. Changes in the sound environment negatively impacted only the performance in a logical reasoning task whereas a significant effect of the sound scenarios was found on self-estimated performance and perceived disturbance for reading comprehension with text memory and mental arithmetic. It is evident that the sonic environment affects not only task performance but also cognitive processing, as estimated by self-reports of the listeners. Visentin and Prodi used self-reports for measuring changes in school-aged children's perceived effort while working in a noisy classroom. It was found that self-ratings were sensitive to the spectro-temporal characteristics of the background noise, but only for a speech perception task, emphasizing the need for further research on a topic which to date is still under-investigated.

EFFECTS OF NOISE AND REVERBERATION ON THE SPEECH COMMUNICATION PROCESS

The speech communication process accounts for premises related to the talker, to the listener, and to the environmental path that links them. As far as the talker perspective is concerned, Sierra-Polanco et al. investigated the effects of aural feedback, reverberation, and noise on speech production. They found that an increase in gain causes a decrease in the voice sound pressure level and a consequent increase in self-reported vocal comfort. Substantial variations in speech level, instead, were not found for reverberation times that varied from 0.07 to 1.90 s at mid frequencies. Considering the listener's perspective, Warzybok et al. investigate the interplay of bottom-up and top-down resources in noise and reverberation. They found that under poor room acoustic conditions, being familiar with the speech material allows for higher speech intelligibility. Thus, the work highlights the influence of higher-level lexical-semantic cues in speech recognition and underscores the limits of conventional tools for assessment even in the common scenario of everyday communication. Taken together, these contributions stress the need to blend the acoustical needs of the listeners with those of the talker to make a step-forward in design practice.

THE ROLE OF SOUNDSCAPE ON THE PERCEPTION OF PUBLIC ENVIRONMENTS

Soundscape, defined by ISO as an acoustic environment as perceived or experienced and/or understood by a person or people, in context, has attracted much attention in both research and practice. Qi et al. conducted an on-site questionnaire investigation at two artificial lakes in Xi'an, China, to explore which blue space characteristics would contribute to a better soundscape and visiting experience. It was found that the eight Perceived Sensory Dimensions of artificial lake spaces, except for social, were positively correlated with soundscape satisfaction, overall satisfaction, soundscape restorativeness, and overall restorativeness. Indoor soundscaping is also important for soundscape studies. Steffens et al. carried out field studies in 12 restaurants in Berlin, investigating whether sound level, reverberation time, and soundscape pleasantness can predict factors associated with overall restaurant quality. It was found that both $L_{A,eq,15}$ and T20 had a significant influence on soundscape pleasantness and eventfulness, and LA,eq, 15 as well as soundscape pleasantness were significant predictors of overall restaurant quality. It is also noted that technologies to control sound field are important too for soundscape creation, where a challenge is to achieve adequate acoustics while maintaining the aesthetics of the space. Cucharero et al. developed a biofiber-based acoustic coating as a feasible solution to improve acoustic environments while preserving the aesthetics of spaces.

SOUND PERCEPTION AS A DRIVER FOR INNOVATIVE ENVIRONMENTAL DESIGN TO FACE THE LATEST SOCIETAL CHALLENGES

The technological evolution of different aspects of our life has posed several challenges to our perception of the everyday environment. Soeta and Onogawa highlighted that although the sound produced by air conditioners has been limited to a comparatively low level, some people may still perceive discomfort. They built a predictive model for the subjective response to the air conditioner sound's quality and levels. A deep knowledge on these types of models could help to assess our perception even in exceptional working conditions. Puglisi et al. showed that perceived noise during the COVID-19 pandemic remote working had a significant effect on working activity and performance. Among the noise sources investigated, 25% of a total of 1,934 of workers recognized the noise generated by people (e.g., talking, moving, calling, listening to music) as the main source of disturbance. A perceptual test may be used to further develop our everyday environment virtualization. Llorca-Bofí and Vorländer highlighted the gap between building simulation and physically based material models. They provided targeted modeling strategies for architects, in both indoor and outdoor demonstrations, for auditory-visual research.

To sum up, the 13 articles in the special issue demonstrate that a novel approach to the acoustics of built environments can be pursued. It is based on the full integration of physical, perceptual, and cognitive features that are elicited in listeners and talkers when interacting with the sonic environment. This framework is promising and can be exploited to improve our knowledge on the impact of sound on the users' activities. In addition, thanks to such novel approaches, the efficacy of the available control strategies can be enhanced, and new ones can be developed too. Together, knowledge and technologies will foster a more enjoyable user experience and improve their quality of life.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Reviewer WG declared a past collaboration with one of the authors JK at time of review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 03 May 2021
doi: 10.3389/fbuil.2021.664442

Check for updates

The Effects of Artificial Lake Space on Satisfaction and Restorativeness of the Overall Environment and Soundscape in Urban Parks

Ying Qi, Xingyue Fang, Tian Gao and Ling Qiu**

Department of Landscape Architecture, College of Landscape Architecture and Arts, Northwest A&F University, Yangling, China

OPEN ACCESS

Edited by:
Jian Kang, University College London, United Kingdom

Reviewed by:
Fangfang Liu, Harbin Institute of Technology, China
Jooyoung Hong, Chungnam National University, South Korea

***Correspondence:**
Tian Gao
tian.gao@nwsuaf.edu.cn
orcid.org/0000-0002-0375-4073
Ling Qiu
qiu.ling@nwsuaf.edu.cn
orcid.org/0000-0002-7021-7235

Specialty section:
This article was submitted to Urban Science, a section of the journal Frontiers in Built Environment

Received: *05 February 2021*
Accepted: *06 April 2021*
Published: *03 May 2021*

Citation:
Qi Y, Fang X, Gao T and Qiu L (2021) The Effects of Artificial Lake Space on Satisfaction and Restorativeness of the Overall Environment and Soundscape in Urban Parks. Front. Built Environ. 7:664442. doi: 10.3389/fbuil.2021.664442

Several studies have proven that soundscape in blue space is conducive to human health and well-being, but few studies have explored which blue space characteristics would contribute to a better soundscape and visiting experience. Therefore, an on-site questionnaire investigation was conducted at two artificial lakes in Xi'an, China. The eight Perceived Sensory Dimensions (PSDs) as a landscape assessment tool were applied to identify the characteristics of artificial lake space in urban parks. The results showed that (1) In artificial lake space, overall environment and soundscape reached a very satisfactory level in general, while the respondents' perceived level of overall restorativeness and soundscape restorativeness as just medium, which indicated that the quality of artificial lake space needs to be improved. (2) According to people's perceptions, artificial lake spaces had the most obvious characteristics of *prospect*, *social* and *space*; *serene* and *nature* were medium; *refuge*, *rich in species,* and *culture* were the least. (3) The eight PSDs of artificial lake space, except for *social*, were positively correlated with soundscape satisfaction, overall satisfaction, soundscape restorativeness, and overall restorativeness. Moreover, among them, *serene* was the most significant characteristic in artificial lake space. These findings could be instructive to the design of urban parks with artificial lakes for improving users' visiting satisfaction and restorativeness.

Keywords: artificial lake space, perceived restorativeness, visiting satisfaction, perceived sensory dimensions (PSDs), soundscape

INTRODUCTION

Blue space, as all visible surface waters in space, includes the marine environment and fresh water such as rivers, lakes, seas, and fountains (Völker and Kistemann, 2011; Foley and Kistemann, 2015; Grellier et al., 2017). There are significant differences in physical characteristics, ecological value, and experience between them. Although people living in inland areas have limited access to the ocean (Kummu et al., 2011), they have more access to freshwater space, which can be divided into natural water space and artificial water space, such as artificial lake space in parks. For inland areas with few natural freshwater resources, the urban blue space is mainly dominated by artificial lake space with certain greenery, which is of great importance for urban inhabitants (Jarvis et al., 2020).

Therefore, it is essential to focus on the construction of blue spaces, like artificial lake spaces, when planning and designing urban parks.

With the ever growing urbanization process, the psychological pressures of urban residents has been increasing sharply, which might lead to several types of physical and mental diseases (Glaser et al., 2000; Wilkins, 2008; Nilsson et al., 2017). Previous studies have claimed the importance of blue space, which is reflected not only in the ecological environment but also in people's physical (Gidlow et al., 2016; Memari et al., 2017) and psychological health (Pasanen et al., 2019; Pearson et al., 2019). Indeed, people that visit green and blue spaces feel happier than those in gray spaces (MacKerron and Mourato, 2013). Nutsford et al. (2016) showed that higher levels of blue space visibility were related to lower psychological stress (Nutsford et al., 2016). Moreover, artificial lakes, as an important type of blue space in urban parks, is closely related to people's life and health interests, but few studies have focused on lakes in urban parks in terms of its well-being aspect. Therefore, it is meaningful to study the artificial lake space, especially its health benefits in the overall environment.

Serenity and calm have been accepted as apparent features in blue space, especially in the open artificial lake space. Hence, as an essential aspect of building peace and serenity, the soundscape in artificial lake space is worth studying and has been proven to be positively related to people's visiting experience. Ma et al. (2021) found that an open artificial lake space provided the maximum soundscape satisfaction and pleasantness in their sample park (Ma et al., 2021). Fan et al. (2021) found that people in the proximity of an open artificial lake space had a relatively high soundscape satisfaction. Furthermore, the soundscape in the artificial lake space affects restorativeness, which contributes mostly to health and well-being (Fan et al., 2021). Patón et al. (2020) investigated 16 kinds of water relating sounds and found that natural water sounds like those from a small stream or pond, can produce relaxation in contrast to obvious artificial water sounds (Patón et al., 2020). As a main kind of blue space in urban parks, artificial lakes possess the potential for satisfaction and health benefits in terms of soundscape. Therefore, in the construction of urban blue spaces, it is necessary to strengthen the soundscape of artificial lake space, to explore the evaluated levels and ways to improve it.

However, mechanisms linked to restoration and the natural environment has often been overlooked (Dzhambov et al., 2018). Despite a few studies that have explored the mechanism of open artificial lake; like stress reduction, thermal comfort improvement, and the promotion of non-water sports activities (Steeneveld et al., 2014; Grassini et al., 2019; Vert et al., 2019), the specific level of satisfaction and restoration (e.g., medium, or better), and the characteristics of urban artificial lake space that can mostly promote human health remain absent. To guide the future planning and design of health-based urban blue spaces, like artificial lake spaces, the specific perceived characteristics related to well-being deserve further exploration. Therefore, a measurement system of perception attributes of an artificial lake space needs to be introduced. Eight perceived sensory dimensions (PSDs) were developed by researchers at the Swedish University of Agricultural Sciences as a classification system that identified the most representative characteristics of nature. The classification system contained eight dimensions, as follows: "*serene*" (e.g., salient and calm), "*natur*e" (e.g., wild and untouched), "*rich in species*" (e.g., many animals and plants), "*space*" (e.g., spacious and free), "*prospect*" (e.g., flat and well-cut lawns with scattered trees), "*refuge*" (e.g., an enclosed and safe place), "*social*" (e.g., entertainment and exhibitions), and "*culture*" (e.g., decorated with fountains and ornamental plants) (Grahn and Stigsdotter, 2010). The eight PSDs have been applied in many studies related to the relationship of urban park characteristics and stress recovery (Grahn and Stigsdotter, 2003; Peschardt and Stigsdotter, 2013; Memari et al., 2017). Considering its wide application in landscape studies, PSDs were used to measure the perceptions of artificial lake space in this study.

At present, there are many indicators to measure spatial restorativeness. Two prevailing theories of restorativeness in the natural environment are widely applied: attentional restoration theory (ART) (Kaplan and Kaplan, 1989; Kaplan, 1995) and stress recovery theory (SRT) (Ulrich et al., 1991). The perceived restorativeness scales (PRS) have been developed according to ART to evaluate people's recovery levels in different environments (Hartig et al., 1997; Fátima et al., 2017), which contains four main components with 16 items in total: Fascination (the attraction of involuntary, effortless attention); Being-Away (a shift away from the present daily routine to a different environment); Compatibility (fit to an individual's planned behavior and environmental demands); and Extent (it can be experienced through immersion in intellectual activities and in physical environments.) (Hartig et al., 1996, 1997). Then Payne (2013) developed an evaluation system for restorative soundscape, incorporating psychological and situational factors based on the PRS and attentional restoration theory (ART)—Perceived Restorativeness Soundscape Scale (PRSS), which can help us assess the restorative aspect of soundscapes (Payne, 2013). PRSS, with a total of 19 items, contains six parts, as follows: Fascination, Being-Away-To, Being-Away-From, Compatibility, Extent (Coherence), and Extent (Scope). At present, the PRSS system is gradually applied in soundscape research (Zhang et al., 2017; Li and Kang, 2019; Zhao et al., 2020). In particular, soundscape as an indispensable part of the overall environment contributes to the construction of PSDs. For example, *serene* in the eight dimensions of PSDs is inseparable from soundscape. The purpose of linking PSDs with PRSS is to explore which relatively important dimensions of PSDs will affect the soundscape restoration, so that designers can improve the design of soundscape recovery from the perspective of PSDs in the future.

Overall, this study explored the relationship between two aspects—PSDs and the restorative potential of artificial lake spaces in urban parks—and identified the specific mechanism of specific artificial lake spaces on rehabilitation potential. The study will help to reasonably evaluate and enhance the health benefits of artificial lake spaces in cities and ultimately contribute to healthy urban and sustainable development. This study's specific objectives were to investigate:

1. What is the perceived evaluation level of satisfaction and restoration of the overall environment and soundscape in artificial lake spaces of urban parks?
2. Representation of the eight PSDs in artificial lake spaces of urban parks.
3. Correlations between the eight PSDs and the satisfaction and restorativeness of the overall environment and soundscape in artificial lake spaces of urban parks.

MATERIALS AND METHODS

Ethics

The studies involving human participants were reviewed and approved by the Ethics Committee of the College of Landscape Architecture and Arts, Northwest A&F University. Written informed consent to participate in this study was provided by the participants.

Study Area

Two typical public urban parks in Xi'an, China, were selected as the study area based on similar sound levels and environmental conditions in both. The field survey was conducted from September to October 2019. The two urban parks with open artificial lakes—Qujiang pond heritage park (opened in 2008) and the Yanming lake wetland park (opened in 2016) were selected as the representative urban artificial lake space. The two parks are freely accessible and popular, and the two lakes are similar in size, gentle in velocity of flow, and slightly eddy. Four sample points were chosen in the perimeter of the hard revetment in each park. The distance of four sample points between the lake edge lines is basically the same (**Figure 1**).

Questionnaire Structure

The questionnaire survey was conducted on-site using an app called "Wen Juan Xing" with participants' agreement. The questionnaire consisted of two sections. The first section of the questionnaire was designed to collect the respondents' perceptions in artificial lake spaces using the eight PSDs, including *serene, nature, rich in species, space, prospect, refuge, social, and culture*. The Likert five-level scale was used with 0 (no feeling) to 5 (very strong feeling). The second section of the questionnaire focused on recovery potential in artificial lake spaces, respectively, soundscape satisfaction, overall satisfaction, soundscape restorativeness, and overall restorativeness. The soundscape satisfaction and overall satisfaction were measured with a five-level Likert scale from 0 (very poor) to 5 (very good). As to soundscape restorativeness, a simplified perceived restorativeness soundscape scale (PRSS) as a subjective questionnaire can comprehensively determine the restoration of soundscapes. Six landscape architecture experts were invited to participate in the research process of the simplified PRSS. Those experts selected two questions in each dimension of PRSS which were considered to be representative and easy to understand. Moreover, "Being-Away-To" and "Extent (Scope)" as two dimensions of PRSS were neglected in this study for the following reasons. First, the "Being-Away-To" and "Being-Away-From" dimensions of the PRSS were developed from the "Being-Away" dimension of the ART theory (Kaplan and Kaplan, 1989; Hammitt, 2000), both of which measure recovery from fatigue due to environmental change. "Extent (Scope)" and "Extent (Coherence)" were developed from the "Extent" dimension according to the ART theory (Kaplan and Kaplan, 1989; Payne, 2013). The "Extent (Scope)" dimension has only one question regarding the scale of the environment ("The sonic environment suggests the size of this place is limitless"), which is not especially relevant to our objectives and was not easily understood by the general respondents in the pre-experiments. Second, the neglect of "Being-Away-To" and "Extent (Scope)" of PRSS caused a unification of dimensions with PRS, which was in favor of respondents understanding and to avoid confusion. That is because respondents should understand the meaning of dimensions for a better comprehension of questions before filling in the questionnaire. Finally, to shorten the investigation time, and to avoid the psychological interference of the long-term investigation on the respondent, a simplified PRSS could be used instead, with Cronbach's alpha of 0.90. Therefore, four subscales, selected as the following metrics, were calculated as a simplified PRSS: Fascination ("I find this sonic environment appealing" and "My attention is drawn to many of the interesting sounds here"); Being-Away-From ("This sonic environment is a refuge from unwanted distractions" and "These voices relieve the pressure of my daily life"); Compatibility ("These sounds relate to activities I like to do" and "I rapidly get used to hearing this type of sonic environment"); Extent (Coherence) ("All the sounds merge to form a coherent sonic environment" and "The sounds I am hearing seem to fit together quite naturally with this place"). As to overall restorativeness, the PRS with 16 items as an instrument for measuring the restorative quality of environment was applied, with Cronbach's alpha of 0.88. There are four parts of PRS according to ART, as follows: Fascination (e.g., It is a fascinating environment); Being-Away (e.g., This is a good experience away from the troubles of real life); Extent (e.g., It is a confusing place); and Compatibility (e.g., I feel I belong here). Both PRSS and PRS were provided in the Chinese version, and each question was answered on a Likert five-point scale in response to "how much do you agree with the statement?" Scores ranged from 1 (not at all) to 5 (completely agree). In particular, the Likert five-point scale represents an evaluation level in terms of both satisfaction and restorativeness, which can be divided into four levels to indicate the level of perception: poor level (1–2), medium to poor level (2–3), medium to good level (3–4), and very good level (4–5).

Field Survey

All respondents were randomly selected among visitors in each park and approached visitors were consented and informed that all answers would be anonymous. The willing participants were then invited to fill the questionnaire in individually using a tablet. Each sampling site was conducted to do a field survey twice over different days to reduce deviation caused by participant selection. Meteorological factors were measured simultaneously from 9:00 a.m. to 5:00 p.m. *via* Kestrel 5500 weather station while

FIGURE 1 | Study area and photographs: **(A)** Location of the Qujiang pond heritage park and the Yanming lake wetland park in Xi'an, China; **(B)** four sample points in the Qujiang pond heritage park; **(C)** four sample points in the Yanming lake wetland park; **(D,E)** two images of the two study sites.

the questionnaire survey was conducted, including temperature (15.45°C, *SD* = 2.58°C), relative humidity (63.35%, *SD* = 11.47%), and wind speed (0.75 m/s, *SD* = 0.66 m/s). In addition, types of sound sources and sound levels can affect the subjective attitude to soundscape (Völker et al., 2018). In the pre-experiment before the field survey, the sound sources in the two sample parks were similar according to the perceptions of the respondents, which indicated that the two parks have a similar soundscape. As to each site's sound levels, it was measured continuously for 8 h, from approximately 9:00 a.m. to 5:00 p.m. *via* Kjaer 2250 class 1 sound level meter (SLM; 1-s logging period). Stable sound conditions were guaranteed at a similar level in each site each day with an average sound level of 64.10 dBA.

Statistical Analysis

A total of 442 respondents (48.4% males) aged 33.65 ± 14.92 completed the survey, and all those involved had self-reported normal audition and vision with 272 in Qujiang Pound Heritage Park and 170 in The Yanming Lake Wetland Park, respectively. First, descriptive statistics (Box line diagram and Bar chart) were used to indicate respondents' visiting experience satisfaction and restorative potential for four variables: soundscape satisfaction, overall satisfaction, soundscape restorativeness, and overall restorativeness.

Secondly, to explore the representative characteristics in the artificial lake space within the eight PSDs, a one-Way ANOVA was conducted. The evaluation level of the eight PSDs served as the dependent variable, and the eight PSDs were used as factors. The SNK method was used to indicate the high or low level of the group according to the scores.

Finally, to examine the relationship between two sets of variables of eight PSDs and the satisfaction and restorativeness, a Canonical Correlation Analysis (CCA) was conducted, which can help in finding how independent variables can predict the

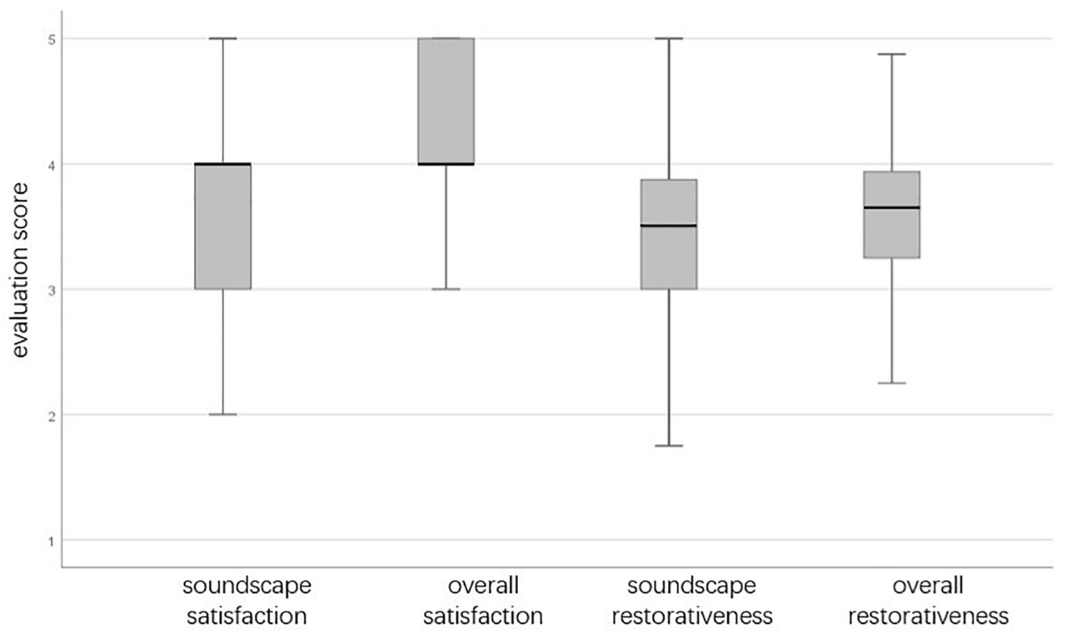


FIGURE 2 | The evaluation score of satisfaction and restorativeness on a five-point scale. Box lines represent extremum and quantile, and the black bold line represents the median.

dependent variables by expressing the correlations between the two sets. In our study, the eight PSDs served as independent variables, while satisfaction and restorativeness of soundscape and overall environment were dependent variables. All statistical analyses were carried out using SPSS 25.0 software.

RESULTS

The Level of Potential Restorativeness in Blue Space

The distribution of soundscape satisfaction (median = 4.00) was relatively scattered, with a score range of 2–5 (**Figure 2**). In contrast, overall satisfaction (median = 4.00) had a more concentrated distribution. It can be concluded that soundscape satisfaction and overall satisfaction in an artificial lake space were relatively higher according to the approximate median score of 4. The score of soundscape restorativeness (median = 3.50) was scattered, with a range of 1.75–5. Overall restorativeness (median = 3.66) was more concentrated than soundscape restorativeness. Obviously, the median score of soundscape restorativeness and overall restorativeness were in the range of 3–4, indicating a medium to good level, which was lower than soundscape and overall satisfaction, with a median of 4, indicating a very good level. In particular, the mean level of overall satisfaction and overall restoration were both higher than those of soundscape.

Perceived Characters in Blue Space

One-way ANOVA was used to analyze the characteristics of artificial lake spaces, suggesting that there were significant differences among them. SNK (Student-Newman-Keuls) divided them into three groups ($p < 0.05$), which showed that *prospect*, *social,* and *space* were the three most vital characteristics in artificial lake spaces, then *serene* and *nature* followed as a medium characteristics. *Refuge*, *rich in species* and *culture,* were the least vital characteristics. It can be concluded that people can look into the distance, carry out social activities, and feel a sense of space around artificial lakes, which are the perceptions of artificial lake spaces (**Table 1**).

Correlations Between the Eight PSDs and Restorative Potential

A canonical correlation was performed to determine which PSDs in artificial lake spaces best predicted how people were satisfied

TABLE 1 | Perceived characteristics in blue spaces.

Eight sensory dimensions	Mean	SD	95% C.I.		Rank
			Upper limit	Lower limit	
Prospect	3.74^a	0.83	3.66	3.81	1
Social	3.72^a	0.86	3.64	3.80	
Space	3.66^a	0.85	3.58	3.74	
Nature	3.53^b	0.93	3.45	3.62	2
Serene	3.51^b	0.88	3.43	3.59	
Culture	3.37^c	1.06	3.27	3.47	3
Rich in species	3.34^c	0.95	3.25	3.43	
Refuge	3.23^c	0.91	3.14	3.31	

^{a,b,c}*Represents different groups.*

and restored. The first ($p < 0.001$) and second ($p = 0.011$) canonical correlation pairs were statistically significant.

However, the degree of explanation of the change in two canonical correlation pairs is listed in **Table 2**, suggesting that the second canonical correlation pair's parameters were much lower since its explained proportion of variance were all lower than 0.1. In contrast, the first canonical correlation pair was more convincing since its explained proportion of variance were all more than 0.2, and the follow-up analysis mainly focused on the first pair canonical correlation pair. The correlation coefficient between two sets of factors obtained for the first pair of typical structures was 0.67 ($p < 0.001$). As depicted in **Figure 3**, there were eight PSDs included in the final model, which contributed differentially (negatively and positively) to eight PSDs; *serene, culture, space, nature,* and *refuge* were the main determinants for the components due to high values of unstandardized coefficients (–0.47, –0.28, –0.21, –0.19, and –0.19, respectively). Overall satisfaction, soundscape restorativeness, and overall restorativeness largely determined the canonical variate (–0.51, –0.58, and –0.81, respectively) based on the dependent set of variables (satisfaction and restorativeness).

DISCUSSION

The Evaluated Level of Satisfaction and Restorativeness in Blue Space

There were more respondents who were very satisfied (scored 4–5 to soundscape satisfaction and overall satisfaction) with the soundscape and overall environment, than those who experienced an excellent restoration (scored 4–5 to soundscape restoration and overall restoration). To be specific, soundscape restorativeness and overall restorativeness had a medium to high score level rather than a very high level. It means that the soundscape restoration and overall restoration level requires the most improvement. The study further focused on improving visiting satisfaction, and restoration in particular, would contribute a lot to realizing the happiness and well-being of blue spaces as artificial lake spaces.

There are many reasons why people were satisfied and obtained health benefits from it. First, people's favoring of open water was one of the main reasons (Deng et al., 2020) because preference and perception evaluation (such as satisfaction evaluation) may be related to the landscape and this was inseparable from people's natural hydrophilicity (Foley, 2011). Some components of preferred landscapes are water, open space, and trees (Kaplan and Kaplan, 1989; Adevi and Grahn, 2012). People consider the blue space as positive and attractive (Völker and Kistemann, 2011). Secondly, it is generally accepted that humans are multisensory (Schwarz, 2013; Nanay, 2018). Therefore, like soundscape, visual landscape is part of the overall environment, and it is also a direct indicator distinguishing between different landscapes. A visual landscape, for example, is one of the direct factors of the overall environmental evaluation, and the evaluation of soundscape will also be affected by the visual landscape. That is consistent with previous studies where Liu et al. (2014) found that overall soundscape evaluation can be affected by the esthetic quality of a visual landscape to a large extend (Liu et al., 2014). Pheasant et al. (2008) found that perceived tranquility linked to both the sound level and percentage of natural visual features including blue space (Pheasant et al., 2008).

The results support that artificial lake spaces can be a relatively good choice for humans' visiting experience and well-being. Previous studies have also shown that blue spaces like lakes were a preferred place for entertainment (Asakawa et al., 2004) and blue water features could provide a restorative landscape (White et al., 2010; Sonntag-Öström et al., 2015), though restoration still requires improvements. In future city constructions, landscape designers and other stakeholders should pay more attention to artificial lakes as one of the most critical public health resources.

TABLE 2 | Proportion of variance explained two pairs of canonical correlation.

Canonical variable	Set 1 by self	Set 1 by Set 2	Set 2 by self	Set 2 by Set 1
1	0.440	0.201	0.623	0.285
2	0.074	0.004	0.063	0.003

Set 1 and 2 mean the first and second canonical correlation pairs, respectively. The second canonical correlation pair was not analyzed in the study due to its very low proportion of explained variance.

Representation of the Eight PSDs in Blue Space

To explore people's perceived characteristics of artificial lake spaces, it is necessary to measure the perception characteristics using the eight PSDs. In the study, *prospect*, *social,* and *space* were important in artificial lake spaces. It means that, when people were on the water bank, they felt a broad vision, coherent and integrated expansive space, according to the meaning of *prospect* and *space*. Furthermore, people often gathered for chatting, sports, or other recreational activities according to the meaning of *social* from the eight PSDs. For the three most strongly perceived attributes, the selected sample site was artificial lake spaces in urban parks, with few aquatic plants on the lake surface and no aquatic trees and facilities to block the sight according to the actual construction of the lake space, as some respondents commented. People preferred to engage in a variety of social activities on the water bank, such as appreciation, being with kids, chatting, etc. (Dinda and Ghosh, 2021), so the PSD of *social* were obvious to be perceived.

Refuge, *rich in species,* and *culture* were less frequently perceived, indicating that artificial lake spaces lacked safety, privacy, diverse plants, and a sense of cultural and historical atmosphere according to the meaning of PSDs. According to the prospect-refuge theory, shelter and open vision are both indispensable as people would feel safe (Appleton, 1976). However, some respondents commented that "The artificial lake here is an open place, but the shore had fewer shelter facilities and plants that provided shade." Moreover, *serene* and *nature* were placed in the middle, indicating that artificial lake spaces were relatively calm, clean, and natural. This result was consistent

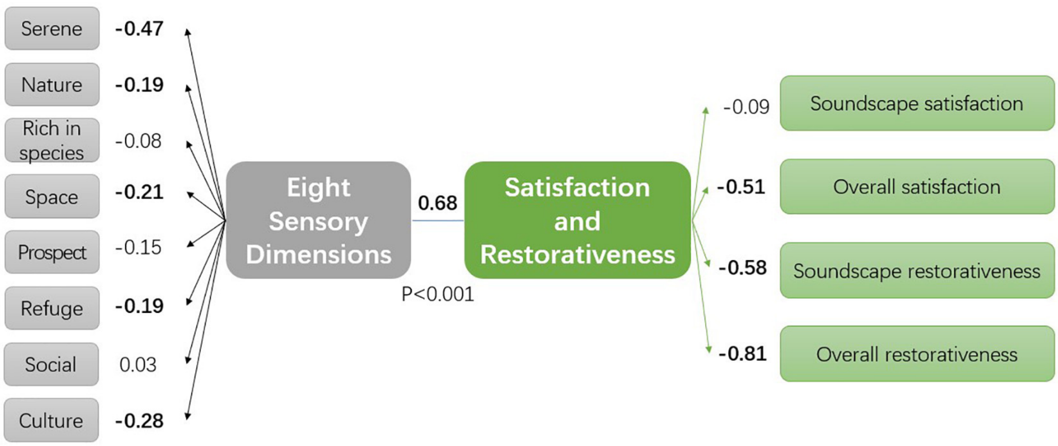


FIGURE 3 | A model of the canonical correlation analysis for the eight sensory dimensions and satisfaction and restorativeness. Included are the value of the first canonical correlation and the values of the significant unstandardized coefficients.

with the previous study, which showed that blue spaces like open artificial lake spaces could provide tranquility to people (Gao et al., 2019; Qiu et al., 2021). On the one hand, water can absorb noise to create a better soundscape; further, human activity sounds were concentrated densely on the water bank, resulting in a decrease of tranquility (Dinda and Ghosh, 2021). Therefore, *serene* was perceived in the middle level. The reason for mid-level *nature*, the urban park lies between a natural environment (high natural environment) and an urban environment (low natural environment) in terms of soundscape (Payne, 2013). Hence, the *nature* level did neither score high nor low.

Correlations Between the Eight PSDs and Restorative Potential

For eight PSDs, *serene, natural, rich in species, space, prospect, refuge, and cultural* except *social* were positively correlated with satisfaction and restorativeness. The first five most significant attributes were *serene*, *cultural*, *space*, *nature,* and *refuge*, which mostly acted on overall restorativeness. According to a previous study, *serene*, *nature,* and *refuge*, with the absence of *rich in species* and *social,* were considered the most representative characteristics of the restorative environment (Memari et al., 2017). *Serene* was the most frequent characteristic people preferred (Grahn and Stigsdotter, 2010). The eight PSDs had a weak relationship with soundscape satisfaction. Considering that the proportion of variance explained in the first pair of typical structures was approximately 20%, other factors may affect the satisfaction and restorativeness, especially the soundscape satisfaction. Given that eight PSDs describe the perception of the total environment, it might not be very proper for soundscape as only a sensory dimension of the overall environment.

To improve the environment of artificial lake spaces in urban parks, PSDs that were positively related to artificial lake space satisfaction and restorativeness should be compared with the currently existing PSDs. Specifically, *nature* and *serene* were in the medial level, and *cultural* and *refuge* were in the lower level, while they were important in restorative potential. In the further landscape promotion of artificial lake spaces, *nature*, *serene*, *cultural,* and *refuge* perceived levels need to be increased. Due to people's hydrophilicity (Foley, 2011) and the importance of open water to well-being (Finlay et al., 2015), it is essential to find the breakthrough point of building artificial lake space.

To find out why *serene* was most effective in artificial lake space, the form of water (moving water or still water) and waterside bank deserves special attention. In a previous study, people felt harmonious and diverse in blue space because they were close to water and the water bank (Burmil et al., 1999). Steinwender et al. (2007) considered wide surfaces, as well as revetment plants, to make good contributions (Steinwender et al., 2007). First, the water itself can provide a high esthetic value component due to its winding revetment, peacefulness, and wide water surface. A wide field of vision is expected in blue spaces (McDougall et al., 2020). Broad vision wins people's favor (Van Berkel et al., 2018). Therefore, a vast and calm water surface can help create a perceived sense of *serenity* and *space*. Second, the waterside bank is also important. In urban parks, people come into contact water through the water bank, which plays a crucial role in the perception and construction of artificial lake spaces. Deng et al. (2020) found their sample lake was inferior to terrain and lawns because the bank had less shade (Deng et al., 2020). A high-quality pathway (Verbić et al., 2016) and an easily accessible closing-water space (McDougall et al., 2020) could improve the overall visiting experience around artificial lake spaces. It can be suggested to landscape planners and decision-makers in relevant departments that water banks should receive full attention when changing perceived attributes of artificial lake spaces.

Therefore, to achieve more *serene, nature, cultural, and refuge* characteristics, planning can include the following: Some rest and viewing facilities can be provided, so that people can stay for a long time to get the benefits of the lakes and create a serene soundscape for more calm. "The increase of cultural facilities will help create a sense of culture," as respondents claimed. Adding well-maintained plants can enhance beauty (Chow, 2013), and plants can enhance *nature* and create *refuge* (Parsons and Daniel, 2002). Therefore, planting of plants should be increased on the lake banks.

Limitations and Future Research

Although blue spaces are indeed beneficial to health, exploring the mechanism of the restorative potential in blue spaces remains significant. The artificial lake space combined with PSDs in this study was just a pilot study. There are still some limitations in the study. First some landscape features of artificial lakes such as the area of water, moving water or still water, people on the water bank or the island, water freezing in winter, etc., and other kinds of blue spaces such as rivers and natural wetlands, are worth further investigations in future studies. Second, demographic group differences, health status, and types of activities should be taken into consideration in future landscape evaluations (Arnberger and Eder, 2011; Lindholst et al., 2013; Carrus et al., 2015). Physiological restoration of artificial lake spaces is also worth further exploring.

CONCLUSION

This study examined the levels of satisfaction and restorativeness of soundscape and overall environment, and the relationship between the PSDs and restorative potential in artificial lake spaces. The main results included that: First, artificial lake spaces had a high level of soundscape satisfaction and overall satisfaction, while restorativeness of soundscape and overall environment still need to be improved, due to limited positive responses; second, *prospect*, *social,* and *space* were the most three obvious PSDs in artificial lake spaces; third, *serene*, *cultural*, *space*, *nature,* and *refuge* primarily affected the restorative potential, especially the overall environment restorativeness. *Serene* was the critical factor for improving the quality of artificial lake spaces in urban parks.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of College of Landscape Architecture and Arts, Northwest A&F University. Written informed consent to participate in this study was provided by the participants.

AUTHOR CONTRIBUTIONS

YQ and XF: methodology, validation, formal analysis, investigation, and visualization. YQ: writing—original draft. TG and LQ: conceptualization, methodology, validation, resources, writing—review and editing, supervision, project administration, and funding acquisition. LQ: formal analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the National Natural Science Foundation of China (grant numbers: 31971720), the Research Fund for Advanced Talents of Northwest A&F University (grant numbers: Z111021501), and Talent Support of Shaanxi Province (grant numbers: A279021715 and A279021830).

REFERENCES

Adevi, A. A., and Grahn, P. (2012). Preferences for Landscapes: A Matter of Cultural Determinants or Innate Reflexes that Point to Our Evolutionary Background? *Landsc. Res.* 37, 27–49. doi: 10.1080/01426397.2011.576884

Appleton, J. (1976). *The Experience of Landscape*. London: Wiley.

Arnberger, A., and Eder, R. (2011). The influence of age on recreational trail preferences of urban green-space visitors: a discrete choice experiment with digitally calibrated images. *J. Environ. Plan. Manag.* 54, 891–908. doi: 10.1080/09640568.2010.539875

Asakawa, S., Yoshida, K., and Yabe, K. (2004). Perceptions of urban stream corridors within the greenway system of Sapporo, Japan. *Landsc. Urban Plan.* 68, 167–182. doi: 10.1016/s0169-2046(03)00158-0

Burmil, S., Daniel, T. C., and Hetherington, J. D. (1999). Human values and perceptions of water in arid landscapes. *Landsc. Urban Plan.* 44, 99–109. doi: 10.1016/s0169-2046(99)00007-9

Carrus, G., Scopelliti, M., Lafortezza, R., Colangelo, G., and Ferrini, F. (2015). Go greener, feel better? The positive effects of biodiversity on the well-being of individuals visiting urban and peri-urban green areas. *Landsc. Urban Plan.* 166, 296–308.

Chow, H.-W. (2013). Outdoor fitness equipment in parks: a qualitative study from older adults' perceptions. *BioMed. Central* 13:1216.

Deng, L., Li, X., Luo, H., Fu, E.-K., Ma, J., Sun, L.-X., et al. (2020). Empirical study of landscape types, landscape elements and landscape components of the urban park promoting physiological and psychological restoration. *Urban For. Urban Green.* 48:126488. doi: 10.1016/j.ufug.2019.126488

Dinda, S., and Ghosh, S. (2021). Perceived benefits, aesthetic preferences and willingness to pay for visiting urban parks: A case study in Kolkata, India. *Int. J. Geoheritage Parks* 2021:007. doi: 10.1016/j.ijgeop.2020.12.007

Dzhambov, A. M., Markevych, I., Hartig, T., Tilov, B., Arabadzhiev, Z., Stoyanov, D., et al. (2018). Multiple pathways link urban green- and bluespace to mental health in young adults. *Environ. Res.* 166, 223–233. doi: 10.1016/j.envres.2018.06.004

Fan, Q.-D., He, Y.-J., and Hu, L. (2021). Soundscape evaluation and construction strategy of park road. *Appl. Acoust.* 174:107685. doi: 10.1016/j.apacoust.2020.107685

Fátima, N., Estefanía, H. F., Stephany, H., and Bernardo, H. (2017). Discrimination of Urban Spaces with Different Level of Restorativeness Based on the Original and on a Shorter Version of Hartig et al.'s Perceived Restorativeness Scale. *Front. Psychol.* 8:1735. doi: 10.3389/fpsyg.2017.01735

Finlay, J., Franke, T., McKay, H., and Sims-Gould, J. (2015). Therapeutic landscapes and well-being in later life: Impacts of blue and green spaces for older adults. *Health Place* 34, 97–106. doi: 10.1016/j.healthplace.2015.05.001

Foley, R. (2011). Performing health in place: the holy well as a therapeutic assemblage. *Health Place* 17, 470–479. doi: 10.1016/j.healthplace.2010.11.014

Foley, R., and Kistemann, T. (2015). Blue space geographies: Enabling health in place. *Health Place* 35, 157–165. doi: 10.1016/j.healthplace.2015.07.003

Gao, T., Song, R., Zhu, L., and Qiu, L. (2019). What Characteristics of Urban Green Spaces and Recreational Activities Do Self-Reported Stressed Individuals Like? A Case Study of Baoji, China. *Int. J. Environ. Res. Public Health* 16:1348. doi: 10.3390/ijerph16081348

Gidlow, C. J., Jones, M. V., Hurst, G., Masterson, D., Clark-Carter, D., Tarvainen, M. P., et al. (2016). Where to put your best foot forward: Psycho-physiological responses to walking in natural and urban environments. *J. Environ. Psychol.* 45, 22–29. doi: 10.1016/j.jenvp.2015.11.003

Glaser, R., Sheridan, J., Malarkey, W. B., MacCallum, R. C., and Kiecolt-Glaser, J. K. (2000). Chronic stress modulates the immune response to a pneumococcal pneumonia vaccine. *Psychosom. Med.* 62, 804–807. doi: 10.1097/00006842-200011000-00010

Grahn, P., and Stigsdotter, U. A. (2003). Landscape planning and stress. *Urban For. Urban Green.* 2, 1–18. doi: 10.1078/1618-8667-00019

Grahn, P., and Stigsdotter, U. K. (2010). The relation between perceived sensory dimensions of urban green space and stress restoration. *Landsc. Urban Plan.* 94, 264–275. doi: 10.1016/j.landurbplan.2009.10.012

Grassini, S., Revonsuo, A., Castellotti, S., Petrizzo, I., Benedetti, V., and Koivisto, M. (2019). Processing of natural scenery is associated with lower attentional and cognitive load compared with urban ones. *J. Environ. Psychol.* 62, 1–11. doi: 10.1016/j.jenvp.2019.01.007

Grellier, J., White, M. P., Albin, M., Bell, S., Elliott, L. R., Gascon, M., et al. (2017). BlueHealth: a study programme protocol for mapping and quantifying the potential benefits to public health and well-being from Europe's blue spaces. *BMJ Open* 7:e016188. doi: 10.1136/bmjopen-2017-016188

Hammitt, W. E. (2000). The Relation between Being Away and Privacy in Urban Forest Recreation Environments. *Environ. Behav.* 32, 521–540. doi: 10.1177/00139160021972649

Hartig, T., Korpela, K., and Evans, G. W. (1996). Validation of a measure of perceived environmental restorativeness. *Goteborg. Psychol. Rep.* 26, 1–64.

Hartig, T., Korpela, K., Evans, G. W., and Grling, T. (1997). A Measure of restorative quality in environments. *Scandinavian Housing Plann. Res.* 14, 175–194. doi: 10.1080/02815739708730435

Jarvis, I., Koehoorn, M., Gergel, S. E., and van den Bosch, M. (2020). Different types of urban natural environments influence various dimensions of self-reported health. *Environ. Res.* 186:109614. doi: 10.1016/j.envres.2020.109614

Kaplan, R., and Kaplan, S. (1989). *The Experience of Nature: A Psychological Perspective*. Cambridge: Cambridge Univ.

Kaplan, S. (1995). The Restorative Benefits of Nature: Toward an Integrative Framework. *J. Environ. Psychol.* 15, 169–182. doi: 10.1016/0272-4944(95)90001-2

Kummu, M., de Moel, H., Ward, P. J., and Varis, O. (2011). How close do we live to water? A global analysis of population distance to freshwater bodies. *PLoS One* 6:e20578. doi: 10.1371/journal.pone.0020578

Li, Z., and Kang, J. (2019). Sensitivity analysis of changes in human physiological indicators observed in soundscapes. *Landsc. Urban Plan.* 190:103593. doi: 10.1016/j.landurbplan.2019.103593

Lindholst, A. C., Dempsey, N., and Burton, M. (2013). Introducing a method for mapping recreational experience. *Proc. Inst. Civil Eng. Urban Des. Plan.* 166, 296–308. doi: 10.1680/udap.11.00012

Liu, J., Kang, J., Behm, H., and Luo, T. (2014). Effects of landscape on soundscape perception: Soundwalks in city parks. *Landsc. Urban Plan.* 123, 30–40. doi: 10.1016/j.landurbplan.2013.12.003

Ma, K. W., Mak, C. M., and Wong, H. M. (2021). Effects of environmental sound quality on soundscape preference in a public urban space. *Appl. Acoust.* 171:107570. doi: 10.1016/j.apacoust.2020.107570

MacKerron, G., and Mourato, S. (2013). Happiness is greater in natural environments. *Glob. Environ. Change* 23, 992–1000. doi: 10.1016/j.gloenvcha.2013.03.010

McDougall, C. W., Quilliam, R. S., Hanley, N., and Oliver, D. M. (2020). Freshwater blue space and population health: An emerging research agenda. *Sci. Total Environ.* 737:140196. doi: 10.1016/j.scitotenv.2020.140196

Memari, S., Pazhouhanfar, M., and Nourtaghani, A. (2017). Relationship between perceived sensory dimensions and stress restoration in care settings. *Urban For. Urban Green.* 26, 104–113. doi: 10.1016/j.ufug.2017.06.003

Nanay, B. (2018). Multimodal mental imagery. *Cortex* 105, 125–134. doi: 10.1016/j.cortex.2017.07.006

Nilsson, M. B., Sun, H., Diao, L., Tong, P., Liu, D., Li, L., et al. (2017). Stress hormones promote EGFR inhibitor resistance in NSCLC: Implications for combinations with β-blockers. *Sci. Transl. Med.* 9:eaao4307. doi: 10.1126/scitranslmed.aao4307

Nutsford, D., Pearson, A. L., Kingham, S., and Reitsma, F. (2016). Residential exposure to visible blue space (but not green space) associated with lower psychological distress in a capital city. *Health Place* 39, 70–78. doi: 10.1016/j.healthplace.2016.03.002

Parsons, R., and Daniel, T. C. (2002). Good looking: in defense of scenic landscape aesthetics. *Landsc. Urban Plan.* 60, 43–56. doi: 10.1016/s0169-2046(02)00051-8

Pasanen, T. P., White, M. P., Wheeler, B. W., Garrett, J. K., and Elliott, L. R. (2019). Neighbourhood blue space, health and well-being: The mediating role of different types of physical activity. *Environ. Int.* 131:105016. doi: 10.1016/j.envint.2019.105016

Patón, D., Delgado, P., Galet, C., Muriel, J., Méndez-Suárez, M., and Hidalgo-Sánchez, M. (2020). Using acoustic perception to water sounds in the planning of urban gardens. *Build. Environ.* 168:106510. doi: 10.1016/j.buildenv.2019.106510

Payne, S. R. (2013). The production of a Perceived Restorativeness Soundscape Scale. *Appl. Acoust.* 74, 255–263. doi: 10.1016/j.apacoust.2011.11.005

Pearson, A. L., Shortridge, A., Delamater, P. L., Horton, T. H., Dahlin, K., Rzotkiewicz, A., et al. (2019). Effects of freshwater blue spaces may be beneficial for mental health: A first, ecological study in the North American Great Lakes region. *PLoS One* 14:e0221977. doi: 10.1371/journal.pone.0221977

Peschardt, K. K., and Stigsdotter, U. K. (2013). Associations between park characteristics and perceived restorativeness of small public urban green spaces. *Landsc. Urban Plan.* 112, 26–39. doi: 10.1016/j.landurbplan.2012.12.013

Pheasant, R., Horoshenkov, K., Watts, G., and Barrett, B. (2008). The acoustic and visual factors influencing the construction of tranquil space in urban and rural environments tranquil spaces-quiet places? *J. Acoust. Soc. Am.* 123, 1446–1457. doi: 10.1121/1.2831735

Qiu, L., Chen, Q., and Gao, T. (2021). The Effects of Urban Natural Environments on Preference and Self-Reported Psychological Restoration of the Elderly. *Int. J. Environ. Res. Public Health* 18:509. doi: 10.3390/ijerph18020509

Schwarz, O. (2013). What Should Nature Sound Like? *Ann. Touris. Res.* 42, 382–401. doi: 10.1016/j.annals.2013.03.001

Sonntag-Öström, E., Stenlund, T., Nordin, M., Lundell, Y., Ahlgren, C., Fjellman-Wiklund, A., et al. (2015). "Nature's effect on my mind" – Patients' qualitative experiences of a forest-based rehabilitation programme. *Urban For. Urban Green.* 14, 607–614. doi: 10.1016/j.ufug.2015.06.002

Steeneveld, G. J., Koopmans, S., Heusinkveld, B. G., and Theeuwes, N. E. (2014). Refreshing the role of open water surfaces on mitigating the maximum urban heat island effect. *Landsc. Urban Plan.* 2021:121.

Steinwender, A., Gundacker, C., and Wittmann, K. J. (2007). Objective versus subjective assessments of environmental quality of standing and running waters in a large city. *Landsc. Urban Plan.* 84, 116–126.

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., and Zelson, M. (1991). Stress Recovery During Exposure to Natural Environments. *J. Environ. Psychol.* 11, 201–230.

Van Berkel, D. B., Tabrizian, P., Dorning, M. A., Smart, L., Newcomb, D., Mehaffey, M., et al. (2018). Quantifying the visual-sensory landscape qualities that contribute to cultural ecosystem services using social media and LiDAR. *Ecosyst. Serv.* 31, 326–335. doi: 10.1016/j.ecoser.2018.03.022

Verbić, M., Slabe-Erker, R., and Klun, M. (2016). Contingent valuation of urban public space: A case study of Ljubljanica riverbanks. *Land Use Policy* 56, 58–67.

Vert, C., Nieuwenhuijsen, M., Gascon, M., Grellier, J., Fleming, L. E., White, M. P., et al. (2019). Health Benefits of Physical Activity Related to an Urban Riverside Regeneration. *Int. J. Environ. Res. Public Health* 16:462.

Völker, S., and Kistemann, T. (2011). The impact of blue space on human health and well-being - Salutogenetic health effects of inland surface waters:

a review. *Int. J. Hyg. Environ. Health.* 214, 449–460. doi: 10.1016/j.ijheh.2011.05.001

Völker, S., Heiler, A., Pollmann, T., Claßen, T., Hornberg, C., and Kistemann, T. (2018). Do perceived walking distance to and use of urban blue spaces affect self-reported physical and mental health? *Urban For. Urban Green.* 29, 1–9.

White, M., Smith, A., Humphryes, K., Pahl, S., Snelling, D., and Depledge, M. (2010). Blue space: The importance of water for preference, affect, and restorativeness ratings of natural and built scenes. *J. Environ. Psychol.* 30, 482–493. doi: 10.1016/j.jenvp.2010.04.004

Wilkins, B. (2008). Review of Stress, coping, and development: An integrative perspective. *Traumatology* 14:149a.

Zhang, Y., Kang, J., and Kang, J. (2017). Effects of Soundscape on the Environmental Restoration in Urban Natural Environments. *Noise Health* 19, 65–72.

Zhao, W., Li, H., Zhu, X., and Ge, T. (2020). Effect of Birdsong Soundscape on Perceived Restorativeness in an Urban Park. *Int. J. Environ. Res. Public Health* 17:5659.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 11 May 2021
doi: 10.3389/fbuil.2021.665332

Acoustic Coatings—A Discreet Way to Control Acoustic Environment

Jose Cucharero [1,2], Tuomas Hänninen [1], Marko Makkonen [1] and Tapio Lokki [2,3]*

[1] Lumir Oy, Vantaa, Finland, [2] Aalto Acoustics Lab, Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University, Espoo, Finland, [3] Department of Computer Science, School of Science, Aalto University, Espoo, Finland

Acoustic comfort is directly related to enhanced well-being and performance of people. A typical challenge faced by architects and acousticians is to achieve adequate acoustics while maintaining the aesthetics of the space and reducing the visual aspects of acoustic materials and elements. In this study, we present a biofiber-based acoustic coating as a feasible solution to improve acoustic environments while preserving the aesthetics of spaces. An acoustic coating is a thin layer of absorption material, but the coating can be sprayed on other sound absorbing structures to make it more effective on a wide frequency range. In addition, this biofiber-based coating acts as a carbon sink during its operating life, thus reducing the carbon footprint of the building. Therefore, the coating is sustainable and is an environmental friendly solution. The absorption properties of the biofiber-based coating are demonstrated in the present study with three case studies, which all had demanding requirements to conceal the acoustic structures.

Keywords: acoustic coating, architectural acoustics, carbon sink, bio-based acoustic materials, acoustic measurements, acoustic design

OPEN ACCESS

Edited by:
Arianna Astolfi,
Politecnico di Torino, Italy

Reviewed by:
Gino Iannace,
University of Campania Luigi Vanvitelli, Italy
Rosario Romano,
University of Naples Federico II, Italy

***Correspondence:**
Tuomas Hänninen
tuomas.hanninen@lumir.fi

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *07 February 2021*
Accepted: *29 March 2021*
Published: *11 May 2021*

Citation:
Cucharero J, Hänninen T, Makkonen M and Lokki T (2021) Acoustic Coatings—A Discreet Way to Control Acoustic Environment. Front. Built Environ. 7:665332. doi: 10.3389/fbuil.2021.665332

1. INTRODUCTION

The acoustic design is included in the plans in many cases at the expense of architectural vision. Traditionally, acoustic structures are suspended ceiling systems or glued sound absorbing panels which might not be visually the most preferred solution. Using panels or suspended structures are often impossible to use in historically significant or otherwise architecturally protected buildings. Acoustic coatings offer a solution to improve acoustic comfort, while maintaining the aesthetics of surfaces, and even be difficult to be visually perceived as acoustical material. The main constraint of acoustic coatings is the limitation of their acoustic properties due to their relatively small thickness. Despite this fact, acoustic coatings can benefit from underlying surfaces so that they can be applied on top of other acoustic materials to improve the acoustic properties of the whole structure and, at the same time, enhance their visual aspects.

Over the past years, the study and development of acoustic eco-materials has received much attention (Arenas and Asdrubali, 2018). Cucharero et al. (2021) studied the influence of the ultrastructure of wood-based pulp fibers on sound absorption. They reported that wood-based pulp fiber foams exhibit similar sound absorption properties as commercial glass wool panels and that pulp fibers of smaller dimensions and processed fiber foams are more effective to produce fibrous sound absorbers. Similar findings regarding fiber dimensions have been reported by other authors (Koizumi et al., 2003; Oldham et al., 2011). Other researchers have reported efficient sound absorption properties of several biomaterials, such as cork (Iannace et al., 2020), bamboo (Koizumi et al., 2003), jute (Oldham et al., 2011), cotton (Oldham et al., 2011), hemp (Oldham et al., 2011; Berardi and Iannace, 2015), kenaf (Lim et al., 2018), coir (Fouladi et al., 2010), fique

(Navacerrada et al., 2014; Berardi and Iannace, 2015), ramie (Yang and Li, 2012). Furthermore, models to predict the sound absorption properties of some natural fibers have also been developed (Berardi and Iannace, 2017). Some researchers have reported cases in which acoustic corrections of a variety of spaces have been successfully accomplished with the help of green acoustic materials (Iannace et al., 2013, 2020).

Acoustical measurements are necessary for the prediction of the most appropriate acoustic treatment. Predictions of room acoustics can be conducted using techniques ranging from analytical expressions to computer simulations. Several analytical expressions, namely, Sabine (Sabine, 1900), Eyring (Eyring, 1930), Millington (Millington, 1932), Cremer (Cremer and Müller, 1982), Kuttruff (Kuttruff, 1994), Fritzroy (Fitzroy, 1959), Neubauer, and Arau-Puchades (Arau-Puchades, 1988), have been developed to predict reverberation time. The accuracy of different analytical expressions has also been studied by many authors. Bistafa and Bradley (2000) reported that the expression developed by Arau-Puchades was the formulae that more accurately predicts reverberation time for different configurations of sound absorbing materials in an unoccupied classroom. The authors also indicated that the most accurate analytical expression depends on the amount and distribution of sound-absorbing material in the room. Astolfi et al. (2008) concluded that the use of Sabine and Eyring expressions is sufficient to predict reverberation time in small classrooms since neither more complex models nor numerical models lead to higher accuracy. Similar findings were reported by Passero and Zannin (2010). In the study conducted by Prawda et al. (2020), the model of Fitzroy provided the most accurate prediction of reverberation time, whereas Sabine formula was reported to predict reverberation time more accurately than other more complex analytical expressions.

In this study, we present three cases where bio-based acoustic coatings have been used to improve the acoustic environment. Cases consist of a traditional small office space, a staircase functioning as an art installation, and a historically significant space at the Supreme Court of Finland. High visual requirements were set in all of the cases for the sound absorbing structures; in practice, the acoustic treatment had to be unnoticeable. Appearance of the spaces were protected for historical reasons in cases I and III, which sets higher requirements for the surface structure of the materials and also structure dimensions. In the staircase, the aim was to produce nearly anechoic acoustic environment to the multistore narrow space made out of concrete with visually unnoticable structures.

The visual evaluation of the spaces was done by the users of the spaces and clients. To evaluate the acoustics of the three premises, reverberation time, T_{30}, and speech clarity, C_{50}, were determined according to ISO 3382-1:2009 before and after installation of the acoustic treatment. Additionally, the effect of the acoustic treatment to reverberation time was predicted using Sabine equation.

In the case studies presented in this study, we demonstrate how bio-based acoustic coating can be used to sustainably increase the comfort of spaces and, in some cases, even decrease the carbon footprint of the building. For example, the most commonly used acoustic material, mineral wool, is very energy intensive, which leads to high carbon footprint (Ruuska, 2013). In contrast, plants are an abundant and sustainable source of fibrous raw material for acoustic products. Cellulose, the main structural component of natural fibers, consist about 49 weight percent (wt%) of carbon. Plants acquire this percent of carbon mainly as carbon dioxide from the air, which is processed into cellulose and other components *via* biosynthesis, with oxygen resulting as a side product. As a rule of thumb, carbon bound in 1 kg of cellulose corresponds to roughly 1.5 kg of atmospheric carbon dioxide. Although the cellulose represents only a fraction of all building materials, their ability to bind atmospheric carbon influences the total carbon footprint of the building. In this study, we speculated the effects of the biofiber-based acoustic coating on the carbon footprint of the space and buildings, as acoustic materials seldom have a positive impact on the total carbon footprint of the buildings. Nevertheless, in this study, we demonstrate that bio-based products could vastly reduce the carbon footprint of acoustic materials.

2. MATERIALS AND METHODS

2.1. Materials

A biofiber-based acoustic coating sprayed on different underlying surfaces was used as the acoustic solution for all the case studies presented in this study. The biofiber-based acoustic coating is manufactured by Lumir Ltd., a Finnish company specializing in bio-based and circular economy acoustic materials. Acoustic coating is sprayed on different kinds of surfaces as 6–8 mm layer. Sound absorption properties of the acoustic coating strongly depends on the underlying surface on top of which the coating is sprayed. **Figure 1** shows the sound absorption coefficients measured for the acoustic coating sprayed on the following underlying surfaces:

- Solution 1: Biofiber-based acoustic coating sprayed on the solid surface without any air cavity left behind.
- Solution 2: Biofiber-based acoustic coating sprayed on a 20 mm glass wool.
- Solution 3: Biofiber-based acoustic coating sprayed on a perforated gypsum board (8 mm side square hole and 20% open area) with an air cavity of 200 mm filled with 50 mm of rock wool.

Sound absorption coefficients were measured according to the ISO 354 standard (ISO 354, 2003) regarding sound absorption measurements in a reverberation room operated by a Finnish Accreditation Service (FINAS) company.

In addition to the acoustic properties, biofiber-based acoustic coating acts as a carbon sink during its operating life, which, under normal conditions, extends to several decades. The coating used in this study contains about 80 wt% cellulosic fibers as the raw material. We can calculate from the density of the coating (1,250 kg/m^3) that the biofibers used in the coating bind ∼1.2 kg of CO_2/m^2 in its fibers. In this way, the installation of this acoustic coating reduces the carbon footprint of buildings, thus contributing to alleviate climate change. **Table 1** presents carbon footprint and carbon uptake data of the acoustic materials used

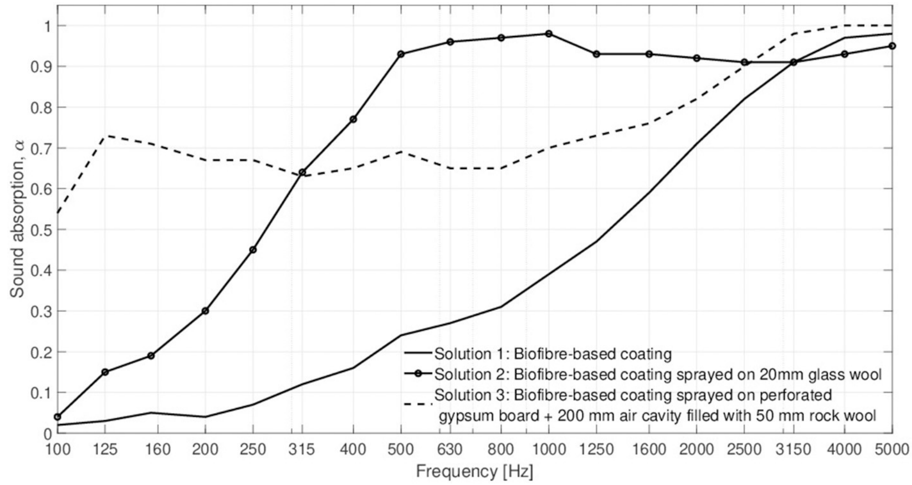


FIGURE 1 | Sound absorption coefficients of biofiber-based acoustic coating sprayed on different base surfaces.

TABLE 1 | Carbon footprint of the building materials used in this study.

Building material	Carbon footprint CO_{2e} g/kg	Carbon uptake CO_{2e} g/kg
Glass wool	3,148	0
Gypsum plasterboard	1,967	0
Biofiber-based acoustic coating	280	1,200

in this study (Ruuska, 2013). The carbon footprint is expressed in terms of carbon dioxide equivalent (CO_{2e}), which is a sum of fossil emissions calculated with the help of the Intergovernmental Panel on Climate Change (IPCC) weighting factors for 100 years (IPCC, 2007). The units are expressed in grams of CO_{2e} per kilogram of product. Carbon uptake values of glass wool and gypsum plasterboard are 0 since they consist of minerals, which inherently do not contain any carbon. The carbon footprint value of the pulp obtained from literature (Sun et al., 2018) is used in calculations for the biofiber-based acoustic coating. Coatings commonly contain binding agents, pigments, and other additives, which can be biobased or inorganic. However, their contribution to the carbon footprint of the acoustic coating is disregarded in this study.

2.2. Methodology

The acoustics of three different premises, including an artwork in a staircase space, a multifunctional room, and an office, have been designed, implemented, and measured. The models and dimensions of the spaces are shown in **Figure 2**. Effect of the designed acoustic treatment on reverberation time was predicted using the reverberation time measurements before the installation of acoustic materials as the starting point for the predictions. Sound absorption coefficients used in the predictions are shown in **Figure 1** and **Table 2**. The acoustic structures in spaces were combinations of biofiber-based acoustic coatings and different base materials. Selection of the base material to the acoustic coating depended on the constrictions set by the owners of the building. Such restrictions were always mainly determined by minimization of the visual impact of acoustic materials and elements. In addition to changes in room acoustics, the environmental impact of the acoustic treatment to the building has been estimated by considering the carbon footprint and uptake of the installed acoustic materials.

Among all the alternative reverberation formulae, the equations of Sabine (Sabine, 1900) and Arau-Puchades (Arau-Puchades, 1988) are used here. The Sabine equation requires uniform distribution of sound absorbing materials in the room, as well as diffused sound field, i.e., equal energy density in all positions of the room, as well as equal probability of sound propagating in all directions (Hodgson, 1996). Obviously, these conditions are not fulfilled in the rooms considered in this study. However, in the acoustic designs reported in the present study, rather than high prediction accuracy, we needed a fast tool to predict reverberation time, and within these terms, Sabine's equation was considered adequate. Additionally, it has been shown by many authors that, in regular shaped rooms, the accuracy of Sabine's equation is not much poorer than that of other more complex analytical expressions or room acoustic software (Bistafa and Bradley, 2000; Astolfi et al., 2008; Passero and Zannin, 2010; Prawda et al., 2020). The predictions obtained by using Sabine's equation are compared with those given by Arau-Puchades' theory. The expression of Arau-Puchades, as opposed to Sabine's theory, takes into account the non-uniform distribution of sound absorbing materials in the room. Furthermore, this expression has been reported to provide greater accuracy than Sabine's equation in spaces with high value of average sound absorption coefficient of room boundaries (Bistafa and Bradley, 2000).

Predictions of reverberation time presented in this study are realized by using measurements of reverberation time before the installation of acoustic materials as the starting point of

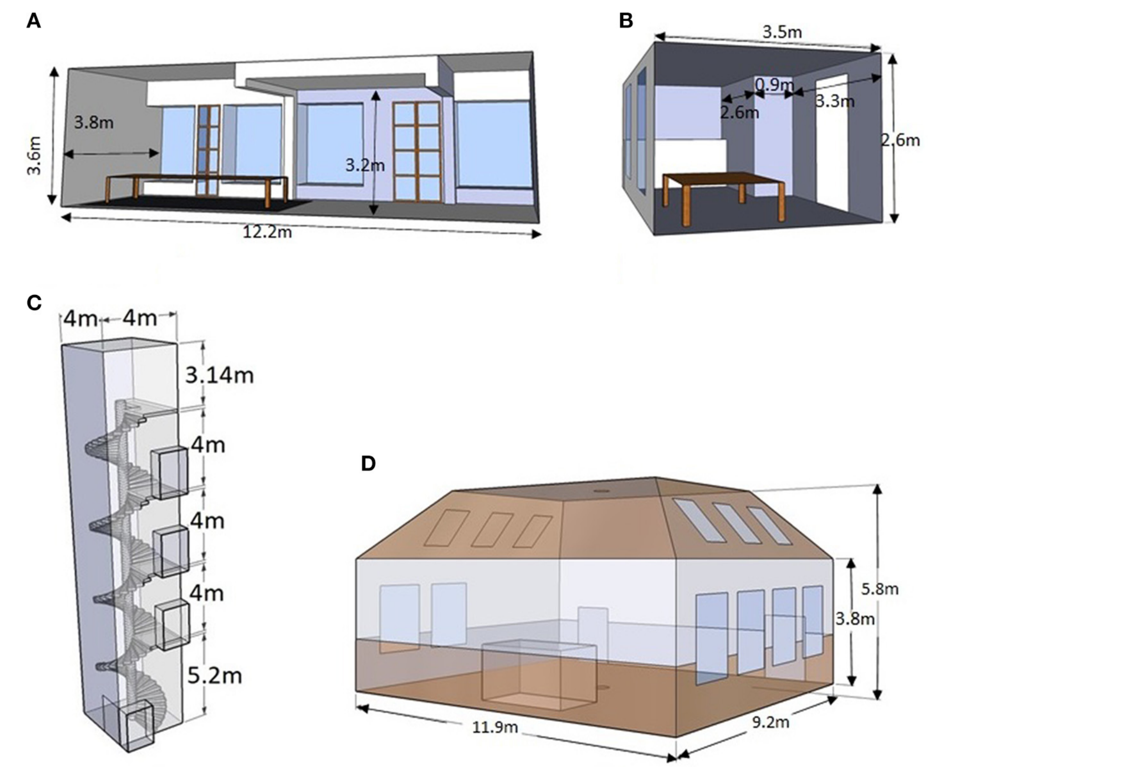


FIGURE 2 | Models and dimensions of the acoustically treated premises: **(A)** The office of Co-founders, **(B)** Co-founder meeting room, **(C)** staircase space, **(D)** a multifunctional room in the Supreme Court.

TABLE 2 | Sound absorption coefficients of building structures used to predict reverberation time through Sabine's theory (Cox and d'Antonio, 2009).

Building material	125 Hz	250 Hz	500 Hz	1,000 Hz	2,000 Hz	4,000 Hz
Painted concrete	0.01	0.01	0.01	0.02	0.02	0.02
Gypsum board with 10 cm air space	0.11	0.13	0.1	0.05	0.04	0.05
Double glazing, 2–3 mm glass, 1 cm gap	0.1	0.07	0.05	0.03	0.02	0.02
Plywood paneling, 1 cm thick	0.28	0.22	0.17	0.09	0.1	0.11
Solid wooden door	0.14	0.1	0.06	0.08	0.1	0.1

predictions, which ensures greater accuracy than applying the formulas of Sabine or Arau-Puchades from scratch. Additionally, the simplicity of Sabine's equation may be a benefit for this prediction procedure, as it requires sound absorption coefficients of only the materials being replaced with the acoustic treatment, whereas Araus-Puchades' theory requires the sound absorption coefficients of all the materials located in the same dimension. In the **Supplementary Material** the procedure followed to predict reverberation time using Sabine and Arau-Puchades formulae has been described.

Evaluation of the acoustics parameters, reverberation time, T_{30}, and speech clarity, C_{50}, were done according to ISO 3382-1 (2009), before and after the installation of acoustic treatment. The software ARTA (Artalabs, 2019) was used to obtain impulse responses using the inverse swept-sine technique (Farina, 2000). The sound source was a Genelec 8030B, and the microphones were omnidirectional 1/4-inch measurement microphones of the model Superlux ECM-999. The Genelec 8030B is not fulfilling the directivity requirements of the ISO 3382 standard; thus, in these measurements, the sound source was always located in a sound reflecting corner to simulate the omnidirectional sound source. In addition, the results of all presented measurement are averages of individual measurements with at least two sound sources and three receiver positions, except for the meeting room shown in **Figure 2B**, where, due to the small dimensions of the room, only one sound source and two receiver positions were considered. Furthermore, all reported before and after measurements were done with the same equipment and in the same source and receiver positions to enable a fair comparison.

2.2.1. Case Study I: The Office of Co-founders

The office of Co-founders Ltd. consisted of two rooms, a meeting room and a common office space, with six workstation around a big table were considered. The meeting room was lightly

TABLE 3 | Summary of room dimensions, acoustically treated surfaces, and average reverberation time before and after the installation of acoustic materials.

	Meeting room	Common office space	Spiral staircase	Multifunctional room
Volume (m^3)	64	176	331	530
S_x, S_y, S_z (m^2)[a]	18.2, 30.7, 36.6	23.2 83.4, 92.7	167.4, 167.4, 63.7	89.6, 69.6, 240.4
Back/front walls area treated (%)	0	0	90	51
Side walls area treated (%)	0	0	98	41
Ceiling/floor area treated (%)	50	47	63	0
Total surface area treated (%)	21	22	89	19
$T_{30(125-5,000Hz)_{before}}$ (s)[b]	0.95	1.23	2.88	1.78
$T_{30(125-5,000Hz)_{after}}$ (s)[b]	0.5	0.76	0.86	1.39
$T_{30(125-5,000Hz)_{Sabine}}$ (s)[b]	0.46	0.63	0.87	1.14
$T_{30(125-5,000Hz)_{Arau-Puchades}}$ (s)[b]	0.28	0.88	0.68	1.16

[a] *S_x, S_y, S_z surface area in the three directions. Front–back walls (S_x), side walls (S_y), and floor–ceiling (S_z).*
[b] *Arithmetic average reverberation time across the third-octave-bands from 125 to 5,000 Hz.*

furnished, and all the surfaces were concrete except one wall, which was covered with acoustic perforated gypsum boards. The common office space was over 3 m high and lightly furnished; all surfaces were concrete, and there were large windows on one side. The users of the space noticed that the most unpleasant acoustic problems were poor speech intelligibility, as well as bad sound insulation between rooms for speech.

The conditions for the acoustic treatment set by the clients were as few visual impacts as possible, brevity in the acoustic design, and installation of acoustic materials. Moreover, the thickness of the acoustic structure to be installed in the ceiling was restricted to 4 cm due to lighting and electrical devices and structural elements found in the ceiling. The proposed acoustic treatment involves:

- Ceiling of the common office space: biofiber-based acoustic coating sprayed on 20 mm glass wool panels glued to the ceiling (LW20 Comf).
- Ceiling of the meeting room: biofiber-based acoustic coating sprayed on 20 mm glass wool panels glued to the ceiling (LW20 Comf).

Acoustical measurements were conducted in the common office space using a total of four sound source–receiver combinations (three source and four receiver positions), and in the meeting room using two sound source–receiver combinations (one source and two receivers). The same source–receiver positions were used for the measurements before and after the installation of acoustic treatment.

2.2.2. Case Study II: Artwork by the Artist Group IC-98

The artist group IC-98 produced an artwork, named *Mare Tranquillitatis*, aimed to create a zone of complete silence that serves as a place of tranquillity and confrontational encounter, as described by the authors. The artwork was created in a five-floor spiral staircase, and the acoustic design of this space was part of the work.

The spiral staircase is located in the School of Business building constructed in 2018 at Aalto University. The dimensions of the staircase space are 4 × 4 × 20.3 m, with a volume of 325 m^3, as shown in **Figure 2C**. All the walls, the ground floor, and the ceiling are painted concrete. The stairs and landings were mosaic concrete and the underneath of the stair-landing was plywood with an air cavity behind it. A concrete pile of 0.5 m of diameter was stranded in the middle of the spiral staircase from the ground floor up to a height of 18.5 m. There were neither sound absorbing materials nor acoustic elements in the space.

There were strict requirements for the acoustic treatment. Changes in the appearance of the staircase space were to be avoided, taking into account the color and the structure of the surfaces. Very thin acoustic structures had to be used as there was very little space between the stairs and walls. Only the walls, ceiling, and underneath of stair-landings were available for acoustic materials. To remain within the set requirements, the following acoustic design was proposed:

- Walls and ceiling: biofiber-based acoustic coating directly sprayed on concrete surfaces.
- Underneath of stair-landings: biofiber-based acoustic sprayed on perforated gypsum board with an air gap of ∼20 cm.

Acoustical measurements were conducted in the five floors of the building, in a total of five sound source–receiver combinations (five source and five receiver positions). The same source–receiver positions were used for the measurements before and after the installation of acoustic treatment.

There are two architecturally exactly similar staircases in the building of the School of Business but acoustic treatment was applied only to one of them. Building users can therefore experience extremely different sonic environments in two architecturally equal spaces.

2.2.3. Case Study III: Multifunctional Space in the Supreme Court of Finland

The Supreme Court of Finland is located in Helsinki in a building that dates back to 1816. The building has gone through several renovations, additions, and modernizations in the past years to provide functional premises for different uses. All renovations

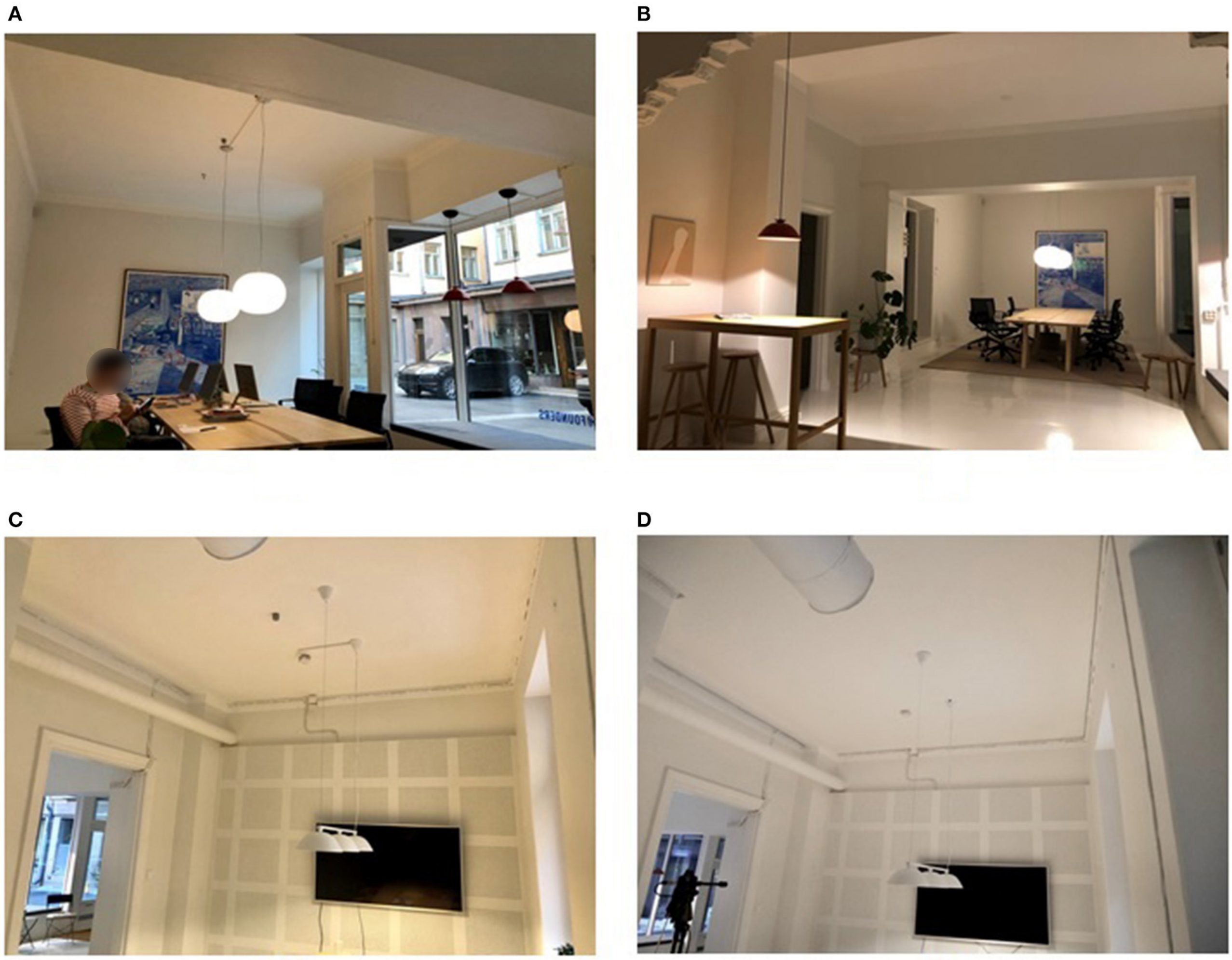


FIGURE 3 | The common office space of Co-founders before **(A)** and after **(B)** the installation of acoustic materials; and the meeting of Co-founders before **(C)** and after **(D)** the installation of acoustic materials.

were always realized by preserving the character of the building and under the supervision of the National Board of Antiquities.

In case study III, a multifunctional premises of the Supreme Court, equipped with the AV system and lightly furnished, is acoustically treated. The room is used for meetings, educational events, and other event purposes. The room has parquet flooring and vaulted ceiling at the borders and is flat in the middle. The height of the hall reaches a maximum of 5.8 m in the middle of the room. The ceiling is covered with wood paneling, and it has six windows in the vaulted area. The side and front walls are painted plastered brick masonry. One half of the back wall is painted plastered brick masonry and the other half is painted plain gypsum board attached to a frame, thus leaving an air activity of 10 cm behind the gypsum board. The inferior area of all the walls is covered with wood paneling. There are two wooden doors in the back wall. The volume of the room is about 530 m^3.

The users of the hall claimed that the main acoustic problem was poor speech intelligibility, even with the AV system due to excessive reverberation. The surfaces available for acoustic treatment were all the areas of the walls without wood paneling. Requirements for the acoustic treatment included keeping the exact same colors of the surfaces and the greatest grade of smootheness to minimize visual impact. Under these requirements, the following acoustic treatment was proposed:

- Front and side walls: biofiber-based acoustic coating directly sprayed on painted plastered brick masonry.
- Painted plastered brick masonry in the back wall: biofiber-based acoustic coating directly sprayed on hard surface.
- Plain gypsum paneling in the back wall: replacement of plain gypsum boards with perforated gypsum boards and biofiber-based acoustic coating sprayed on the perforated gypsum board (∼90 cm air-gap behind the perforated boards).

Acoustical measurements before the installation of acoustic treatment were conducted using a total of three sound source–receiver combinations. After installation of acoustic materials,

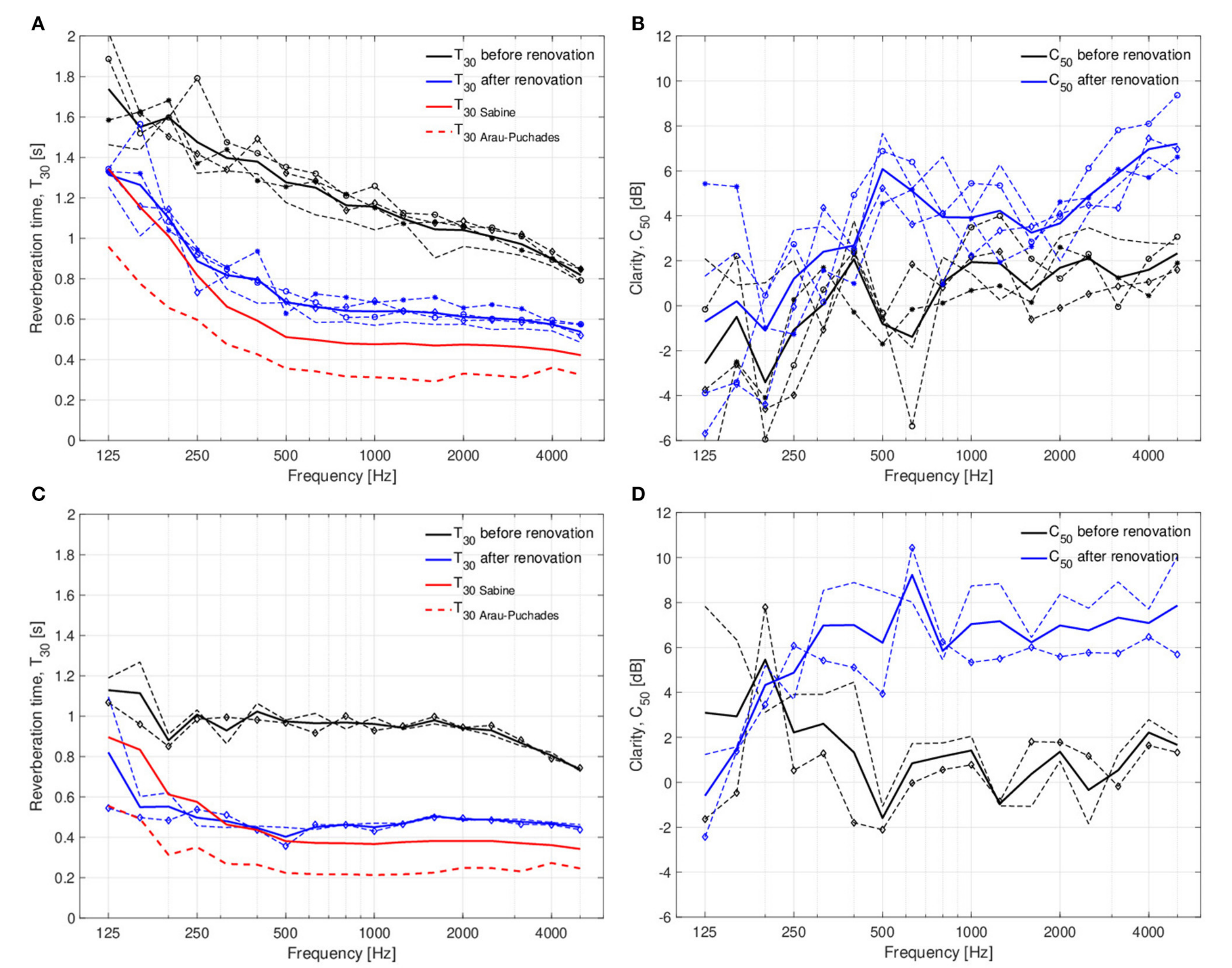


FIGURE 4 | Measurement results of T_{30} and C_{50} before and after the installation of acoustic treatments in the Co-founder spaces. Common office space **(A,B)**, and meeting room **(C,D)**. Reverberation time predictions according to Sabine's and Arau-Puchades theories are included in **(A)**.

measurements were taken using the same sound source–receiver positions with an additional sound source position; thus, a total of six sound source–receiver combinations were used.

3. RESULTS AND DISCUSSION

Table 3 presents a summary of the geometrical dimensions of the room, acoustically treated surface area between the parallel surfaces, and total surface area treated in the room, as well as average measured reverberation times and reverberation time predictions across the third-octave-bands from 125 to 5,000 Hz.

3.1. The Office of Co-founders

Employees of Co-founders Ltd. felt working in their office space unsuitable for efficient working due to poor acoustics. Space was not suitable for having more than one ongoing conversation between people or in phone simultaneously. Office is located along the one of the first "modern" shopping streets in Helsinki, and the owner of the estate wanted to preserve the original appearance of the space. **Figure 3** presents photographs of the Co-founder office before (**Figures 3A,C**) and after (**Figures 3B,D**) the installation of the acoustic treatment. The seamlessly applied acoustic coating on glass wool panels glued to the ceiling resulted in a smooth and seamless structure that has the same color as that of the underlying painted concrete. Thus, the aesthetics of the room have been successfully preserved.

Acoustic measurements of T_{30} and C_{50} before and after the installation of acoustic treatment reflects remarkable improvement of acoustic conditions in both spaces, as shown in **Figure 4**. In the common office space, reverberation time decreased significantly, reaching values of 0.6 s at frequencies above 250 Hz, 0.8 s at 250 Hz, and 1.35 s at 125 Hz. The mild change in reverberation time at 125 Hz octave band results from the restriction of 4 cm of the thickness of the acoustic

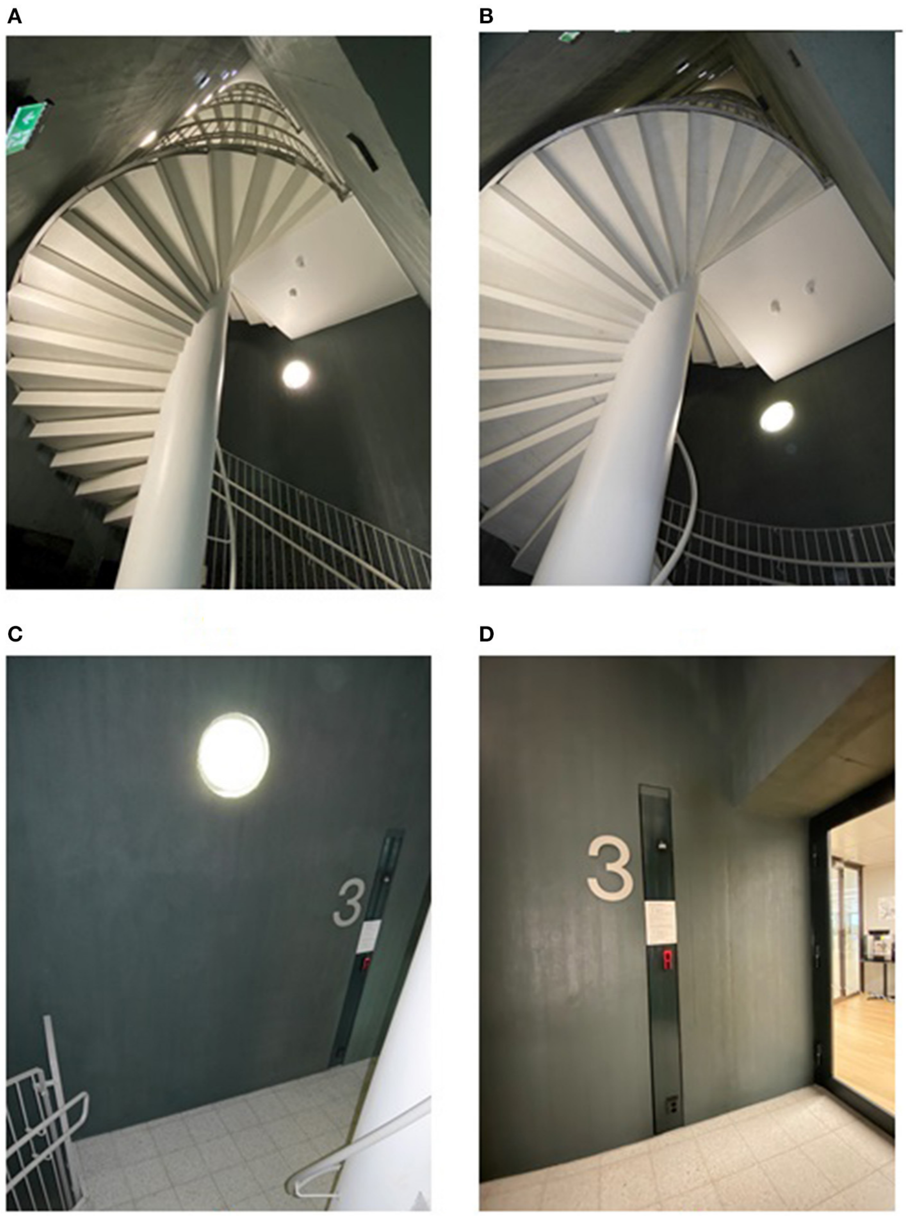


FIGURE 5 | Staircase space before **(A,C)** and after **(B,D)** the installation of acoustic materials.

structure in the ceiling due to structural elements and electrical and lighting devices. In the case of the meeting room, the acoustic treatment reduced reverberation time to 0.4–0.6 s at frequencies above 125 Hz and 0.8 s at 125 Hz. Speech clarity is also enhanced in both rooms between 2 and 6 dB at frequencies above 125 Hz. Deviations between the individual measurements from the common office space may be caused by the fact that there were three different sound source positions among the four measurements, which together with the varying surface materials as well as furniture in the room may influence the ratio between early and late reflections. The acoustics of the meeting room belongs now to class A acoustic comfort according to the Finnish standard SFS-5907 (SFS 5907, 2002). Employees of Co-founders Ltd. expressed that their office felt cozy after the acoustic improvements and that no visible changes in the appearance of the rooms could be observed.

Predicted reverberation time according to Sabine's equation shows good agreement with the reverberation time measured after the installation of acoustic treatment. The slight differences between the predicted and the measured reverberation time, [0–0.2] s, is probably related to the unfulfilled diffuse field conditions required by Sabine's theory, which is mainly attributed to the non-uniform distribution of sound absorbing materials in the room, as well as to the high average absorption coefficient

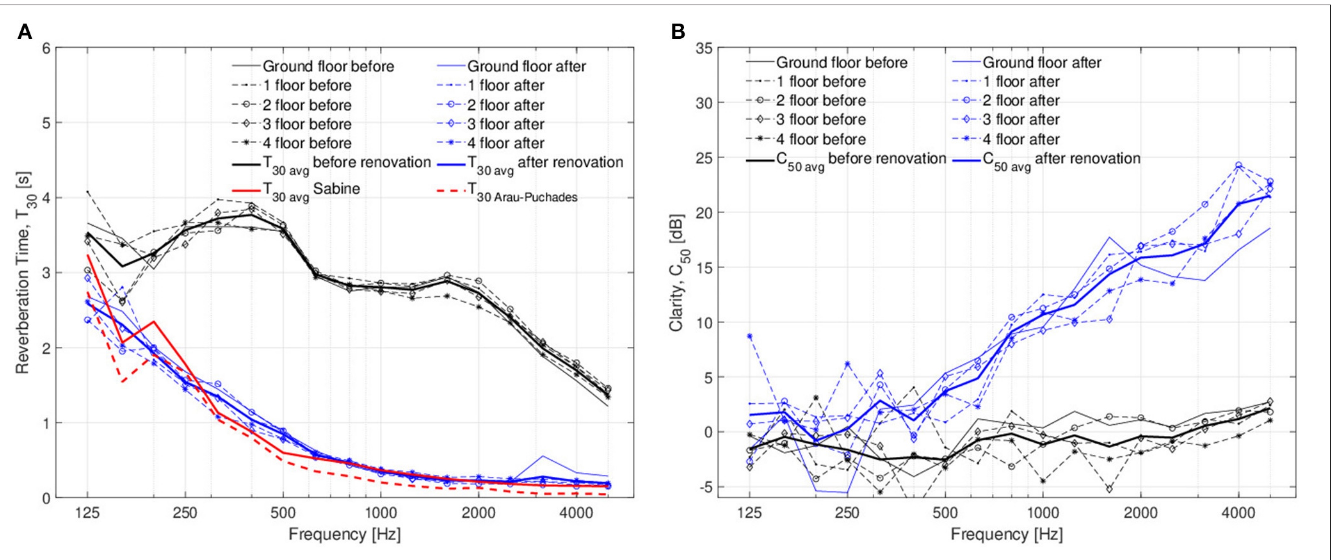


FIGURE 6 | Measurement results of T_{30} **(A)** and C_{50} **(B)** before and after the installation of acoustic treatment in the spiral staircase. Reverberation time predictions according to Sabine's and Arau-Puchades theories are included in **(A)**.

FIGURE 7 | The historical room in the Supreme Court, before **(A)** and after **(B)** the acoustic treatment.

of the room boundaries. Surprisingly, the prediction given by Arau-Puchades is considerably more inaccurate than Sabine's prediction in both rooms, although according to theory, Arau-Puchades equation should be more adequate for the sound field conditions given in these rooms. This result could be partially attributed to the sound absorption coefficients given to the concrete floor, which, if treated with some coating or painting, could indeed be a more reflective surface than that of concrete. The sound absorption coefficients of the floor surface are only used in the reverberation time predictions by Arau-Puchades theory. They are not needed in the predictions by Sabine's theory. In addition, as shown in **Table 3**, the total surface area of the acoustically treated rooms is 21% for the meeting room and 22% for the common office space. This value may be still too low to get the full benefit from Arau-Puchades equation.

3.2. Artwork by the Artist Group IC-98

Figure 5 presents photographs of the spiral staircases before and after the installation of acoustic materials. The acoustic coating sprayed on all the concrete walls, ceilings, and underneath of stair-landings has minimal visual impact but huge effects in the sonic environment of the space. Acoustic measurements of T_{30} and C_{50} were conducted in the five floors of the building, averaging the results in each floor. The results are shown in **Figure 6**. T_{30} after installation of acoustic treatment is significantly reduced, especially at mid-frequencies and high

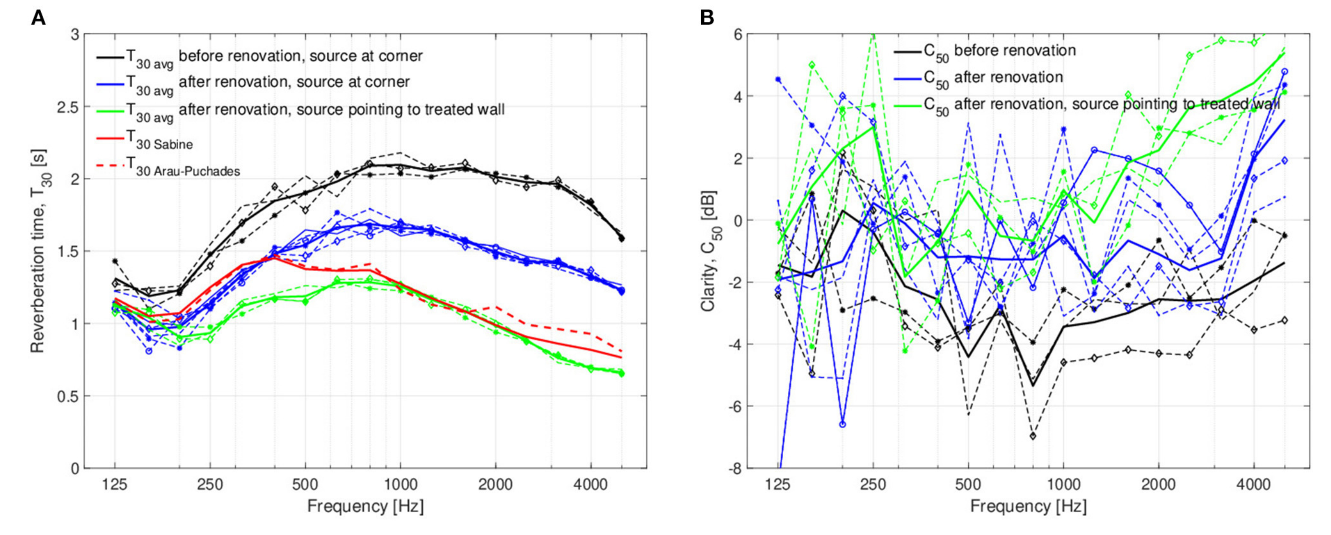


FIGURE 8 | Measurement results of T_{30} **(A)** and C_{50} **(B)** before and after the installation of acoustic treatment in the historical room in the Supreme Court. Reverberation time predictions according to Sabine's and Arau-Puchades theories are included in **(A)**.

frequencies, where the biofiber-based acoustic coating sprayed on hard surface provides greater sound absorption. T_{30} is also reduced to some extent at frequencies below 500 Hz. This is attributed to the perforated gypsum panels sprayed with biofiber-based acoustic coating that are installed on the stair-landings. This significant difference in T_{30} at frequencies below and above 500 Hz causes an interesting effect for speech, as one could clearly hear immediate attenuation of consonant sounds and emphasis of vowel sounds. This is also reflected in **Figure 6B**, where it can be seen that speech clarity increases with frequency, with an improvement of ∼15–20 dB at frequencies above 2,000 Hz.

Both Sabine and Arau-Puchades reverberation time predictions showed high accuracy, with Sabine's theory presenting slightly more accurate predictions. The high accuracy of Sabine's theory in this space could be expected as the sound absorbing materials were uniformly distributed in all the surfaces of the space. However, such high average absorption coefficient of boundaries is not suitable for Sabine's theory (Sabine, 1900). On the contrary, high average absorption coefficients of room boundaries are beneficial for Arau-Puchades theory (Bistafa and Bradley, 2000).

Such an extreme use of acoustic materials in a space where long and loud reverberation is expected resulted in a baffling effect, according to designers of the art installation. Tinted acoustic coating conceals the sound absorbing surface in plain sight, which was one of the objectives.

3.3. Multifunctional Space in the Supreme Court of Finland

Figure 7 presents photographs of the multifunctional space before and after the installation of acoustic treatment. In this case, the acoustic coating was finished with care to make the surface as smooth as possible. As shown in the figures, only the upper part of the wall could be treated, as all other surfaces and the whole ceiling are made of wood.

Results from acoustical measurements, T_{30} and C_{50}, are shown in **Figure 8**. The average results from speech clarity measurements indicate an improvement of C_{50} of around 2–4 dB at frequencies above 250 Hz. However, individual clarity measurements show remarkable fluctuations at different frequencies, as well as significant deviations between measurements. On the contrary, such fluctuations are not reflected in reverberation time measurements. Such fluctuations could be attributed to the fact that only 19 % of the total surface area of the room was acoustically treated. The reverberation time could not be lowered as much as it would be required due to the small area that was possible to cover with acoustic coating. However, a deftly audible result was obtained. The client commented that they could not see the acoustical treatment at all; thus, visually, the renovation was successful. Moreover, the users of the room have reported that now the sound system of the room is improving the speech intelligibility as it did not help at all before the renovation.

The sound sources of the sound system are located on the acoustically treated walls, and they are pointing to the opposite treated walls. In an effort to understand the reasons for the remarkable enhancement on sound quality of the sound system in the hall, we measured T_{30} and C_{50} with the sound source pointing to the acoustically treated wall at a height of 1.7 m. The results are shown in **Figure 8** in green color. The results show a reduction of over 0.5 s of T_{30} with respect to the reverberation time measurements taken with the sound source located in the corner. In this measurement, the directivity of the sound source is playing a crucial role, as most of the sound energy at mid-frequencies and high frequencies is directed toward the treated wall, resulting in greater energy dissipation of first reflections and more significant reduction of reverberation time. This second set of acoustical measurements with the source pointing to the treated wall is an excellent example on why, when using directive sound sources for acoustical measurements, the source has to be pointing to a corner with reflective surface, preferably a 3D corner, to distribute sound energy all around the room and approximate, to the extent possible, the response of omnidirectional sound sources.

In this room, Sabine's and Arau-Puchades theories provide similar predictions of reverberation time. Interestingly, at frequencies above 500 Hz, they are in good agreement with the reverberation time measured for the sound source pointing to the treated wall, whereas below 500 Hz, the predictions agree with the reverberation time measured for the source located in the corner. However, considering the adequate position of the sound source being the reflective corner, both prediction methods considerably overestimate reverberation time at mid and high frequencies. The failure of both prediction methods may be attributed to the irregular shape of the ceiling, as well as the large dimensions of the space.

3.4. Environmental Viewpoint

The biofiber-based acoustic coating used in all of the cases increased the bound atmospheric carbon of the buildings. The most efficient way to use such carbon negative material would be applying it directly to the existing structures, for example, the staircase walls in case II (biofiber-based acoustic coating). The carbon footprint of this structure would be about −980 CO_{2e} g/m^2. The use of thick and tinted acoustic coating decreases the need of other building materials, such as filler and paints, which also decreases the carbon footprint; however, this has not been taken into consideration in the calculations.

In many cases, however, the sound absorption properties of the coating are not enough and other acoustic materials required to be used as a base for coating. The carbon negative footprint of the coating decreases the carbon footprint of the whole acoustic solutions. By coating the 20 mm 95 kg/m^3 glass wool, the carbon footprint of the structure is decreased by 16 % (5,000 CO_{2e} g/m^2). For coated 12.5 mm 760 kg/m^3 gypsum, the same effect is merely 5 % (17,706 CO_{2e} g/m^2).

In individual spaces, the decrease of the carbon footprint by biofiber-based acoustic coating might seem insignificant, especially in presented case studies, as the rooms are quite small. However, if we consider a small office building with 15,000 m^2 of acoustic surface, the reduction of the carbon footprint of such a building by the coating would be 14.7 tons CO_{2e}, which equals to emissions of flying around the earth for about 2.5 times (Ciers et al., 2019).

4. CONCLUSIONS

This article presents three case studies in which the acoustical conditions were significantly improved with sustainable biofiber-based acoustic coatings. The acoustic structures consisted of acoustic coatings applied directly on an existing non-acoustic surface and combined structures with glass wool or perforated gypsum. Moreover, the acoustic treatment was implemented so that it is visually unnoticeable, which is often an important aspect, e.g., in historical buildings or in some special purpose rooms. The results of the case studies surprised the clients as they noticed the large change in acoustics but no visual changes.

The clients frequently demanded the solutions that have the minimal carbon footprint. In fact, European Union is aiming to be climate neutral by 2050 with net zero greenhouse gas emissions (GGE). This goal is unachievable merely by reducing the GGE, and industries need to find ways to produce carbon negative solutions for current products. Although acoustic products are responsible for merely a fraction of the total carbon footprint of the building, they play an important role in achieving zero carbon building. We foresee that bio-based acoustic absorption materials are one of the most important sustainable solutions for sound absorption. The presented case studies proved that acoustical treatment can be done while respecting the visual aesthetics with environment friendly implementation.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

JC, TH, and MM: design of the renovations. JC: acoustical measurements. JC, TH, and TL: writing and editing the text. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from Lumir Oy (Finland) and the cooperation of Lumir personnel Kari Kammiovirta and Hannes Hynninen. The members from artist group IC-98, Patrik Söderlund and Visa Suonpää, are also acknowledged for their excellent work and contribution to this study. The authors would like to express their gratitude to Co-Founder Ltd., the Supreme Court of Finland, and Aalto University for their permission to publish the results presented in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2021.665332/full#supplementary-material

REFERENCES

Arau-Puchades, H. (1988). An improved reverberation formula. *Acta Acust. United Acust*. 65, 163–180.

Arenas, J. P., and Asdrubali, F. (2018). "Eco-materials with noise reduction properties," in *Handbook of Ecomaterials*, eds L. M. T. Martínez, O. V. Kharissova, and B. I. Kharisov (Cham: Springer), 3031–3056. doi: 10.1007/978-3-319-48281-1_137-1

Artalabs (2019). *Arta Software*. Available online at: www.artalabs.hr/a (accessed February 5, 2021).
Astolfi, A., Corrado, V., and Griginis, A. (2008). Comparison between measured and calculated parameters for the acoustical characterization of small classrooms. *Appl. Acoust*. 69, 966–976. doi: 10.1016/j.apacoust.2007.08.001
Berardi, U., and Iannace, G. (2015). Acoustic characterization of natural fibers for sound absorption applications. *Build. Environ*. 94, 840–852. doi: 10.1016/j.buildenv.2015.05.029
Berardi, U., and Iannace, G. (2017). Predicting the sound absorption of natural materials: best-fit inverse laws for the acoustic impedance and the propagation constant. *Appl. Acoust*. 115, 131–138. doi: 10.1016/j.apacoust.2016.08.012
Bistafa, S. R., and Bradley, J. S. (2000). Predicting reverberation times in a simulated classroom. *J. Acoust. Soc. Am*. 108, 1721–1731. doi: 10.1121/1.1310191
Ciers, J., Mandic, A., Toth, L. D., and Op't Veld, G. (2019). Carbon footprint of academic air travel: a case study in Switzerland. *Sustainability* 11:80. doi: 10.3390/su11010080
Cox, T., and d'Antonio, P. (2009). *Acoustic Absorbers and Diffusers: Theory, Design and Application*. London: Taylor and Francis group.
Cremer, L., and Müller, H. A. (1982). *Principles and Applications of Room Acoustics*, Vol. 1. Essex: Applied Science publishers.
Cucharero, J., Ceccherini, S., Maloney, T., Lokki, T., and Hänninen, T. (2021). Sound absorption properties of wood-based pulp fibre foams. *Cellulose* 28, 4267–4279. doi: 10.1007/s10570-021-03774-1
Eyring, C. F. (1930). Reverberation time in "dead" rooms. *J. Acoust. Soc. Am*. 1, 217–241. doi: 10.1121/1.1915175
Farina, A. (2000). "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *The 108th Audio Engineering Society Convention* (Paris: Audio Engineering Society). paper 5093, 23 p.
Fitzroy, D. (1959). Reverberation formula which seems to be more accurate with nonuniform distribution of absorption. *J. Acoust. Soc. Am*. 31, 893–897. doi: 10.1121/1.1907814
Fouladi, M. H., Nor, M. J. M., Ayub, M., and Leman, Z. A. (2010). Utilization of coir fiber in multilayer acoustic absorption panel. *Appl. Acoust*. 71, 241–249. doi: 10.1016/j.apacoust.2009.09.003
Hodgson, M. (1996). When is diffuse-field theory applicable? *Appl. Acoust*. 49, 197–207. doi: 10.1016/S0003-682X(96)00010-2
Iannace, G., Ciaburro, G., Guerriero, L., and Trematerra, A. (2020). Use of cork sheets for room acoustic correction. *J. Green Build*. 15, 45–55. doi: 10.3992/1943-4618.15.2.45
Iannace, G., Trematerra, A., and Trematerra, P. (2013). Acoustic correction using green material in classrooms located in historical buildings. *Acoust. Aust*. 41, 213–218.
IPCC (2007). *Climate Change 2007: The Physical Science Basis: Summary for Policymakers*. Assessment Report of the Intergovernmental Panel on Climate Change. Geneva: IPCC.
ISO 3382-1 (2009). *Measurement of Room Acoustic Parameters–Part 1: Performance of Spaces*. Geneva: International Organization for Standardization.
ISO 354 (2003). *Acoustics–Measurements of Sound Absorption in a Reverberation Room*. Geneva: International Organization for Standardization.
Koizumi, T., Tsuijuchi, N., and Adachi, A. (2003). The development of sound absorbing materials using natural bamboo fibers and their acoustic properties. *J. INCE Jpn*. 2, 210–216. doi: 10.11372/souonseigyo1977.27.210
Kuttruff, H. (1994). "Sound decay in enclosures with non-diffuse sound field," in *Proceedings of the Wallace Clement Sabine Centennial Symposium* (Cambridge MA), 5–7.
Lim, Z., Putra, A., Nor, M., and Yaakob, M. (2018). Sound absorption performance of natural Kenaf fibres. *Appl. Acoust*. 130, 107–114. doi: 10.1016/j.apacoust.2017.09.012
Millington, G. (1932). A modified formula for reverberation. *J. Acoust. Soc. Am*. 4, 69–82. doi: 10.1121/1.1915588
Navacerrada, M. A., Díaz, C., and Fernández, P. (2014). Characterization of a material based on short natural fique fibers. *BioResources* 9, 3480–3496. doi: 10.15376/biores.9.2.3480-3496
Oldham, D. J., Egan, C. A., and Cookson, R. D. (2011). Sustainable acoustic absorbers from the biomass. *Appl. Acoust*. 72, 350–363. doi: 10.1016/j.apacoust.2010.12.009
Passero, C. R. M., and Zannin, P. H. T. (2010). Statistical comparison of reverberation times measured by the integrated impulse response and interrupted noise methods, computationally simulated with Odeon software, and calculated by Sabine, Eyring and Arau-Puchades' formulas. *Appl. Acoust*. 71, 1204–1210. doi: 10.1016/j.apacoust.2010.07.003
Prawda, K., Schlecht, S. J., and Välimäki, V. (2020). "Evaluation of reverberation time models with variable acoustics," in *Proceedings of the 17th Sound and Music Computing Conference* (Torino), 145–152.
Ruuska, A. (2013). Carbon footprint for building products. ECO2 data for materials and products with the focus on wooden building products. *VTT Technol*. 115 p. Available online at: https://cris.vtt.fi/en/publications/carbon-footprint-for-building-products-eco2-data-for-materials-an
Sabine, W. (1900). "Reverberation: Introduction", *Am. Arch.* [Reprinted in W. C. Sabine, Collected Papers on Acoustics (London: Harvard University Press, 1922), pp. 3–68.]
SFS 5907 (2002). *SFS 5907. Acoustic Classification of Spaces in Buildings*. Helsinki: Finnish Standards Association SFS.
Sun, M., Wang, Y., Shi, L., and Klemeš, J. J. (2018). Uncovering energy use, carbon emissions and environmental burdens of pulp and paper industry: a systematic review and meta-analysis. *Renew. Sustain. Energy Rev*. 92, 823–833. doi: 10.1016/j.rser.2018.04.036
Yang, W., and Li, Y. (2012). Sound absorption performance of natural fibers and their composites. *Sci. China Technol. Sci*. 55, 2278–2283. doi: 10.1007/s11431-012-4943-1

Conflict of Interest: JC, TH, and MM were employed by company Lumir Oy.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 13 May 2021
doi: 10.3389/fbuil.2021.659115

Check for updates

Multidimensional Psychological Evaluation of Air Conditioner Sounds and Prediction *via* Correlation Parameters

Yoshiharu Soeta[1]* *and Ei Onogawa*[2]

[1]*Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Osaka, Japan,* [2]*Research and Innovation Center, Mitsubishi Heavy Industries, Aichi, Japan*

Air conditioners are regarded as a major source of noise in built environments. Although noise control technology has reduced the sound produced by air conditioners to a comparatively low level, some people may still feel that certain aspects of the sound quality lead to discomfort. Indeed, both the sound level and the sound quality of an air conditioner can affect user's acoustic comfort. The aim of this study was to determine the factors that significantly influence the subjective response to the sound of air conditioners. We assessed the A-weighted equivalent continuous sound pressure level (L_{Aeq}) and factors extracted from the autocorrelation function (ACF) and interaural cross-correlation (IACF). Subjective loudness, sharpness, and annoyance were evaluated using a paired comparison method. Multiple regression analyses were performed using a linear combination of L_{Aeq}, the ACF factors, IACF factors, and assessment of their standard deviations. The multiple regression analyses indicated that L_{Aeq}, the delay time of the first maximum peak, the width of the first decay of the ACF, and the magnitude and width of the IACF could predict subjective responses to air conditioner sounds.

Keywords: air conditioner, correlation parameter, loudness, sharpness, annoyance

Edited by:
Louena Shtrepi,
Politecnico di Torino, Italy

Reviewed by:
Jin Yong Jeon,
Hanyang University, South Korea
Lily M. Wang,
University of Nebraska — Lincoln, United States
Valtteri Hongisto,
Turku University of Applied Sciences, Finland

****Correspondence:***
Yoshiharu Soeta
y.soeta@aist.go.jp

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *27 January 2021*
Accepted: *30 April 2021*
Published: *13 May 2021*

Citation:
Soeta Y and Onogawa E (2021) Multidimensional Psychological Evaluation of Air Conditioner Sounds and Prediction via *Correlation Parameters. Front. Built Environ. 7:659115. doi: 10.3389/fbuil.2021.659115*

INTRODUCTION

Electrical appliances and mechanical equipment such as air conditioners, refrigerators, and washing machines are regarded as main noise sources in built environments. Air conditioners are widely used in residential houses and offices and are generally in operation for long periods. Therefore, many efforts have been focused on reducing the sound pressure level (SPL) of these devices during operation. As a result, the SPLs of the devises are now relatively low (Ayr et al., 2001; Tang and Wong, 2004). However, some people may still feel annoyed by certain aspects of the sound quality, even when the SPL of simulated noises in residential houses is low (Oliva et al., 2017; Hongisto et al., 2019). Therefore, both the SPL and sound quality of an air conditioner are important for acoustic satisfaction.

Techniques for evaluating noise in built environments have been developed with an emphasis on frequency characteristics. Sound communication (SC) curves have been proposed to evaluate office noise including that created by air conditioners (Beranek, 1956). Similar to SC curves, noise criteria (NC) curves have also been developed (Beranek, 1957). The curves include octave bands from 63 to 8,000 Hz and are still widely used. Balanced noise criterion (NCB) curves have been proposed as a modified version of NC curves that consider spectral imbalances (Beranek, 1989). Room criteria (RC) curves and the revised version, termed RC Mark II curves, have also been developed to assess the spectrum balance and low-frequency vibrations of noise produced by heating, ventilating, and air conditioning systems (Blazier, 1981; Blazier, 1997). Room noise criterion (RNC) curves, which have been proposed to fill the gap between NCB and RC, consider the effect of temporal variations in low frequency sounds (Schomer, 2000). Noise measurements and questionnaire surveys in offices indicated that, when compared with several other indices, including the NC, NCB, RC Mark II, and RNC, the A-weighted equivalent SPL (L_{Aeq}) is the best index for evaluating subjective auditory sensations (Ayr et al., 2003).

The proposed noise indices mainly focus on the energy of sounds in terms of the frequency characteristics. Considering the characteristics of the human auditory system and the results of a large number of psychoacoustic experiments, psychoacoustic factors, such as loudness, sharpness, and roughness have been proposed for evaluating noise (Zwicker and Fastl, 1999). Previous studies have evaluated the relationships between psychoacoustic factors and subjective responses to air conditioner noises in a built environment (Lee et al., 2017; Soeta and Shimokura, 2017; Lee and Wang, 2018; Lee and Wang, 2020) and a vehicle (Leite et al., 2009; Yoon et al., 2012; Nakasaki et al., 2013; Wagner et al., 2014; Soeta et al., 2016). The results indicated that psychoacoustic factors are significant predictors of subjective responses.

As with other psychoacoustic factors, autocorrelation function (ACF) and interaural cross-correlation (IACF) factors have been proposed based on the results of psychological and physiological experiments (Soeta and Ando, 2015). The results indicated that ACF factors were significantly correlated with subjective preference and annoyance ratings for air conditioner noises in a built environment (Soeta and Shimokura, 2017) and a vehicle (Soeta et al., 2016). Analytical approaches using the ACF and IACF are advantageous in that they are based on human cerebral function, describe basic temporal sensations, such as loudness and pitch (Ando, 2009; Soeta and Ando, 2015), and have predictive power that is equivalent to that of psychoacoustic factors (Soeta et al., 2016; Soeta and Shimokura, 2017).

Tonal noises generated by air conditioners can be annoying (Landström, et al., 1991; Landström, et al., 1994; Ryherda and Wang, 2008). Several indices, such as the prominence ratio and tonal audibility, have been proposed to quantify the prominence, or tonality, of a tone (Lee et al., 2017; Lee and Wang, 2018; Lee and Wang, 2020). The analytical approach using the ACF can be used to quantify the perception of tonality. The peak amplitude of the ACF, ϕ_1, is related to the bandwidth of a sound. The envelope decay of the ACF, τ_e, reflects the degree to which a sound has repetitive components.

The semantic differential method has been widely used to measure affective content (Osgood et al., 1957). In many cases, three main dimensions can be obtained regardless of the object type and the cultural background of the participants (Osgood, 1960). A systematic literature review confirmed that the three main dimensions of sound are *Evaluation*, *Potency*, and *Activity* (Ma et al., 2018). *Evaluation* refers to general human judgment, *Potency* is the degree of sensitivity to magnitude, and *Activity* is the sensation of the temporal and spectral patterns of a sound. When the three perceptual dimensions of air conditioner noise were extracted (Susini et al., 2004), they correlated with the spectral contents, subjective loudness, and spectral centroid.

The aim of this study was to determine the ACF and IACF factors that were most dominant in the subjective responses to air conditioner sounds. We dealt with three main perceptual dimensions of sound: loudness as *Potency*, sharpness as *Activity*, and annoyance as *Evaluation*. The ACF and IACF are analysis methods based on the processing of temporal patterns of neural activities in the auditory system (Cariani and Delgutte, 1996; Saberi et al., 1998). This method could be helpful in improving the sound quality of air conditioners during the manufacturing process because it can be used to obtain information about problematic noise pitches and the spectral centroid of noise.

METHODS

Analysis of Air Conditioner Sounds

We used a binaural microphone (BHS I, HEAD Acoustics) to measure sounds generated by three outlet units and one inlet unit of split-type air conditioners in an anechoic room. The number of compressor revolutions was set to 0–106 revolutions per second. The number of fan revolutions was set to 225–1,170 revolutions per minute. The outdoor unit was placed on the floor. The indoor unit was placed at a height of 1.8 m. The microphone was installed at a height of 1.6 m and a distance of 1.0 m for the outdoor unit and 3.3 m for the indoor unit. The experimental setup is shown in **Figure 1**. For all measurements, the generated sound was recorded via an analog-to-digital converter (SQuadrigaII, HEAD Acoustics) with a sampling rate of 48kHz and a resolution of 32 bits.

Factors determined from the ACF and IACF have been proposed for evaluating environmental noise and sound quality (Ando, 2009; Soeta and Ando, 2015). To determine the ACF and IACF factors in the present study, the normalized IACF of the signals recorded at the microphones representing the left and right ears, $p_l(t)$ and $p_r(t)$, respectively, as a function of the running step, s, was defined by

$$\phi_{lr}(\tau) = \phi_{lr}(\tau; s, T) = \left(\frac{\Phi_{lr}(\tau; s, T)}{\sqrt{\Phi_{lr}(\tau; s, T)\Phi_{lr}(\tau; s, T)}} \right) \tag{1}$$

where

$$\Phi_{lr}(\tau; s, T) = \frac{1}{2T} \int_{s-T}^{s+T} p_l'(t) p_r'(t+\tau) dt. \tag{2}$$

When the signal recorded at one ear was used, **Eq. 1** defines the normalized ACF. $2T$ is the integration interval and $p'(t) = p(t) * s_e(t)$. $s_e(t)$ is the ear sensitivity, which, for convenience, represents the impulse response of an A-weighted filter (Ando, 2009; Soeta and Ando, 2015). The normalized ACF and IACF were calculated using the geometric mean of the energy at s and the energy at $s+\tau$.

L_{Aeq} was determined from the A-weighted $p(t)$ as a function of the running step, s. L_{Aeq} was calculated using

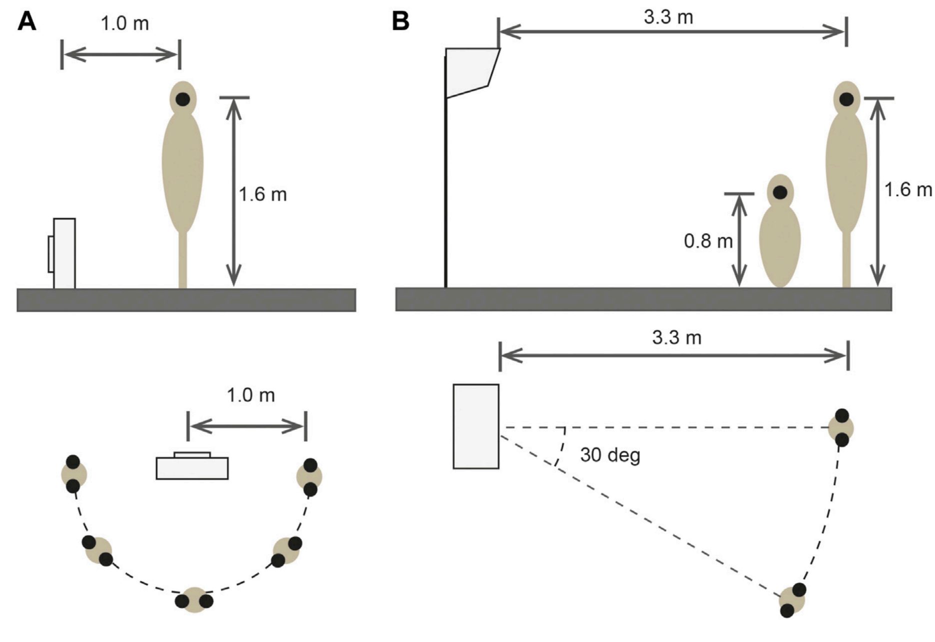


FIGURE 1 | Installation of **(A)** outlet and **(B)** inlet air conditioner units in an anechoic room. The binaural microphone was placed at the ear position on the mannequin.

$$L_{Aeq}(s,T) = 10\log\Phi_{ll(rr)}(0;s,T) \tag{3}$$

This indicates that the ACF includes L_{Aeq} as one of its factors. The other ACF factors determined from the normalized ACF are shown in **Figure 2A**. τ_1 is defined as the delay time of the first maximum peak and related to the perceived pitch. ϕ_1 is defined as the amplitude of the first maximum peak and related to the perceived pitch strength (Ando 2009; Soeta and Ando 2015). Higher values of τ_1 and ϕ_1 mean that the pitch of the sound is lower and stronger, respectively. The ϕ_1 value is related to the bandwidth of a sound and increases as the bandwidth of a sound narrows. The effective duration of the ACF, τ_e, was defined by the ten-percentile delay of the envelope of the normalized ACF and represents a repetitive component including the sound source itself (Ando, 2009). The τ_e values for a pure tone and white noise are ∞ and almost zero, respectively. Sharply filtered bandpass noises have been found to have larger τ_e values compared with loosely filtered bandpass noises (Soeta et al., 2004). The other ACF factor, the width of the first decay, $W_{\phi(0)}$, was defined using the delay time interval at a normalized ACF value of 0.5. $W_{\phi(0)}$ is equivalent to the spectral centroid (Soeta and Ando, 2015). Higher values of $W_{\phi(0)}$ mean that the sound contains more low frequency components.

The interaural cross-correlation coefficient (IACC) is linked to the subjective diffuseness and apparent source width (Ando, 2009), and was defined by.

$$IACC(s,\ T) = \left|\phi_{lr}(\tau;\ s,\ T)\right|_{\max},\ |\tau| \le 1\,[ms] \tag{4}$$

When the IACC is higher, a listener perceives a narrower sound image. The other IACF factors determined from the normalized IACF are shown in **Figure 2B**. τ_{IACC} is the interaural delay time at which IACC was defined and related to the sense of direction at low frequencies (Ando 2009). W_{IACC} is the width of the IACF defined by the interval of the delay time at a value of δ below the IACC. W_{IACC} depends on the frequency composition of the signals and is related to the apparent source width (Ando 2009).

To evaluate the noise characteristics both quantitatively and qualitatively, we calculated L_{Aeq}, τ_1, ϕ_1, $W_{\phi(0)}$, τ_e, IACC, τ_{IACC}, and W_{IACC} as a function of time. The integration interval, *2T*, was 500 ms and the running step, *s*, was 1 ms in all calculations. The analysis was performed with A Matlab-based analysis program (Mathworks, Natick, MA).

Subjective Assessments

Fifteen stimuli were selected from the measured air conditioner noise samples based on the distribution of the ACF and IACF factors. **Table 1** summarizes the mean ACF and IACF factors for the fifteen selected stimuli. **Figures 3**, **4** show the one-third octave band spectra and A-weighted sound pressure level with an integration time of 125 ms for the stimuli used in this study. The stimuli were presented to participants binaurally using a headphone amplifier (HDVD800, Sennheiser, Germany) and headphones (HD800, Sennheiser, Germany). Each stimulus was 2.0 s long and included a 0.1 s rise and fall ramp. Previous studies have indicated that participants can judge the loudness (Wright, 1947), sharpness (Hoechstetter et al., 2016), and annoyance (Hiramatsu et al., 1978) for sounds that are only 100 ms in duration. Thus, we considered 2 s to be sufficient for evaluating loudness, sharpness, and annoyance. The participants listened to the stimuli while sitting in a soundproof room with an ambient temperature of 22–25 degrees. All stimuli were presented

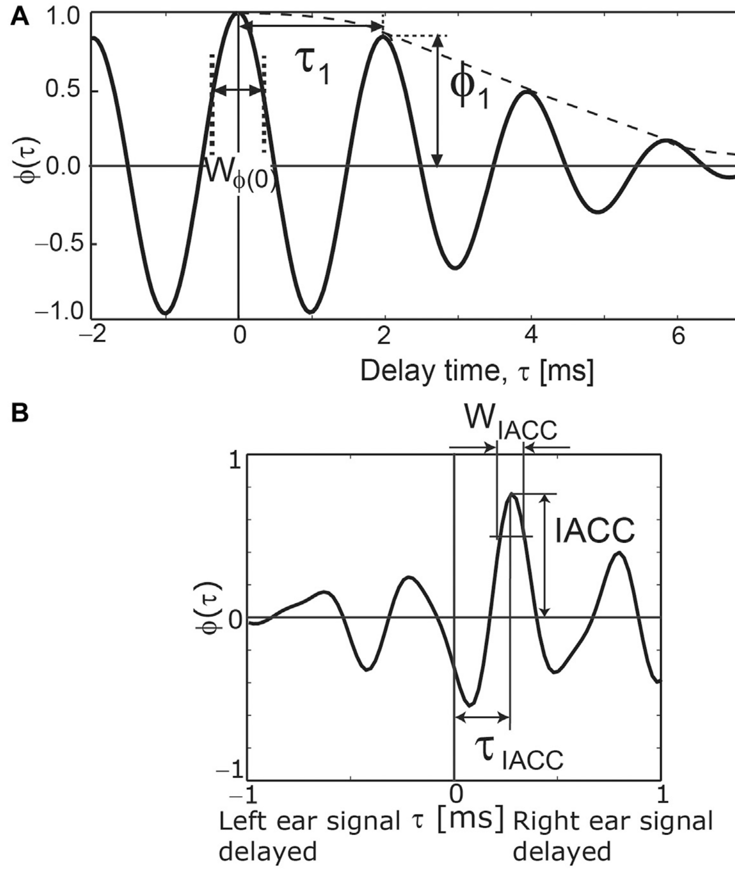


FIGURE 2 | (A) Definitions of the ACF factors, τ_1, ϕ_1, τ_e, and $W_{\phi(0)}$. **(B)** Definitions of the IACF factors, IACC, τ_{IACC}, and W_{IACC}.

TABLE 1 | Range of mean ACF and IACF values obtained from air-conditioner sounds used in the subjective assessments.

	L_{Aeq} (dB)	τ_1 (ms)	ϕ_1	τ_e (ms)	$W_{\phi(0)}$ (ms)	IACC	τ_{IACC} (ms)	W_{IACC} (ms)
Range	55.8–71.8	1.2–11.7	0.14–0.93	12.8–9,978.4	0.35–0.54	0.68–0.97	-0.07–0.10	0.15–0.23

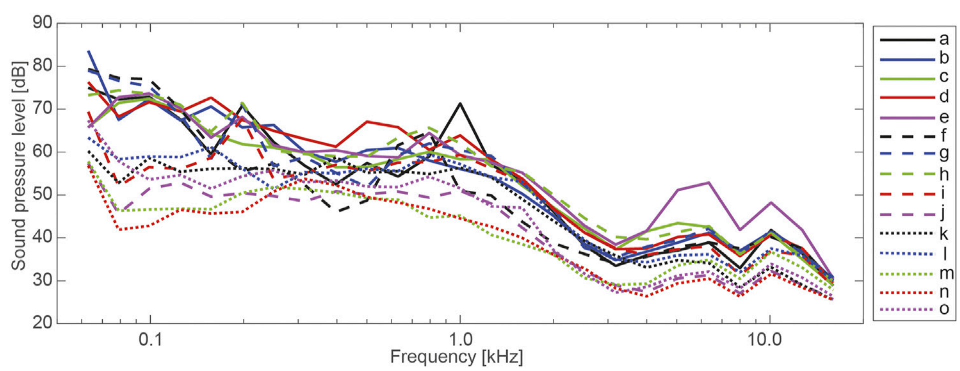


FIGURE 3 | Measured one-third octave band spectra for 15 sounds used in the subjective assessments.

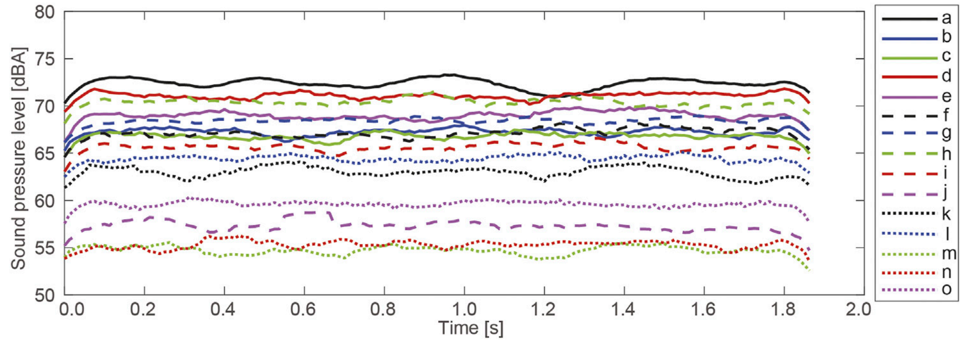


FIGURE 4 | Measured A-weighted sound pressure level [dB] with an integration time of 125 ms for 15 sounds used in the subjective assessments.

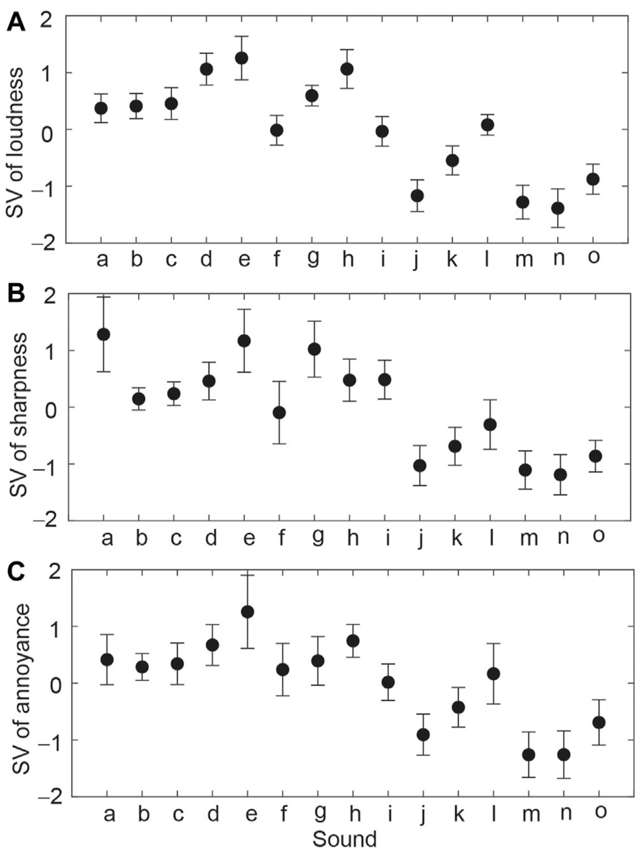


FIGURE 5 | Scale values of **(A)** loudness, **(B)** sharpness, and **(C)** annoyance for each participant. The symbols indicate the mean values and the error bars indicate the standard deviations.

at the same L_{Aeq} ± 0.2dB as the actual measured noises. L_{Aeq} was verified using a dummy head microphone (KU100, Neumann, Germany) and a sound calibrator (Type 4,231, Brüel and Kjær, Denmark).

We selected subjective loudness as *Potency*, subjective sharpness as *Activity*, and subjective annoyance as *Evaluation* to reflect the three perceptual dimensions (Ma et al., 2018).

TABLE 2 | Correlation coefficients among subjective loudness, annoyance, and sharpness.

	Loudness	Sharpness	Annoyance
Loudness	1.00		
Sharpness	0.74**	1.00	
Annoyance	0.83**	0.77**	1.00

Subjective loudness, sharpness, and annoyance caused by air conditioner sounds were evaluated to clarify the effects of the ACF and IACF factors on each subjective response. Participants between 20 and 54 years of age (median age of 23.0 years) with normal hearing and no history of neurological diseases took part in the experiments. Fifteen participants (11 men) took part in the sharpness and annoyance experiment. Eight out of the fifteen (6 men) participated in the loudness experiment. Seven participants (4 men) took part in the loudness experiment only. According to our previous studies, we considered the involvement of at least ten participants to be necessary to ensure sufficient statistical power (Soeta et al., 2016; Soeta and Shimokura 2017; Soeta and Kagawa 2020). The normality of the scale values of loudness, sharpness, and annoyance was tested using the Shapiro-Wilk test (Shapiro and Wilk 1965). The results indicated scale values of loudness, sharpness, and annoyance except for one stimulus (f) were normally distributed. Informed consent was obtained from each participant after the key elements of the study was explained. The study protocol was approved by the ethics committee of the National Institute of Advanced Industrial Science and Technology (AIST) of Japan.

In Scheffé's method (Scheffé, 1952), one combination is assigned to each participant for comparison. In the modified Scheffé's method, a pairwise comparison is performed between one iteration with one participant and another iteration with a different participant (Sato, 1985; Nagasawa, 2002). In our experiment, all combinations of pairs (i.e., 105 pairs (N(N−1)/2, N = 15) were presented in random order for each participant, and the presentation order within each pair was randomized. The silent interval between the stimuli was 1.0 s long. After the presentation of each pair, the participants were asked to judge which stimulus from each pair was louder, sharper, or more

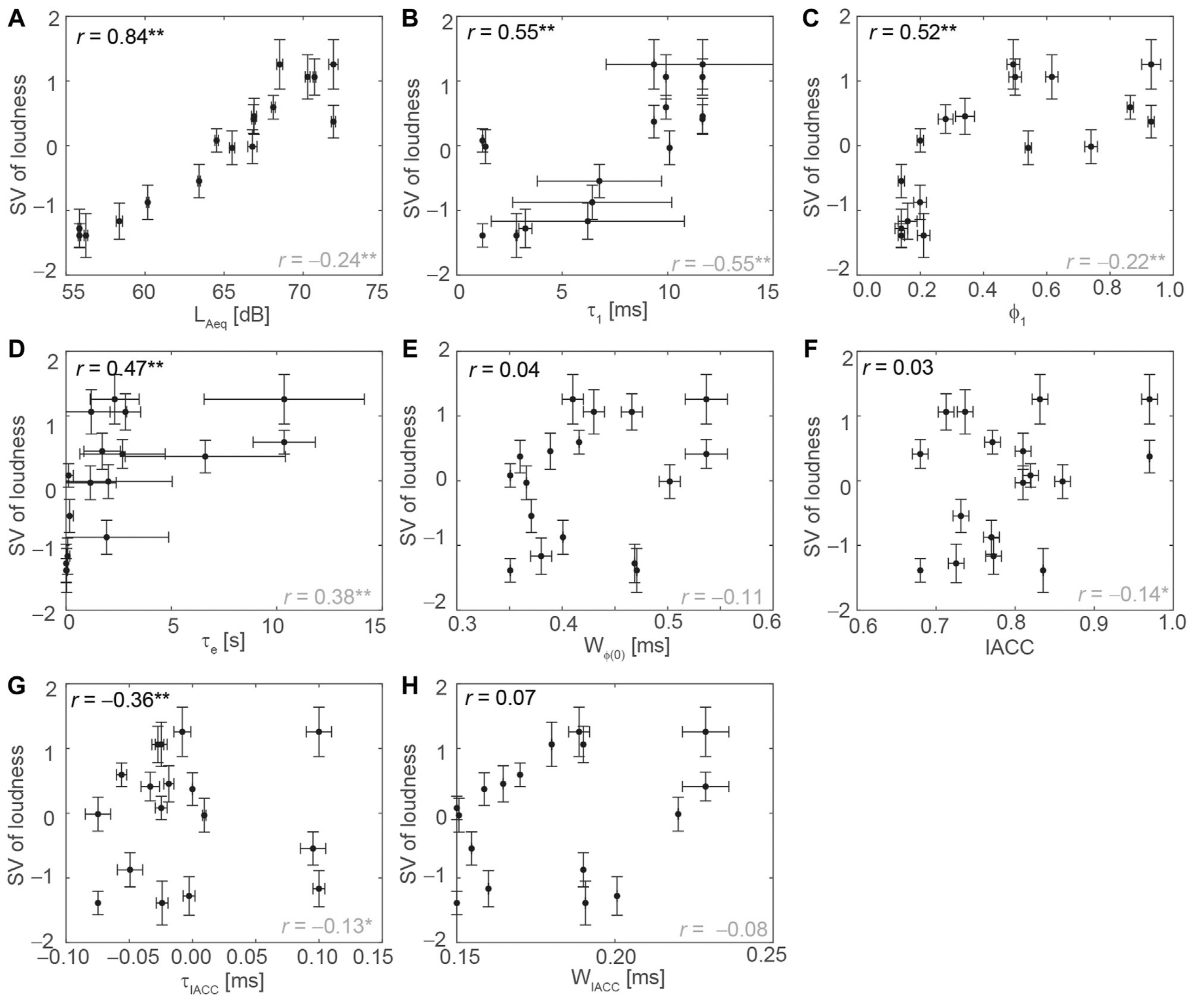


FIGURE 6 | Relationship between correlation parameters and the scale values of loudness. **(A)** L_{Aeq}, **(B)** τ_1, **(C)** ϕ_1, **(D)** τ_e, **(E)** $W_{\phi(0)}$, **(F)** IACC, **(G)** τ_{IACC}, and **(H)** W_{IACC}. The symbols indicate the mean values and the error bars indicate the SDs. The correlation coefficients of the mean and SDs are shown in black and gray, respectively. Asterisks represent the level of significance, i.e., $^{**}p < 0.01$, $^{*}p < 0.05$.

annoying using a seven-point scale. Judgements were made using one of seven statements. For example, in the case of loudness, participants were asked to select one of the following: I perceived sound i as strongly louder than sound j (3 points); I perceived i as moderately louder than j (2 points); I perceived i as slightly louder than j (1 point); I perceived the loudness of the two sounds to be equal (0 point); I perceived j as slightly louder than i (−1 point); I perceived j as moderately louder than i (−2 points); I perceived j as strongly louder than i (−3 points). The averaged values were calculated and defined as scale values (SVs) of loudness. An analysis of variance (ANOVA) was then carried out on the results of the paired comparison experiments (Sato, 1985; Nagasawa, 2002).

To calculate the effects of ACF and IACF characteristics on participant loudness, sharpness, and annoyance, multiple regression analyses were carried out using a linear combination of the mean ACF and IACF factors and their standard deviations (SDs) as predictive variables. The outcome variables were the SVs of loudness, sharpness, and annoyance for all participants. Stepwise selection of the predictive variables was applied by successively adding or removing variables. The step criteria applied for entry and removal were based on the statistical significance level of the F-value, which was set at 0.05 and 0.10, respectively. Predictive variables with a variance inflation factor of 3.0 or more were excluded to avoid multicollinearity. The analyses were performed with SPSS software (SPSS version 22.0, IBM Corp., NY).

RESULTS

The ANOVA for the scale values revealed that the main effect (i.e., the differences between the stimuli) was statistically significant

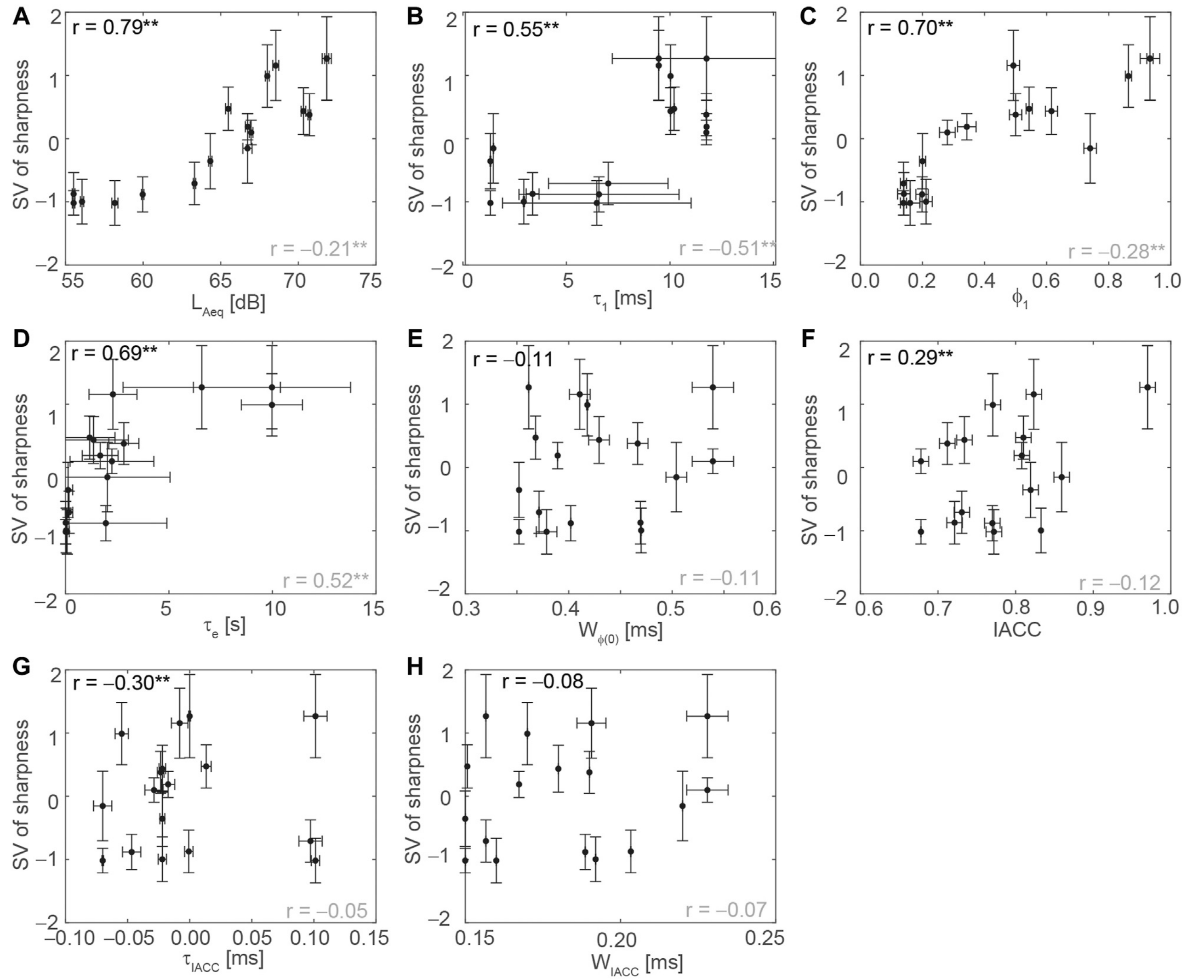


FIGURE 7 | Relationship between correlation parameters and the scale values of sharpness. **(A)** L_{Aeq}, **(B)** τ_1, **(C)** ϕ_1, **(D)** τ_e, **(E)** $W_{\phi(0)}$, **(F)** IACC, **(G)** τ_{IACC}, and **(H)** W_{IACC}. The symbols indicate the mean values and the error bars indicate the SDs. The correlation coefficients of the mean and SDs are shown in black and gray, respectively. Asterisks represent the level of significance, i.e., $^{**}p < 0.01$, $^{*}p < 0.05$.

(F (14, 2,834) = 581.22, $p < 0.001$, for loudness, F (14, 2,834) = 586.84, $p < 0.001$, for sharpness, and F (14, 2,834) = 390.80, $p < 0.001$, for annoyance). **Figure 5** shows the scale values for loudness, sharpness, and annoyance. Loudness and annoyance exhibited a similar tendency. The correlation coefficients among loudness, sharpness, and annoyance are shown in **Table 2**. Loudness was highly correlated with sharpness and annoyance although they are proposed as independent psychological dimensions (Ma et al., 2018). This might have been caused by the relatively narrow range of physical parameters produced by a small number of air conditioners (three outlet units and one inlet unit).

Figures 6–8 show the relationship between each ACF/IACF factor and loudness, sharpness, and annoyance scores, respectively. Scale values of loudness were highly correlated with L_{Aeq} (r = 0.84, $p < 0.01$), τ_1 (r = 0.55, $p < 0.01$), ϕ_1 (r = 0.52, $p < 0.01$), and the SD of τ_1 (r = –0.55, $p < 0.01$). Scale values of sharpness were highly positively correlated with L_{Aeq} (r = 0.79, $p < 0.01$), ϕ_1 (r = 0.70, $p < 0.01$), and τ_e (r = 0.69, $p < 0.01$). Scale values of annoyance were highly positively correlated with L_{Aeq} (r = 0.79, $p < 0.01$) and ϕ_1 (r = 0.52, $p < 0.01$).

The final models of the multiple linear regression analysis and the standardized partial regression coefficients were as follows:

$$SV_{loudness} \approx a_1 + 0.82L_{Aeq} - 0.24IACC - 0.23SD\_L_{Aeq} - 0.14SD\_\tau_1 \quad (5)$$

$$SV_{sharpness} \approx a_2 + 0.52L_{Aeq} + 0.28\tau_e - 0.14\,W_\phi(0) - 0.19SD\_\tau_1 - 0.17SD\_IACC \quad (6)$$

$$SV_{annoyance} \approx a_3 + 0.79L_{Aeq} - 0.20SD\_L_{Aeq} \quad (7)$$

The model was statistically significant (F (4, 220) = 269.97, $p < 0.001$, for loudness (F (5, 219) = 121.80, $p < 0.001$, for sharpness, F

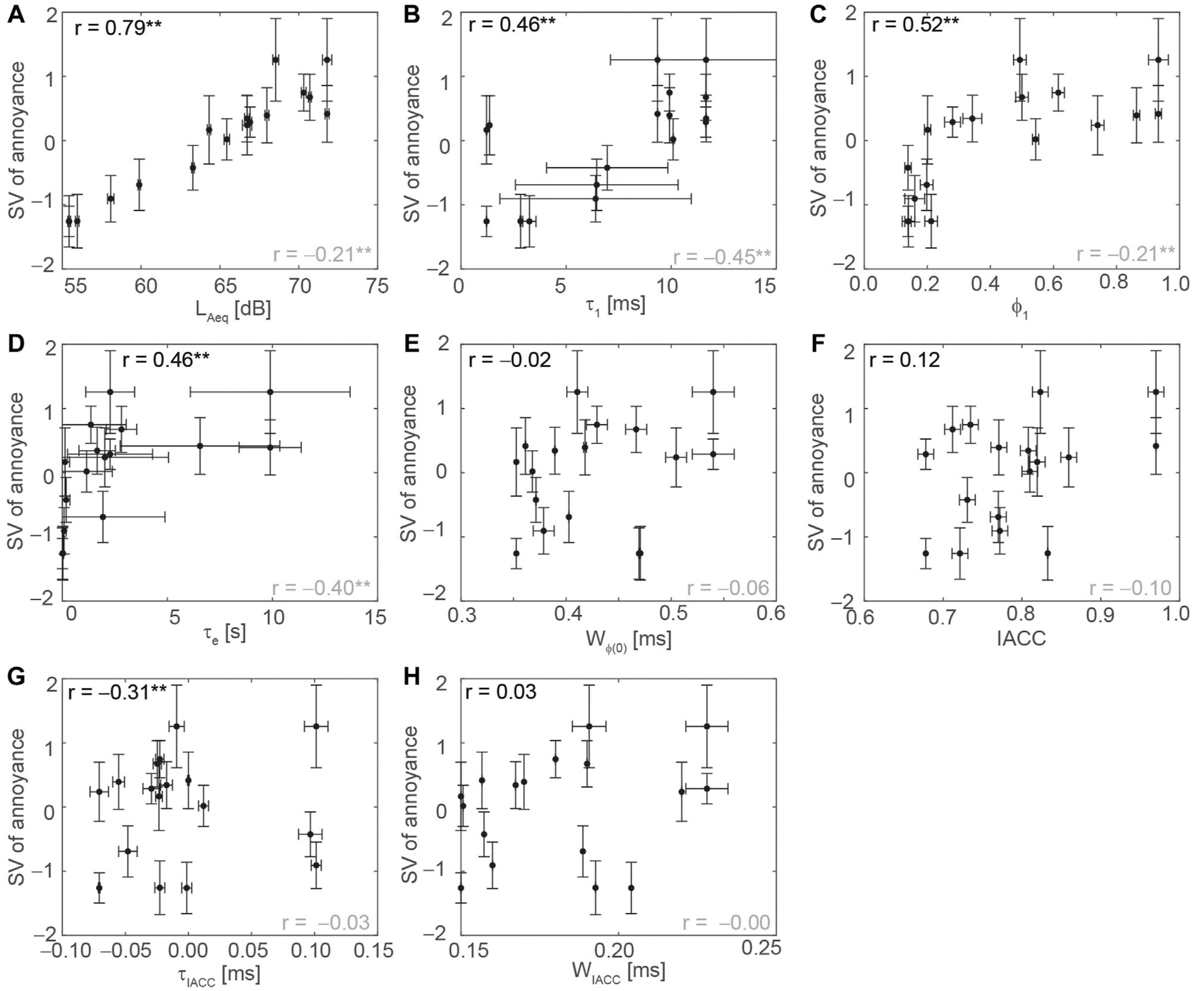


FIGURE 8 | Relationship between correlation parameters and the scale values of annoyance. **(A)** L_{Aeq}, **(B)** τ_1, **(C)** ϕ_1, **(D)** τ_e, **(E)** $W_{\phi(0)}$, **(F)** IACC, (G) τ_{IACC}, and **(H)** W_{IACC}. The symbols indicate the mean values and the error bars indicate the SDs. The correlation coefficients of the mean and SDs are shown in black and gray, respectively. Asterisks represent the level of significance, i.e., $^{**}p < 0.01$, $^{*}p < 0.05$.

(2, 222) = 224.49, $p < 0.001$, for annoyance), and the adjusted coefficient of determination, R^2, was 0.83 for loudness, 0.73 for sharpness, and 0.67 for annoyance.

DISCUSSION

L_{Aeq} has been found to be a consistently significant factor influencing annoyance of air conditioner sounds (Ayr et al., 2001; Ayr et al., 2003). The multiple linear regression analysis showed that the energy-index of L_{Aeq} was the significant factor influencing the perception of loudness, sharpness, and annoyance of air conditioner sounds. The regression coefficients were all positive, suggesting that higher L_{Aeq} values are associated with louder, sharper, and more annoying sounds. A previous study indicated that ϕ_1 was a significant factor and L_{Aeq} was not a significant factor influencing annoyance (Soeta and Shimokura, 2017), which is not consistent with the present finding. A possible reason for this discrepancy might be the differing L_{Aeq} range between the two studies. Specifically, the present study had a higher and broader range of L_{Aeq} values. The effect of L_{Aeq} may have been much greater than that of ϕ_1 in the present study.

The temporal variation in the energy-index of L_{Aeq}, denoted as the SD of L_{Aeq}, was also a significant factor in predicting loudness and annoyance. This is consistent with previous findings regarding loudness (Soeta and Kagawa, 2020) and annoyance (Fujii et al., 2002; Sato et al., 2007; Jeon and Sato, 2008; Gille et al., 2017; Soeta and Kagawa, 2020), and confirms that not only L_{Aeq}, but also the temporal variation of L_{Aeq}, has a large influence on subjective response. Although the partial coefficients for the SD of L_{Aeq} were positive in previous studies, they were negative in this study. Further, the SDs of L_{Aeq} were much smaller than those in previous studies (Fujii et al., 2002; Jeon and Sato, 2008; Gille et al., 2017; Soeta and Kagawa, 2020). The large differences in temporal variation might have an influence on the subjective responses.

The sharpness of a sound is determined by the balance of high frequency and low frequency components (Zwicker and Fastl, 1999), such that sounds with more high frequency components are perceived to be sharper. We expected that the $W_{\phi(0)}$ would be negatively correlated with subjective sharpness in the present study, and found this to be the case. This indicates that $W_{\phi(0)}$ is a significant predictor of characteristics in the *Activity* dimension, which is consistent with previous findings regarding airplane noise (Soeta and Kagawa, 2020). The ACF factor, τ_e, shows the degree to which a sound has repetitive components. In this study, τ_e was a significant factor in predicting sharpness with a positive partial coefficient, suggesting that the sharpness of the frequency bandwidth might determine whether sounds are perceived as sharp.

The binaural index, IACC, was a significant factor in predicting loudness, with a negative regression coefficient. This suggests that air conditioner sounds with lower IACC values, which have wider sound images (Ando, 2009), could be perceived as louder. This is consistent with the previous findings regarding airplane noise (Soeta and Kagawa, 2020). In addition, previous studies have indicated that IACC is a significant predictor of annoyance for floor impact sounds (Jeon and Sato, 2008; Jeon et al., 2009). This suggests that IACC could be a significant predictor of subjective evaluations of sounds.

CONCLUSION

We analyzed multidimensional psychological responses to air conditioner sounds to determine the factors that significantly influence subjective perceptions of loudness, sharpness, and annoyance in this context. The results indicated that the L_{Aeq}, τ_1, and the temporal variation of τ_1, among other factors, significantly influenced subjective responses. This indicates that factors influencing the ACF and IACF are useful indices for the evaluation of air conditioner sounds.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The ethics committee of the National Institute of Advanced Industrial Science and Technology (AIST). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Both authors have been involved in the design, experiment, and analysis of this study.

FUNDING

This work was partly supported by a Grant-in-Aid for Scientific Research (B) (Grant No. 18H03324) from the Japan Society for the Promotion of Science.

ACKNOWLEDGMENTS

We thank Sydney Koke, MFA, from Edanz Group (https://en-author-services.edanz.com/ac) for editing a draft of this manuscript.

REFERENCES

Ando, Y. (2009). *Auditory and Visual Sensations*. New York, NY: Springer. doi:10.1007/b13253

Ayr, U., Cirillo, E., and Martellotta, F. (2001). An Experimental Study on Noise Indices in Air Conditioned Offices. *Appl. Acoust.* 62, 633–643. doi:10.1016/S0003-682X(00)00072-4

Ayr, U., Cirillo, E., Fato, I., and Martellotta, F. (2003). A New Approach to Assessing the Performance of Noise Indices in Buildings. *Appl. Acoust.* 64, 129–145. doi:10.1016/S0003-682X(02)00075-0

Beranek, L. L. (1956). Criteria for Office Quieting Based on Questionnaire Rating Studies. *J. Acoust. Soc. America* 28, 833–852. doi:10.1121/1.39824310.1121/1.1908489

Beranek, L. L. (1957). Revised Criteria for Noise in Buildings. *Noise Control* 3, 19–27. doi:10.1121/1.2369239

Beranek, L. L. (1989). Balanced Noise-criterion (NCB) Curves. *The J. Acoust. Soc. Am.* 86, 650–664. doi:10.1121/1.398243

Blazier, W. E. (1981). Revised Noise Criterion for Application in the Acoustical Design and Rating of HVAC Systems. *Noise Control. Eng. J.* 16, 64–73. doi:10.3397/1.2832172

Blazier, W. E. (1997). RC Mark II: A Refined Procedure for Rating the Noise of Heating, Ventilating, and Air Conditioning (HVAC) Systems in Buildings. *Noise Control. Eng. J.* 45, 243–250. doi:10.3397/1.2828446

Cariani, P. A., and Delgutte, B. (1996). Neural Correlates of the Pitch of Complex Tones. I. Pitch and Pitch Salience. *J. Neurophysiol.* 76, 1698–1716. doi:10.1152/jn.1996.76.3.1698

Fujii, K., Atagi, J., and Ando, Y. (2002). Temporal and Spatial Factors of Traffic Noise and its Annoyance. *J. Temp Des. Archit Environ.* 2, 33–41.

Gille, L.-A., Marquis-Favre, C., and Weber, R. (2017). Aircraft Noise Annoyance Modeling: Consideration of Noise Sensitivity and of Different Annoying Acoustical Characteristics. *Appl. Acoust.* 115, 139–149. doi:10.1016/j.apacoust.2016.08.022

Hiramatsu, K., Takagi, K., Yamamoto, T., and Ikeno, J. (1978). The Effect of Sound Duration on Annoyance. *J. Sound Vibration* 59, 511–520. doi:10.1016/S0022-460X(78)80131-X

Hoechstetter, M., SautterGabbert, J.-M. U., Gabbert, U., and Verhey, J. L. (2016). Role of the Duration of Sharpness in the Perceived Quality of Impulsive Vehicle Sounds. *Acta Acustica united with Acustica* 102, 119–128. doi:10.3813/AAA.918929

Hongisto, V., Saarinen, P., and Oliva, D. (2019). Annoyance of Low-Level Tonal Sounds - A Penalty Model. *Appl. Acoust.* 145, 358–361. doi:10.1016/j.apacoust.2018.09.023

Jeon, J. Y., and Sato, S. (2008). Annoyance Caused by Heavyweight Floor Impact Sounds in Relation to the Autocorrelation Function and Sound Quality Metrics. *J. Sound Vibration* 311, 767–785. doi:10.1016/j.jsv.2007.09.029

Jeon, J. Y., Lee, P. J., Kim, J. H., and Yoo, S. Y. (2009). Subjective Evaluation of Heavy-Weight Floor Impact Sounds in Relation to Spatial Characteristics. *J. Acoust. Soc. Am.* 125, 2987–2794. doi:10.1121/1.3081390

Landström, U., Kjellberg, A., Söderberg, L., and Nordström, B. (1991). The Effects of Broadband, Tonal and Masked Ventilation Noise on Performance, Wakefulness and Annoyance. *J. Low Frequency Noise, Vibration Active Control.* 10, 112–122. doi:10.1177/026309239101000402

Landström, U., Kjellberg, A., SÖDerberg, L., and Nordström, B. (1994). Measures against Ventilation Noise - Which Tone Frequencies Are Least and Most Annoying? *J. Low Frequency Noise, Vibration Active Control.* 13, 81–88. doi:10.1177/026309239401300301

Lee, J., and Wang, L. M. (2018). Development of a Model to Predict the Likelihood of Complaints Due to Assorted Tone-In-Noise Combinations. *J. Acoust. Soc. Am.* 143, 2697–2707. doi:10.1121/1.5036731

Lee, J., and Wang, L. M. (2020). Investigating Multidimensional Characteristics of Noise Signals with Tones from Building Mechanical Systems and Their Effects on Annoyance. *J. Acoust. Soc. Am.* 147, 108–124. doi:10.1121/10.0000487

Lee, J., Francis, J. M., and Wang, L. M. (2017). How Tonality and Loudness of Noise Relate to Annoyance and Task Performance. *Noise Cont. Eng. J.* 65, 71–82. doi:10.3397/1/376427

Leite, R. P., Paul, S., and Gerges, S. N. Y. (2009). A Sound Quality-Based Investigation of the HVAC System Noise of an Automobile Model. *Appl. Acoust.* 70, 636–645. doi:10.1016/j.apacoust.2008.06.010

Ma, K. W., Wong, H. M., and Mak, C. M. (2018). A Systematic Review of Human Perceptual Dimensions of Sound: Meta-Analysis of Semantic Differential Method Applications to Indoor and Outdoor Sounds. *Building Environ.* 133, 123–150. doi:10.1016/j.buildenv.2018.02.021

Nagasawa, S. Y. (2002). Improvement of the Scheffé's Method for Paired Comparisons. *KANSEI Eng. Int.* 3, 47–56. doi:10.5100/jje.27.Supplement_22810.5057/kei.3.3_47

Nakasaki, R., Hasegawa, H., Kasuga, M., Kobayashi, R., Hasegawa, H., and Kasuga, M. (2013). Evaluation of Air-Conditioning Sounds in a Vehicle to Determine Thermal Feelings Using Psychoacoustic Parameters. *Acoust. Sci. Tech.* 34, 159–165. doi:10.1250/ast.34.159

Oliva, D., Hongisto, V., Hongisto, V., and Haapakangas, A. (2017). Annoyance of Low-Level Tonal Sounds - Factors Affecting the Penalty. *Building Environ.* 123, 404–414. doi:10.1016/j.buildenv.2017.07.017

Osgood, C. E., Suci, G., and Tannenbaum, P. H. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.

Osgood, C. E. (1960). The Cross-Cultural Generality of Visual-Verbal Synesthetic Tendencies. *Behav. Sci.* 5, 146–169.

Ryherd, E. E., and Wang, L. M. (2008). Implications of Human Performance and Perception under Tonal Noise Conditions on Indoor Noise Criteria. *J. Acoust. Soc. America* 124, 218–226. doi:10.1121/1.2932075

Saberi, K., Takahashi, Y., Konishi, M., Albeck, Y., Arthur, B. J., and Farahbod, H. (1998). Effects of Interaural Decorrelation on Neural and Behavioral Detection of Spatial Cues. *Neuron* 21, 789–798. doi:10.1016/S0896-6273(00)80595-4

Sato, S., You, J., and Jeon, J. Y. (2007). Sound Quality Characteristics of Refrigerator Noise in Real Living Environments with Relation to Psychoacoustical and Autocorrelation Function Parameters. *J. Acoust. Soc. Am.* 122, 314–325. doi:10.1121/1.2739440

Sato, S. (1985). *Statistical Method of Sensory Test*. Tokyo, Japan: JUSE Press.

Scheffé, H. (1952). An Analysis of Variance for Paired Comparisons. *J. Am. Stat. Assoc.* 47, 381–400. doi:10.2307/2281310

Schomer, P. D. (2000). Proposed Revisions to Room Noise Criteria. *Noise Control. Eng. J.* 48, 85–96. doi:10.3397/1.2827970

Shapiro, S. S., and Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52, 591–611. doi:10.1093/biomet/52.3-4.591

Soeta, Y., and Ando, Y. (2015). *Neurally Based Measurement and Evaluation of Environmental Noise*. Tokyo, Japan: Springer. doi:10.1007/978-4-431-55432-5

Soeta, Y., and Kagawa, H. (2020). Three Dimensional Psychological Evaluation of Aircraft Noise and Prediction by Physical Parameters. *Building Environ.* 167, 106445. doi:10.1016/j.buildenv.2019.106445

Soeta, Y., and Shimokura, R. (2017). Sound Quality Evaluation of Air-Conditioner Noise Based on Factors of the Autocorrelation Function. *Appl. Acoust.* 124, 11–19. doi:10.1016/j.apacoust.2017.03.015

Soeta, Y., Maruo, T., and Ando, Y. (2004). Annoyance of Bandpass-Filtered Noises in Relation to the Factor Extracted from Autocorrelation Function (L). *J. Acoust. Soc. Am.* 116, 3275–3278. doi:10.1121/1.1782931

Soeta, Y., Nakagawa, S., Kamiya, Y., and Kamiya, M. (2016). Subjective Preference for Air-Conditioner Sounds inside a Car in Summer and Winter. *J. Ergon.* 6, 170. doi:10.4172/2165-7556.1000170

Susini, P., McAdams, S., Winsberg, S., Perry, I., Vieillard, S., and Rodet, X. (2004). Characterizing the Sound Quality of Air-Conditioning Noise. *Appl. Acoust.* 65, 763–790. doi:10.1016/j.apacoust.2004.02.003

Tang, S. K., and Wong, M. Y. (2004). On Noise Indices for Domestic Air Conditioners. *J. Sound Vibration* 274, 1–12. doi:10.1016/S0022-460X(03)00567-4

Wagner, V., Enigk, H., Beitz, T., and Kallus, K. W. (2014). Subjective and Objective Evaluation of the Air Conditioning Sound. *J. Ergon.* 4, 131. doi:10.4172/2165-7556.1000131

Wright, H. N. (1947). The Growth of Auditory Sensation. *J. Acoust. Soc. Am.* 19, 584–591. doi:10.1121/1.1916525

Yoon, J.-H., Yang, I.-H., Jeong, J.-E., Park, S.-G., and Oh, J.-E. (2012). Reliability Improvement of a Sound Quality Index for a Vehicle HVAC System Using a Regression and Neural Network Model. *Appl. Acoust.* 73, 1099–1103. doi:10.1016/j.apacoust.2012.02.018

Zwicker, E., and Fastl, H. (1999). *Psychoacoustics: Facts and Models*. Berlin, Germany: Springer. doi:10.1007/978-3-662-09562-1 CrossRef Full Text

Conflict of Interest: EO was employed by Mitsubishi Heavy Industries.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 13 May 2021
doi: 10.3389/fbuil.2021.676009

Junk Food or Haute Cuisine to the Ear? – Investigating the Relationship Between Room Acoustics, Soundscape, Non-Acoustical Factors, and the Perceived Quality of Restaurants

Jochen Steffens[1,2]*, *Tobias Wilczek*[2] *and Stefan Weinzierl*[2]

[1]*Düsseldorf University of Applied of Sciences, Institute of Sound and Vibration Engineering (ISAVE), Düsseldorf, Germany,* [2]*Technische Universität Berlin, Audio Communication Group, Berlin, Germany*

OPEN ACCESS

Edited by:
Arianna Astolfi, Politecnico di Torino, Italy

Reviewed by:
Catherine Guastavino, McGill University, Canada
Jooyoung Hong, Chungnam National University, South Korea
Jin Yong Jeon, Hanyang University, South Korea

***Correspondence:**
Jochen Steffens
jochen.steffens@hs-duesseldorf.de

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *04 March 2021*
Accepted: *27 April 2021*
Published: *13 May 2021*

Citation:
Steffens J, Wilczek T and Weinzierl S (2021) Junk Food or Haute Cuisine to the Ear? – Investigating the Relationship Between Room Acoustics, Soundscape, Non-Acoustical Factors, and the Perceived Quality of Restaurants. Front. Built Environ. 7:676009. doi: 10.3389/fbuil.2021.676009

Sound and music are well-studied aspects of the quality of experience in restaurants; the role of the room acoustical conditions, their influence on the visitors' soundscape evaluation and their impact on the overall customer satisfaction in restaurants, however, has received less scientific attention. The present field study therefore investigated whether sound pressure level, reverberation time, and soundscape pleasantness can predict factors associated with overall restaurant quality. In total, 142 persons visiting 12 restaurants in Berlin rated relevant acoustical and non-acoustical factors associated with restaurant quality. Simultaneously, the A-weighted sound pressure level ($L_{A,eq,15}$) was measured, and the reverberation time in the occupied state ($T_{20,occ}$) was obtained by measurements performed in the unoccupied room and a subsequent calculation of the occupied condition according to DIN 18041. Results from linear mixed-effects models revealed that both the $L_{A,eq,15}$ and $T_{20,occ}$ had a significant influence on soundscape pleasantness and eventfulness, whereby the effect of $T_{20,occ}$ was meditated by the $L_{A,eq,15}$. Also, the $L_{A,eq,15}$ as well as soundscape pleasantness were significant predictors of overall restaurant quality. A comprehensive structural equation model including both acoustical and non-acoustical factors, however, indicates that the effect of soundscape pleasantness on overall restaurant quality is mediated by the restaurant's atmosphere. Our results support and extend previous findings which suggest that the acoustical design of restaurants involves a trade-off between comfort and liveliness, depending on the desired character of the place.

Keywords: comfort, gastronomy acoustics, soundscape, room acoustics, reverberation time, restaurant quality, restaurant acoustics, noise

INTRODUCTION

Anecdotal evidence suggests that the acoustic conditions in many catering establishments are problematic, particularly due to high noise levels. Indeed, a recent survey among 13,000 Americans identified noise as the most bothersome irritation in restaurants across the US (24%), followed by poor service (23%), high prices (12%), and parking problems (10%) (Herklots, 2018). This finding is corroborated by studies demonstrating detrimental effects of unwanted sound on mood, overall satisfaction and intended revisit (Novak et al., 2010) as well as on the ability to communicate and the actual behavior in the room (Navarro and Pimentel, 2007; Meng et al., 2018). That is, restaurant guests offended by noise levels are more likely to leave prematurely, and less likely to return to that venue. While owners and managers often seem to be aware of such effects, they are reluctant to use room acoustics measures not only because of the costs but because they fear they will compromise their restaurants' liveliness by room-acoustical modifications. Or as simply put by a chef-entrepreneur: 'the second worst thing to a restaurant that is too noisy is a restaurant that is too quiet.' (Lindborg, 2016, p. 309). These concerns are understandable, given the positive effects of particular soundscapes, for instance including musical, natural, or human sounds, or the masking effect of background sounds, ensuring the privacy of one's conversation in a restaurant (e.g., Astolfi and Filippi, 2004; Tarlao et al., 2021). In this paper, we therefore aim to clarify further the contribution of (room-) acoustical parameters, soundscape evaluation, and non-acoustical factors to overall quality evaluation of restaurants.

As noted by Lindborg (2015), restaurants are vibrant social places whose design is subject to competing requirements. Here, various practical demands and decisions driven by visual design can lead to room-acoustical disadvantages. For example, the need to attract customers can lead to large windows toward the street, increasing the reverberation time. This increase can lead to a decrease of speech intelligibility and thus to a perceived need to raise one's voice, resulting in an overall loudness increase in the restaurant known as the Lombard effect (Lombard, 1911; Junqua et al., 1999). Similar effects can be attributed to the need to give an impression of cleanliness, leading to hard and acoustically reflective floors and tabletop materials (Lindborg, 2015).

The acoustical quality requirements in rooms differ depending on a room's use and are set in national standards such as the German DIN 18041 (DIN German Institute for Standardization, 2016; Nocke, 2017). For example, the requirements for classrooms including hearing-impaired students have been set high, whereas short-term stay areas, such as restaurants, have not. However, the recommendations for the acoustic design of gastronomic rooms are sometimes far apart. For example, while the DIN 18041 suggests a reverberation time of 0.85 s for a restaurant with a volume of 300 m^2 and a ceiling height of 2.8 m, Rindel (2018) recommends a much shorter reverberation time of 0.38 s for the same room.

Surprisingly, research in the hospitality industry has paid scant attention to the acoustics of restaurants and its perception (Ryu and Han, 2011; Ponnam and Balaji, 2014). Studies dealing with the influence of the acoustic environment on patrons have predominantly considered music as a relevant factor (e.g., Caldwell and Hibbert, 2002); only a few studies also considered a possible influence of ambient noise (Antun, et al., 2010; Bitner, 1992). One study explicitly focusing on acoustical comfort in restaurants surveyed 11 restaurants and 825 patrons, and obtained acoustical comfort by the four parameters Privacy, Comfort, Quietude and Communication and also measured reverberation time and background noise level as acoustical factors (Battaglia, 2014). The results show that reverberation time predicted each one of the four comfort parameters; also, comfort was predicted by the sound pressure level. Battaglia (2014) concludes that reverberation times of 0.5–0.7 s are within the optimal range for perceived acoustical comfort. However, his study did not include reverberation times under 0.5 s, so the question remains whether a further lowering of the reverberation time might increase acoustical comfort.

In general, findings from consumer and soundscape research and music and environmental psychology suggest a high potential of soundscape design on food taste, atmosphere, and overall restaurant quality. For example, North (2012) showed that the taste of wine as rated by restaurant visitors can reflect background music's emotional connotation. Another study by Yan and Dando (2015) let participants taste multiple concentrations of solutions of five prototypic tastants, during conditions with or without broad-spectrum auditory stimulation, simulating airline cabin noise. Their results revealed that sweetness was rated significantly lower under noisy conditions, while no difference in intensity rating was observed for salty, sour, and bitter tastants. In contrast, umami was rated higher under noisy conditions. Further sound-taste correspondences were observed by Crisinel et al. (2012), who found that identical cinder toffees were rated significantly more bitter, while listening to a soundtrack connoted with bitterness than when listening to a soundtrack connoted with sweetness.

Regarding the atmospheric effect of sound, North and Hargreaves (1998) observed a positive effect of background music on a student cafeteria's general atmosphere. For instance, playing classical music positively affected visitors' willingness to spend more money on the products offered. Caldwell and Hibbert (2002) tested the effect of the tempo and preference of music being played on patrons' behaviors and found that the enjoyment of the music positively predicted total money spent, enjoyment of dining, and intention to return and to recommend; by contrast, musical tempo did not show any significant effects.

Concerning the relative contribution of various influencing factors on overall restaurant evaluations, the Attribute-Value theory assumes that consumers rate services such as a restaurant meal in terms of a set of attributes (Kassarjian and Robertson, 1991). Each of these attributes has a certain level of importance to the customer, which can vary considerably by market segment. For example, some consumers might be attracted by a restaurant's low price, while others by a restaurant's upscale image and its food quality. The overall value is weighed up according to the individual importance of attributes and finally leads to

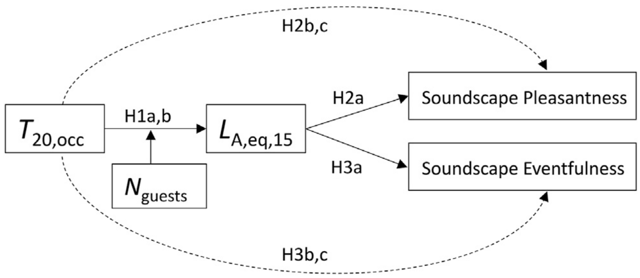


FIGURE 1 | Hypothesized relationships between object- and subject-centred acoustical variables obtained in our study.

deciding which restaurant is chosen by the consumer (Johns and Pine, 2002).

However, within the foodservice research, there is no consensus on the definition and importance of the individual quality dimensions. In their literature review, Johns and Pine (2002) found the principal dimensions to be choice and quality of food and drinks, the price for value, service quality, atmosphere, location and convenience. By contrast, in their theoretical work, Antun et al. (2010) only identified three general quality domains: the restaurant's food, service and atmosphere, whereby the authors added a social and a healthfulness dimension. Regarding the relative importance of those dimensions, research has provided conflicting evidence. Food quality is mainly considered the most critical factor for restaurant diners. For example, Pettijohn et al. (1997) found that food quality, cleanliness, price and value have the greatest impact on the customer's perception in fast food restaurants. These factors were recognized as the fundamental requirements, while atmosphere and menu variety had a lesser impact on customer satisfaction. By contrast, other scholars argue that the perception of service quality is the most decisive factor for diners' intentions to return (Blose et al., 2019).

Again, we believe that the literature has not yet paid enough attention to potential (non-musical) acoustical factors, particularly concerning the question of which acoustical factors contribute to overall restaurant quality, and how large the potential effect is compared to the above-discussed non-acoustical factors. Therefore, we conducted a field study in 12 Berlin restaurants to triangulate research from room acoustics, soundscape, as well as consumer and hospitality research. Our study considered subject-centred measures on soundscape evaluation and non-acoustical restaurant quality dimensions as well as object-centred acoustical parameters in terms of the A-weighted equivalent continuous sound pressure level measured over a 15 min time interval ($L_{A,eq,15}$) and the reverberation time in the occupied state ($T_{20,occ}$).

To clarify the interrelation between the two object-centred acoustical parameters, we first tested whether reverberation time influenced the $L_{A,eq,15}$ beyond the higher gain that naturally comes with more reverberation, indicating a change in communication behavior (H1a). Moreover, we expected that this effect would be moderated by the number of patrons in the restaurant (H1b).

In line with previous findings (e.g., Battaglia, 2014; Gozalo et al., 2015), we further hypothesized that the $L_{A,eq,15}$ would negatively predict soundscape pleasantness (H2a). Similarly, we expected a negative influence of $T_{20,occ}$ on soundscape pleasantness (H2b). Similar effects of the acoustical parameters on soundscape eventfulness were assumed, however, with an inverse effect direction (H3a, H3b). However, based on H1, we expected that the effect of $T_{20,occ}$ on soundscape pleasantness and eventfulness would be mediated by the $L_{A,eq,15}$ (H2c, H3c). The above-mentioned hypotheses regarding the relationship between object- and subject-centred acoustical variables are illustrated in **Figure 1**.

Finally, regarding the influence of acoustical parameters, we hypothesized that the $L_{A,eq,15}$ (H4a) and soundscape pleasantness (H4b) would predict overall restaurant quality. It was expected that these effects remain significant even when controlling for other influencing factors, such as atmosphere, food quality, and service. This assumption was tested by establishing a comprehensive structural equation model (SEM) that predicts overall restaurant quality and considers the respective acoustical and non-acoustical factors and potential interrelationships.

METHOD

Sample

To address our research questions, we conducted a field survey in 12 randomly selected restaurants in Berlin. Randomly selected patrons dining in the particular restaurant were asked to fill out the questionnaire during a period of three to 4 h on a regular service day. Depending on the manager's preference, they were approached either by the author or the restaurant's staff. Participants filled out the questionnaire on a tablet PC.

Eight to 17 guests per restaurant filled out the questionnaire, resulting in a total of 142 participants (mean age: 34.7 years, SD = 13.0). Fifty-one participants were male, 88 were female, two 'divers', and one person preferred not to disclose their gender. Eighty-seven participants had an academic degree, 40 had a general qualification for university entrance, 11 had a general certificate for secondary education, and four persons had no official education certificate. For 56.7% of the patrons, it was their first visit to the respective restaurant, whereas 28.7% and 14.6% of them reported repeated or regular visits, respectively.

Design and Measures

The questionnaire consisted of four sections and 55 items (see **Supplemental Material**):

- 10 person-related items (age, gender, education, noise sensitivity, hearing impairment, mealtime, visitation motifs, frequency of visits)
- 23 restaurant quality items (10 quality items each as 'importance' and 'performance', willingness to recommend the restaurant and repeat visit, recommendations)
- 16 soundscape items
- 6 personality traits items (Extraversion and Neuroticism)

TABLE 1 | Restaurant attributes and requirements according to the DIN 18041-3:2016 and Rindel (2018).

ID	Type	T20,m,empty [s]	T20,m,occup,80% [s]	$L_{A,eq,15}$ [dB(A)]	Capacity	Volume V [m^3]	Area a [m^2]	N_{guests}
1	French cuisine	0.70	0.48	n.a	47	152	54	15
2	Swiss Fondue	0.53	0.44	64.3	72	392	119	6–21
3	Steaks	0.80	0.56	75.0	36	142	44	19–29
4	Restaurant-Café mix	1.01	0.69	73.3	75	354	111	15–28
5	Italian cuisine	0.64	0.49	69.6	70	302	97	16–33
6	Lifestyle/healthy food	n.a	n.a	70.9	37	n.a	n.a	5–10
7	Indian cuisine	0.54	0.43	58.9	44	202	61	5–11
8	Breakfast and Brunch	1.03	0.90	74.1	23	338	80	3–8
9	Indian cuisine	0.70	0.56	76.7	72	430	119	33–50
10	Indian cuisine	0.74	0.50	73.5	58	194	63	21–41
11	Hip Brunch/Café	0.68	0.54	75.6	33	184	56	21–28
12	German cuisine	0.97	0.64	71.2	195	793	240	73–96

Note: *n.a. (not available) refers to missing values due to technical or organizational issues.*

Eight Soundscape parameters (i.e., pleasant, chaotic, vibrant, uneventful, calm, annoying, eventful, monotonous) were assessed using a self-translated German version of the ISO/FDIS 12913–2 soundscape standard. Robust values for soundscape pleasantness and eventfulness were then calculated using the formulas proposed in the ISO standard. Restaurant quality and visitation motives were obtained by a self-translated German version of the questionnaire by Ponnam and Balaji (2014). Finally, the Big Five personality traits Extraversion and Neuroticism were measured by the German short inventory BFI-S (Gerlitz and Schupp, 2005).

Acoustical Measurements and Restaurant Attributes

For the L_{Aeq15}' measurements, a NTI XL2 acoustic analyzer with class 1 measurement microphone (M2210) was used. The acoustical scenes were recorded in first-order ambisonics format using a Sennheiser Ambeo VR Mic to allow for acoustical simulation of the restaurant soundscape in the laboratory. The microphones were both placed in the guest room, close to a regular table at head height.

As room-acoustical measurements in the occupied state would have led to a considerable disturbance of the guests and employees and a threat to the ecological validity of the soundscape assessments, they were conducted in empty condition outside the restaurants' opening hours. This was done using a self-constructed omnidirectional source (with reverse-horn principle), 'DBX DriveRack RTA-M' microphone, 'Focusrite Scarlett 2i2' interface and 'Room EQ Wizard 5.19' software. The measurement signal was a logarithmic sweep with a length of 256 k samples at a sampling rate of 44.1 kHz. Cooling aggregates and other noise sources were switched off where possible. To obtain values of $T_{20,m,empty}$, third-octave band measurements from 125 to 4000 Hz were arithmetically averaged. The reverberation time in occupied state $T_{20,occ}$ was then calculated for the occupied condition at the moment of the questionnaire's completion according to DIN 18041-3:2016 (2016) and used for further statistical analyses. Therefore, also the number of guests was assessed by manual count approximately every 15 min. Depending on the fluctuation speed, the count interval was shortened to five or extended to 30 min. Occupancy between these measurement intervals was estimated through linear interpolation. Restaurant attributes including the averaged $L_{A,eq,15}$, the $T_{20,occ}$ for 80% occupancy, and further room attributes are presented in **Table 1**.

Procedure

Guests were asked to fill out the questionnaire during 3 to 4 h on a regular service day. Depending on the manager's preference, they were approached by the second author or the restaurant's staff. They filled out the questionnaire on a tablet PC provided by the authors or on their own smartphone using the browser-based platform LimeSurvey. $L_{A,eq,15}$ measurements were conducted during the questionnaire distribution, and measurements of the room acoustics were performed before or after opening hours under empty conditions.

Data Analysis

$L_{A,eq,15}$ measurements failed in one restaurant (ID 1) due to technical problems, and one restaurant declined the room acoustical measurement (ID 6). Also, one restaurant (ID 8) turned out to be a mixture of food service, hotel lobby, and café with different activities and affordances compared to a classical restaurant. It therefore behaved differently than the rest of the sample, and was thus considered an outlier and excluded from the analysis.

For each participant who visited one of the remaining restaurants used in our analysis, the timestamp of the questionnaire transfer was assigned to the respective $L_{A,eq,15}$. Similarly, the present occupancy

TABLE 2 | Results of the Principal component analysis–Varimax-rotated factor loadings of the non-acoustical restaurant quality items on the observed three underlying dimensions.

Item	Product	Atmosphere	Service
Culinary Quality	0.82	0.13	0.19
Choice	0.71	0.20	0.25
Price for value	0.69	0.01	0.15
Plating	0.63	0.31	0.08
Interior Design	0.02	0.86	0.16
Ambience	0.17	0.75	0.00
Image	0.36	0.60	0.20
Friendliness	0.28	0.04	0.87
Availability	0.19	0.21	0.85

and the particular restaurant's room acoustical measures were assigned to each questionnaire.

The fact that soundscape evaluations were performed by multiple patrons in different restaurants, resulted in data with a two-level structure: evaluations (persons) nested within restaurants. Therefore, several linear mixed-effects models (LMM) taking into account the two-level structure were computed to test our hypotheses. These models included a random intercept for each restaurant, while $L_{A,eq,15}$, reverberation time, and soundscape pleasantness constituted fixed effects. To test for a non-linear, U-shaped relationship, the LMM was calculated with $T_{20,occ}$ and $T_{20,occ}^2$ as fixed effects. Significance tests were carried out with Type III tests of fixed effects via Satterthwaite's degrees of freedom method. Marginal R^2 were computed to obtain the variance in the respective dependent variable explained by the fixed effects (Nakagawa et al., 2013).

Before testing our hypotheses, a principal component analysis (PCA) on the non-acoustical quality items in the questionnaire was carried out, using orthogonal Varimax rotation (see **Table 2**). The Kaiser-Meyer-Olkin measure of sampling adequacy (= 0.74) and Bartlett's test of sphericity (χ^2 [36] = 449.75.2; $p < 0.001$) indicated substantial correlations amongst items to warrant a PCA. The scree-plot and the Kaiser criterion suggested a three-factor solution that explained 65% of the overall variance. Based on the factor loadings, the three resulting factors were named 'product', 'atmosphere', and 'service'. Also, a factor 'overall restaurant quality' was established employing a CFA, utilizing the input variables 'willingness to recommend restaurant', 'repeat visit', and 'recommend'.

To combine the results of the CFA with classical regression approaches and to test for direct and indirect effects between the manifest acoustical and latent non-acoustical variables, a comprehensive model on overall restaurant quality was then computed using Structural Equation Modeling (SEM) approach (Kline, 2011). Also here, the different restaurants served as cluster variable in the analysis.

Statistical analyses were carried out using R and R Studio, including the packages lavaan (Rosseel, 2012), lme4 (Bates et al., 2015), lmertest (Kuznetsova et al., 2017) and ggplot2 (Wickham, 2016).

TABLE 3 | Table of fixed effects from a linear-mixed-effects model predicting the A-weighted sound pressure level through reverberation time ($T_{20,occ}$) and the number of restaurant guests (N_{Guests}).

	b	*SE*	*df*	*t*	*p*
(Intercept)	0.22	0.25	7.747	0.879	0.406
$T_{20,occ}$	0.59	0.19	17.517	3.137	0.006
N_{guests}	−0.72	0.13	15.387	−5.411	<0.001
$T_{20,occ}$ * N_{guests}	0.84	0.17	66.334	5.052	<0.001

Acoustical parameters and soundscape dimensions.

RESULTS

Relationship Between Acoustical Parameters and Number of Restaurant Guests

In the first step, the expected relationship between the A-weighted sound pressure level $L_{A,eq,15}$ (dependent variable) and the reverberation time $T_{20,occ}$ (independent variable, H1a), and the moderating effect of the number of guests N_{Guests} (H1b) were investigated. A linear mixed-effects model (LMM) revealed both significant main effects of $T_{20,occ}$ and N_{Guests}, and an interaction effect of $T_{20,occ} \times N_{Guests}$, providing empirical support for both hypotheses (see **Table 3**). The model explains 53.4% of the variance of the $L_{A,eq,15}$ values.

To clarify the relationship between the two soundscape dimensions, we computed a first LMM which revealed that soundscape eventfulness significantly predicted soundscape pleasantness, F (1,157.6) = 2.65, β= −0.176, p = 0.03. The low amount of shared variance ($R^2_{marginal}$ = 0.032), however, warrants the independent consideration of the two dimensions in the course of the following analyses.

Concerning the relationship between acoustical parameters and soundscape dimensions, a second LMM revealed that the A-weighted sound pressure level constituted a significant negative predictor of soundscape pleasantness, confirming H2a, F (1.17.7) = 9.07, β= −0.034, $p < 0.01$. The $L_{A,eq,15}$ explained 11.4% of the overall variance of the dependent variable ($R^2_{marginal}$ = 0.114). The relationship between the two variables as obtained in the different restaurants is depicted in **Figure 2**.

Also, a third LMM utilizing $T_{20,occ}$ as independent variable showed a significant influence of reverberation time on soundscape pleasantness, confirming H2b, F (1.12.9) = 5.87, β= −0.040, p = 0.03. The $T_{20,occ}$ explained 8.5% of the overall variance of the dependent variable ($R^2_{marginal}$ = 0.085). The relationship between the two variables as obtained in the different restaurants is depicted in **Figure 3**.

This effect of $T_{20,occ}$, however, disappeared when controlling for $L_{A,eq,15}$, suggesting no impact of reverberation beyond the amplification of the sound level given for physical reasons (H2c). Indeed, a mediation analysis confirmed our hypothesis, as illustrated in **Figure 4**. Following the approach by Preacher

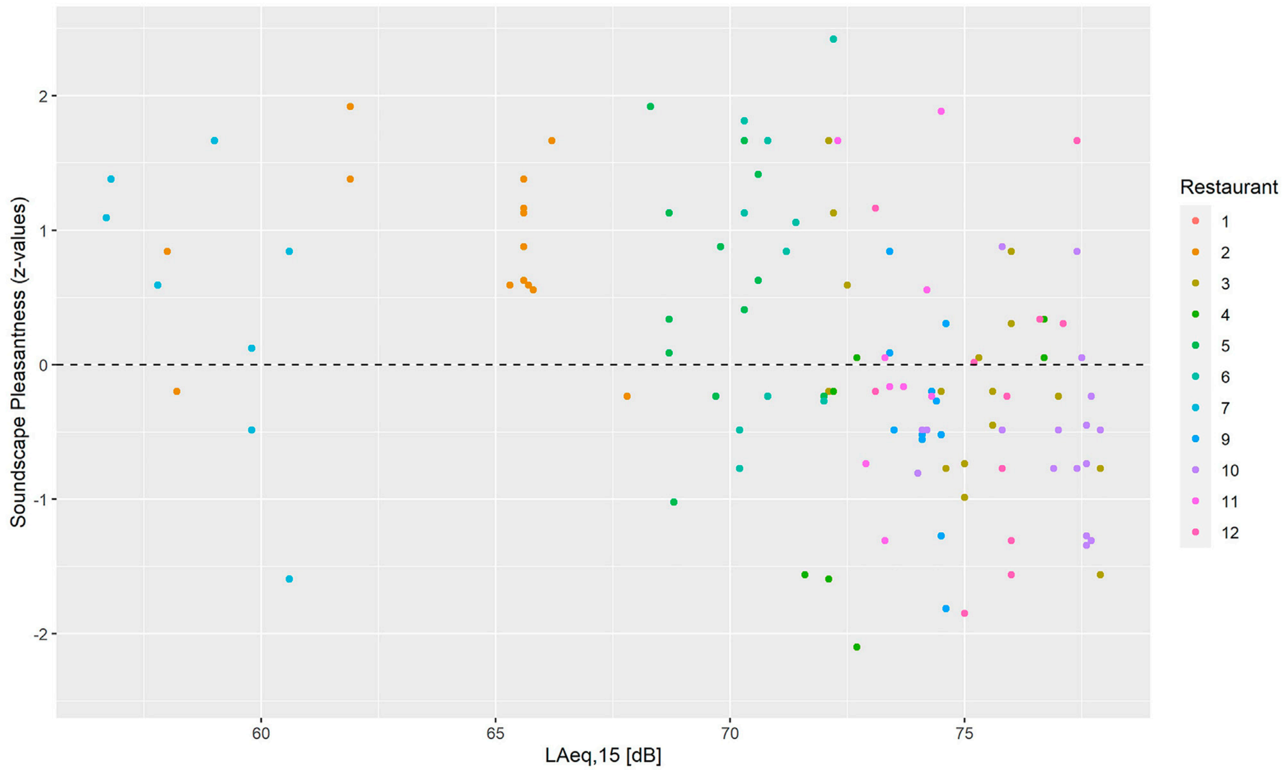


FIGURE 2 | $L_{A,eq,15}$ values and z-standardized soundscape pleasantness judgements.

and Hayes (2004), we applied a bootstrapping approach for significance testing. Here, a bias-corrected bootstrapped confidence interval with 10,000 samples did not include zero for the effect of $T_{20,occ}$ on $L_{A,eq,15}$, 95% CI [0.079 1.439], and of $L_{A,eq,15}$ on soundscape pleasantness, 95% CI [−0.488–0.077], whereas it included zero for the non-significant direct effect of $T_{20,occ}$ on soundscape pleasantness, 95% CI [−0.559, 0.014].

Regarding the effect of acoustical parameters on soundscape eventfulness, another LMM confirmed our hypothesis H3a that the A-weighted sound pressure level positively predicts soundscape eventfulness, F (1,18.5) = 4.71, β= 0.50, $p < 0.01$, explaining 24.1% of the overall variance of the dependent variable ($R^2_{marginal} = 0.241$). By contrast, reverberation time did not show a significant effect on soundscape eventfulness. This was true in case of a model that only included the linear term of $T_{20,occ}$, F (1, 17.2) = 1.69, $\beta = -0.35$, $p = 0.21$, but also for a model that tested for a non-linear, quadratic effect of reverberation time while controlling for a potential linear effect, ${T_{20,occ}}^2$, F (1, 49.3) = 0.00, $\beta = 0.13$, $p = 0.97$; $T_{20,occ,}$ F (1, 54.0) = 0.02, $\beta = -0.49$, $p = 0.89$. Therefore, our results failed to reject the null hypothesis associated with H3b.

Notwithstanding, due to the above shown, significant association between $L_{A,eq,15}$ and $T_{20,occ}$ (H1a), we conducted another mediation analysis which suggests an indirect effect of reverberation time on soundscape eventfulness, mediated by the A-weighted sound pressure level (see **Figure 5**). Again, the bias-corrected bootstrapped confidence intervals did not include zero for the significant effects, $T_{20,occ} \rightarrow L_{A,eq,15}$, 95% CI [0.079 1.439], $L_{A,eq,15}$ → soundscape eventfulness, 95% CI [0.427 0.706]. By contrast, they included zero for the non-significant direct effect of $T_{20,occ}$ on soundscape eventfulness, 95% CI [−0.432, 0.026].

Acoustical Parameters and Overall Quality Ratings

In the next step, we tested our fourth hypothesis stating that acoustical parameters, namely the sound pressure level (H4a) and reverberation time (H4b), would also predict overall restaurant quality. Therefore, we computed two further LMMs which revealed that the $L_{A,eq,15}$ negatively predicted overall restaurant quality, F (1.20.3) = 7.27, $\beta = -0.33$, $p = 0.01$, confirming H4a. Here, $L_{A,eq,15}$ explained 9.6% of the variance of overall restaurant quality as rated by the patrons ($R^2_{marginal}$ = 0.096). Again, results did not provide any empirical support for a direct effect of reverberation time, $T_{20,occ}$: F (1, 10.4) = 0.91, $\beta = -0.16$, $p = 0.36$, thus failing to reject the null hypothesis associated with H4b.

Finally, we conducted a third mediation analysis, which also suggests an indirect effect of reverberation time on overall restaurant quality, mediated by the A-weighted sound pressure level (see **Figure 6**). Again, the bias-corrected bootstrapped confidence intervals did not include zero for

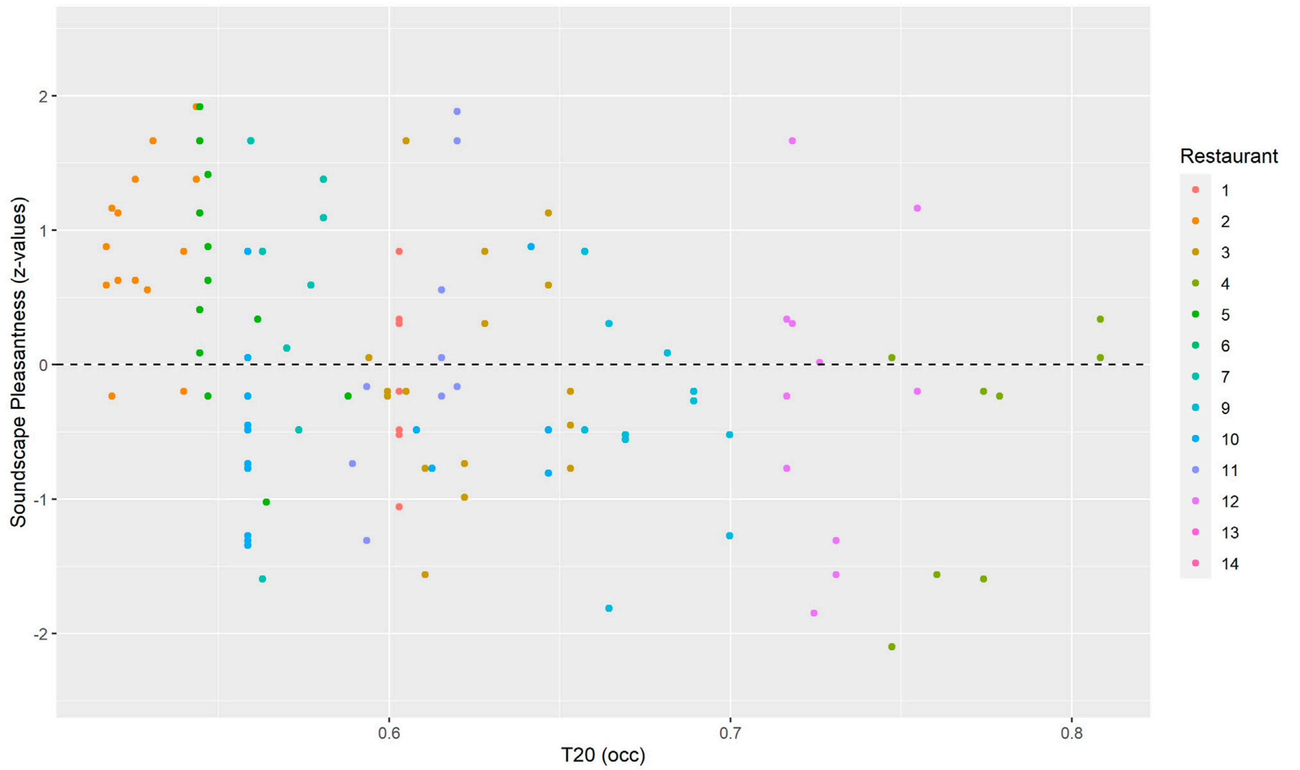


FIGURE 3 | $T_{20,occ}$ values and z-standardized soundscape pleasantness judgements.

the significant effects, $T_{20,occ} \rightarrow L_{A,eq,15}$, 95% CI [0.079 1.439], $L_{A,eq,15} \rightarrow$ overall restaurant quality, 95% CI [-0.437–0.116]. By contrast, they included zero for the non-significant direct effect of $T_{20,occ}$ on overall restaurant quality, 95% CI [--0.354, 0.402].

The following analysis aimed to integrate object- and subject-centred acoustical parameters as well as non-acoustical quality factors in a comprehensive structural equation model (SEM) predicting overall restaurant quality. Concerning the non-acoustical quality factors, we utilized the items and the factor structure obtained by the PCA (see Section **Data Analysis**). We assumed that all three non-acoustical quality factors 'atmosphere', 'service', and 'product' measured by their assigned items would positively contribute to overall restaurant quality. Based on the results obtained in the previous sections, we further expected a direct influence of the A-weighted sound pressure level on both soundscape pleasantness and restaurant quality (H4a). Finally, we hypothesized a direct influence of soundscape pleasantness on restaurant quality (H4b). Conducting a first SEM model revealed insufficient model fits, $\chi^2(78) = 859.3$, $p < 0.01$, RMSEA = 0.097, SRMR = 0.120, CFI = 0.897 (robust measures). Accordingly, modification indices (mi) were obtained, which suggested removing the direct effect of soundscape pleasantness on restaurant quality (mi = 19.3) and instead adding a regression path to 'atmosphere' (mi = 24.0). Also, adding a path from 'product' to soundscape pleasantness was indicated (mi = 13.8). As all three steps were considered useful and are supported by the literature (e.g., North and Hargreaves, 1998; Antun et al., 2010; Yan and Dando, 2015), we computed a modified SEM. According to thresholds reported by Kline (2011), the SEM yielded acceptable to good model fit indices, $\chi^2(78) = 859.3$, $p < 0.01$, RMSEA = 0.072, SRMR = 0.095, CFI = 0.942. The model paths are illustrated by **Figure 7**. The final model thus confirms H4a regarding the effect of the $L_{A,eq,15}$. By contrast, it does not provide empirical support for the assumed direct effect of soundscape pleasantness on restaurant quality (H4b); it rather suggests that this effect is mediated by the atmosphere of a restaurant.

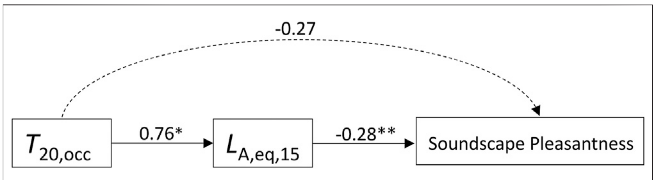


FIGURE 4 | Mediation analysis on acoustical parameters predicting soundscape pleasantness. Note: The estimates represent standardized regression coefficients β, and the asterisks indicate the significance of the respective effects, $^*p < 0.05$, $^{**}p < 0.01$, as obtained by the SEM routine of the Lavaan Package in R/RStudio.

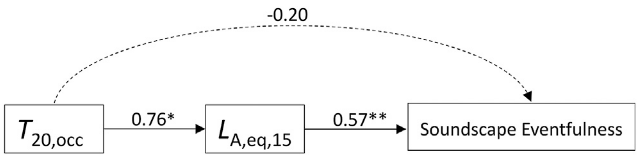


FIGURE 5 | Mediation analysis on acoustical parameters predicting soundscape eventfulness. Note: The estimates represent standardized regression coefficients β, and the asterisks indicate the significance of the respective effects, *$p < 0.05$, **$p < 0.01$, as obtained by the SEM routine of the Lavaan Package in R/RStudio.

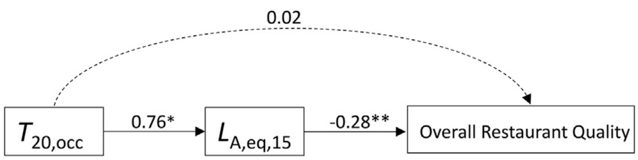


FIGURE 6 | Mediation analysis on acoustical parameters predicting overall restaurant quality. Note: The estimates represent standardized regression coefficients β, and the asterisks indicate the significance of the respective effects, *$p < 0.05$, **$p< 0.01$, as obtained by the SEM routine of the Lavaan Package in R/RStudio. Structural equation model including acoustical and non-acoustical parameters.

Soundscape Perception, Restaurant Choice, and Personality

Finally, we explored whether the two obtained Big Five dimensions Extraversion and Neuroticism which have been shown to predict soundscape evaluations in previous studies (e.g., Steffens et al., 2017) would be associated with pleasantness and eventfulness judgements or the choice of a particular restaurant. Here, no LMM revealed a significant relationship between soundscape and Big Five dimensions (all ps > 0.05). By contrast, an ANOVA with the different restaurants (nominal scale) as independent and the patrons' extraversion scores (interval scale) as dependent variable revealed that the extraversion scores significantly differ across restaurants, F (1,12) = 2.96, $p < 0.01$, $R^2 = 0.194$. This finding suggests that individuals choose restaurants, which fit and/or support their personality. However, no such an effect was shown for the Big Five dimension 'Neuroticism' ($p = 0.38$).

DISCUSSION

In our study, we investigated the relationship between object- and subject-centred acoustical factors and non-acoustical quality parameters as well as their contribution to the overall evaluation of restaurants. One major aim of our study was to investigate whether soundscape pleasantness and eventfulness in restaurants can be predicted by acoustical measures, namely the A-weighted sound pressure level and the reverberation time in the occupied state. Results confirmed our assumptions that loudness-associated measures and room-acoustical conditions significantly predict both soundscape pleasantness and eventfulness. The negative effect of (unwanted) loud sounds on soundscape pleasantness is widely in line with previous studies (e.g., Novak et al., 2010; Herklots, 2018). However, they contradict findings by Tarlao et al. (2021) who observed a positive influence of sound level on soundscape pleasantness, but who also dealt with generally lower sound levels and music as predominant sound source in the restaurant. This discrepancy indeed suggests non-linear effects in terms of an inverted U-shape and an optimal sound level in the mid-range. Further investigating this optimal medium level depending on the predominant sound source and the character of the restaurant should thus be subject to future research.

Results further indicate that the reverberation time is effective only through the increase in sound level it triggers. This amplification can lead to patrons increasingly raising their speech level to ensure intelligibility, resulting in a negative feedback loop of amplification. This assumption is supported by the fact that, in our study, the effect of reverberation time on the $L_{A,eq,15}$ was moderated by the number of patrons in the restaurant.

The second major aim was to study the influence of subject- and object-centred acoustical parameters on overall restaurant quality while considering other relevant non-acoustical parameters. The results of the SEM suggest that the $L_{A,eq,15}$ constitutes a (direct) negative predictor of overall restaurant quality. This finding corroborates previous findings by Battaglia (2014) and highlights a potential sensory annoyance caused by high-noise restaurant environments leading to a worsened overall experience. Moreover, the final SEM suggests that soundscape pleasantness affects overall restaurant quality only indirectly, mediated by the atmosphere. This finding emphasizes the need to design actively the restaurant's acoustical atmosphere beyond pure loudness reduction and to consider non-musical sounds when researching restaurant atmosphere.

Conversely, the results of the linear mixed-effects model suggest that, analogously to the above-mentioned effect on soundscape evaluation, the effect of reverberation time on restaurant quality was mediated by the A-weighted sound pressure level. This finding slightly contradicts previous results by Battaglia (2014) and Rindel (2018) who observed direct effects of reverberation time on restaurant quality, but who also did not control for potential mediation effects.

Based on the findings of object-centred acoustical parameters on soundscape and restaurant quality assessments, we argue that restaurants dealing with high sound pressure levels should consider room acoustical treatment. Here, personal communication with some Berlin restaurant owners who declined participation revealed that many managers are aware of acoustical problems, but they do not want to draw their guests' attention to it. Notwithstanding, many restaurants, which would be considered problematic in the light of our results, seem to be running successful businesses. Hence, poor acoustics might not necessarily predict a lack of commercial success, but it is a decisive influencing factor of perceived overall quality and atmosphere. Of

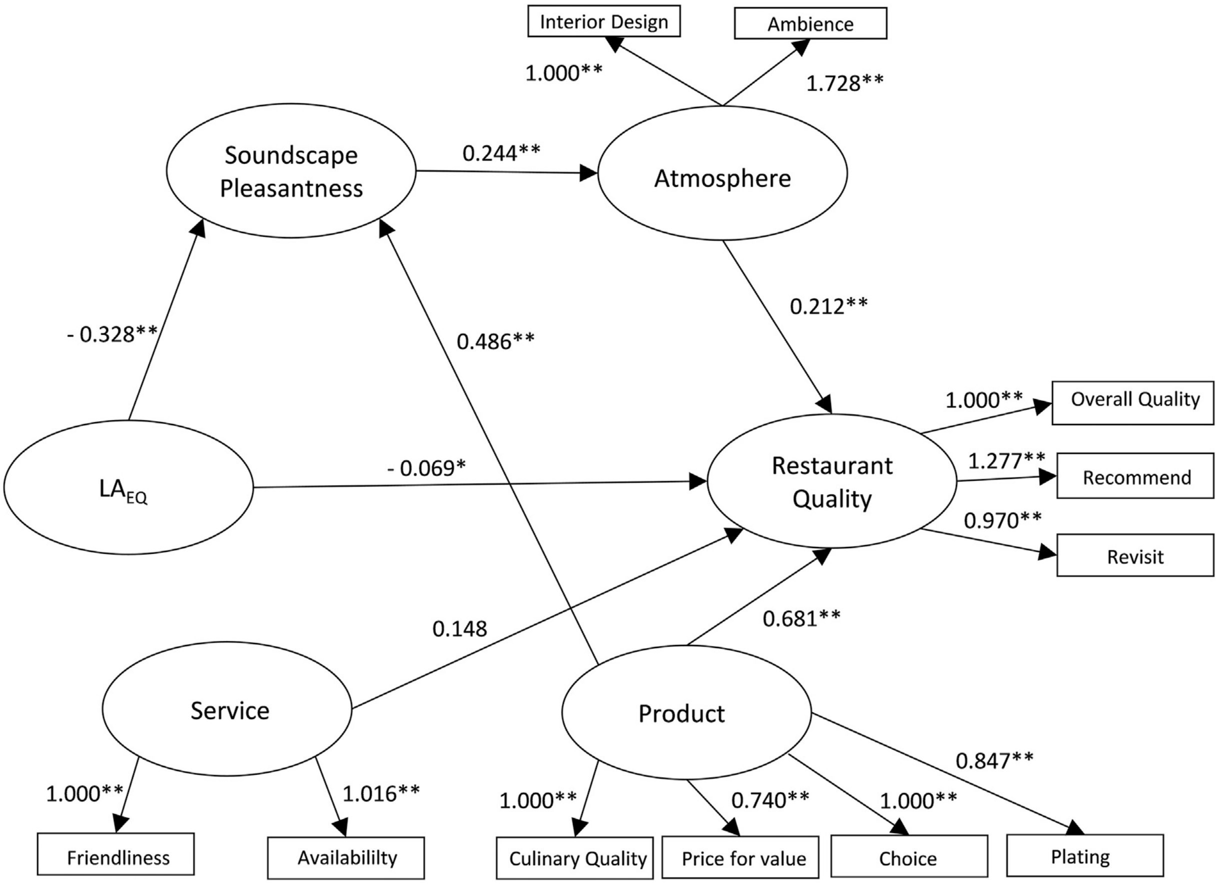


FIGURE 7 | Structural equation model predicting restaurant quality, including acoustical and non-acoustical factors. Note: The estimates represent standardized regression coefficients β, and the asterisks indicate the significance of the respective effects, *$p < 0.05$, **$p < 0.01$, as obtained by the SEM routine of the Lavaan Package in R/RStudio.

course, whether the acoustic atmosphere is an essential attribute mainly depends on the restaurant's market segment. Does it deliver services or experiences? If selling service is prioritized, it may be more goal-oriented to invest resources in high quality of food and service. This approach may be adequate in fast-food services, bistros, or take-aways. If, however, the guest is meant to undergo a memorable experience during their visit, as indicated by a shift from service to experience economy (Pine and Gilmore, 1998), attention needs to be drawn to atmospherics, covering all the senses, including the auditory one. As suggested by our results, the preference for certain restaurants and its atmospherics might also be governed by person-related variables, such as Extraversion. Here, it could be assumed that an extraverted person might feel more comfortable in a lively as opposed to a calm restaurant environment. However, more research is needed to elaborate further on such relationships.

A couple of limitations associated with the study design have to be addressed. For example, our sample is biased by a non-random self-selection process. In particular, participants could only be picked from a group of people who chose to dine at one of the 12 restaurants. That is, persons disliking busy restaurants with high noise levels might be systematically underrepresented in our study. Moreover, as we could not manipulate any variables in this field study, we could only observe correlations, not allowing to draw definite causal directions. Thus, some observed effects could also be interpreted oppositely. For example, patrons who rate the overall quality of the restaurant highly might be more relaxed and happy, which might result in an overall more positive soundscape compared to an acoustical environment created by annoyed patrons. Thus the observed effects, as well as the established structural equation model, should be validated in the course of multiple laboratory studies, for instance, in a virtual environment. Finally, due to limited resources and time restrictions related to the questionnaire, we could not obtain more potentially relevant variables, such as behavioral measures (e.g., duration of stay or money spent) or acoustical and non-acoustical parameters, for example, speech intelligibility, privacy, the visual design, or the 'hipness' of a restaurant. Also, we did not consider psychoacoustical metrics or audio features which might be better suited to predict pleasantness and quality evaluations. However, increasing the number of variables in our models would also increase the required sample size to a considerable degree. Here, we experienced that convincing restaurants to participate is not an easy task, because, understandably, managers do not want to disturb their guests nor do they want to draw their attention to a potentially flawed aspect of their restaurant. This also led to limited statistical power, particularly related to observing small effects of the room acoustics beyond pure sound-amplification processes.

Notwithstanding, our results have demonstrated that high sound levels negatively affect both the pleasantness of the soundscape and the overall quality of a restaurant. As an 'amplifier' for the sound level in the room, also the reverberation time of the room has a significant influence on both variables. Thus, our study clearly demonstrates the value of an acoustical design of restaurants and the associated investments. Such design should be guided by the desired atmosphere, which our study found to be critical in determining how strongly and in what way soundscape pleasantness affects the perceived overall quality of a restaurant.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JS, TW, and SW designed the study. TW collected data. JS and TW developed and performed the statistical analysis. JS and TW wrote the first draft of the manuscript. All authors interpreted data, reviewed, and edited the manuscript, and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank all Berlin restaurants for taking part in the study and for allowing us to take the acoustical measurements.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2021.676009/full#supplementary-material

REFERENCES

Antun, J. M., Frash, R. E., Costen, W., and Runyan, R. C. (2010). Accurately Assessing Expectations Most Important to Restaurant Patrons: The Creation of the DinEX Scale. *J. Foodservice Business Res.* 13 (4), 360–379. doi:10.1080/15378020.2010.524539

Astolfi, A., and Filippi, M. (2004). Good Acoustical Quality in Restaurants: a Compromise between Speech Intelligibility and Privacy. Proceedings of the International Congress on Acoustics (ICA), Kyoto, Japan.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using Lme4. *J. Stat. Softw.* 67 (1). doi:10.18637/jss.v067.i01

Battaglia, P. L. (2014). Achieving Acoustical Comfort in Restaurants. in Proceedings of the 168th Meeting of the Acoustical Society of America. Indianapolis, IN, United States, 27–31 October 2014, 15001.

Bitner, M. J. (1992). Servicescapes: The Impact of Physical Surroundings on Customers and Employees. *J. Marketing* 56 (2), 57–71. doi:10.2307/1252042

Blose, J. E., Frash, R. E., and DiPietro, R. B. (2019). A Top Box Analysis of DinEX to Optimize Restaurant Resources. *Int. J. Hospitality Tourism Adm.* 20 (2), 156–181. doi:10.1080/15256480.2017.1359732

Caldwell, C., and Hibbert, S. A. (2002). The Influence of Music Tempo and Musical Preference on Restaurant Patrons' Behavior. *Psychol. Mark.* 19 (11), 895–917. doi:10.1002/mar.10043

Crisinel, A.-S., Cosser, S., King, S., Jones, R., Petrie, J., and Spence, C. (2012). A Bittersweet Symphony: Systematically Modulating the Taste of Food by Changing the Sonic Properties of the Soundtrack Playing in the Background. *Food Qual. Preference* 24 (1), 201–204. doi:10.1016/j.foodqual.2011.08.009

DIN German Institute for Standardization (2016). *Acoustic Quality in Rooms - Specifications and Instructions for the Room Acoustic Design*. DIN 18041-3: 2016).

Gerlitz, J.-Y., and Schupp, J. (2005). Zur Erhebung der Big-Five-basierten Persoenlichkeitsmerkmale im SOEP: [On the measurement of the Big Five-based personality traits in the SOEP]. *DIW Res. Notes* 4, 2005.

Gozalo, G. R., Carmona, J. T., Morillas, J. B., Vílchez-Gómez, R., and Escobar, V. G. (2015). Relationship between Objective Acoustic Indices and Subjective Assessments for the Quality of Soundscapes. *Appl. Acoust.* 97, 1–10. doi:10.1016/j.apacoust.2015.03.020

Herklots, T. (2018). Zagat Releases 2018 Dining Trends Survey. Available at: https://zagat.googleblog.com/2018/01/zagat-releases-2018-dining-trends-survey.html (Accessed February 1, 2021).

Kassarjian, H. H., and Robertson, T. (1991). *Perspectives in Consumer Behavior*. 4th ed. (Englewood Cliffs (NJ): Prentice-Hall).

Johns, N., and Pine, R. (2002). Consumer Behaviour in the Food Service Industry: a Review. *Int. J. Hospitality Manage.* 21 (2), 119–134. doi:10.1016/s0278-4319(02)00008-7

Junqua, J.-C., Fincke, S., and Field, K. (1999). The Lombard Effect: a Reflex to Better Communicate with Others in Noise. In 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99, (2083–2086). IEEE. doi:10.1109/icassp.1999.758343

Kline, R. B. (2011). "Principles and Practice of Structural Equation Modeling," in *Methodology in the Social Sciences*. 3rd ed. (New York: Guilford Press).

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* 82 (13), 1–26. doi:10.18637/jss.v082.i13

Lindborg, P. (2016). A Taxonomy of Sound Sources in Restaurants. *Appl. Acoust.* 110, 297–310. doi:10.1016/j.apacoust.2016.03.032

Lindborg, P. (2015). Psychoacoustic, Physical, and Perceptual Features of Restaurants: A Field Survey in Singapore. *Appl. Acoust.* 92, 47–60. doi:10.1016/j.apacoust.2015.01.002

Lombard, E. (1911). Le signe de l'elevation de la voix [The sign of voice raising]. *Ann. Mal. De L'oreille Du Larynx* 37, 101–119.

Meng, Q., Zhao, T., and Kang, J. (2018). Influence of Music on the Behaviors of Crowd in Urban Open Public Spaces. *Front. Psychol.* 9, 2763. doi:10.3389/fpsyg.2018.00596

Nakagawa, S., Schielzeth, H., and O'Hara, R. B. (2013). A General and Simple Method for obtainingR2from Generalized Linear Mixed-Effects Models. *Methods Ecol. Evol.* 4 (2), 133–142. doi:10.1111/j.2041-210x.2012.00261.x

Navarro, M. P. N., and Pimentel, R. L. (2007). Speech Interference in Food Courts of Shopping Centres. *Appl. Acoust.* 68 (3), 364–375. doi:10.1016/j.apacoust.2006.01.006

Nocke, C. (2017). German DIN 18041 Acoustic Quality in Rooms. *The J. Acoust. Soc. America* 141 (5), 3686. doi:10.1121/1.4988020

North, A. C., and Hargreaves, D. J. (1998). The Effect of Music on Atmosphere and Purchase Intentions in a Cafeteria1. *J. Appl. Soc. Pyschol* 28 (24), 2254–2273. doi:10.1111/j.1559-1816.1998.tb01370.x

North, A. C. (2012). The Effect of Background Music on the Taste of Wine. *Br. J. Psychol.* 103 (3), 293–301. doi:10.1111/j.2044-8295.2011.02072.x

Novak, C. C., La Lopa, J., and Novak, R. E. (2010). Effects of Sound Pressure Levels and Sensitivity to Noise on Mood and Behavioral Intent in a Controlled Fine Dining Restaurant Environment. *J. Culinary Sci. Tech.* 8 (4), 191–218. doi:10.1080/15428052.2010.535756

Pettijohn, L. S., Pettijohn, C. E., and Luke, R. H. (1997). An Evaluation of Fast Food Restaurant Satisfaction. *J. Restaurant Foodservice Marketing* 2 (3), 3–20. doi:10.1300/j061v02n03_02

Pine, B. J., and Gilmore, J. H. (1998). Welcome to the Experience Economy. *Harv. Bus Rev.* 76 (4), 97–105.

Ponnam, A., and Balaji, M. S. (2014). Matching Visitation-Motives and Restaurant Attributes in Casual Dining Restaurants. *Int. J. Hospitality Manage.* 37, 47–57. doi:10.1016/j.ijhm.2013.10.004

Preacher, K. J., and Hayes, A. F. (2004). SPSS and SAS Procedures for Estimating Indirect Effects in Simple Mediation Models. *Behav. Res. Methods Instr. Comput.* 36 (4), 717–731. doi:10.3758/bf03206553

Rindel, J. H. (2018). Suggested Acoustical Requirements for Restaurants, Canteens, and Cafeterias. In Baltic-Nordic Acoustics Meeting 2018, Reykjavik, Iceland, 15–18, April 2018.

Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling and More. Version 0.5–12 (BETA). *J. Stat. Softw.* 48 (2), 1–36. doi:10.18637/jss.v048.i02

Ryu, K., and Han, H. (2011). New or Repeat Customers: How Does Physical Environment Influence Their Restaurant Experience? *Int. J. Hospitality Manage.* 30 (3), 599–611. doi:10.1016/j.ijhm.2010.11.004

Steffens, J., Steele, D., and Guastavino, C. (2017). Situational and Person-Related Factors Influencing Momentary and Retrospective Soundscape Evaluations in Day-To-Day Life. *J. Acoust. Soc. America* 141 (3), 1414–1425. doi:10.1121/1.4976627

Tarlao, C., Fernandez, P., Frissen, I., and Guastavino, C. (2021). Influence of Sound Level on Diners' Perceptions and Behavior in a Montreal Restaurant. *Appl. Acoust.* 174, 107772. doi:10.1016/j.apacoust.2020.107772

Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data analysisUse R!*. New York: Springer.

Yan, K. S., and Dando, R. (2015). A Crossmodal Role for Audition in Taste Perception. *J. Exp. Psychol. Hum. Perception Perform.* 41 (3), 590–596. doi:10.1037/xhp0000044

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 18 May 2021
doi: 10.3389/fbuil.2021.666152

Changes of Voice Production in Artificial Acoustic Environments

Tomás Sierra-Polanco[1], *Lady Catherine Cantor-Cutiva*[2,3], *Eric J. Hunter*[4] *and Pasquale Bottalico*[1]*

[1]Department of Speech and Hearing Science, University of Illinois Urbana-Champaign, Champaign, IL, United States, [2]Department of Collective Health, Universidad Nacional de Colombia, Bogota, Colombia, [3]Program of Speech-Language Pathology, Universidad Manuela Beltrán, Bogota, Colombia, [4]Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, MI, United States

The physical production of speech level dynamic range is directly affected by the physiological features of the speaker such as vocal tract size and lung capacity; however, the regulation of these production systems is affected by the perception of the communication environment and auditory feedback. The current study examined the effects of room acoustics in an artificial setting on voice production in terms of sound pressure level and the relationship with the perceived vocal comfort and vocal control. Three independent room acoustic parameters were considered: gain (alteration of the sidetone or playback of one's own voice), reverberation time, and background noise. An increase in the sidetone led to a decrease in vocal sound pressure levels, thus increasing vocal comfort and vocal control. This effect was consistent in the different reverberation times considered. Mid-range reverberation times (T30 ≈ 1.3 s) led to a decrease in vocal sound pressure level along with an increase in vocal comfort and vocal control, however, the effect of the reverberation time was smaller than the effect of the gain. The presence of noise amplified the aforementioned effects for the variables analyzed.

Keywords: room acoustics, gain, background noise, reverberation time, speech level, vocal comfort, vocal control

OPEN ACCESS

Edited by:
Nicola Prodi,
University of Ferrara, Italy

Reviewed by:
Davide Borelli,
University of Genoa, Italy
Costantino Carlo Mastino,
University of Cagliari, Italy

***Correspondence:**
Pasquale Bottalico
pb81@illinois.edu

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *09 February 2021*
Accepted: *20 April 2021*
Published: *18 May 2021*

Citation:
Sierra-Polanco T, Cantor-Cutiva LC, Hunter EJ and Bottalico P (2021) Changes of Voice Production in Artificial Acoustic Environments. Front. Built Environ. 7:666152. doi: 10.3389/fbuil.2021.666152

INTRODUCTION

Vocal communication involves multiple physiologic (oral and aural) and cognitive systems. From the perspective of production, the regulation of speech level is primarily affected by physiological features of the speaker such as vocal tract size, vocal fold length, and lung capacity (Riede and Brown, 2013). This production regulation is affected by perceived communication demands, such as communication partners or communication environment, sense of vocal comfort, and applied vocal effort. For example, speech level and speech style can be partner-specific such as speaking to a child (Rowe, 2008) or to someone with a perceived hearing loss (Krause and Braida, 2004). Another example was presented by Lane and Tranel (1971) where aspects of auditory feedback such as background noise, altered sidetone (amplified playback of one's own voice), hearing loss, and room acoustics were described. The alteration in auditory feedback can modify vocal parameters, such as Sound Pressure Level (SPL), and can modify the talker's perception of vocal comfort and vocal control (Pelegrín-García and Burnskog, 2012; Bottalico et al., 2015). These parameters may be modified by the implementation of artificial settings delivered by headphones with the goal of increasing vocal comfort and control while decreasing vocal effort in occupational voice users such as teachers and call center operators. All of these are affected by the relationship between voice production and hearing sensitivity (Hunter et al., 2006) and how the auditory system and auditory feedback play a fundamental role in voice production including the perception of effort and comfort.

Vocal effort has been defined as the perceived exertion of a vocalist to a perceived communication scenario (Hunter et al., 2020). Changes in vocal effort have been shown to be correlated with other vocal adaptations such as vowel modifications, along with changes in vocal fundamental frequency, dB SPL, spectral tilt, and speech rate (Berardi, 2015; McKenna and Stepp, 2018; Berardi, 2020). Even though vocal effort changes associate with a range of vocal production parameters, radiated speech level seems to be the primary production parameter related to vocal effort (ISO 9921, 2003) even when speech level is being controlled for (McKenna and Stepp, 2018).

Changes in vocal effort (as measured using vocal production metrics) or communication environment can affect vocal comfort. Vocal comfort can be defined as a subjective attribute that is directly correlated to the positive evaluation of the room for speech production and to the perceived support. Vocal comfort has been shown to be negatively correlated to the feeling of having to raise the voice and to the tiredness after speaking for a long time period in the room (Pelegrín-García et al., 2014; Cipriano et al., 2017). A questionnaire investigation showed that voice comfort is more closely related to the perceived noise annoyance than to the perceived room reverberance. Vocal comfort is related to all aspects that reduce vocal effort (Titze, 1999; Titze, 2000). It appears to decrease with the speaker's perceived fatigue and the sensation of needing to increase the voice level (Pelegrín-García and Brunskog, 2012). Previous research in classroom settings showed that the vocal comfort increases with the perception of the classroom as being good to speak in and with the perceived support and enhancement, while it decreases with the perceived exhaustiveness of speaking in a classroom during a lesson and with the sensation of having to increase the voice level (Pelegrín-García and Brunskog, 2012).

The alteration in auditory feedback can also modify the perception of a communication scenario, thus affecting voice production, vocal comfort, and the perception of vocal control. Vocal control can be defined as the capacity to self-regulate vocal production, e.g., SPL, fundamental frequency, and resonance. The sensation of control relates to the ability to adjust the voice consciously. In a communication environment, in general, speakers try to control their voice production in order to increase speech intelligibility. For example, while considering a communication partner with hearing limitations, a talker (deliberately or inadvertently) uses "clear speech" (Krause and Braida, 2004; Ferguson and Kewley-Port, 2007). This type of speech has been characterized by a slower speech rate, a wider range of fundamental frequency, and a higher temporal modulation index than conversational speech (Bottalico et al., 2016a). Likewise, when talking in a noisy environment, people tend to raise the level of their voice in order to maintain understandable communication (Lombard, 1911). The maximization of intelligibility, clarity, vocal comfort and control, and the minimization of vocal effort and fatigue, should be the priority of any professional talker (Bottalico et al., 2016a).

Growing evidence suggests that there is an association between vocal production level and external auditory feedback. External auditory feedback consists of the external path between mouth and ears and is strongly influenced by the acoustics of the environment where the speaker is speaking. Such environmental effects are room noise, vocal amplification of one's own voice, and, room reverberation.

A commonly experienced external auditory effect that directly impacts vocal production level is that of elevated room noise, or the Lombard Reflex or Effect (Lombard, 1911; Junqua, 1993). For example, Yiu and Yip (2016) recorded a monologue passage for twenty-four vocally healthy young adults (12 men and 12 women, aged 19–22 years) using an Ambulatory Phonation Monitor (APM model 3,200) under three natural environment conditions in a randomized order. These conditions were: a quiet room (clinic room, mean 35.5 dBA, ranged from 34 to 37 dBA), a room with moderate noise level (clinic corridor, mean 54.5 dBA, ranged from 53 to 56 dBA), and a room with high noise (a pantry room with a noisy exhaust fan, mean 67.5 dBA, ranged from 66 to 69 dBA). The results showed significant increases in mean voice level and self-reported vocal effort in the high-noize environment than in the other two conditions.

Vocal level was shown to be affected by the reverberation time of the room (Black, 1950), and by the level at which a speaker perceived his/her own voice, as well as the level of the background noise (Siegel and Pick, 1974). More recently, studies have added further details to these and other factors such as speaker-listener distance and acoustic characteristics of the room and/or of the communication channel (Black, 1950; Pelegrín-García et al., 2011; Bottalico et al., 2015; Bottalico et al., 2016a; Bottalico et al., 2017a; Bottalico et al., 2017b; Bottalico, 2017). Pelegrín-García et al. (2011) found that voice level decreased as reverberation time increased, while Black (1950) reported that greater vocal intensity was found in less reverberant rooms than in more reverberant rooms. This is common even in extreme reverberation conditions (Rollins et al., 2019).

Furthermore, external auditory feedback can be artificially altered by modifying the playback of one's own voice (i.e., sidetone alteration). In a study of the effect of sidetone alteration on voice levels by increasing the sidetone gain of 20 dB, Siegel and Pick (1974) found a ratio of change in the voice level of 0.15 dB/dB. This ratio increased to 0.21, 0.30, and 0.34 dB/dB when speech-spectrum noise was added during the experiment at 60, 70, and 80 dB, respectively.

Recent investigations on speech adjustments were related to an increase of external auditory feedback (Bottalico et al., 2015) and to reverberation times (Bottalico et al., 2016b). The above mentioned showed that the effect of reflective panels, placed close to the speaker, had a decrease of about 1 dB in voice level, which was observable in rooms with different reverberation times and in different speech styles.

In summary, previous research suggests that voice level, vocal comfort, and vocal control vary 1) when the gain level of external auditory feedback increases and 2) under different reverberant conditions. These variations could be also affected by the presence of noise. The perceived vocal comfort was lower in rooms with very low or very high reverberation time.

Nevertheless, to better understand how speech adjusts to room acoustics, it is necessary to have control of the acoustical parameters. This can be facilitated by creating virtual acoustics scenarios.

To explore this topic, the current study examined the effects of room acoustics in a virtual setting on vocal SPL, and self-reported vocal comfort and control. Three independent room acoustic parameters were considered: gain (alteration of the sidetone), reverberation time (T30), and background noise. This relationship was stated to better understand how these independent and dependent variables relate to each other in simulated scenarios. As we have mentioned, previous studies have been performed in real scenarios, which are not malleable nor changeable, but fixed. By having simulated scenarios, this study proposes a wide range of possibilities that could be infinitely modified, in a simple way, on its initial parameters for independent variables. The main research questions of this study were based on the following statements regarding relationships between:

(1) Voice level variations and participant's gain level of external auditory feedback (sidetone or self-amplification).
(2) Vocal comfort (and control) responses and participant's gain level of external auditory feedback (sidetone or self-amplification).
(3) Voice level variations and different simulated T30 of rooms where participants are speaking.
(4) Vocal comfort (and control) response and different simulated T30 of rooms where participants are speaking.
(5) Finally, if there are such effects:
(5a) Voice level variations and the presence or absence of noise.
(5b) Vocal comfort (and control) and the presence or absence of noise.

Hence, the present work is aimed to provide contributions on how acoustical environments affect voice production in terms of objective measurements such as SPL, but also in terms of perceptual measurements such as self-reported vocal comfort and vocal control.

MATERIALS AND METHODS

The speech of 30 talkers was recorded in fourteen different virtual acoustical scenarios of external auditory feedback, including three gain levels and three T30, each of them with and without the presence of speech-shaped noise. The participants' speech was recorded with a microphone placed at a fixed distance of 15 cm from the mouth. A preliminary calibration procedure of the microphone was performed at the beginning of the recording session per participant. The calibration level was set to 94 dB at 1 kHz. The recordings were performed in a soundproof double-walled Whisper Room (interior dimensions: 226 × 287 cm and h = 203 cm). T30 was measured for mid-frequencies to be 0.07 s in the soundproof room and background noise equal to 25 dB(A). The speech signals were processed to calculate SPL.

Participants

This study was conducted with approval from and in accordance with the policies of the Office of Protection of Research Subject at the University of Illinois at Urbana Champaign (IRB 18179). Thirty participants (17 females and 13 males) participated in this experiment. All the participants were Native American English-speaking young adults (age 19–32 years old; mean age 23 years), with self-reported normal speech and hearing, and no reported or observable upper respiratory infection on the day of the recording. In general, none of them reported hearing conditions. 26 participants reported that their primary ethnicity was "Caucasian," two were "Asian," and two "Hispanic-Latino." Four of them reported being eventual smokers. Five reported voice training in the past, such as singing lessons, and four reported a history of speech or language therapy in their childhood.

Instructions and Conditions

The participants were instructed to read aloud the first 6 sentences of "The Rainbow Passage," a standardized text in English (Fairbanks, 1960), under fourteen different virtually simulated acoustic conditions. Each task had a duration of about 27 s of reading. Before the measurements, each participant was presented with the printed passage to familiarize themselves with it.

The fourteen virtually simulated acoustic conditions were: a reference condition (no gain, no reverberation) and the result of all possible combinations of two gain levels of the external auditory feedback (+5 and +10 dB) and three different T30. The six aforementioned conditions were presented with and without speech-shaped noise added. The order of administration of the fourteen scenarios was randomized to provide an equal distribution of any (short-term) vocal discomfort across all the tasks, as well as to control for any unknown confounding variables relating to the task order.

Participants answered two questions after each task of the experiment: 1) How comfortable was it to speak in this condition? And 2) How well were you able to control your voice in this condition? These questions were worded in a manner consistent with the relevant ISO standard (ISO 28802, 2012) and administered immediately after exposure to the noise conditions in each task. Participants responded by making a vertical tick on a continuous horizontal line of 100 mm in length on a visual analog scale; this scale was provided on paper. The score was measured as the distance of the tick from the left end of the line. The extremes of the lines were 'not at all' (left) and 'extremely' (right).

Equipment

The speech material was recorded by a frequency response Class 1 microphone placed at a fixed distance of 15 cm from the mouth (M2211, NTi Audio, Tigard, OR, United States). The microphone was calibrated at the beginning of the recording session per participant using a Class 1 Sound Calibrator NTi Audio (Tigard, OR, United States) with automatic atmospheric pressure compensation (ref 94 dB ± 0.2 dB at 1 kHz ±1%). The microphone output was split into two lines: the first for direct

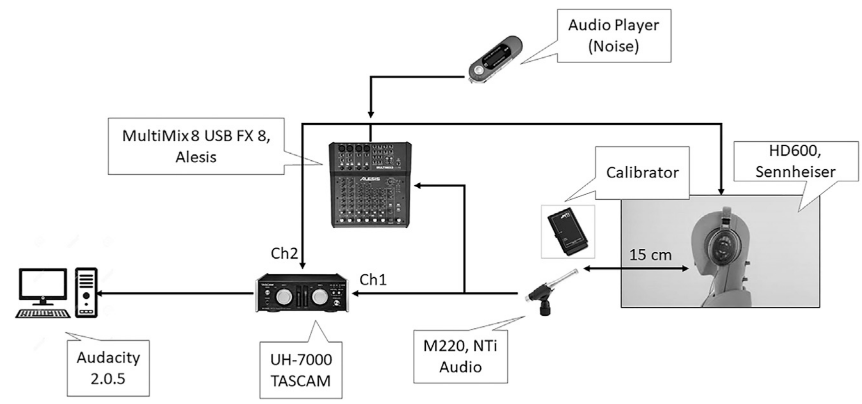


FIGURE 1 | Equipment set-up for experimental measurements.

recording and the second for creating the virtual acoustic environment. The direct digital recording sampled at 44.1 kHz was recorded using an external soundboard (UH-7000 TASCAM, Teac Corporation, Montebello, CA, United States) connected to a personal computer (PC) running Audacity 2.0.5 (SourceForge, La Jolla, CA). For the virtual environment, the direct microphone output was combined, in half of the conditions, with speech-shaped noise using a digital mixer (MultiMix 8 USB FX 8, Alesis, Cumberland, RI, United States). The voice signal was digitally processed to add reverberation using a real-time effect processor of the digital mixer and played back to the participant using open headphones (HD600, Sennheiser, Wedemark, Germany). The delay between the uttered voice and its transmission through the processing loop (i.e., Alesis digital mixer) and back to the participant's headphones was measured to be lower than 5 ms. This value is below the range between 16 and 26 ms threshold which is considered a noticeable echo (Lezzoum et al., 2016). The disposition of the equipment is depicted in **Figure 1**.

Room Acoustic Parameters

Room acoustic T30 conditions (ISO 3382-2, 2008) of the virtual scenarios were obtained from impulse responses (IRs) calculated with the convolution method. An exponential sweep signal was emitted by the mouth of a Head and Torso Simulator (HATS, GRAS 45BB KEMAR). The sweep was captured by the microphone, real-time processed, played back with open headphones, and finally recorded by the ears of the HATS. The recorded sweep was deconvolved with the emitted sweep inverted on the time axes, obtaining the IR, as exposed in the appendix by Pelegrín-García and Brunskog (2012).

The average T30 for combined 500 Hz and 1 kHz octave bands, were determined for the Whisper Room and each of the 3 simulated environments (ISO 3382-2, 2008). It was 0.07 s in Whisper Room T30 condition, 1.13 s in Low T30 condition, 1.39 s in Medium T30 condition, and 1.90 s in High T30 condition. The measured values of T30 for the Whisper Room and the three simulated conditions between 125 and 8 kHz are given in **Table 1**. To manipulate the level of external auditory feedback, three different gain factors were introduced in the real-time processor. These gain factors were chosen with the goal of obtaining a difference between the voice level measured at the ears in the air (with no sidetone modification) and the voice level measured at the ears position after the real-time processor, equal to 0, 5, and 10 dB.

In 7 out of the 14 tasks performed by each participant, speech-shaped noise was added to the real-time processor with the same power. The power level was set to obtain an A-weighted equivalent level averaging both ears of about LAeq = 70 dB(A) at the ears of the talker (measured with the HATS). This level was chosen among the one used by Siegel and Pick (1974) to stimulate the variation in the voice level with the sidetone alteration without excessive noise exposure for the participants. The values per octave band for background noise conditions, with and without speech-shaped noise, are reported in **Table 1**.

Voice Processing and Statistical Analysis

Analysis of the speech parameters was performed with Matlab R2017a (MathWorks, Natick, MA, United States). For each of the 14 tasks, a time history of A-weighted SPL was obtained from recorded speech. The time information associated with time histories (which typically ranged from 0 to 30 seconds within a task) was be retained for inclusion in the statistical analysis.

Statistical analysis was conducted using R Studio (version 1.2.5033). Linear Mixed-Effects (LME) models were fitted by restricted maximum likelihood (REML). Random effects terms were chosen based on variance explained. A random effect is referred to as a factor that may affect the outcome but does not have main relevance. The selection of random effects is based on taking out the variance associated with a specific factor, due to low interest in its effect. Thus, it is used as a random factor to remove variance. Models were selected based on the Akaike information criterion (Akaike, 1998; the model with the lowest value being preferred) and the results of likelihood ratio tests (a significant result indicating that the more complex of the two nested models in the comparison is preferred) and were built using lme4, lmerTest, and multcomp packages. Tukey's post-hoc pair-wize comparisons (Multiple Comparisons of

TABLE 1 | T30 measured in Whisper Room conditions and 3 simulated environments (Low, Medium, and High) per octave band. Background noise conditions with and without speech-shaped noise spectrum per octave band.

	125 Hz	250 Hz	500 Hz	1 KHz	2 KHz	4 KHz	8 KHz
T30 whisper room (s)	0.164	0.129	0.079	0.061	0.064	0.054	0.048
T30 low (s)	0.512	0.821	1.071	1.191	0.922	0.799	0.016
T30 medium (s)	1.318	1.279	1.383	1.403	1.351	1.270	1.161
T30 high (s)	1.763	1.721	1.965	1.835	1.371	1.163	0.884
Background noise (dB)	34.0	26.7	16.0	14.8	13.4	15.5	16.4
Speech-Shaped Noise (dB)	59.8	62.0	66.6	60.3	64.2	59.7	55.4

The measurements were performed with the HATS.

TABLE 2 | Mean values and standard error (se) for the variable SPL in dB(A), perceived vocal comfort, and control in %, for the 14 conditions.

T30	Gain	Noise	SPL (dB)	se	Comfort (%)	se	Control (%)	se
Whisper room	0	No noise	73.3	0.06	79.1	3.90	85.4	2.70
Whisper room	0	Speech-shaped	76.7	0.06	54.7	4.22	61.7	4.35
Low	5	No noise	72.8	0.06	85.0	2.86	86.4	2.67
Low	5	Speech-shaped	76.0	0.06	65.5	4.56	73.7	4.21
Low	10	No noise	72.0	0.06	81.2	3.24	82.1	3.17
Low	10	Speech-shaped	74.4	0.06	69.5	4.37	76.5	3.87
Medium	5	No noise	72.9	0.06	82.8	3.77	84.6	3.88
Medium	5	Speech-shaped	75.8	0.06	70.0	3.89	75.8	3.87
Medium	10	No noise	71.8	0.06	76.7	4.43	84.8	3.12
Medium	10	Speech-shaped	74.0	0.06	79.0	3.33	83.1	3.20
High	5	No noise	73.1	0.06	82.3	3.44	85.7	2.84
High	5	Speech-shaped	76.0	0.06	67.8	3.88	75.4	3.75
High	10	No noise	71.6	0.06	81.0	3.68	81.9	3.41
High	10	Speech-shaped	73.9	0.06	77.2	3.83	81.4	3.65

Means: Tukey Contrasts) were performed to examine the differences between all levels of the fixed factors of interest. These are pair-wize z tests, where the z statistic represents the difference between an observed statistic and its hypothesized population parameter in units of the standard deviation. The p-values for these tests were adjusted using the default single-step method (Hothorn et al., 2008). The LME output includes the estimates of the fixed effects coefficients, the standard error associated with the estimate, the degrees of freedom (df), the test statistic (t), and the p-value. The Satterthwaite method is used to approximate degrees of freedom and calculate p-values.

RESULTS

Six Linear Mixed Effects (LME) models were run, two LME for each of the three different response variables: SPL, vocal comfort, and vocal control. The first of the two sets of LME models focused on gain as a fixed effect, while the second on T30. Both of them considered the effect of noise and gender as a fixed factor. The results section is divided into two subsections: 1) effects of gain and noise on SPL, vocal comfort, and vocal control, and 2) effects of T30 and noise on SPL, vocal comfort, and vocal control. **Table 2** summarizes the outcomes for the 14 conditions.

Effects of Gain

Effects of Gain and Noise on SPL

A Linear Mixed Effects (LME) model was run with the response variable SPL [in dB(A)]. This model, reported in **Table 3**, has the following fixed factors 1) gain, 2) noise, 3) gender, and 4) the interaction of gain and noise. The random effects were 1) T30, 2) chronological task order, 3) time (where time was measured in ms for each participant overall assessment), and 4) identification number of each participant. The reference levels used in the models were: 0 dB for gain, background without speech-shaped noise (No Noise) for noise condition, and female for gender.

The estimates of standard deviation for time as a random effect was 1.26 dB(A), for participant identification number was 2.78 dB(A), for order was 0.20 dB(A), and for T30 was 0.07, whereas the residual standard deviation was 6.49 dB(A). The mean variation in SPL from 0dB to 5 dB of gain for no noise added condition, was −0.31 dB(A), while it was −1.41 dB(A) from 0dB to 10 dB. As shown in **Figure 2**, when the speech-shaped noise is added, overall, the voice SPL increases 3.49 dB(A). When noise was added, the differences from 5 to 10 dB to the reference level (0 dB), were −0.78 and −2.65 dB(A), respectively. Since the gender was statistically significant, **Figure 2** differentiate among females and males where, generically, males were louder than females by 2.93 dB(A).

Post-hoc comparisons were made considering the effect of gain and its interaction with noise. These comparisons confirmed that, overall, SPL measured in 0 dB of gain condition was higher

TABLE 3 | LME models fit by REML for the response variable SPL and the fixed factors 1) gain, 2) noise, 3) gender, and the interaction between gain and noise.

	Estimate	Std. Error	df	t value	*p*-value	
(Intercept)	71.76	0.68	31.7	104.89	<0.001	***
Gain 5	−0.31	0.10	5.9	−3.00	0.024	*
Gain 10	−1.41	0.10	5.9	−13.66	<0.001	***
Noise speech-shaped	3.49	0.08	176,742.7	45.12	<0.001	***
Gender male	2.93	1.02	30.0	2.86	0.008	**
Gain 5: noise speech-shaped	−0.47	0.09	184,158.1	−5.23	<0.001	***
Gain 10: noise speech-shaped	−1.24	0.09	186,146.1	−13.93	<0.001	***

*Signif. codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1, '' < 1.*
The reference levels were: 0 dB for gain, without speech-shaped noise (No Noise) for noise condition, and female for gender.

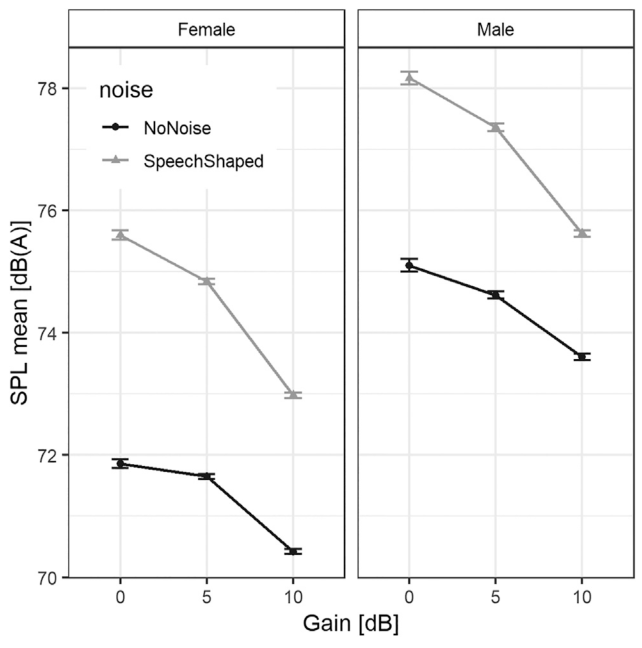


FIGURE 2 | Mean SPL in dB(A) vs. room gain regarding noise conditions and gender. The error bars indicate standard error.

than both, that in the condition with 5 dB of gain (−0.54 dB(A), SE = 0.09, $z = -5.87$, $p< 0.001$) and in 10 dB of gain condition (−2.03 dB(A), SE = 0.09, $z = -21.95$, $p < 0.001$), whereas the difference between 5 and 10 dB gain conditions was −1.49 dB(A) (SE = 0.03, $z = -46.98$, $p < 0.001$).

Effects of Gain and Noise on Vocal Comfort

One LME model was run with the response variable self-reported vocal comfort in % (0 = 'not at all comfortable,' 100 = 'extremely comfortable') and the fixed factors 1) gain, 2) noise, 3) gender, and 4) the interaction between gain and noise. The random effects were 1) T30, 2) chronological task order, and 3) participant. The output of this model is reported in **Table 4**. The estimate of standard deviation for participant as a random effect was 14.58%, for order was 2.56% and for T30 was 0.00%, whereas the residual standard deviation was 14.74%. The mean increase in self-reported vocal comfort from 0 to 5 dB of gain was 3.55%, while it was 0.27% from 0 to 10 dB; in the conditions without noise added. As shown in **Figure 3**, the vocal comfort decreased by −25.31% when the speech-shaped noise was added. For these conditions, when the noise was added, the mean increase in self-reported vocal comfort from 0dB to 5 dB of gain was 13.43%, while it was 20.77% from 0dB to 10 dB.

Post-hoc comparisons confirmed that, overall, the vocal comfort measured in the condition with 0 dB of gain was lower than that in both the condition with 5 dB of gain (8.50%, SE = 2.26, $z = 3.76$, $p < 0.001$), and the condition with 10 dB of gain (10.52%, SE = 2.26, $z = 4.66$, $p < 0.001$). Furthermore, the vocal comfort reported in the condition with 10 dB of gain was 2.02% higher than that in the condition with 5 dB of gain (SE = 1.59, $z = 1.27$, $p = 0.406$).

Effects of Gain and Noise on Vocal Control

The analysis of vocal control was similar to vocal comfort. One LME model was run with the response variable self-reported vocal control in % (0 = 'not at all controlled,' 100 = 'extremely controlled') and the fixed factors 1) gain, 2) noise, 3) gender, and 4) the interaction between gain and noise. The random effects were 1) T30, 2) chronological task order, and 3) participant. The output of the model is reported in **Table 5**. The estimate of standard deviation for participants as a random effect was 13.47%, for order was 2.71 and 0.00% for T30, whereas the residual standard deviation was 13.20%. The mean decrease in self-reported vocal control from 0 to 5 dB of gain was 0.45%, while it was 3.01%, from 0 to 10 dB; in the conditions without noise added. As shown in **Figure 4**, the vocal control decreased by 24.19% when the speech-shaped noise was added. For these conditions, when the noise was added, the mean increase in self-reported vocal control from 0 to 5 dB of gain was 13.28%, while it was 18.87% from 0 to 10 dB.

Post-hoc comparisons regarding the interactions between gain and noise confirmed that, overall, the vocal control measured in the condition with 0 dB of gain was lower than that in both the conditions with 5 dB of gain (6.42%, SE = 2.03, $z = 3.17$, $p = 0.004$) and the condition with 10 dB of gain (7.93%, SE = 2.02, $z = 3.92$, $p < 0.001$). Furthermore, the vocal control reported in the condition with 10 dB of gain was 1.51% higher than that in the condition with 5 dB of gain (SE = 1.43, $z = 1.06$, $p = 0.533$).

Effects of Reverberation Time (T30)

Effects of Reverberation Time and Noise on SPL

One LME model was run with the response variable SPL (in dB(A)). This model has 1) T30, 2) noise, 3) gender, and 4) the

TABLE 4 | LME models fit by REML for the response variable self-reported comfort and the fixed factors 1) gain, 2) noise, 3) gender, and 4) the interaction between gain and noise.

	Estimate	Std. Error	df	*t* value	*p*-value	
(Intercept)	80.40	4.67	61.2	17.23	<0.001	***
Gain 5	3.55	3.18	367.4	1.12	0.265	
Gain 10	0.27	3.18	367.1	0.08	0.993	
Noise speech-shaped	−25.32	3.91	369.4	−6.48	<0.001	***
Gender male	−1.75	5.64	27.0	−0.31	0.759	
Gain 5: noise speech-shaped	9.88	4.51	368.7	2.19	0.029	*
Gain 10: noise speech-shaped	20.50	4.50	367.4	4.55	<0.001	***

*Signif. codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1, '' < 1*
The reference levels were: 0 dB for gain, without speech-shaped noise (No Noise) for noise condition, and female for gender.

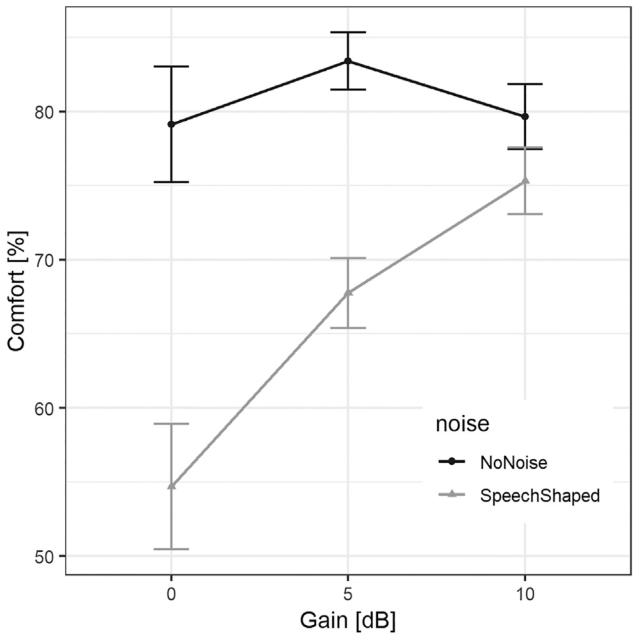


FIGURE 3 | Mean self-reported vocal comfort in % (0 = 'not at all comfortable', 100 = 'extremely comfortable') across participants per gain level. The error bars indicate ± standard error.

interaction between T30 and noise as fixed factors, as reported in **Table 6**. The random effects were 1) gain, 2) chronological task order, 3) time (where time was measured in ms for each participant overall assessment), and 4) identification number of each participant. The reference levels used in this model were: Low T30, background without speech-shaped noise (No Noise) for noise conditions, and female for gender.

The differences among T30 conditions were more pronounced in the noise added conditions. The estimates of standard deviation for time as a random effect was 1.25 dB(A), for participant was 2.82 dB(A), for order was 0.26 dB(A), and for gain was 0.85 dB(A), whereas the residual standard deviation was 6.43 dB(A). As shown in **Figure 5**, the mean variation in SPL from the Low T30 to Medium T30 was −0.08 dB(A), and a variation of −0.01 dB(A) from Low T30 to High T30, without noise added. When artificial speech-shaped noise was present, voice SPL increases 2.80 dB(A). For noise added conditions, the differences were −0.25 and −0.23 dB(A) for Low T30 vs. Medium T30 and Low T30 versus High T30, respectively. Generically, males were louder than females by 2.93 dB(A).

Post-hoc comparisons including interaction between T30 and noise confirmed that, overall, SPL measured in Low T30 condition was higher than that in the condition with Medium T30 (−0.17 dB(A), SE = 0.04, $z = -4.28$, $p < 0.001$) and in the condition with High T30 (−0.12 dB(A), SE = 0.04, $z = -3.01$, $p = 0.007$), whereas the difference between the condition with Medium T30 and High T30 was 0.05 dB(A) (SE = 0.04, $z = 1.22$, $p = 0.443$).

Effect of Reverberation Time and Noise on Vocal Comfort

To analyze the effects of T30 on vocal comfort, another LME model was run with the response variable self-reported vocal comfort (in %) and the fixed factors 1) T30, 2) noise, 3) gender, and 4) the interaction between T30 and noise. The random effects were 1) gain, 2) chronological task order, and 3) participant. The output of this model is reported in **Table 7**. The estimate of standard deviation for participant as a random effect was 15.08%, for order was 2.66%, and for gain was 0.85%, whereas the residual standard deviation was 14.05%. The mean decrease in self-reported vocal comfort, without noise added, from Low T30 to Medium T30 was 3.47%, while it was 1.15% from Low T30 to High T30. As shown in **Figure 6**, when the artificial speech-shaped noise was present, the vocal comfort decreased by 16.01%. For noise added conditions, there was an increase of comfort when T30 factors were higher than Low T30, 7.67% for Medium T30, and 5.48% for High T30. Generically, males' comfort was lower than females by 2.78%, with no statistical significance.

Post-hoc comparisons regarding interaction between T30 and noise confirmed that, overall, the vocal comfort measured in Low T30 condition was lower than that in Medium T30 condition (2.10%, SE = 1.86, $z = 1.13$, $p = 0.498$) and High T30 condition (2.16%, SE = 1.88, $z = 1.15$, $p = 0.483$), whereas the difference between Medium T30 and High T30 was 0.06% (SE = 1.86, $z = 0.03$, $p = 0.999$). None of these comparisons were statically significant.

TABLE 5 | LME models fit by REML for the response variable Control and the fixed factors 1) gain, 2) noise, 3) gender, 4) the interaction between gain and noise.

	Estimate	Std. Error	df	*t* value	*p*-value	
(Intercept)	86.39	4.28	59.9	20.19	<0.001	***
Gain 5	−0.45	2.85	366.3	−0.16	0.874	
Gain 10	−3.01	2.85	366.1	−1.06	0.291	
Noise speech-shaped	−24.19	3.51	368.4	−6.90	<0.001	***
Gender male	−1.19	5.20	27.0	−0.23	0.821	
Gain 5: noise speech-shaped	13.73	4.04	367.6	3.40	<0.001	***
Gain 10: noise speech-shaped	21.88	4.04	366.3	5.42	<0.001	***

*Signif. codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1, '' < 1.*
The reference levels were: 0 dB for gain, without speech-shaped noise (No Noise) for noise condition, and female for gender.

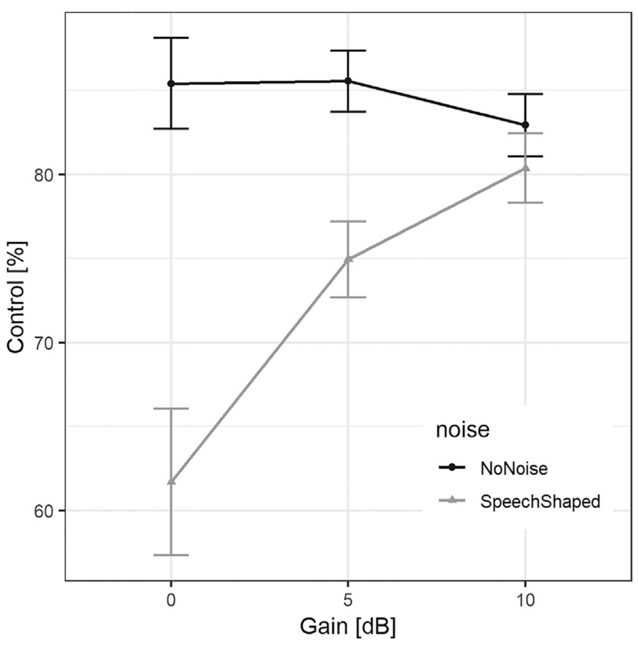


FIGURE 4 | Mean self-reported vocal control in % across participants per gain level. The error bars indicate ±standard error.

Effect of Reverberation Time and Noise on Vocal Control

To analyze the effects of T30 on vocal control, a final LME model was run with the response variable self-reported vocal control in % (0 = 'not at all controlled,' 100 = 'extremely controlled') and the fixed factors 1) T30, 2) noise, 3) gender and 4) the interaction between T30 and noise. The random effects were 1) gain, 2) chronological task order, and 3) participant. The output of this model is reported in **Table 8**. The estimate of standard deviation for participant as a random effect was 13.83%, for order was 2.54%, and for gain was 0.00%, whereas the residual standard deviation was 12.79%. The mean decrease in self-reported vocal control, without noise added, was 0.16% from Low T30 to Medium T30, while it was 0.56% from Low T30 to High T30. As shown in **Figure 7**, when the artificial speech-shaped noise was present, the vocal comfort decreased by 9.17%. For noise added conditions, there was an increase of control when T30 factors were higher than Low T30, 4.37% for Medium T30, and 3.07% for High T30. Generically, males' control was lower than females by 2.06%, with no statistical significance.

Post-hoc comparisons regarding interaction between T30 and noise confirmed that, overall, the vocal control measured in the condition Low T30 was lower than that in Medium T30 condition (2.10%, SE = 1.70, $z = 1.24$, $p = 0.430$) and High T30 condition (1.25%, SE = 1.71, $z = 0.73$, $p = 0.745$), whereas the control was 0.85% lower in High T30 than Medium T30 (SE = 1.69, $z = -0.50$, $p = 0.871$). None of these comparisons were statically significant.

DISCUSSION

In this study, several acoustics scenarios have been virtually created by modifying the external gain, as well as reverberation time, and by adding speech-shaped noise on the overall external auditory feedback. The speech adjustments in terms of SPL, and self-reported vocal comfort and control were measured in the aforementioned virtual scenarios.

Effect of Noise and Gender

Overall, the mean SPL at 15 cm from the mouth was measured as 75.26 and 72.48 dB(A) for conditions with and without speech-shaped noise, respectively. The equivalent level of the speech-shaped noise was 70 dB(A), while the background noise in the whisper room was 25 dB(A) when no noise was added. The increase in SPL when artificial noise was added is consistent with the Lombard effect (Lombard, 1911), which refers to the tendency of speakers to raise their voice in order to be understood in noisy environments. As a result of adding background noise, the perceived vocal comfort and control decreased by 12.06 and 9.05%, respectively. This decrease in self-reported vocal comfort and control, when noise was added, confirmed the tendency showed by Bottalico et al. (2015) in real rooms. Bottalico et al. (2015) showed that the differences for comfort and control on normal voice production were estimated to be 11.1 and 9.4% lower when noise was added. Even if gender was a statistically significant factor in the regulation of voice SPL [i.e., males were louder than females by 2.93 dB(A)], vocal comfort and control were not statistically different between gender in the two sets of conditions with and without noise.

TABLE 6 | LME models fit by REML for the response variable SPL and the fixed factors 1) T30, 2) noise, 3) gender, and 4) the interaction between T30 and noise.

	Estimate	Std. Error	df	*t* value	*p*-value	
(Intercept)	70.96	0.92	8.6	77.20	<0.001	***
T30 medium	−0.08	0.06	147,858.3	−1.40	0.162	
T30 high	−0.01	0.06	141,383.0	−0.11	0.911	
Noise speech-shaped	2.80	0.06	165,453.1	50.77	<0.001	***
Gender male	2.93	1.04	29.6	2.82	0.008	**
T30 medium: noise speech-shaped	−0.17	0.08	148,129.7	−2.21	0.026	*
T30 high: noise speech-shaped	−0.22	0.08	161,267.1	−2.89	0.004	**

*Signif. codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1, '' < 1.*
The reference levels were: Low T30, without speech-shaped noise (No Noise) for noise condition, and female for gender.

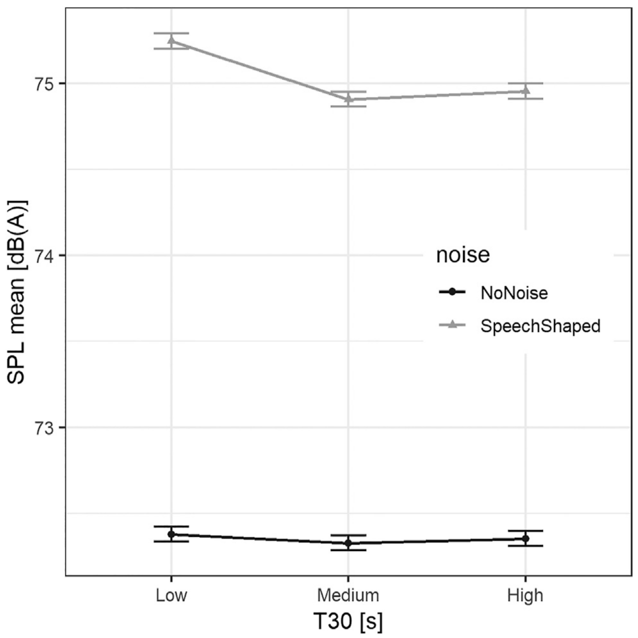


FIGURE 5 | Mean SPL in dB(A) vs. T30 regarding noise conditions. The error bars indicate ± standard error.

Effect of Gain

Regarding the effect of sidetone alteration (an alteration of the level at which a person is perceiving his/her own voice), when the gain factor was increased, participants decreased their voice SPL while reporting a higher level of vocal comfort and control. When the sidetone was increased, by 5 and 10 dB, participants showed a statistically significant decrease in their vocal SPL of 0.54 and 2.03 dB(A), respectively, confirming the results of Siegel and Pick (1974). Siegel and Pick (1974) conducted four different experiments which concluded that when the sidetone is increased by 10 dB there was a decrease of voice SPL within the range of 1.0–5.8 dB, with a mean estimate of 3.5 dB. Regarding vocal comfort and control, differences were statistically significant only in the conditions in which speech-shaped noise was added. In the different sidetone conditions related to this study, the results showed that the lowest levels of vocal comfort and control were reported in the condition without alteration of the sidetone (i.e., 0 dB of gain). Vocal comfort increased by 8.50 and 10.52%, for gain 5 and 10 dB, while control increased by 6.42 and 7.93%, for gain 5 dB and 10 dB, respectively, setting 0 dB as gain reference. This could have important implications for professional voice users. Specifically, many of these professionals who use electroacoustic systems for the playback of their own voice (like singers, broadcasters, or call center operators) may benefit from increasing the level of their monitors/headphones for an increase in the perceived vocal comfort and control and a decrease in the vocal effort. However, it is necessary to be careful not to increase the feedback level over the limit that may induce hearing loss.

Effect of Reverberation Time

According to the results presented in this study and comparing them with other studies cited in this discussion, we hypothesized that there is a trend for individuals to react differently while speaking in "middle-range" reverberation times, considering "middle-range" within the values that are explicitly cited on each study conditions. This does not pretend to concretely assess specific quantitative values or ranges of values (high or low), but a relationship within three or more different reverberation times (ordered by levels) when compared in the same experiment. Our opinion pertains to how "middle-range" reverberation times guide participants to improve their own comfort and control (along with the decrease of SPL). Nonetheless, it is important to remind that the reverberation times for this study were measured from the oral-binaural impulse response recorded by the HATS, rather than using the standardized method following the ISO 3382-2.

Following the former, these differences were 0.17 and 0.12 dB(A) lower for Medium T30 and High T30, as compared to Low T30, respectively. These results are similar to trends shown in previous studies (Bottalico and Astolfi, 2012; Puglisi et al., 2014; Durup et al., 2015; Puglisi et al., 2017), where voice SPL was presented to lower down in medium ranges of reverberation times. The medium-range conditions for reverberation times on those experiments were within 0.7 and 0.9 s, which indicate that there is a trend for higher comfort and control when reverberation conditions are in-between a range of values, i.e., by comparing higher and/or lower values of reverberation times with medium-range conditions. Whether for gain conditions the variations in SPL were substantial, the effect of variations of T30 were rather small.

TABLE 7 | LME models fit by REML for the response variable self-reported vocal comfort and the fixed factors 1) T30, 2) noise, 3) gender, and 4) the interaction between T30 and noise.

	Estimate	Std. Error	df	*t* value	*p*-value	
(Intercept)	84.24	4.36	37.7	19.31	<0.001	***
T30 medium	−3.47	2.65	311.7	−1.31	0.191	
T30 high	−1.15	2.66	311.9	−0.43	0.666	
Noise speech-shaped	−16.01	2.64	308.5	−6.05	<0.001	***
Gender male	−2.78	5.83	27.0	−0.48	0.637	
T30 Medium: Noise Speech-Shaped	11.13	3.75	311.7	2.97	0.003	**
T30 High: Noise Speech-Shaped	6.62	3.73	308.8	1.77	0.077	

*Signif. codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1, '' < 1.*
The reference levels were: Low T30, without speech-shaped noise (No Noise) for noise condition, and female for gender.

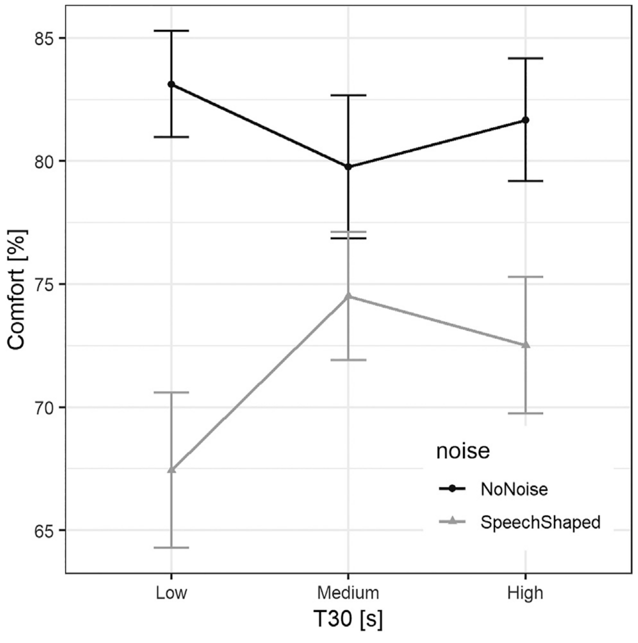


FIGURE 6 | Mean self-reported vocal comfort in % (0 = 'not at all comfortable', 100 = 'extremely comfortable') across participants per T30. The error bars indicate ± standard error.

The effect of reverberation time in self-reported vocal comfort had a similar trend, regarding Medium T30. Comfort was 2.10 and 2.16% higher for Medium T30 and High T30 than that from Low T30. Also, there was an increase in control for Medium T30 and High T30 of 2.10 and 1.25% higher with respect to Low T30. This might indicate a greater comfort and control for middle-range reverberation times, as opposed to Bottalico et al. (2016a) study, where an increase of comfort is shown in an anechoic and a reverberant room, as opposed to a semi-reverberant room. Bottalico et al. (2016a) reported 3.4% higher comfort for anechoic and 0.8% higher comfort for a reverberant room, both compared with the semi-reverberant room. Similar behavior was shown in that same study for control, where it was 4.5% higher for anechoic and 3.9% for reverberant than that from the semi-reverberant. It is important to point out that in Bottalico et al. (2016a), the authors presented two voice styles (normal and loud) and calculate averages among those two voice styles to give estimates on ΔSPL magnitudes. These findings could lead to a misunderstanding on self-reported vocal parameters because in the loud style, the voice intensity was higher and, consequently, the reflected sound was more intense.

More investigation is expected on this topic to build up a better understanding of how T30 is affecting (or even if it is affecting at all) voice production in a meaningful way. However, the variations in SPL suggest that lower vocal demands were experienced by talkers in Medium T30 conditions (T30 = 1.39 s).

Limitations and Future Directions

Some limitations of this study were the lack of ecological validity, i.e., the fact that laboratory conditions were virtually simulated, not realistic. Moreover, the use of standardized reading material instead of spontaneous speech avoids variations in phonation time, which may represent a limitation on evaluating self-reported vocal comfort and control. Furthermore, due to participants being American English native speakers, hinders the generalization of these results to speakers in other languages and/or in other forms of spoken English.

In the future, studies on simulated environments could have an increase in the levels of reverberation time, gain, and background noise. By broadening the range of reverberation times, gain levels, and noise conditions might show up further recommendations about acoustical conditions that would maximize voice comfort and control while minimizing SPL and voice effort. Finally, it is important to point out that adding other acoustical objective measurements would be useful for better understanding the variations on voice comfort and control, such as speed rate of speech and frequency of utterances, which are directly related to the movement of the vocal folds, thus with voice effort and fatigue.

CONCLUSION

The aim of this study was to evaluate the effect of external auditory feedback, such as reverberation time, altered sidetone (i.e., gain level), and background noise. The external auditory

TABLE 8 | LME models fit by REML for the response variable self-reported vocal control and the fixed factors 1) T30, 2) noise, 3) gender, and 4) the interaction between T30 and noise.

	Estimate	Std. Error	df	*t* value	*p*-value	
(Intercept)	85.32	3.96	40.1	21.53	<0.001	***
T30 medium	−0.16	2.41	312.6	−0.07	0.946	
T30 high	−0.56	2.43	312.9	−0.23	0.816	
Noise speech-shaped	−9.18	2.41	309.5	−3.81	<0.001	***
Gender male	−2.07	5.35	27.0	−0.39	0.702	
T30 medium: noise speech-shaped	4.53	3.42	312.5	1.32	0.186	
T30 high: noise speech-shaped	3.63	3.40	309.8	1.07	0.285	

*Signif. codes: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1, '' < 1.*
The reference levels were: Low T30, without speech-shaped noise (No Noise) for noise condition, and female for gender.

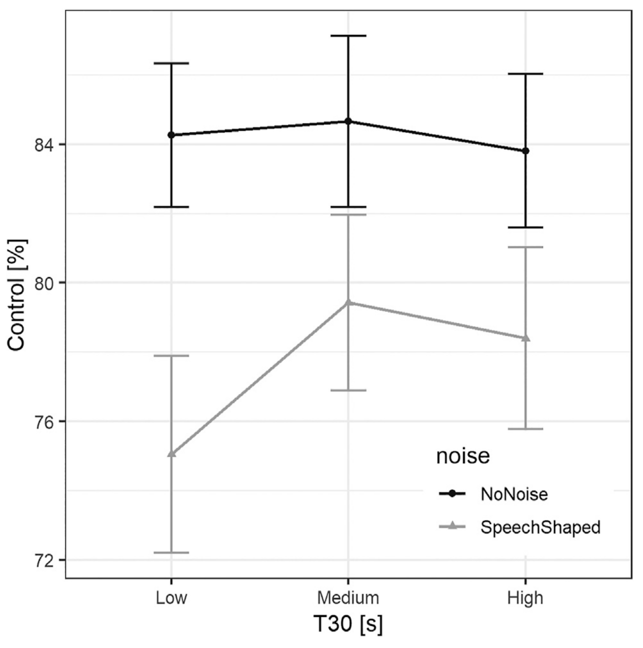


FIGURE 7 | Mean self-reported vocal control in % across participants per T30. The error bars indicate ± standard error.

feedback was modified by changing the sidetone with three levels of gain (0, 5, and 10 dB), these changes showed that an increase in the sidetone led to a decrease of SPL and an increase in self-perception of voice comfort and control. This information is important because it can guide vocal health promotion actions helping to decrease the occurrence of voice disorders and improve speakers´ voice-related quality of life. For instance, among occupational voice users, such as teachers and call center operators, considering their high risk of developing voice disorders associated with their working conditions (Pelegrín-García and Brunskog, 2012; Cantor-Cutiva et al., 2013; Bottalico et al., 2015; Bottalico et al., 2016a; Cantor-Cutiva and Burdorf, 2016; Bottalico et al., 2017a; Banks et al., 2017; Cipriano et al., 2017; Cantor-Cutiva et al., 2019; Carrillo-Gonzalez et al., 2019), it is determinant to identify specific elements that can help to improve "healthy" occupational voice use. Therefore, knowing that sidetone may help to decrease SPL and increase self-perceived voice comfort and control, speech and language pathologists at the workplaces may train occupational voice users using sidetone to strengthen voice comfort and control and reduce occupational voice misuse.

In addition, results on Medium T30 being associated with the highest voice comfort and control (along with lowest SPL), when speech-shaped noise was added are also interesting. At the workplaces, professionals from Safe and Health at Work may consider these results for designing "safe" workplaces (classrooms, call center rooms, schools) for "healthy" occupational voice use. In this way, the intervention actions would start in the environment and not in the workers, which is suggested in the hierarchy of controls (Castro, 2003).

Finally, all the experiments conducted in this study were based on simulated acoustical environments, which represents a great step forward in the development of alternative techniques to performs research on voice production and sound propagation.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the database consists of speech recordings, which are considered identifiable data. Requests to access the datasets should be directed to Pasquale Bottalico, pb81@illinois.edu.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Office of Protection of Research Subject at the University of Illinois at Urbana Champaign (IRB 18179). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

PB conceived and designed the study. TS and PB performed the experiments and analyze the data. TS and PB wrote the paper. EH and LC-C participate in revising the paper.

REFERENCES

Akaike, H. (1998). "Information Theory and an Extension of the Maximum Likelihood Principle," in *Selected Papers of Hirotugu Akaike* (New York, NY: Springer), 199–213. doi:10.1007/978-1-4612-1694-0_15

Banks, R. E., Bottalico, P., and Hunter, E. J. (2017). The Effect of Classroom Capacity on Vocal Fatigue as Quantified by the Vocal Fatigue Index. *Folia Phoniatr Logop* 69 (3), 85–93. doi:10.1159/000484558

Berardi, M. L. (2015). *Quantification of Vocal Effort in Varying Acoustical Environments [Honors]*. Provo, UT: Brigham Young University.

Berardi, M. L. (2020). *Validation and Application of Experimental Framework for the Study of Vocal Fatigue*. [Doctoral dissertation]. [East Lansing, MI]: Michigan State University. doi:10.25335/epf5-c186

Black, J. W. (1950). The Effect of Room Characteristics upon Vocal Intensity and Rate. *The J. Acoust. Soc. Am.* 22 (2), 174–176. doi:10.1121/1.1906585

Bottalico, P., Astolfi, A., and Hunter, E. J. (2017a). Teachers' Voicing and Silence Periods during Continuous Speech in Classrooms with Different Reverberation Times. *J. Acoust. Soc. Am.* 141 (1), EL26–EL31. doi:10.1121/1.4973312

Bottalico, P., and Astolfi, A. (2012). Investigations into Vocal Doses and Parameters Pertaining to Primary School Teachers in Classrooms. *J. Acoust. Soc. America* 131 (4), 2817–2827. doi:10.1121/1.3689549

Bottalico, P., Graetzer, S., and Hunter, E. (2016b). Effect of Reverberation Time on Vocal Fatigue. *Proc. Int. Conf. Speech Prosody* 2016, 494–497. doi:10.21437/SpeechProsody.2016-101

Bottalico, P., Graetzer, S., and Hunter, E. J. (2016a). Effects of Speech Style, Room Acoustics, and Vocal Fatigue on Vocal Effort. *J. Acoust. Soc. America* 139 (5), 2870–2879. doi:10.1121/1.4950812

Bottalico, P., Graetzer, S., and Hunter, E. J. (2015). Effects of Voice Style, Noise Level, and Acoustic Feedback on Objective and Subjective Voice Evaluations. *J. Acoust. Soc. Am.* 138 (6), EL498–EL503. doi:10.1121/1.4936643

Bottalico, P., Passione, I. I., Graetzer, S., and Hunter, E. (2017b). Evaluation of the Starting Point of the Lombard Effect. *Acta Acustica United Acustica* 103 (1), 169–172. doi:10.3813/AAA.919043

Bottalico, P. (2017). Speech Adjustments for Room Acoustics and Their Effects on Vocal Effort. *J. Voice* 31 (3), e1–392. doi:10.1016/j.jvoice.2016.10.001

Cantor Cutiva, L. C., Vogel, I., and Burdorf, A. (2013). Voice Disorders in Teachers and Their Associations with Work-Related Factors: a Systematic Review. *J. Commun. Disord.* 46 (2), 143–155. doi:10.1016/j.jcomdis.2013.01.001

Cantor-Cutiva, L. C., Bottalico, P., and Hunter, E. (2019). Work-related Communicative Profile of Radio Broadcasters: a Case Study. *Logopedics Phoniatrics Vocol.* 44 (4), 178–191. doi:10.1080/14015439.2018.1504983

Carrillo-Gonzalez, A., Camargo-Mendoza, M., and Cantor-Cutiva, L. C. (2019). Relationship between Sleep Quality and Stress with Voice Functioning Among College Professors: a Systematic Review and Meta-Analysis. *J. Voice* 13, 141. doi:10.1016/j.jvoice.2019.11.001

Castro, A. B. (2003). 'Hierarchy of Controls': Providing a Framework for Addressing Workplace Hazards. *AJN Am. J. Nurs.* 103 (12), 104. doi:10.1097/00000446-200312000-00030

Cipriano, M., Astolfi, A., and Pelegrín-García, D. (2017). Combined Effect of Noise and Room Acoustics on Vocal Effort in Simulated Classrooms. *J. Acoust. Soc. America* 141 (1), EL51–EL56. doi:10.1121/1.4973849

Cutiva, L. C. C., and Burdorf, A. (2016). Work-related Determinants of Voice Complaints Among School Workers: An Eleven-Month Follow-Up Study. *Am. J. Speech Lang. Pathol.* 25 (4), 590–597. doi:10.1044/2016_AJSLP-14-0191

Durup, N., Shield, B., Dance, S., and Sullivan, R. (2015). An Investigation into Relationships between Classroom Acoustic Measurements and Voice Parameters of Teachers. *Building Acoust.* 22 (3-4), 225–241. doi:10.1260/1351-010X.22.3-4.225

Fairbanks, G. (1960). *"The Rainbow Passage," in Voice and Articulation Drillbook*. 2nd Edition. New York: Harper & Row, 124–139.

Ferguson, S. H., and Kewley-Port, D. (2007). Talker Differences in Clear and Conversational Speech: Acoustic Characteristics of Vowels. *J. Speech Lang. Hear. Res.* 50 (5), 1241–1255. doi:10.1044/1092-4388(2007/087

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biom. J.* 50 (3), 346–363. doi:10.1002/bimj.200810425

Hunter, E. J., Cantor-Cutiva, L. C., Van Leer, E., Van Mersbergen, M., Nanjundeswaran, C. D., Bottalico, P., et al. (2020). Toward a Consensus Description of Vocal Effort, Vocal Load, Vocal Loading, and Vocal Fatigue. *J. Speech Lang. Hear. Res.* 63 (2), 509–532. doi:10.1044/2019_JSLHR-19-00057

Hunter, E. J., Švec, J. G., and Titze, I. R. (2006). Comparison of the Produced and Perceived Voice Range Profiles in Untrained and Trained Classical Singers. *J. Voice* 20 (4), 513–526. doi:10.1016/j.jvoice.2005.08.009

ISO 28802 (2012). *Ergonomics of the Physical Environment—Assessment of Environments by Means of an Environmental Survey Involving Physical Measurements of the Environment and Subjective Responses of People*. Geneva: International Organization for Standardization.

ISO 3382-2 (2008). *Acoustics - Measurement of Room Acoustic Parameters—Part 2: Reverberation Time in Ordinary Rooms*. Geneva: International Organization for Standardization.

ISO 9921 (2003). *Ergonomics - Assessment of Speech Communication*. Geneva: International Organization for Standardization.

Junqua, J. C. (1993). The Lombard Reflex and its Role on Human Listeners and Automatic Speech Recognizers. *J. Acoust. Soc. Am.* 93 (1), 510–524. doi:10.1121/1.405631

Krause, J. C., and Braida, L. D. (2004). Acoustic Properties of Naturally Produced Clear Speech at Normal Speaking Rates. *J. Acoust. Soc. America* 115 (1), 362–378. doi:10.1121/1.1635842

Lane, H., and Tranel, B. (1971). The Lombard Sign and the Role of Hearing in Speech. *J. Speech Hearing Res.* 14 (4), 677–709. doi:10.1044/jshr.1404.677

Lezzoum, N., Gagnon, G., and Voix, J. (2016). Echo Threshold between Passive and Electro-Acoustic Transmission Paths in Digital Hearing Protection Devices. *Int. J. Ind. Ergon.* 53, 372–379. doi:10.1016/j.ergon.2016.04.004

Lombard, E. (1911). "Le signe de l'elevation de la voix," [The sign of voice rise] in. *Ann. des Maladies de L'Oreille du Larynx* 37 (2), 101–119.

McKenna, V. S., and Stepp, C. E. (2018). The Relationship between Acoustical and Perceptual Measures of Vocal Effort. *J. Acoust. Soc. America* 144 (3), 1643–1658. doi:10.1121/1.5055234

Pelegrín-García, D., Brunskog, J., and Rasmussen, B. (2014). Speaker-oriented Classroom Acoustics Design Guidelines in the Context of Current Regulations in European Countries. *Acta Acustica united with Acustica* 100 (6), 1073–1089. doi:10.3813/AAA.918787

Pelegrín-García, D., and Brunskog, J. (2012). Speakers' Comfort and Voice Level Variation in Classrooms: Laboratory Research. *J. Acoust. Soc. America* 132 (1), 249–260. doi:10.1121/1.4728212

Pelegrín-García, D., Smits, B., Brunskog, J., and Jeong, C.-H. (2011). Vocal Effort with Changing Talker-To-Listener Distance in Different Acoustic Environments. *J. Acoust. Soc. America* 129 (4), 1981–1990. doi:10.1121/1.3552881

Puglisi, G. E., Astolfi, A., Cantor Cutiva, L. C., and Carullo, A. (2017). Four-day-follow-up Study on the Voice Monitoring of Primary School Teachers: Relationships with Conversational Task and Classroom Acoustics. *J. Acoust. Soc. Am.* 141 (1), 441–452. doi:10.1121/1.4973805

Puglisi, G. E., Astolfi, A., Pavese, L., and Carullo, A. (2014). Long-term Vocal Parameters in Semi-anechoic and Reverberant Rooms. *Forum Acusticum* 14, 7–12. doi:10.3403/30218742

Riede, T., and Brown, C. (2013). Body Size, Vocal Fold Length, and Fundamental Frequency–Implications for Mammal Vocal Communication. *Nova Acta Leopoldina NF* 111 (380), 1–20. doi:10.1007/3-540-32504-2_10

Rollins, M. K., Leishman, T. W., Whiting, J. K., Hunter, E. J., and Eggett, D. L. (2019). Effects of Added Absorption on the Vocal Exertions of Talkers in a Reverberant Room. *J. Acoust. Soc. Am.* 145 (2), 775–783. doi:10.1121/1.5089891

Rowe, M. L. (2008). Child-directed Speech: Relation to Socioeconomic Status, Knowledge of Child Development and Child Vocabulary Skill. *J. Child. Lang.* 35 (1), 185–205. doi:10.1017/S0305000907008343

Siegel, G. M., and Pick, H. L., Jr (1974). Auditory Feedback in the Regulation of Voice. *J. Acoust. Soc. Am.* 56 (5), 1618–1624. doi:10.1121/1.1903486

Titze, I. R. (2000). *Principles of Voice Production (Second Printing)*. Iowa City, IA: National Center for Voice and Speech, 361–366.

Titze, I. R. (1999). Toward Occupational Safety Criteria for Vocalization. *Logopedics Phoniatrics Vocol.* 24, 49–54. doi:10.1080/140154399435110

Yiu, E. M.-L., and Yip, P. P. S. (2016). Effect of Noise on Vocal Loudness and Pitch in Natural Environments: an Accelerometer (Ambulatory Phonation Monitor) Study. *J. Voice* 30 (4), 389–393. doi:10.1016/j.jvoice.2015.05.016

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

METHODS
published: 04 June 2021
doi: 10.3389/fbuil.2021.687237

Multi-Detailed 3D Architectural Framework for Sound Perception Research in Virtual Reality

Josep Llorca-Bofí and Michael Vorländer*

Institute for Hearing Technology and Acoustics, RWTH Aachen University, Aachen, Germany

The presentation of architectural design in simulation environments demands strong 3D modeling abilities. Architects usually demonstrate presentation skills that mostly address researchers in the building simulation field. However, there is still a gap between the architect's deliverable and the contextual scenario for overarching research purposes, mainly caused by the lack of knowledge in the areas where research disciplines overlap. This dilemma is particularly present in the practice of 3D modeling for sound perception research in virtual reality since the building modelers must also gather diverse pieces of knowledge into a contained scenario: ranging from sound sources, sound propagation models to physically based material models. Grounded on this need, this article presents a comprehensive framework, defined by the visual and acoustic cues—geometries, materials, sources, receivers, and postprocessing—on one side and three levels of detail on the other. In this way, very specific research application needs can be covered, as well as a modular concept for future modeling demands. The interconnection between every model element is particularly designed, enabling the assembly among different modalities at different levels of detail. Finally, it provides targeted modeling strategies for architects, depicted in one indoor and one outdoor demonstration for auditory-visual research.

Keywords: virtual reality, auralization, visualization, architectural representation, soundscapes, CAD

OPEN ACCESS

Edited by:
Arianna Astolfi,
Politecnico di Torino, Italy

Reviewed by:
Francesco Aletta,
University College London, United Kingdom
Jooyoung Hong,
Chungnam National University, South Korea

****Correspondence:***
Josep Llorca-Bofí
josep.llorca@akustik.rwth-aachen.de

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *29 March 2021*
Accepted: *10 May 2021*
Published: *04 June 2021*

Citation:
Llorca-Bofí J and Vorländer M (2021) Multi-Detailed 3D Architectural Framework for Sound Perception Research in Virtual Reality. Front. Built Environ. 7:687237. doi: 10.3389/fbuil.2021.687237

INTRODUCTION

The practice of 3D modeling for research purposes in the field of virtual acoustics presents a typical dilemma as the tasks demand certain abilities from varied artistic fields (Calderon et al., 2006; Boeykens and Liège, 2011) and also require knowledge on the scientific field under research. In other words, it is common to find researchers conducting tests in the field of virtual acoustics who do not have skills in 3D modeling research scenarios. By contrast, there is also a large community of professional 3D drawers—including architects, graphic designers, and gaming developers—who find it difficult to deliver 3D data that fulfills the demands of researchers. This usually leads to two types of solutions: researchers forgo the possibilities of 3D modeling by constructing a simplified model, with all the time and effort constraints included; or, when they do collaborate, only 20% of researchers undertake close collaboration with 3D modeling experts to obtain the expected environment, with implications for the time and effort a project requires (Thery et al., 2019).

Virtual acoustic environments have demonstrated versatility in various research areas, as they allow easy manipulations of experimental test conditions or simulated acoustic scenes. Although the evolution of auditory and cognitive models is constantly pursued (Søndergaard and Majdak 2013;

Relaño-Iborra et al., 2019), listening experiments are still considered to be the gold standard (Brinkmann et al., 2019, Pausch and Fels 2019), usually necessitating a defined 3D environment. For easy manipulation of experimental conditions, it is desirable that acoustic conditions, types, positions, and the orientations of the involved sound sources, as well as the order of examined conditions, can be changed without physical modifications of the laboratory.

In the following, the person concerned with the task is called a "modeler." During the 3D preparation workflow of the research demands, modelers can easily lose track of what is necessary. In particular, the application of virtual acoustic environments (VAEs) with full control over virtual sound sources, playing back definable source signals, such as speech, music, technical, or synthesized signals, requires specific implementations in the 3D scenario (Vorländer 2020). The validity and reproducibility of this approach are further increased by the integration of source directivity (Monson et al., 2012; Shabtai et al., 2017), generic or individual human binaural data (Xie 2013; Thiemann and van de Par 2019), simulation of room acoustics (Vorländer 1989; Naylor 1993; Dalenbäck 1996), Doppler shifts in case of moving sources or receivers (Strauss, 1998; Wefers and Vorländer 2018), and diffraction filtering for urban environments (Svensson et al., 1999; Tsingos et al., 2001). These requirements fall within the competence of the acoustic specialist, who implements the corresponding filters on the 3D model but there are also prerequisites for the acoustic 3D model, such as the acoustic characterization of surfaces based on absorption and scattering properties (Vorländer 2020), the sizing of the objects in accordance with the target wavelengths (Pelzer and Vorländer 2010), and the extension of the numerical acoustics mesh arranged for the calculation of wave effects (Marburg 2002). Despite these are crucial decisions when modeling a scenario, not much research has yet been done on the automatic simplification of CAD models from details toward a specific acoustically relevant resolution (Vorländer 2020).

Despite the gap between what modelers know and what researchers demand, the modelers are also equipped with modeling strategies that mostly address researchers in architectural design and analysis: 3D modeling tasks are present in almost all fields of architectural practice and research. This is the reason why every architecture student, professional architect, interior designer, or building engineer can be considered as a potential "modeler." The use of 2D and 3D computer-aided design (CAD) tools is mandatory in order to communicate among professionals, and it was progressively introduced in the architecture student's curriculum (Clayton et al., 2002). On the research side, the simulation of different aspects of buildings is the common way to examine in detail the behavior of different architectural designs. Structural design based on a 3D wire-frame model, for example, provides optimized and fast tools for dimensioning the building structure (Hasan et al., 2019). Other fields, such as network approaches, use architectural simulation at the city scale in correlation with social events, such as busking (Clua et al., 2020). In the small and building scale, some finite element methods allow the calculation of a building structure, considering the interaction of the different parts as a unique mesh (Roca et al., 2010; Castellazzi et al., 2017). Modeling tasks are also useful in the simulation of hygrothermal interactions with the building envelope (Künzel et al., 2005) or in the generation of solar envelopes to improve the building comfort (de Luca et al., 2021), as well as in other diverse fields of building research. Additionally, the modeling software normally used by "modelers" is coveted for simulation demands. As examples, they are usually trained to define 3D geometries in software like AutoCAD[1], Rhinoceros3D,[2] Blender,[3] 3ds Max,[4] Maya,[5] or SketchUp[6]. Other software like Unity[7] and UnrealEngine[8] are also used as engines for virtual reality, offering powerful platforms for real-time architectural visualization. Last, the modeling paradigm has evolved in a way that the integration of geometrical data with additional information makes the task of the "modeler" desirable for building research. Thus, the common modeling technique that originated with CAD as a system was to automate the task of drafting 2D geometry. The emergence of 3D CAD initially focused almost entirely on creating geometry in support of visualization. More recently, object-oriented CAD systems replaced 2D symbols with building elements (objects) capable of representing the behavior of common building elements. These building elements can be displayed in multiple views, as well as having nongraphic attributes assigned to them, such as acoustic material properties, structural properties, assigned assembly rules, or parametric dimensions. Capturing these relationships and behaviors is just not possible in the previous CAD paradigm. Building information modeling (BIM) provides a single logical, consistent source for all information associated with the building. The knowledge of the aforementioned modeling techniques is what makes the "modeler" a valid actor to contribute to the acoustics research workflow.

In the field of architectural acoustics, practitioners and consultants use different commercial software with 2D and 3D layouts for room acoustics simulation: CATT-Acoustic,[9] Ease[10], and Odeon,[11] or Max[12]. All of these examples require a geometrical definition of the studied spaces as a list of points defining faces and faces defining volumes. The modeling strategies highly depend on each software, which needs

[1]https://www.autodesk.com/products/autocad/overview
[2]https://www.rhino3d.com/
[3]https://www.blender.org/
[4]https://www.autodesk.es/products/3ds-max/overview
[5]https://www.autodesk.com/products/maya/overview
[6]https://www.sketchup.com/
[7]https://unity.com/
[8]https://www.unrealengine.com/
[9]https://www.catt.se/
[10]https://ease.afmg.eu/
[11]https://odeon.dk/
[12]https://cycling74.com/

professional expertise. Additionally, other noncommercial simulation software such as Pachyderm[13] (van der Harten 2013), RAVEN,[14] Razr,[15] or rtSOFE,[16] developed for research purposes have a specific geometric definition for rooms. This range of software possibilities, including the respective modeling strategies they require, makes the approach difficult for the "modeler" and enlarges the gap between them and the researcher, as stated before. As observed in some examples, CATT-Acoustic demands a list of points in Cartesian coordinates, which can be extracted from a DXF format or usually defined in a .txt file. On the other hand, Raven also runs as a plugin in Sketchup and Rhinoceros3D, taking the geometry from there. Finally, EASE can import model information using the AutoCAD DXF file format after simplification of complex components. The three different software tools presented three modeling procedures that have to be carefully taken into account before the simulation work is done. For this reason, there is a need to gather necessary modeling knowledge that will be useful for potential "modelers."

As a starting point for similar future optimizations of the modeling workflows, this contribution presents a multi-detailed 3D architectural framework, divided into modules and ready to be used in virtual acoustic research applications, which include audio and/or visual modalities. It defines three levels of detail, which cover three different levels in the architectural approach: a low level of detail, as far-away scales such as urban, neighborhood, or landscape; a high level of detail, for close scales such as indoor or furniture-focused environments or; and a medium level of detail, for intermediate scales. Those three levels are defined for each of the acoustical and visual cues: geometry, materials, sources, and receivers. This framework provides the chance to combine different levels and different cues for a desired researcher-defined scenario. First, the general requirements regarding hardware and software requirements are described. Based on these requirements, design aspects and implementation the implementation of these designs are presented in detail. Finally, the article presents two demonstrations of an outdoor and an indoor environment, respectively.

MATERIALS AND EQUIPMENT

General Considerations

The sciences of architectural visual and acoustic simulation share some strategies when representing reality, but they also differ in several principles (Monedero 2015a; Monedero 2015b). The distinction emerges from the different physical behavior of light and sound phenomena and how they interact with the built environment. From the architectural point of view, it is useful to understand which are those physical phenomena in the real world and how the models simulate them. This requires a fundamental knowledge of concepts such as the generation, propagation, interaction, and perception of sound and light. After the description of those phenomena by experimental observation—what commonly generated numerical descriptions of reality over decades in research—the task of deriving analytical models might be useful for the prediction of new situations. This is where virtual reality may become a laboratory for research in architecture: by understanding the behavior of the human perception in the built environment, one can predict the perceived cues in a new architectural design before building it.

This is not always the common path for architects. Although it seems obvious that fundamental analyses of real scenarios are essential before trying to simulate them by virtual means, experience tells that this is not so. Evidence shows also that most of the 3D modelers with an architectural background, use simulation techniques before understanding their theoretical basis, relying on the techniques as feedback tools for different design options (Bouchlaghem et al., 2005; Attia et al., 2009; Thery et al., 2019). In addition, prerequisites in the fields of visual simulation are different to acoustic simulations. For this reason, the proposed framework allows for a comprehensive approach to both the visual and the acoustic models, applying the prerequisites in a language as close as possible to the architectural modeling tradition.

The prerequisites for simulated visual environments have been established in the field of computer graphics. They follow four broad subfields: geometry, or ways to represent and process surfaces; animation, ways to represent and manipulate motion; rendering, algorithms to reproduce light transport; and an imaging or processing acquisition, ways to reproduce the visual characteristics of objects (Foley et al., 1997). First, the geometry representation models imply the mathematical definition of space. This urges the modeler to define the environment objects through CAD platforms, which are commonly used by architects and modelers. Second, the animation representation methods range from camera movement until the creation of avatars. Since architecture elements are normally represented as static objects, the knowledge of camera properties and viewer movement are of the most relevance here. Where a deeper understanding is needed, it is rendering the properties of scenes. This requires the use of lighting units, lighting distribution, and types of artificial lighting characteristics, as well as a material definition of surfaces, scattering, and shading properties from the modeler side. Finally, the human interaction systems mainly cover the imaging prerequisites. They range from monitor screens; projections; 360 enveloping scenarios, such as "CAVE's"; or head-mounted displays (such as VR headsets); and other manual controllers, such as joysticks or controllers, which need a specific hardware setup knowledge.

The prerequisites for simulated auditory environments are defined in the field of virtual acoustics and, under the term "Auralization." They include three main components: a sound generation model, a sound transmission model, a signal processing, and sound reproduction (Vorländer 2020). The

[13]https://www.food4rhino.com/app/pachyderm-acoustical-simulation
[14]https://www.akustik.rwth-aachen.de/
[15]http://medi.uni-oldenburg.de/razr/
[16]https://www.ei.tum.de/aip/startseite/

sound generation model implies the characterization of sound sources, in terms of spatial directivities and temporal domain. This urges the modeler to locate the sources in a plausible way as well as design the coherent degrees of freedom for their movements. The sound propagation model takes care of the propagation, reflection, scattering, and diffraction phenomena inside of rooms or through the environment, as well as the effects of possible structural transmission effects in building structures. This second block requires the ability to detect which are the relevant polygons that will affect those phenomena, such as large polygonal surfaces close to the listener or diffraction edges from two adjacent polygons. Finally, the audio signal processing and sound reproduction can be performed *via* loudspeakers or headphones with a big range of binaural techniques. This demands detailed attention in the implementation of the 3D model in the laboratory.

Up to this point, significant differences can be noted between approaches to acoustic and visual simulations. The main difference shows up when confronting four to three main prerequisites. This makes the comprehensive understanding complex. Whereas visual simulation tradition considers "geometry" and "rendering" as two separate fields for the simulation workflows, the acoustic simulation framework devotes a sole model for propagation calculation—including geometry and material definition (Schröder, 2011). Since the architectural modeling is often object based—meaning that each object is defined by a separate 3D object (e.g., a pillar is geometrically defined by a prism and materially defined by a texture)—the separation between geometry and material definitions seems more adequate for an architectural framework. This distinction is included in the present work.

For the construction of the scenarios, the software platforms are applied as follows. SketchUp[17] is used for the geometrical definition of scenario in both acoustic and visual terms. SketchUp is also used for the acoustic material characterization of the scene, whereas 3D Studio Max provides a suitable environment to define the mapping for the visual material characterization. Unreal Engine 4 is used for the visual animation, rendering, and imaging processes. RAVEN, working as a plugin with SketchUp or Matlab,[18] is used for the definition of sound sources and sound receivers as well as for performing the sound propagation simulation. The complete setup can be designed to run in Windows with Unreal Engine 4 or higher, SketchUp, and Matlab installed. The system requirements depend on each software. Having set these components, the proposed framework is ready to be defined.

Design Preparation and Guidelines

The use of the framework for upcoming 3D models requires some design preparations. The distinction between the visual model and the acoustical model starts already in the preparatory considerations.

Regarding the visual model, at least the following requirements must be addressed:

- Extension of the model, in accordance with the test necessities.
- Extension of the test area for subjects, which determines the modeling extension of the visual model and whether far objects need to be modeled or just included in a spherical image around the viewer.
- Level of detail from the near objects until the distant ones, which determines the modeling load and strategy.
- Modeling technique, derived from the previous points, such as object-based or using photogrammetric techniques.
- Quality of photographic data, especially for texture and material definitions.
- Visual rendering strategies and craftsmanship, which can be learned and gained after experience.

Regarding the acoustic model, at least the following aspects must be covered:

- Extension of the model, considering the simulation of free-field sound propagation as well as reflection, scattering, and diffraction.
- Extension of the test, paying special attention to the subjects' closest area, where the finer definition of the acoustic meshes (polygons) will play an important role.
- Face (polygon) count control, assuring an efficient calculation.
- Level of detail of the meshes, balancing plausibility and computing effort.
- Quality of anechoic data.
- Quality of directivity properties of sound sources and receivers.
- Acoustic rendering strategies and craftsmanship, which can be achieved after having gained experience.

Finally, the coincidence check between both models must be assured.

THE FRAMEWORK

The framework is presented as a double-entry table. On the vertical axis, the types of cues are divided into two main groups: "visual" and "acoustic." Those are also divided into "geometry," "materials," "sources," and "receivers." On the horizontal axis, the cues are divided into three levels of detail: "low," "medium," and "high." **Table 1** shows an overview of the framework. In the following, every module is explained.

Visual Definitions

Geometries

The simplest geometric definition of the visual environment is a **Sphere** mesh. Spheres are centered at a fixed position which coincides with fixed viewer and listener positions. The sphere can be defined in CAD software as a polygonal mesh, with normals facing the center of the sphere. Spheres are mapped with a spherical wrap. The mapping of the sphere can be matched with a 360° texture or an HDRI sky.

The medium definition is a **Welded** mesh. Those are complex meshes, which can be defined as simplicial complexes that might be produced *via* 3D photogrammetric techniques or with laser

[17]https://www.sketchup.com

[18]https://de.mathworks.com/products/matlab.html

TABLE 1 | Multi-detailed 3D architectural framework.

Cues	Definitions	1	2	3
		Level of detail		
		Low	Medium	High
Visual	Geometries	**Sphere**	**Welded**	**Object**
	Materials	**Monochrome**	**Color**	**PBR**
	Lighting	**Global**	**Focal**	**Directional**
	Viewers	**Point**	**Path**	**Area**
Acoustic	Geometries	**Sphere**	**Cloth**	**Objects**
	Materials	**Absorption**	**Absorption + Scattering**	**Absorption + GeoShape**
	Sound sources	**Omni**	**Static**	**Dynamic**
	Listeners	**Omni**	**Static**	**Dynamic**

scan (Remondino 2011; Douglass et al., 2015). The mapping of such meshes follows the true orthophoto method (DTM or DSM), which is usually included at the end of the modeling chain in common software like Photoscan.

The highest level of detail is defined by **Object** meshes. They define individual spatial objects (such as doors, floors, walls, etc.) as individual 3D objects, mainly defined by elementary polyhedron such as cubes, prisms, cylinders, or pyramids (Arens et al., 2005). They are generally mapped with a cubic mapping, with dimensions of 2m × 2m × 2m, centered on the origin. If any of those meshes require specific mapping, the corresponding mapping is included.

Materials

The lowest material definitions of the visual environment are **Monochrome** materials. These provide a homogenous rendering output. These materials are instances of a single color material. All possible complex material definitions are substituted by this material. The base color is predefined as white (RGB: 255,255,255).

The medium definitions are **Color** materials. These provide an abstract rendering output. They define every material by a different color *via* a "base color" map definition. The base color is predefined, but it can be redefined by the user in the Material Editor.

The highest definitions are **PBR** materials. These provide a realistic rendering output. They are defined by several parameters used by the physically based rendering (PBR) techniques, including "Base color," "Metallicity," "Reflection," "Roughness," and "Normal" (Greenberg 1999). Those parameters are defined whether by constant values or by maps. The maps are predefined, but they can be redefined by the user in the Material Editor. The orientation and size of the texture are defined by the cubic mapping of the meshes. The quality of the graphical data is crucial for the final visual output. Important texture requisites are seamless textures, color-balanced, and high-resolution photographs. The material definitions are rendered in **Figure 1**.

Lighting

The lowest light source level of detail in the visual environment is defined by **Global** lights. This renders a diffuse global illumination triggered in all directions by a white environment casting shadows on all the objects.

The medium level is defined by **Focal** lights. They render artificial illumination defining the source in one point (or collection of points), with a specific directivity in the 3D environment. They do not illuminate the whole scenario but by optimizing the illumination on several areas and saving lighting resources. Those lights are composed of point lights, spot lights, or rectangular lights, with defined values such as intensity, attenuation radius, or light color.

The highest level is defined by **Directional** lights. They render illumination defining the source in one direction. All the lighting rays are parallel to that direction. The most used application of those lights is the "sun light." Attached to the sun, there might be a "sky light," rendering diffuse light, colored after the sun's position, a "sky sphere" displaying a sky representation that includes sky color and clouds, and an "atmospheric fog rendering" adding humidity effects. The ligthing definitions are rendered in **Figure 2**.

Viewers

The lowest viewer level of detail in the visual environment is defined by **Point** positions. They locate the viewer in a fixed position where rotation of the head is allowed.

The medium level is defined by a **Path**. The viewer is allowed to walk through a specified line, including head rotation.

The highest level is defined by an **Area**. The viewer is allowed to walk freely inside a specified area, including head rotation.

Acoustic Definitions

Geometries

The lowest geometric level of detail of the acoustic cues is defined by **Effects**. They perform artificial reverberation effects for defined volumes in the model. They cannot even be considered as geometric definitions, since those effects rely on synthetic reverberation tails, calculated after several parameters such as "absorption" or "size of room." Advanced methods such as RAZR also include other perceptual features such as clarity and localization and adapt the characteristics of an equivalent rectangular space to achieve the intended perceptual result.

The medium level is defined by **Cloth** meshes. Those meshes are triangulated networks made of vertices and edges. The mesh is a continuous object, presenting no empty triangles, with a maximum of 400 triangles, fixed for computation fluency. For specific testing regions, specific welded meshes are optimized. The optimized meshes contain finer resolution areas (close to the testing areas) and coarser definition of the net for the rest of the

TABLE 2 | Suggested module combinations for demonstrators (A) outdoor noise evaluation and (B) classroom speech intelligibility.

Cues	Definitions	Outdoor noise evaluation			Classroom speech intelligibility		
		L	**M**	**H**	**L**	**M**	**H**
Visual	Geometries	■	■				●
	Materials		■				●
	Lighting	■					●
	Viewers	■					●
Acoustic	Geometries		■				●
	Materials	■				●	
	Sound sources			■		●	
	Listeners		■			●	

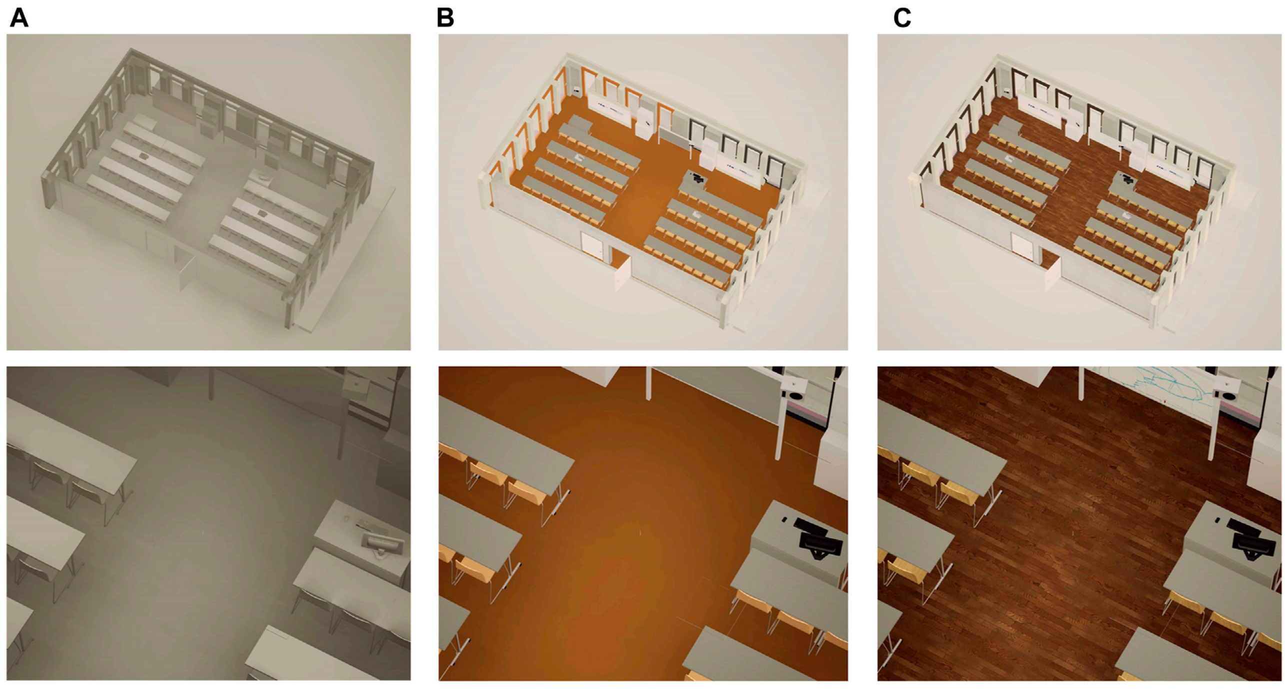


FIGURE 1 | Levels of detail for the material definitions of the visual cues: **(A) monochrome** materials, white; **(B) color** materials; **(C) pbr** materials.

model. The edges disposition responds to the diffraction effects, whereas the triangle disposition supports the reflection effects. This geometric definition is especially effective for outdoor scenarios since the combination of diffraction and reflection of sound is required. Noise mapping for environmental simulation and prediction in cities, dwelling, or rural areas normally requires this kind of input meshes. Examples of commercial software in this field are Soundplan,[19] CadnaA,[20], and Mithra[21]. In the abovementioned software, topography can be easily defined by importing CAD or GIS (Geographical Information System) formats. Other open-source tools for real-time outdoor sound simulation, such as virtual acoustics (VA),[22] the same input information was used.

The highest level is defined by **Object** meshes. Those objects are independent geometries characterized by acoustic material properties. The sizing of the objects corresponds with the target wavelengths. The extension of the scenario is arranged for reflection, scattering, and diffraction calculations. There is a material assignment for each object. This geometric definition is especially effective for indoor scenarios. This type of geometric definition is normally used in software such as RAVEN, ODEON, EASE, or CATT-Acoustic. The geometric definitions are rendered in **Figure 3**.

[19]https://www.soundplan.eu/de/

[20]https://www.datakustik.com/products/cadnaa/cadnaa/

[21]https://www.geomod.fr/fr/geomatique-modelisation-3d/mithrasound/

[22]http://www.virtualacoustics.org/

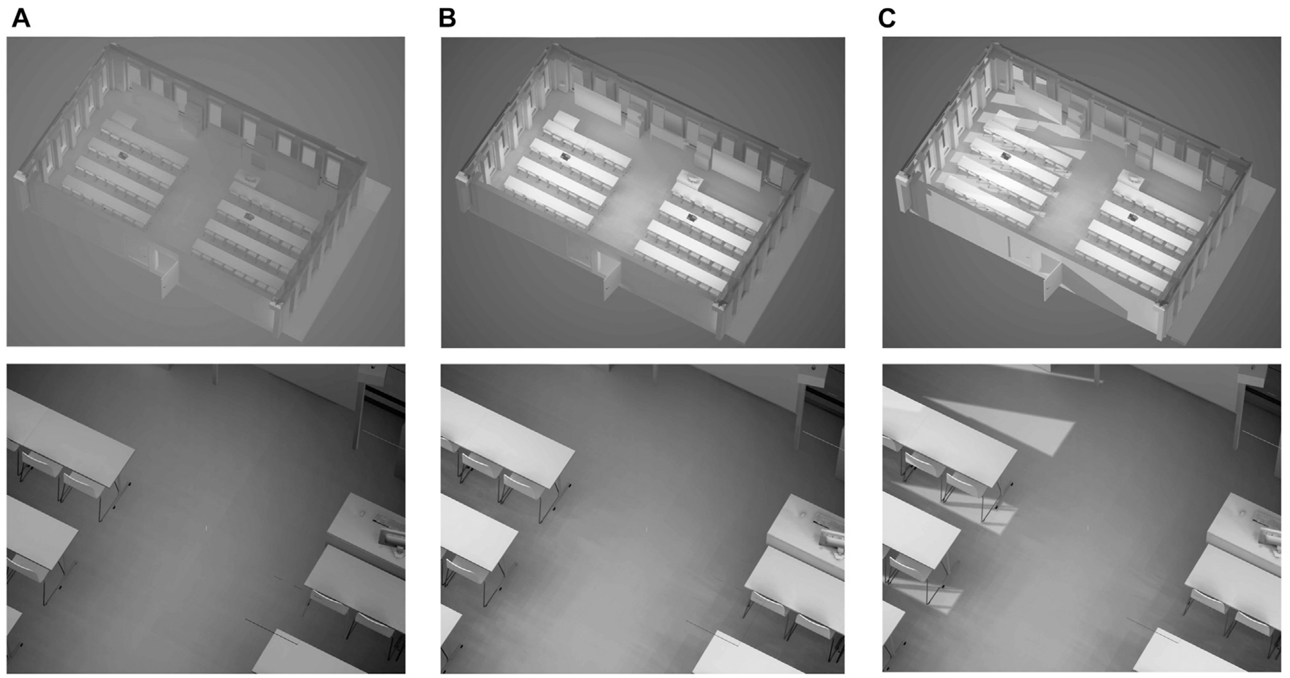


FIGURE 2 | Levels of detail for the lighting definitions of the visual cues: **(A) global** lighting; **(B) focal** lighting, as two rectangular lights on the table regions, together with global lighting; **(C) directional** lighting as sunlight, together with focal and global lighting.

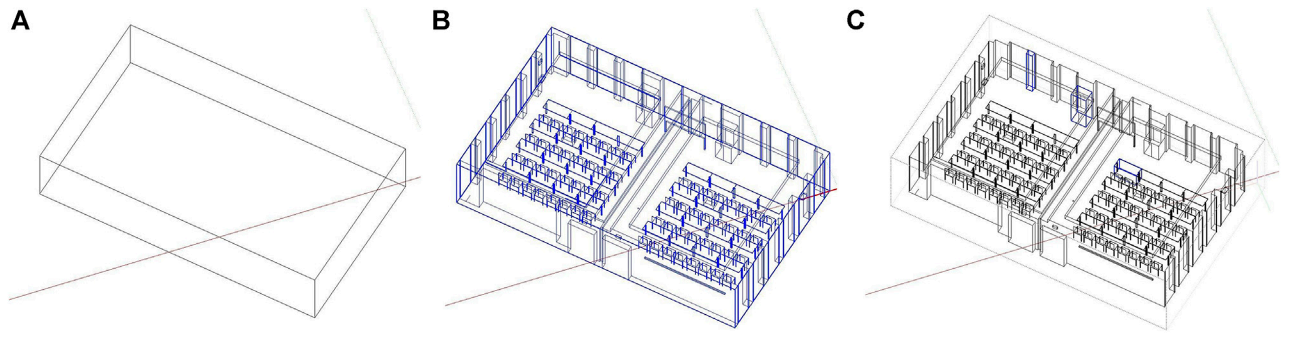


FIGURE 3 | Levels of detail for the geometric definitions of the acoustic cues: **(A) effects**, from a given volume; **(B) cloth** mesh as one single net; **(C) objec**t meshes, as independent geometries.

Materials

The lowest material level of detail of the acoustic cues is defined by **Absorption** properties. They contain only absorption coefficients but not scattering information of the materials.

The medium level is defined by **Absorption + Scattering** materials. They contain absorption and scattering coefficients in different frequency bands.

The highest level is defined by **Absorption + GeometricShape.** This material definition is only used for numerical model calculation. No simulation with geometrical acoustics is possible with this definition as it would violate the condition of short wavelengths compared with geometric details.

Sources

The lowest level of detail for the sources in the acoustic scenario is defined by **Omni** sound sources. They are defined as points in the three-dimensional space. The sources are characterized as omnidirectional sources with a uniform spatial radiation pattern, meaning that they radiate the sound in constant intensity toward all spatial directions.

The medium level is defined by **Static** sound sources. They are defined as fixed points in the three-dimensional space. To simulate the spatial properties of the source signals, the directivity of the source must be known. The directivity function reveals the frequency-dependent amplitude for every spatial direction.

FIGURE 4 | Visual definitions for **(A)** geometry, **(B)** materials, and **(C)** final view from the receiver position of Demonstrator A.

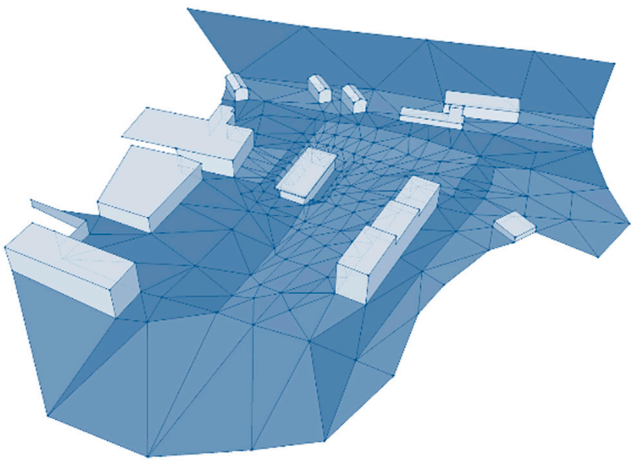


FIGURE 5 | Model for the acoustic definitions. The three blue tonalities refer to the acoustic materials assigned to the mesh: **(A)** hard-buildings, reflective, and diffuse; **(B)** soft-grass and moderately absorptive; **(C)** hard-asphalt and reflective.

The highest level is defined by **Dynamic** sound sources. They are defined as points in the three-dimensional space, receiving certain degrees of freedom such as moving along a path or free movement. Those sound sources are also provided with directivity functions, in the most complex case, also including signal-dependent directivities.

Listeners

The lowest level of detail for the listeners in the acoustic scenario is defined by **Omni** receivers. They are defined as points in the three-dimensional space. The receivers are characterized as omnidirectional receivers, with a constant frequency response for all directions of the space. These receiver points may be used as measurement locations in the room, for further comparison of room acoustic parameters.

The medium level is defined by **Static** receivers. They are defined as points in three-dimensional space. To simulate binaural responses, human listeners are characterized by a frontal direction with a general or individual head-related transfer function (HRTF). This function characterizes how an ear receives a sound from a point in space, affected by the size and shape of the head, ears, ear canal, and other aspects (Blauert 1997).

The highest level is defined by **Dynamic** receivers. They are defined as a point in the three-dimensional space able to move within a specified trajectory, or within a restricted area. Those receivers are also characterized by a frontal direction connected to an individual or standardized head-related transfer function on each point of the trajectory.

ANTICIPATED RESULTS AND APPLICATIONS

The main goal of the presented framework is to provide modules to be combined in a final 3D scenario. Since the final scenario is intended to be used in perception research in the form of audio–visual tests, two extreme applications are explained here. However, the combinations between the modules can provide suitable scenarios for other cases. The three study cases are just examples and not definitive final setups. **Table 2** depicts the applications presented and their module combinations.

Demonstrator A: Outdoor Noise Evaluation

Tests done in the area of psychoacoustics in noise evaluation are relevant for noise impact in residential or educational areas (Janssens et al., 2008; Soeta and Kagawa 2020). As a prerequisite, the test designer will decide which are the visible and audible areas in the test in order to define the extension of the visual and acoustic models. Since the evaluation of the noise is done in this case after the study of several cognitive or emotional aspects, such as preference, attention to response, or digit span

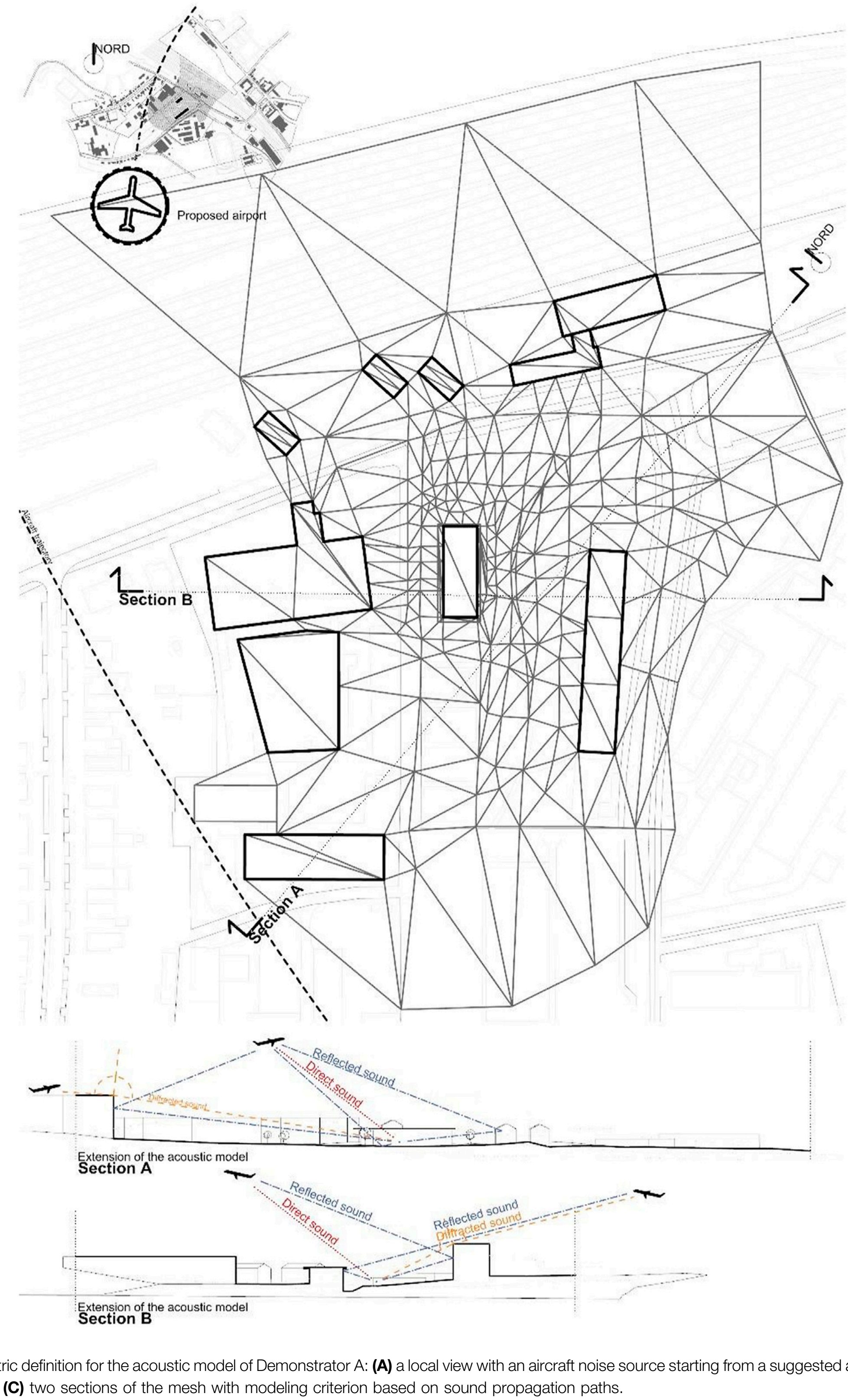


FIGURE 6 | Geometric definition for the acoustic model of Demonstrator A: **(A)** a local view with an aircraft noise source starting from a suggested airport location; **(B)** mesh definition; **(C)** two sections of the mesh with modeling criterion based on sound propagation paths.

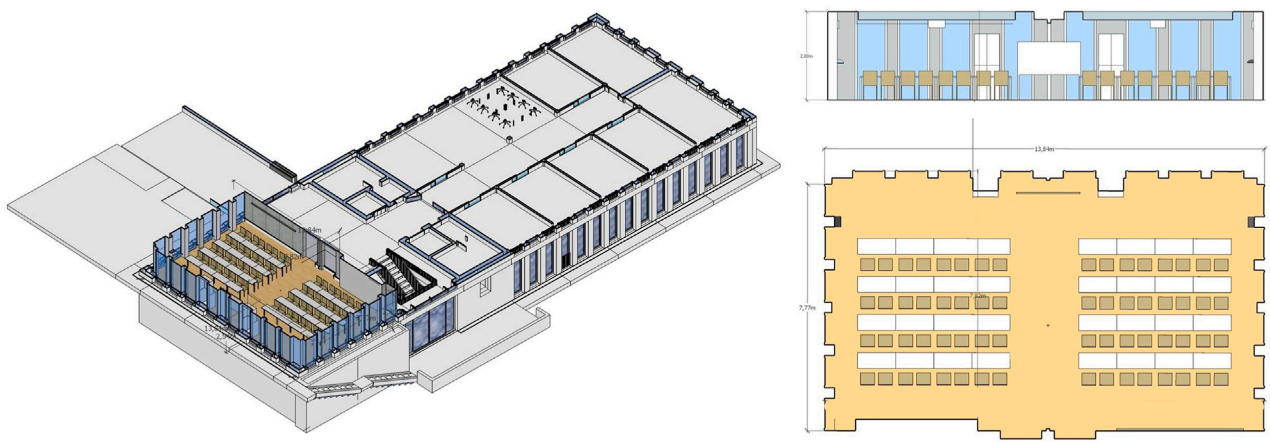


FIGURE 7 | Different views of the acoustic model of the classroom for Demonstrator B. The classroom belongs to a bigger building, which is not taken into account for this study.

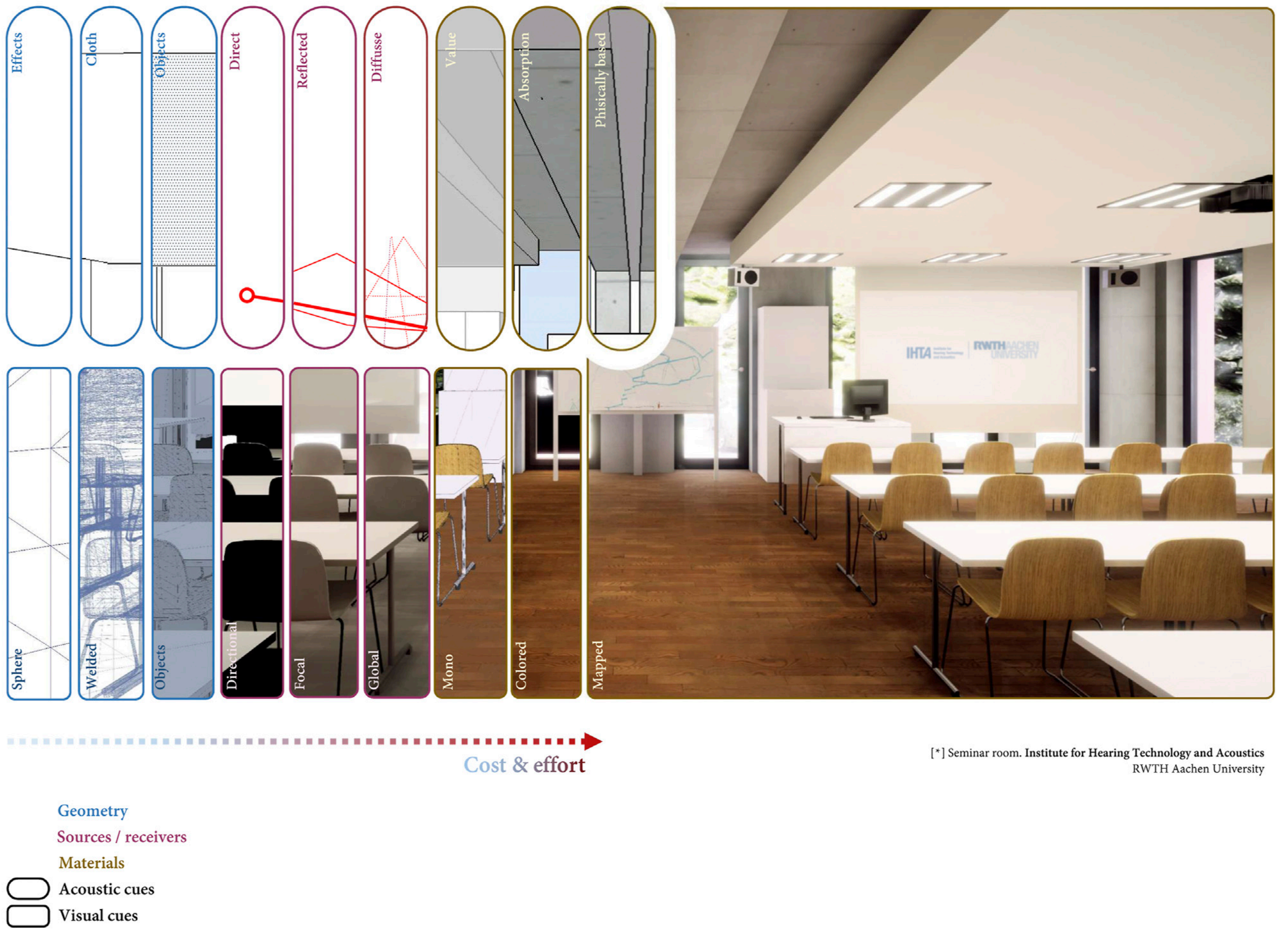


FIGURE 8 | The complete framework modules displayed for the demonstrator “indoor classroom speech intelligibility” as used in a usual modeling workflow. Geometry, sources, and receivers, and material definitions are represented for each acoustic and visual cue. The order of them follows the common modeling and simulation workflow: geometry, basic lighting, material definitions, and final rendering, as well the cost and effort implied in the process. The same logic applies for outdoor scenarios.

tests (Keller and Sekuler 2015; Shu and Ma 2019) the listener-viewer will be seated in a fixed position, with the rotation and slight head movement as the only degrees of freedom. The 3D data is availabe at the open access database (IHTApark, 2021). The suggested module combination is as follows:

- *Visual geometry definition*: a **welded** mesh for the close field and a **sphere** mesh for the far field (fig X). The **welded** mesh is generated by photogrammetry, after 462 photographs on site, after mesh generation by triangulation. The photogrammetric generation has been conducted with the educational version of the software Photoscan.[23] The **sphere** covering the far field is centered on the receiver position with a radius of 2,000 m. The visual definitions are rendered in **Figure 4**.
- *Visual material definition*: the welded mesh is **mapped** with the graphical information of the photographs. The process is generated also with Photoscan, following the true orthophoto algorithms (Remondino 2011). The Sphere is mapped with an HDRI sky using the cubemap texture from the same photographs set.
- *Visual lighting definition*: **global** lighting from the material definition.
- *Visual viewers definition*: **dynamic** viewers within a 2 m × 2 m area, with head movement allowed.
- *Acoustic geometry definition*: a **cloth** mesh made of 493 triangles (See **Figure 5**). The sizing of the triangles grows according to the distance to the receiver, which is located at the center of the mesh. The smallest triangle, which is close to the receiver presents an area of 10.5 m^2 and contains a circle of 3.3 m diameter. Therefore, this triangle acts as a rigid wall for frequencies around 100 Hz. This is the lower frequency limit of the model, which is valid for perceptual applications, regarding the human auditory range. Every half-edge of the mesh is connected with his opposite half-edge, assuring that the mesh contains no wholes. The mesh extension responds to three criteria (See **Figure 6**):
    - *Direct sound*: the model extension covers the “visible” sound source positions (in red), like cars on the “visible” roads. The “nonvisible“ roads are neglected.
    - *Reflected sound*: the model contains the geometry in charge of sound bouncing from the floor, neighbor façades, and ceilings (in blue). Every reflection will almost duplicate the received energy, particularly reflected sound from hard surfaces like façades.
    - *Diffracted sound*: the model contains the edges that enable calculation of the diffracted sound paths (in orange), coming from “nonvisible” sound sources. Those edges are considered both on building corners or terrains.
- *Acoustic material definition*: consisting of **absorption coefficients** is included. Three different acoustic materials are defined, corresponding to hard-buildings, hard-floor, and soft-floor (see Figure 4.1).
- *Sound source definitions*: as **dynamic**—aircraft and cars.
- *Listener definitions*: as static position from which the listener evaluates the scene. That position allows free head rotation.

An example for auralization in an application of soundscape research (ISO 12913) is given by the Institute for Hearing Technology and Acoustics (IHTA-Institute for Hearing Technology and Acoustics, RWTH Aachen University, 2021). Here, the reference scene is captured with photogrammetry technique, as a baseline for the visual rendering. The auditory reference event is recorded with an Ambisonics microphone. In postprocessing, additional sound sources and additional buildings such as dwellings or detached houses can be added to the virtual scene. One application can be found in the work by Lihoreau and colleagues (Lihoreau et al., 2006) which is focused on outdoor sound propagation modeling models under different atmospheric conditions; or (Dreier and Voränder, 2021), on aircraft noise application of simulations.

Demonstrator B: Indoor Classroom Speech Intelligibility

In this type of environment, the close environment and the details around the listener and viewer play an important role. Due to the high level of detail demanded in the visual cues, all definitions are set to the highest requirements. Typical perception experiments include cognitive tests of work, learning performance, or selective attention (Reynolds, 1992). This unique setup requires powerful hardware. Whereas in the previous application, not many GPU and CPU resources were demanded for the visual model, in this one, the smooth performance of the scenario will require a well-equipped machine. As a rule of thumb, experience shows that current gaming PCs are well equipped for such tasks. Regarding the acoustic cues, moving properties of the subject may be restricted to an area and sound sources may be fixed in position too. The suggested module combination is as follows:

- *Visual material definition*: **PBR** materials, including
    - Albedo, defined by an orthophotography of a real-world material,
    - Metallicity, defined as an integer between 1 and 0, meaning 1 as metal and 0 as nonmetal,
    - Roughness, defined by gray-scale photography or as an integer between 1 and 0, meaning 1 as a diffuse and 0 as a specular surface;
    - ambient occlusion, defined with Normal maps,
    - All textures are freely available at https://www.textures.com/. The mapping UVs of the geometry is generally set as a prismatic projection for all objects. Special objects, such as the chairs, the computer screen, and the keyboard are custom mapped with the “UV unwrapping” technique, in 3Ds Max software, educational version.
- *Visual lighting definition*: **focal**, **directional**, and **global** lighting are included. Two rectangular lights (4m × 4m

[23]https://www.agisoft.com/

each) are located at each side of the classroom. A directional light in the direction of the Sun generates natural light and direct rays coming from the left windows of the room. A sky light generates diffuse illumination for the whole scene. The sky light takes the light colors according to the day time.

- *Viewers definition*: **dynamic**, they can freely move around the model. No collision is included at the moment. The camera contains postprocessing filtering of Image Tint, Vignetting, and Exposure control set to constant.
- *Acoustic geometry definition*: independent modeled **objects**. Those are prisms and rectangles. The smallest polygon area is 1 m^2, corresponding to the chairs sit (See **Figure 7**).
- *Acoustic material definition*: both **absorption** and **scattering** definitions are included. There are four materials defined: floor-wood, windows-glass, wall-concrete, and ceiling-concrete.
- *Sound sources definition*: set to **static**. Sound sources are 4 loudspeakers and a talker at the frontal desk.
- *Listeners definition*: set to **dynamic**. The listener can freely move around the classroom and sit down on a chair.
- *Visual geometry definition*: independent modeled **objects**. Those are prisms, rectangles, and special geometries for specific objects, such as the chairs. Objects are differentiated among them in the modeling hierarchy. The criterion of differentiation is the material definition, meaning that one object corresponds to one material.

A complete view of the framework is presented in **Figure 8**. The 3D data is availabe at the open access database (IHTAclassroom, 2021).

DISCUSSION

The elaboration of audio-visual 3D models for sound perception research requires manual work. The wide range of modeling techniques, modalities, and software revisited here demonstrate this fact. However, a methodical way of connecting them is possible. The presented framework divides the environment chunks according to what has been done in previous research and provides a way to combine them. Despite this, the definition of the modules seems evident after all, and the combination among them is a useful help to the potential "modelers."

It is noteworthy that there is modeling freedom in each module. This enables the model to be personalized, within certain restrictions. In other words, this gives the assurance that no matter the fine details, the module is kept assembled with the rest of the framework. Therefore, the replicability of this method is assured in the assembly and module definitions, rather than the finer details, according to the purpose of the present work.

An unexpected consideration emerged related to the match between acoustic and visual simulations. It appeared when analyzing the postprocessing techniques normally used in the photography and the film industry, as part of which the treatment of light through cameras is filtered with numerous methods (like lens correction, tinting, or spectral correction of color). These processes can be compared to the techniques used by an audio engineer at the mixing console. The final decision is not to include them in the framework since they fall within the artistic work with multiple variations escaping from the controlled values for laboratory conditions.

When it comes to the result of the auditory-visual representation, one might ask whether or not this representation is correct in the sense of ecological validity. If the scenario exists already, a reference measurement or recording can be done for comparison. This was studied in a comprehensive manner by Brinkmann and colleagues (Brinkmann 2017; Brinkmann et al., 2019). The results point out that the representation is almost authentic with best-matched input data, in terms of a nondistinguishable auditory perception of realism. The auditory impressions in comparison of real and virtual spaces are, hence, similar and characteristic for the spaces, although not identical in an A-B comparison. Nobody, however, could identify which one was the auralization and which one the recording.

In the case of pure prediction, this picture changes. Blind input data quality and modeling quality determine the uncertainties of the auditory impression at the end. It was shown by Vorländer (2013) that research efforts must be intensified in the field of robust characterization of acoustic material properties.

The software tools used in this work are examples without any restriction to be replaced by other tools. Nevertheless, all typical steps and important considerations in the workflow were explained by this set of typical software tools. The main conclusion of this work is that modeling techniques in the visual representation of architecture and acoustics follow different approaches and different strategies when implementing models in various levels of detail. Their comprehensive and combined development in creative processes is not harmonized yet. The categorization schemes as listed in **Table 2** may achieve transparency in terms of the definitions and interpretations of both visual and auditory aspects, and possibly in the future, the development of a cross-modality approaches to modeling in architectural acoustics.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the Open Science repository ZENODO under the following permanent links: Demonstrator A database. IHTApark: https://zenodo.org/record/4629760. Auralization of demonstrator A: 'Auralization of virtual aircrafts in real scenes' https://www.akustik.rwth-aachen.de/go/id/dzhe/lidx/1. Demonstrator B database. IHTAclassroom: https://zenodo.org/record/4629716. Auralization of demonstrator B: https://www.youtube.com/watch?v=I-pYDMtxFtM.

AUTHOR CONTRIBUTIONS

Both authors contributed to the conception of the manuscript. The first author wrote the manuscript and generated the study cases. The second author edited the text. Both authors contributed to the article and approved the submitted version.

FUNDING

The first author is grateful for the financial support of the PAAD project (https://www.akustik.rwth-aachen.de/cms/Technische-Akustik/Forschung/~kbfdp/Research-in-architectural-design/lidx/1/) under a Junior Principal Investigator Fellowship at RWTH Aachen University.

ACKNOWLEDGMENTS

The authors thank Lukas Aspöck for modeling the information of Demonstrator B.

REFERENCES

Arens, C., Stoter, J., and van Oosterom, P. (2005). Modelling 3D Spatial Objects in a Geo-DBMS Using a 3D Primitive. *Comput. Geosciences* 31 (2), 165–177. doi:10.1016/j.cageo.2004.05.013

Attia, S., Beltrán, L., De Herde, A., and Hensen, J. (2009). Architect Friendly: a Comparison of Ten Different Building Performance Simulation Tools. Proceedings of the 11th International Builidng Performance Simulation Association Conference and Exhibition. 204-211. URL: http://hdl.handle.net/2268/167578.

Blauert, J. (1997). *The Psychophysics of Human Sound Localization*. Harvard MA: The MIT Press.Spatial Hearing.

Boeykens, S., and Liège, France. (2011). Using 3D Design Software, BIM and Game Engines for Architectural Historical Reconstruction. Proceedings of the 14th international conference on Computer Aided Architectural Design Futures, 493-509. , 2011 Liège: Les Éditions de l'Université de Liège. ISBN: 978-2-8745-6142-9.

Bouchlaghem, D., Shang, H., Whyte, J., and Ganah, A. (2005). Visualisation in Architecture, Engineering and Construction (AEC). *Automation in Construction* 14 (3), 287–295. doi:10.1016/j.autcon.2004.08.012

Brinkmann, F., Aspöck, L., Ackermann, D., Lepa, S., Vorländer, M., and Weinzierl, S. (2019). A Round Robin on Room Acoustical Simulation and Auralization. *The J. Acoust. Soc. America* 145, 2746–2760. doi:10.1121/1.5096178

Brinkmann, F. (2017). *Binaural Processing for the Evaluation of Acoustical Environments*. Berlin, Germany: . Technische Universität Berlin, 2017 . [PhD thesis] Available at: https://depositonce.tu-berlin.de/handle/11303/9454.

Calderon, C., Nyman, K., and Worley, N. (2006). The Architectural Cinematographer: Creating Architectural Experiences in 3D Real-Time Environments. *Int. J. Architectural Comput.* 4 (4), 71–90. doi:10.1260/147807706779398953

Castellazzi, G., D'Altri, A. M., de Miranda, S., and Ubertini, F. (2017). An Innovative Numerical Modeling Strategy for the Structural Analysis of Historical Monumental Buildings. *Eng. Structures* 132 (1), 229–248. doi:10.1016/j.engstruct.2016.11.032

Clayton, M. J., Warden, R. B., and Parker, T. W. (2002). Virtual Construction of Architecture Using 3D CAD and Simulation. *Automation in Construction* 11 (2), 227–235. doi:10.1016/S0926-5805(00)00100-X

Clua, Á., Llorca-Bofí, J., and Psarra, S. (2020). Urban Opportunities and Conflicts Around Street Musicians: the Relationship between the Configuration of Public Space and Outdoor Acoustics in Ciutat Vella, Barcelona. *J. Urban Des.* 25 (5), 561–589. doi:10.1080/13574809.2019.1699398

Dalenbäck, B. I. L. (1996). Room Acoustic Prediction Based on a Unified Treatment of Diffuse and Specular Reflection. *J. Acoust. Soc. America* 100, 899–909. doi:10.1121/1.416249

de Luca, F., Dogan, T., and Sepúlveda, A. (2021). Reverse Solar Envelope Method. A New Building Form-Finding Method that Can Take Regulatory Frameworks into Account. *Automation in Construction* 123, 103518. doi:10.1016/j.autcon.2020.103518

Douglass, M., Lin, S., and Chodoronek, M. (2015). The Application of 3D Photogrammetry for In-Field Documentation of Archaeological Features. *Adv. Archaeol. Pract.* 3 (2), 136–152. doi:10.7183/2326-3768.3.2.136

Dreier, C., and Vorländer, M. (2021). Aircraft noise - Auralization-based assessment of weather dependent effects on loudness and sharpness. *J. Acoustic. Soc. Am.* (in press).

Foley, J. D., van Dam, A., Feiner, S. K., and Hughes, J. K. (1997). *Computer Graphics: Principles and Practice*. Boston: Addison-Wesley.

Greenberg, D. P. (1999). A Framework for Realistic Image Synthesis. *Commun. ACM* 42 (8), 44–53. doi:10.1145/2F310930.31097010.1145/310930.310970

Hasan, A. M. M., Torky, A. A., and Rashed, Y. F. (2019). Geometrically Accurate Structural Analysis Models in BIM-Centered Software. *Automation in Construction* 104, 299–321. doi:10.1016/j.autcon.2019.04.022

IHTA-Institute for Hearing Technology and Acoustics, RWTH Aachen University (2021). Auralization of Virtual Aircrafts in Virtual Scenes. Available at: https://www.akustik.rwth-aachen.de/go/id/dzhe/lidx/1 (Accessed April 22, 2021).

IHTAclassroom (2021). Multi-detailed 3D Architectural Model for Sound Perception Research in Virtual Reality. Available at: https://zenodo.org/record/4629716 (Accessed April 22, 2021).

IHTApark (2021). Multi-detailed 3D Architectural Model for Sound Perception Research in Virtual Reality. Available at: https://zenodo.org/record/4629760 (Accessed April 22, 2021).

Janssens, K., Vecchio, A., and Van der Auweraer, H. (2008). Synthesis and Sound Quality Evaluation of Exterior and interior Aircraft Noise. *Aerospace Sci. Technol.* 12 (1), 114–124. doi:10.1016/j.ast.2007.10.002

Keller, S., and Sekuler, R. (2015). Memory and Learning with Rapid Audiovisual Sequences. *J. Vis.* 15, 15. doi:10.1167/2F15.15.710.1167/15.15.7

Künzel, H. M., Holm, A., Zirkelbach, D., and Karagiozis, A. N. (2005). Simulation of Indoor Temperature and Humidity Conditions Including Hygrothermal Interactions with the Building Envelope. *Solar Energy* 78 (4), 554–561. doi:10.1016/j.solener.2004.03.002

Lihoreau, B., Gauvreau, B., and Bérengier, M. (2006). Outdoor sounds propagation modeling in realistic environments: application of coupled parabolic and atmospheric models. *J. Acoustic. Soc. Am.* 120, 110. doi:10.1121/1.2204455

Marburg, S. (2002). Six Boundary Elements Per Wavelength: Is that Enough? *J. Comp. Acous.* 10, 25–51. doi:10.1142/S0218396X02001401

Monedero, J. (2015a). *Simulación visual de la iluminación: teoría, técnicas, análisis de casos*. Barcelona: Iniciativa Digital Politècnica. URL: http://hdl.handle.net/2117/80463.

Monedero, J. (2015b). *Simulación visual de materiales: teoría, técnicas, análisis de casos*. Barcelona: Iniciativa Digital Politècnica. URL: http://hdl.handle.net/2117/82368.

Monson, B. B., Hunter, E. J., and Story, B. H. (2012). Horizontal Directivity of Low- and High-Frequency Energy in Speech and Singing. *J. Acoust. Soc. America* 132, 433–441. doi:10.1121/1.4725963

Naylor, G. M. (1993). ODEON-another Hybrid Room Acoustical Model. *Appl. Acoust.* 38, 131–143. doi:10.1016/0003-682X(93)90047-A

Pelzer, S., and Vorländer, M. (2010). Frequency- and Time-dependent Geometry for Real-Time Auralizations. Proc. 20th International Congress on Acoustics. Sydney, Australia. Url: https://www.researchgate.net/publication/266335423_Frequency-and_Time-dependent_Geometry_for_Real-time_Auralizations.

Relaño-Iborra, H., Zaar, J., and Dau, T. (2019). A Speech-Based Computational Auditory Signal Processing and Perception Model. *J. Acoust. Soc. America* 146 (5), 3306–3317. doi:10.1121/1.5129114

Remondino, F. (2011). Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sensing* 3 (6), 1104–1138. doi:10.3390/rs3061104

Reynolds, R. E. (1992). Selective Attention and Prose Learning: Theoretical and Empirical Research. *Educ. Psychol. Rev.* 4, 345–391. doi:10.1007/BF01332144

Roca, P., Cervera, M., Gariup, G., and Pela', L. (2010). Structural Analysis of Masonry Historical Constructions. Classical and Advanced Approaches. *Arch. Computat Methods Eng.* 17, 299–325. doi:10.1007/s11831-010-9046-1

Schröder, D. (2011). *Physically Based Real-Time Auralization of Interactive Virtual EnvironmentsPhD Dissertation*. Germany: RWTH Aachen University. Available at: http://publications.rwth-aachen.de/record/50580/files/3875.pdf.

Shabtai, N. R., Behler, G., Vorländer, M., and Weinzierl, S. (2017). Generation and Analysis of an Acoustic Radiation Pattern Database for Forty-One Musical Instruments. *J. Acoust. Soc. Am.* 141, 1246–1256. doi:10.1121/1.4976071

Shu, S., and Ma, H. (2019). Restorative Effects of Classroom Soundscapes on Children's Cognitive Performance. *Ijerph* 16 (2), 293. doi:10.3390/2Fijerph1602029310.3390/ijerph16020293

Soeta, Y., and Kagawa, H. (2020). Three Dimensional Psychological Evaluation of Aircraft Noise and Prediction by Physical Parameters. *Building Environ.* 167, 106445. doi:10.1016/j.buildenv.2019.106445

Søndergaard, P., and Majdak, P. (2013). "The Auditory Modelling Toolbox," in *The Technology of Binaural Listening*. Editor J. Blauert (Heidelberg: Springer Berlin), 33–56.

Strauss, H. (1998). *Implementing Doppler Shifts for Virtual Auditory Environments*. 104th Conv Audio Enginr Soc, Amsterdam. preprint 4687. Available at: http://www.aes.org/e-lib/browse.cfm?elib=8493.

Svensson, U. P., Fred, R. I., and Vanderkooy, J. (1999). An Analytic Secondary Source Model of Edge Diffraction Impulse Responses. *J. Acoust. Soc. America* 106, 2331–2344. doi:10.1121/1.428071

Thery, D., Boccara, V., and Katz, B. F. G. (2019). Auralization Uses in Acoustical Design: a Survey Study of Acoustical Consultants. *J. Acoust. Soc. America* 145, 3446–3456. doi:10.1121/1.5110711

Thiemann, J., and van de Par, S. (2019). A Multiple Model High-Resolution Head-Related Impulse Response Database for Aided and Unaided Ears. *EURASIP J. Adv. Signal. Process.* 2019. doi:10.1186/s13634-019-0604-x

Tsingos, N., Funkhouser, T., Ngan, A., and Carlbom, I. (2001). Modeling Acoustics in Virtual Environments Using the Uniform Theory of Diffraction. In: SIGGRPAH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques. 545–552. doi:10.1145/383259.383323

van der Harten, A. (2013). Pachyderm Acoustical Simulation: towards Open-Source Sound Analysis. *Archit Des.* 83, 138–139. doi:10.1002/ad.1570

Vorländer, M. (2020). *Auralization*. Switzerland: Springer Nature. doi:10.1007/978-3-030-51202-6

Vorländer, M. (2013). Computer Simulations in Room Acoustics: Concepts and Uncertainties. *J. Acoust. Soc. America* 133 (3), 1203–1213. doi:10.1121/1.4788978

Vorländer, M. (1989). Simulation of the Transient and Steady-state Sound Propagation in Rooms Using a New Combined ray-tracing/image-source Algorithm. *J. Acoust. Soc. America* 86, 172–178. doi:10.1121/1.398336

Wefers, F., and Vorländer, M. (2018). Flexible Data Structures for Dynamic Virtual Auditory Scenes. *Virtual Reality* 22, 281–295. doi:10.1007/s10055-018-0332-9

Xie, B. (2013). *Head-related Transfer Function and Virtual Auditory Display*. Plantation: J. Ross Publishing. doi:10.1121/1.4799319 CrossRef Full Text

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 14 June 2021
doi: 10.3389/fbuil.2021.688836

Sound Quality Characteristics of Importance for Preschool Children's Perception and Wellbeing After an Acoustic Intervention

Kerstin Persson Waye and Jonas Karlberg*

School of Public Health and Community Medicine, Occupational and Environmental Medicine, Gothenburg University, Gothenburg, Sweden

In Sweden preschool-age children spend most of their waking hours at preschool. At this age children undergo substantial physical and mental development and their hearing sensations may not be comparable to those of an adult. The preschool sound environment is loud and highly intermittent, and the acoustic may not be supportive for young children's hearing, and wellbeing. This article describes an acoustic intervention among seven preschools, and comparisons with three reference preschools. The intervention included changing floor mats to plastic mats designed to reduce impact sounds, adding damping cushions under chairs, change of ceiling absorbers and, in some rooms, addition of wall absorbers. The effect of the intervention was studied using a previously developed interview protocol, "Inventory of Noise and Children's Health," in combination with sound level and room acoustic measurements. A total of 61 children aged 4–6 years were interviewed before the intervention, and 56 after. A reduction of the sound levels in a range of LAeq 1.2–3.8 dB for meal/craft rooms and play rooms were found for the intervention preschools using stationary noise level meters, while this was not found for the reference preschools. The reverberation time (T20) decreased slightly after the interventions. The average room frequency response for the two room types tended to be more flat after the interventions. Further investigations are needed to see its importance for the perceived acoustic quality. The results linking children's perception and response to the measured reduction in sound levels, confirmed an association between reduced sound levels after the acoustic intervention and a 30% reduction in stomach ache, as well as in children's perceptions of scraping, screeching sounds. Children's perceptions of these sounds were further associated with important oral communication outcomes. Children's bodily sensations of sounds were also associated with psychoacoustic symptoms and wellbeing. Despite the seemingly modest reduction in sound level, the acoustic intervention was indeed perceived and reported on by the children. Future studies should pay more attention to how a supportive preschool acoustic environment should be optimised and acoustically described to take preschool-age hearing and perception into account.

Keywords: acoustic intervention, child, perception, reaction, psychosomatic symptoms, INCH questionnaire, noise, high frequency sounds

Edited by:
Jian Kang,
University College London,
United Kingdom

Reviewed by:
Hui Ma,
Tianjin University, China
Zihe Gao,
Zhengzhou University, China

***Correspondence:**
Kerstin Persson Waye
Kerstin.persson.waye@amm.gu.se

Specialty section:
This article was submitted to
Indoor Environment,
a section of the journal
Frontiers in Built Environment

Received: *31 March 2021*
Accepted: *26 May 2021*
Published: *14 June 2021*

Citation:
Persson Waye K and Karlberg J (2021) Sound Quality Characteristics of Importance for Preschool Children's Perception and Wellbeing After an Acoustic Intervention.
Front. Built Environ. 7:688836.
doi: 10.3389/fbuil.2021.688836

INTRODUCTION

Preschool-age children are in a phase of life which involves substantial physical and mental development. This is also the age period where most of the neural development takes place (Tamburlini et al., 2002). In Sweden, a great majority of children aged 1–5 years spend most of their waking hours at preschool, hence this environment is of prominent importance for their health and wellbeing. Given that the critical period for optimal formation of linguistic skills is also within this early age period, ranging from about 12 to 36 months (Kuhl, 2010), the acoustic environment in which language is communicated and understood is highly relevant. However, preschool premises are often not acoustically optimized for specific preschool requirements supporting children's perception of sounds, learning, and communication. The preschool sound environment is intermittent and unpredictable, with one Swedish study showing more than 80 one-second logged events exceeding 85 A-weighted decibel (dBA) during hours of preschool activity (Sjodin et al., 2012). Other stationary measurements of noise levels in Turkey found that the sound levels during the meal period was slightly higher than the period of play (Gokdogan and Gokdogan 2016). Five minutes A-weighted equivalent noise levels (LAeq) during the meal times ranged from 60 to 81 dB for the group of 3 years old, 67–82 dB for the 4 years old and 69–85 dB for the group of 5 years old children. Similar noise levels were recorded in public preschools on Iceland where 89% of the samples were above 70 dB LAeq and 43% above 75 dB LAeq (Jonsdottir et al., 2015). High sound levels were also recorded in a preschool in Germany, showing 8 h equivalent noise levels from stationary measurements of 71 dB LAeq, and average dosimeter levels on teachers of 80 dB LAeq (Eysel -Gosepath et al., 2010). Dosimeter measurements on children show that children tend to be exposed to even higher sound levels. An average of 154 measurements on children show levels of 84 dB LAeq during their time spent indoors, with maximum A-weighted noise levels, Fast time weighting (LAFmax) up to 118 dB (Persson Waye et al., 2011); the latter exceeds the permissible maximum levels for the occupational environment in Sweden (AFS 2005:16, 2005).

Preschool children's hearing and auditory perception differs from that of adults (Fels, 2008). Using anthropometric data from children and subsequent simulation, Fels (2008) was able to show that the diffraction and reflection properties of the head, pinna, and torso (the head-related transfer functions; HRTF) in children are not comparable to those in adults. The HRTFs play a major role when it comes to localizing sounds, as understanding speech in a room under noisy conditions is tightly dependent on the directivity pattern of the head. The HRTF of a child up to the age of seven amplifies the frequencies around 6 kHz, and the ear canal of a child further adds to this frequency amplification, which may be compared to an adult whose HRTF and ear canal leads to amplification around 3 kHz. It has further been found using an auditory oddball paradigm and event-related brain potentials (ERPs) that changes of pitch, but not loudness, evoked ERPs among 6–9-year-olds when the sound was irrelevant to their current task (Sussman and Steinschneider, 2011). For slightly older children (10–12 years), both pitch and loudness evoked a response in similar task situations. This may indicate that neural processing of sound frequency and sound intensity develop during different phases of a child's development, and that neural processing of frequency precedes that of intensity. Initial support for certain sounds and characteristics playing a key role for small children was also found in a qualitative study of 36 preschool children aged 4–5 (Persson Waye et al., 2013), where uncontrollable sounds and distressing sounds (i.e., angry yelling and scraping, screeching sounds) were experienced as both physically and emotionally painful.

Surprisingly little attention has been paid to the differences between adults and small children in hearing and hearing function, and the possible consequences for attention, learning, and wellbeing. The large body of research in this field has focused on the impact of transportation noise on children's cognitive functions, with the most consistent results found for exposure to aircraft noise and impaired reading comprehension and long-term memory (Clark and Paunovic, 2018). Fewer studies have been targeted at noise created by the inhabitants of classrooms or preschools, even though this so-called "babble noise" usually contributes to a higher indoor noise level than transportation noise, and has been found to negatively affect primary school children's performance on verbal tasks (Dockrell and Shield, 2006). A high babble noise and or poor reverberation would be particularly destructive for young children's word recognition up to the age of 12–15 years as they require better acoustics than adults (Neuman et al., 2010). Apart from its effect on cognition, noise may affect young children's emotional wellbeing, although there have been very few studies on this topic. One exception is a study by Klatte et al. (2010) which found an association between higher reverberation time in the class room (indicative of poor acoustics) and young children's reporting of a greater disturbance from indoor noise ("My classmates often behave noisily" "Our teachers often reprove us for silence") as well as reporting poorer relationships with their teachers and classmates, less motivation and less social integration. Another exception is a recent study by Astolfi et al. (2019). Interviews with 6–7 years old pupils were compared to a wide range of room acoustic measures and noise level measurements. They found that children rated themselves as less happy in classrooms characterised as having bad acoustics. Interestingly, it was also found that reported wellbeing and noise disturbance seemed to depend on the pupils being happy or not, calling for the need of a longitudinal study design. Some further guidance on how noise may affect children emotionally and behaviourly can be derived using preschool teachers' perspectives. As part of a large survey on preschool teachers' occupational environment and their health, we included one question asking if teachers judged preschool noise to affect children's behaviour and if so, how. The analysis showed that 82% of them considered noise to affect children's behaviour (Persson Waye et al., 2019). A content analysis of the free text provided by the nearly 4,000 preschool teachers that answered yes on the posed question showed that the most common observations were children vocalizing to be heard, followed by children being distracted, unfocussed, angry, sad, exhausted, and withdrawn.

Acoustic interventions performed in preschools are typically aimed at reducing the reverberation time, reducing speech interference, or improving clarity. The reverberation time (T or T60) measures the time it takes for a sound to decay by 60 dB (ISO 3382-2, 2008), and is hence a measure of how quickly a sound in the room is attenuated. A longer reverberation time would increase the background noise level and possibly the levels from the children. The length of the reverberation time depends on scattering objects and how absorbent the surfaces are in the room. The Swedish standard (SS 25268:2007+T1, 2017) suggests that the reverberation time (T20) in preschools should be 0.5–0.6 s depending on the room function. Clarity (C50) and Deutlichkeit (D50) describe the ratio of early (before 50 ms) and late reverberation energy, with more early energy meaning less alteration of the direct sound. High values of clarity are typically considered to make the spoken sound more easily perceived, but scientific evidence is ambiguous and especially lacking for children. Less influence of secondary reflexes may also have the beneficiary effect of reducing irrelevant noise. As children and adults tend to increase their voice when speaking in high background noise, referred to as the Lombard effect (Lane and Tranel 1971), a reduction of the secondary reflexes should in principle reduce the need for children to raise their voices. This is important as it is commonly reported that small children raise their voice to be heard in noisy environment (Lindström et al., 2010; Persson Waye et al., 2019). The so-called inverse Lombard effect is though still not well understood and evidence on the mechanism remains inconclusive. Speech transmission index (STI) is also frequently used as a measure of speech intelligibility (IEC 60268-16, 2011). In brief, a test signal is emitted that resembles speech including modulated frequencies, and the signal in the receiver position is compared to the original and analysed with reference to the modulation depth.

Few studies have investigated the effects of acoustic interventions and their possible benefits for children's learning or wellbeing. Acoustic interventions in schools and preschools typically include fitting absorbents, changing floor carpets, and fitting chairs with noise-reducing cushions. One study evaluated whether the fitting of sound absorbent panels in the ceilings of four classrooms with poor acoustical design in two preschools/kindergarten would affect 3–5 years old children's cognitive performance, linguistic skills and measure of helplessness. Measurements were done before and one year after the fitting and showed an improvement in letter and number recognition, language skills as rated by the teacher, and a reduced susceptibility to induced helplessness using un unsolvable jigsaw puzzle (Maxwell and Evans, 2000). Another study evaluated the combined effect of reducing the external noise from the train and fitting the classrooms facing the rail by absorbent panels in the ceilings (Bronzaft 1981). Comparisons were made before and after the interventions with pupils in classrooms facing a quiet side. Before the intervention the reading scores were significantly lower for the pupils in the noisy classrooms as compared to the quieter classrooms, but 1 year after the interventions this difference had disappeared.

The study presented here evaluated the effectiveness of an intervention aimed at improving the acoustical qualities in preschools in terms of reduced noise levels, and examined the effects of the intervention on children's perceptions and reactions assessed before and after the intervention. The study originate from a framework derived on the basis of focus group discussions with children (Dellve et al., 2013). During 11 focus groups, children recruited from five preschools were interviewed about their perception of sound in the preschool situation, their understanding of the source of this sound, and their perceived reactions at both the emotional and the physiological level. The results formed the basis for the development and validation of a questionnaire known as Inventory of Noise and Children's Health (INCH) (Persson Waye et al., 2013). The present study used INCH to measure the children's perceptions and reactions.

The aim of this study was to study the association between an acoustic intervention and children's perceptions of and reactions to sounds in their preschool environment. Specifically, we aimed to study the association between perceptions, reactions, and symptoms in relation to the sound environment both in terms of objective measurements and as perceived by the children.

MATERIALS AND METHODS

Selection and Recruitment

In the period from October 2006 to October 2009, children aged four to five and their parents were recruited from seven preschools in Mölndal, Sweden, where interventions were undertaken with the purpose of improving the acoustical qualities in the preschools. In total, 63 children and 59 parents filled out the questionnaire before and after the intervention. The response rates ranged from 80% in the parents to 98% in the children. This article reports on the data from the children and the acoustic measurements; parental data will be reported elsewhere. Two of the children fell outside the age range of 4–5 years and were excluded from further analysis, resulting in a study population of 61 children. Data from both the pre- and post-measurements were available for 56 of them.

Ethics

The study was carried out in accordance with the Declaration of Helsinki on Biomedical Research Involving Human Subjects, and was approved by the ethics committee in Gothenburg (ref: 670-06). Parental consent was obtained in advance for the children's participation, but regardless of parental consent, no child was made to take part against their own will.

Acoustic Interventions

Acoustic interventions included changing floor mats from traditional plastic to plastic mats designed to reduce impact sounds, adding damping cushions under chairs, and installing sound-absorbing tiles on the ceilings and some of the walls. Table tops had already been changed to acoustically soft material before the intervention. All of the mentioned acoustic interventions were in accordance with SS 25268:2007 T1 (2017). The expected effect of the absorbers was a moderate reduction of the A-weighted

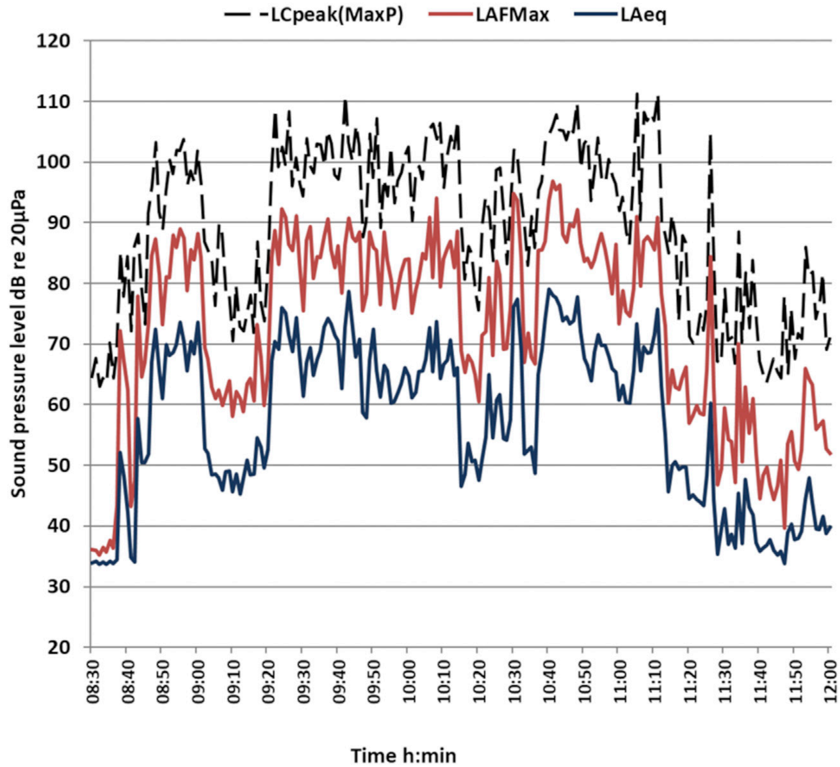


FIGURE 1 | Typical example of the distribution of sound pressure levels dB (C-weighted peak levels LCpeak, LAFMax, LAeq) in a play room in the morning, measured as 30-s equivalent noise levels.

equivalent sound level in the range of 3 dB, while the changes in table tops and floor mats and the addition of damping cushions were expected to mainly lead to a decrease in the level of contact sounds resulting from, for example, plates and glasses being put down on the table or chairs being pulled over the floor. These latter types of sounds would normally not be of large importance for the overall A-weighted sound level in a preschool, but could still be of importance for high-frequency contact sounds in the acoustic environment.

The room acoustic parameters reverberation time (T20), Clarity (C50), Deutlichkeit (D50), Centre time (Ts) and speech transmission index (STI) were evaluated in three random preschools before and after acoustic interventions (ISO 3382-1, 2009). Two room types were studied; "meal/craft room" and "play room."

Monitoring of the Sound Environment

Noise levels were measured during the course of 1 week, 1 month before the intervention and again 3 months after the intervention, using stationary measurements and personal dosimeters. The purpose of measurements using the stationary sound level meters was to gain an overall estimate of the sound levels in the room from the activities, and to see how the acoustic interventions possibly affected the sound levels. Sound levels in the room used for crafts and eating meals (meal/craft room) and the room for organised activity or play (play room) were measured for 1–2 days per room using a stationary sound level meter type I (Bruel and Kjaer 2260) equipped with a ½ inch microphone. The system was calibrated before and after the measurement week and before and after every movement between measurement positions using an acoustic calibrator (Bruel and Kjaer 2231). The microphone was placed 0.5 m from the ceiling at a position where the activity noise would be representative for that room. **Figure 1** shows a typical example of a measured sound environment in one of the preschool play rooms.

The measured 30-s equivalent sound level varied greatly over time during the course of a day, depending on the activities and presence of children and staff in these rooms. In order to cope with this highly variable condition, we adopted a method of analysing the periods when the rooms were occupied. Periods when the rooms were occupied were first identified by asking the staff about the time periods for meals, craft, and play, and then including these periods for each respective room while excluding periods spent elsewhere, for example doing outdoor activities. Using these definitions, the occupied periods were 08:30–11:30 for the play room and 11:30–12:30 and 14:00–15:15 for the meal/craft room; in the analyses, the two meal/craft room periods were combined. We then analysed minute by minute the length of a measurement period for respective room for the levels to become stable, this means that the cumulative period analysed was not deviating significantly from the proceeding period established using a *t*-distribution based 95% CI. The logarithmic mean values were calculated using the formula stated in Nordtest Method NT ACOU 115 (2003) section 10.4. The theoretical approach was that these periods would be long enough to not show deviation if longer measurement periods would be made at any random interval. The hence derived sound levels would in other words not have been significantly different if we have included a longer time period. In our case, stable periods were derived after an average of 17–22 min depending on the room. The stable periods were determined when the deviation between the different periods were less than 0.5 dB. Using these time periods, we calculated percentiles of time for sound levels in the room. In this study, we judged that sound periods exceeding the 50% level (LAeq50%) would reflect levels when the room is inhabited. For more details of the method, see Nordtest Method NT ACOU 115 (2003). The sound levels for these time periods were averaged for the room types and used as a base for the statistical analyses of the effect of the intervention.

Individual noise exposures were obtained from two children at a time, using dosimeters (SPARK 705 +), and analysed using BLAZE 5.06 software, type II. The purpose was to gain an estimate of the children's and personnel's noise exposure when being indoors. The settings used were 30-s averaging intervals, a gain of 30 dB, and a range of 43–113 dB. Only time periods when the staff and children were participating in their usual indoor preschool activities were used for these personal measurements (referred to in the following as time indoors, Ti). The same procedure was undertaken in the reference schools where no interventions took place. The total numbers of dosimeter measurements at times I (pre-intervention) and II (post-intervention) were 61 and 55 for the study group children, 66 for the study group staff on both occasions. For the reference groups there were 18 children and 18 staff at time I and 20 children and 23 for staff at time II.

Children's Perceptions of Sounds

Perception of sounds was measured by means of a questionnaire (INCH) developed through focus group interviews with 3–5-year-old preschoolers (Dellve et al., 2013) and subsequently validated and presented in Persson Waye et al., (2013). The interviews were performed by one research assistant within our research group A Agge. She was trained in interviewing children by a special care pedagogue that carried out the focus group interviews in the preceding study (Dellve et al., 2013). After parental consent and in agreement with the preschool teachers, one child at a time was interviewed in a room at the preschool premises. The child was initially asked general questions on what colours he/she would like at the preschool, and if he/she liked the dining room and the play room and what sounds he/she normally heard at the preschool. After this introductory conversation the child was asked how frequently he or she heard sounds from the three sound sources found to be most relevant from the analyses of focus groups (Dellve et al., 2013): yelling and angry children, strong and loud sounds, and scraping and screeching sounds. Answers were indicated on a five-point Likert scale (ranging from "almost never" to "very often") presented as five circles increasing in size and including 1–5 dots.

Using the same scale of circles, the children were asked how often they experienced the teachers yelling or shouting when talking. For the analyses, a bipolar scale was created with scores ≤3 recorded as 0 and scores >3 as 1.

Bodily perception of sounds was indicated by pointing at various parts of the body of a child-like figure with neutral bodily and facial expression, shown in isolation but similar to the middle figure on the scale in **Figure 2**. The answer was recorded for all three sounds separately, but to increase the robustness of the analysis, any physical perception of any sound was used; this was scored as 1 and 0, with the latter indicating no bodily perception of any sound.

Children's Reactions to Sounds

Aspects of reaction were measured using the following wording: "How do you feel when you hear (the sounds of angry, yelling children) (loud and strong sounds) (scraping and screeching sounds)?" Answers were indicated on a bipolar visual scale representing figures drawn with different facial and bodily expressions ranging from glad/safe to sad/afraid (sad reaction) and from kind/friendly to angry/irritated (angry reaction), respectively. The scale of the sad reaction is given in **Figure 2**.

The reaction was recoded to neutral position (code 3) for those children who indicated on the previous question on perception that they "almost never" heard the sound. In the analyses, scores ≤3 were coded as 0 and scores >3 as 1.

Symptoms and Wellbeing

Psychosomatic symptoms that may be related to noise among children were elicited using the question: "During the last few days at preschool have you had a (tummy ache) (headache)?" The prevalence of hoarse throat was measured using the same question. Answers were given on a 5-graded scale with circles increasing in size from "never" to "often." Finally, a question was asked about general wellbeing, using figures similar to those used for sad/happy reaction (**Figure 2**). As above, a bipolar scale was created with scores ≤3 recorded as 0 and scores >3 as 1.

Procedure

An overview of the design and procedure is given in **Table 1**. Children in the intervention preschools were interviewed 1 month before (time I) and 3 months after the intervention (time II). Times I and II were chosen to include spring and autumn, which are similar with regard to daylight and time spent outdoors. The periods therefore did not include December and January, which are the darkest and coldest period, and the summer months of May and August. In order to diminish the risk of inter-rater variance as much as possible, the interviews were performed by two trained persons. The children were asked questions in a structured way, and presented with visual representations of scales on show cards. When the child was not able to answer the question, they were not prompted to do so. In the reference preschool, children were not interviewed; only sound levels were measured.

Study Population

The children in the studied preschools typically arrived before breakfast, which was at 08.00, and were collected by their parents or similar between 16.00 and 18.00. **Table 2** shows the

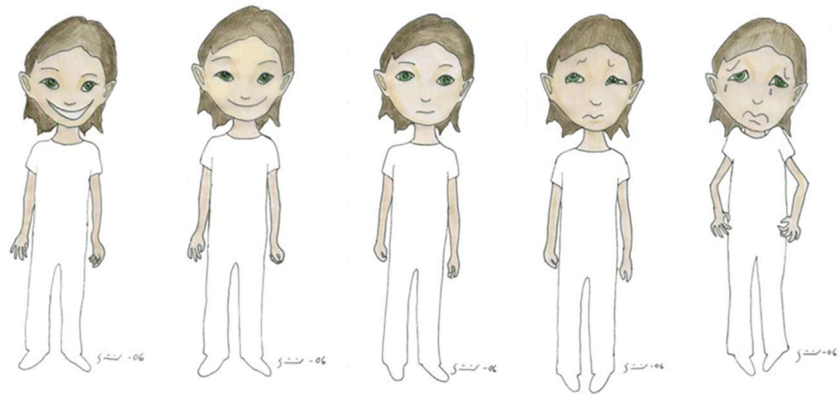


FIGURE 2 | Visual representation of the scale measuring sad reaction, with the furthest left position labelled happy/safe and the furthest right position labelled sad/afraid. The neutral position was used in a separate figure for children to score the bodily reaction.

TABLE 1 | Study design and the time measurements in relation to the intervention, in the intervention preschools and the reference preschools.

Preschool	Autumn 2006	Spring 2007	Autumn 2007	Spring 2008
i1	Time I	Time II	—	—
i2	Time I	Time II	—	—
i3	Time I	Time II	—	—
i4	—	—	Time I	Time II
i5	—	—	Time I	Time II
i6	—	—	Time I	Time II
i7	—	—	Time I	Time II
c1	—	—	Time I	Time II
c2	—	—	Time I	Time II
c3	—	—	Time I	Time II

i1, i2, ... denotes the seven preschools that were part of the intervention. The time of the intervention is illustrated with a vertical arrow. Time (I) denotes the first measurements, which were conducted in the intervention schools one month before the intervention, including children's response, room acoustics, sound measurements (dosimeters and stationary), and voice (reported elsewhere). Time (II) denotes the second set of measurements, which were conducted in the intervention schools three months after the intervention. c1, c2, and c3 are the reference preschools, where only sound level measurements (dosimeter and stationary) were conducted at times I and II.

TABLE 2 | Study population before (Time I) and after (Time II) the intervention.

Number of respondents	Children, *n* (response rate, %)	Time I	Time II
		61 (91)	56 (90)
Gender	Girls	48%	49%
	Boys	52%	51%
Age	4 years	52%	33%
	5 years	48%	49%
	6 years	—	8%

distribution of age and gender of the children included in the analysis of the interviews. There was a high participation rate among the children both before and after and there was no difference in response rate between preschools. The children included in the before and after study were fairly well distributed over gender and age groups. All children aged 4–6 years were asked to participate in the interviews; the number who participated ranged from 4 to 15 per preschool. Only children who took part in both the before and after study were included in the analyses.

Supplementary Table S1 gives the numbers of children attending the preschools every day during the study periods; as shown, the average numbers of children were fairly similar across periods.

Statistical Analyses

Equivalent noise levels from the stationary measurements in the two room categories in each preschool before and after the intervention, as well as the dosimeter measures, are presented using descriptive statistics. The dosimeter levels will have been most strongly influenced by the direct sounds from the child's own voice, the voices of others, and activities, and to a lesser degree by indirect sounds from the room; they will therefore mainly be used to describe the child's sound exposure. The noise levels from the stationary microphones, on the other hand, include both direct and indirect (reflected) sound signals from activities in the meal/craft room and play room, and could be hypothesised to be affected by the interventions. In the statistical analyses of whether the intervention affected the noise in the rooms, we compared sound levels occurring more than 50% of the time (LAeq50%) as this would reflect levels when the room is inhabited.

Generalised estimating equation (GEE) logistic regression models were applied to analyse the associations between different outcomes and relevant explanatory variables while accounting for potential confounders and the repeated measures due to the intervention (before-after design) Compound symmetry structure was used for the working correlation matrix (structure: exchangeable). Three models were built. Model I included equivalent stationary noise levels (before and after) as primary explanatory variable, and children's sound perception, perception of teachers' vocal behaviour, bodily sensations, sound source reactions, and symptoms as dependent variables. As only one of the sound sources (scraping and screeching sound) was found to be significantly associated with the change of sound level, only this source was included in further analyses. Model II adopted children's perceptions of

TABLE 3 | Mean equivalent A-weighted noise levels, exceeded 50% of the time (LAeq 50%)* dB measured in the meal/craft room and play room at the intervention i) and reference (c) preschools.

	Meal/craft room			Play room		
	Time 1	Time II	Difference (95%CI)	Time I	Time II	Difference (95%CI)
Intervention preschools n **= 6** [b]	69.1	67.9	**1.2 (0.55—1.83)**[a]	69.3	65.6	*3.8 (–0.08—7.58)*[a]
Reference preschools n **= 3**	67.6	67.5	0.04 (–2.08–2.16)[a]	67.2	66.9	0.3 (–2.74–1.54)[a]

Statistical significant difference is indicated in bold, and borderline indicated in italic. 95% CI means 95% confidence interval.
[a]*calculated based on sound level periods of 17–22 min.*
[b]*missing values from one preschool.*

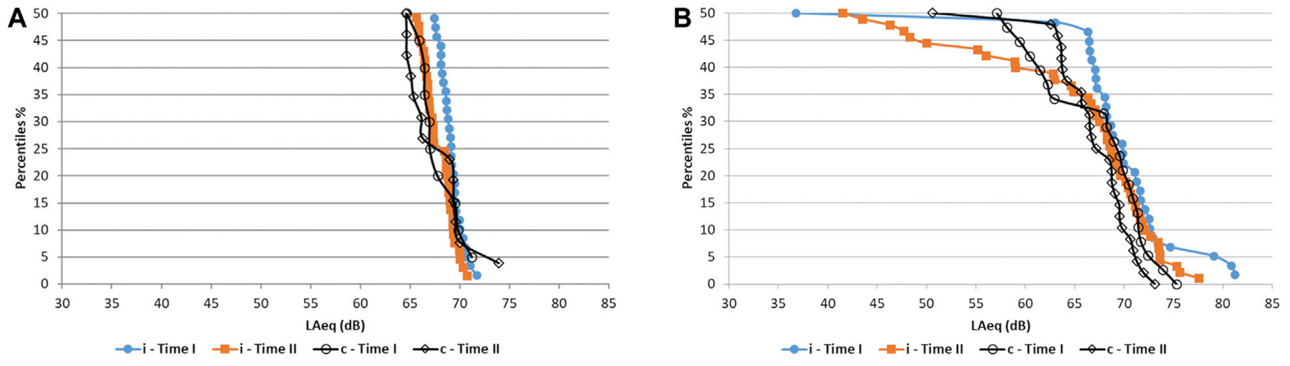


FIGURE 3 | LAeq values from stationary measurements before (time I) and after (time II) interventions. Each point represent an equivalent value calculated for a time interval of 17–22 min, depending on the room. Sound levels exceeding the 50% level, LAeq50% are shown. Intervention, "i," and reference, "c," for meal/craft room **(A)** and play room **(B)**.

sounds as an indicator of perceived noise, and perceptions of teacher's vocal behaviour, sound source reactions, and symptoms as outcomes. Finally, Model III adopted bodily sensations as an explanatory or independent variable for reported symptoms. Dependent variables were recalculated into binary variables by giving scale values 1, 2, and 3 a score of 0, 4, and 5 a score of 1. Scale values of 4 and 5 are interpreted as "often" and "very often" (perception and symptoms) or as "much" and "very much" (for reaction and wellbeing). For bodily reaction, any bodily reaction to any sound was given a value of 1, and no reaction to any sound a value of 0.

Potential confounding variables such as age and gender were included in all models. The results of regression analyses are given as odds ratios (OR) with 95% confidence intervals (CI). The OR in the presented models denotes an increased or decreased odds of the binary outcomes in relation to a change of the independent variable. If the odds is below 1 and the 95% confidence intervals do not include 1, it shows a significantly reduced perception, reaction or symptom at time II as compared to time I. All statistical analyses were performed using version 20.0 of IBM SPSS Statistics for Windows (IBM Corp, Armonk, NY, United States), applying two-tailed tests and a 5% level of significance.

RESULTS

Indoor Noise Levels

The mean equivalent A-weighted noise levels measured in the meal/craft room and play room are given in **Table 3**.

For the intervention preschools the mean point estimate of the difference in LAeq 50% in the meal/craft room at time II was rather modest or 1.2 dB, but statistically significant from time I. In the play room the mean point estimate for the intervention preschools was greater, but did not fully reach statistical significance (p = 0.059, t-test), as also indicated by the 95% CI (**Table 3**). Importantly, there were no statistical differences of mean point estimates the LAeq 50% levels between times I and II in the reference preschools. In **Figures 3A,B** the equivalent A-weighted noise levels for each averaged time interval of between 17 and 22 min, used to calculate equivalent levels exceeding 50% of the time, are shown for each room type.

The average dosimeter levels for children in the intervention group were 85 dB LAeq (95% CI 83.0–86.0) before and 83 dB LAeq (95% CI 82.2–84.6) after the intervention; the corresponding figures for children in the reference group were 84 dB LAeq (95% CI 82.8–85.2) and 84 dB LAeq (95% CI 82.3–86.5) at times I and II, respectively. There was no difference in dosimeter levels between the intervention and the reference groups. Neither did we see any clear differences between weekdays. **Table 4** shows the 5, 25 and 50% percentile dosimeter sound levels for the intervention and reference group combined. Dosimeter levels for the staff were on average 6–8 dB LAeq lower than those for the children in both groups (data not shown), and their exposure was significantly different from the children's exposure (Student's t-test $p < 0.001$). In **Figures 4A,B** the measured dosimeter values for children are shown, each dot or square represent one daily measurement. The obtained values for LAeq and LAFmax are shown together with the limit value for equivalent levels 8 h (LAeq 85 dB) and limit value for maximum

TABLE 4 | Percentiles of dosimeter sound levels LAeqᵢ corresponding to time spent indoors from all intervention 1) and reference (c) preschools at time I and II.

Percentiles (%)	Time I				Time II			
	$LAeq_i$ dB		LAFmax dB		$LAeq_i$ dB		LAFmax dB	
	Intervention	References	Intervention	References	Intervention	References	Intervention	References
5	90	87	121	122	91	91	121	121
25	87	85	120	120	86	88	120	120
50	85	84	119	119	83	85	118	120

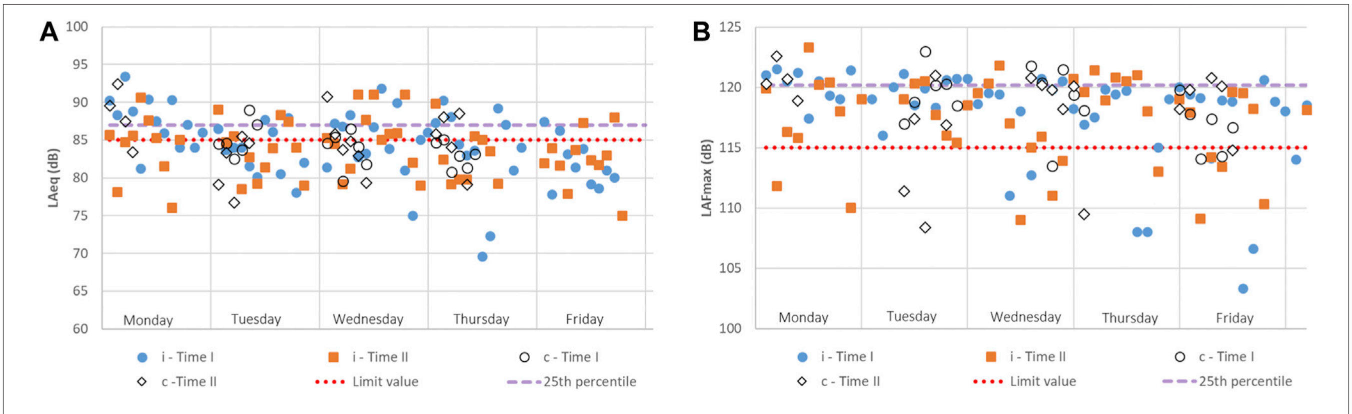


FIGURE 4 | Measured dosimeter values per weekday for time I and time II together with 25th percentile, which is based on the whole data-set. In **(A)** the values for LAeq are shown together with the Limit value (85 dB). In **(B)** the values for LAFmax are shown together with the Limit value (115 dB). Intervention marked with "i" and reference schools "c."

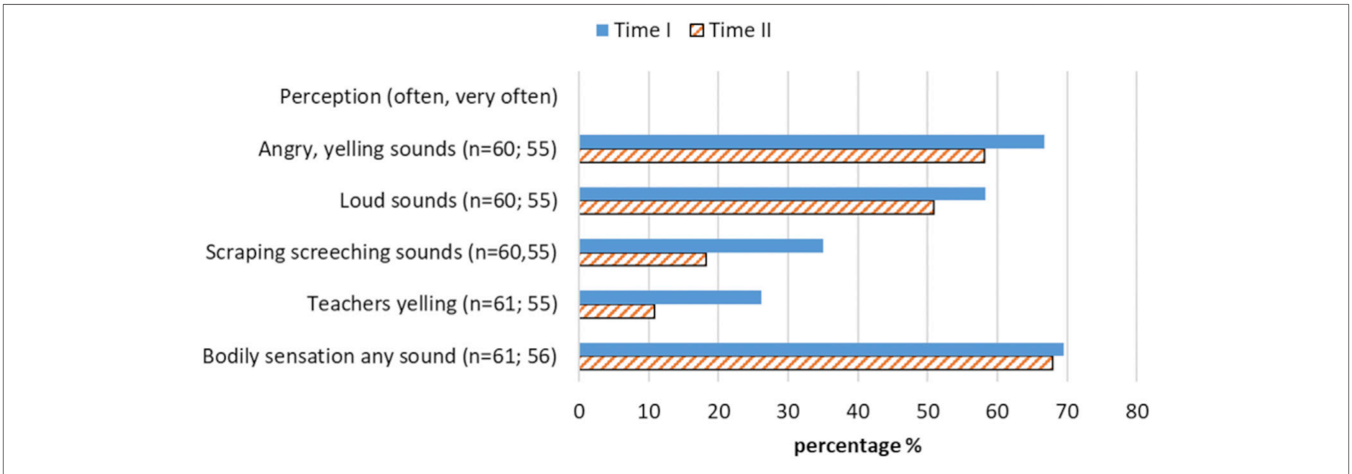


FIGURE 5 | Prevalence of children's perception of angry, yelling sounds, loud sounds, scraping screeching sounds, teachers yelling, and feeling a bodily sensation of any sound "often" and "very often." Blue bars denote before (time I) and orange bars with diagonal lines denote after (time II), the intervention.

levels (LAFmax 115 dB), respectively. As can be seen in **Figure 3A**, there is some variation of the equivalent sound level between children's exposure, but a clear majority are exposed to sound levels above LAeq 80 dB during their time spent indoors. For the maximum levels most children are at some time during the day exposed to events exceeding LAFmax 115 dB. According to the work environment statue (AFS 2005:16, 2005) noises above the lower action limit (80 LAeq 8h) dB) require attention while levels above the limit (85LAeq 8h dB or 115 LAFmax dB) require actions such as a requirement for hearing protection.

Children's Response

Prevalence of children's perception of angry, yelling sounds, loud sounds, scraping screeching sounds, teachers yelling, and feeling a bodily sensation of any sound "often" and "very often" at time I and II are given in **Figure 5**.

As shown in **Figure 5**, the proportion of children hearing various sounds tended to be higher before the intervention, with more than 65% of the children reporting hearing angry, yelling sounds often or very often, around 55% hearing loud sounds, and about a third hearing scraping, screeching sounds and the teacher yelling often or very often.

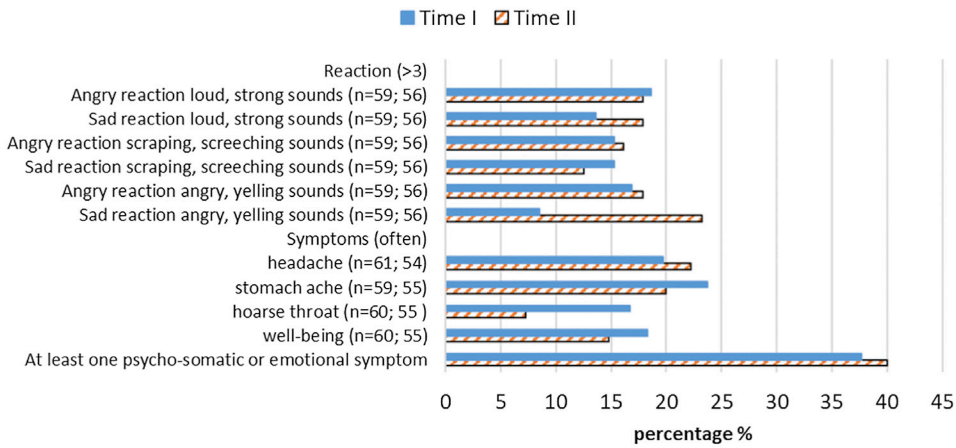


FIGURE 6 | Prevalence of children's angry or sad reactions of four or five on the scale and symptoms reported often. Blue bars denote before (time I) and orange bars with diagonal lines denote after (time II), the intervention.

TABLE 5 | Associations between the explanatory variables of measured or perceived noise exposure, provided by the three analytical models, and children's perceptions, reactions, and bodily symptoms.

Dependent variable	Model I[a]	Model II[b]	Model III[c]
	Exp B 95% CI (Exp B)	Exp B 95% CI (Exp B)	Exp B 95% CI (Exp B)
Perception of yelling sounds	n.s	—	—
Perception of loud sounds	n.s	—	—
Perception of scraping sounds	0.69 (0.55–0.86)***	—	—
Sad reaction to scraping sounds	n.s	0.33 (0.09–1.17) p < 0.10	—
Angry reaction to scraping sounds	n.s	0.37 (0.15–0.89)*	—
Perception of teacher yelling	0.80 (0.62–1.04) p < 0.10	0.36 (0.14–0.94)*	—
Hoarse throat	n.s	0.32 (0.11–0.91)*	Not applicable
Stomach ache	0.71 (0.54–0.93)*	n.s	0.22 (0.07–0.78)*
Headache	n.s	n.s	0.15 (0.04–0.66)*
Wellbeing	n.s	n.s	0.21 (0.05–0.89)*
Bodily perception of any sound	n.s	n.s	—

[a]*Model I: stationary levels in the meal/craft room as independent variable.*
[b]*Model II: perception of scraping screeching sounds as independent variable.*
[c]*Model III: bodily perception of any sound as independent variable.*
** p < 0.05; ** p < 0-01; *** p < 0.001; n. s = not significant.*

Remarkably, close to 70% reported feeling some sounds in various parts of their body, with very similar prevalences in the before and after conditions.

Prevalence of children's angry or sad reactions to loud strong sounds, scraping screeching sounds and angry, yelling sounds, given a value of four or five on the scale as well as symptoms being reported to occur "often" at time I and II are given in **Figure 6**.

The pattern of reactions to sounds seemed more ambiguous, with most reactions tending to be higher after the intervention (**Figure 6**). Conversely, most symptoms seemed to be slightly lower after the intervention, with hoarse voice having the largest prevalence reduction.

Results from the GEE logistic regression for all three models are given in **Table 5**. Model I included the change of stationary noise levels before and after measured in the meal/craft room, as levels in the play room were not found to be statistically significantly explanatory factors between time I and time II. Age and gender were not found to have a significant impact on any outcome.

Model I showed that a change in noise levels in the meal/craft room was associated with a 31% reduction in children's perception of scraping and screeching sounds with the odds being 0.69, and a near 30% reduction in the frequency of reported stomach ache (odds of 0.71). There was also a tendency toward a 20% reduction in children reporting occurrences of the teacher yelling or calling out with a raised voice, but this did not reach statistical significance (p = 0.093). The perceptions of the other sound characteristics were not significantly affected by the change of sound level. Model II showed that a reduction of the perception of scraping and screeching sounds per se was associated with a 63% reduction of reporting of anger in reaction to these sounds with the odds being 0.37, while differences in sad reactions did not reach statistical significance (p = 0.086). A reduced perception of scraping and screeching sounds was associated with a 64% reduction in children's reporting teachers yelling or calling out

TABLE 6 | Calculated ratio of area under the curve for averaged room frequency responses per room type.

	Normalised area under curve		
	Time I	Time II	Ratio
Meal/craft room	11,737	6,519	0.56
Play room	11,889	7,740	0.65

TABLE 7 | Averaged values and 95%CI for C50 and D50 per room type.

	C50 (dB)			
	Time I	Time II	Δ	JND
Meal/craft room	10 (95% CI 9.0–10.5)	13 (95% CI 12.5–14.3)	4	1
Play room	8 (95% CI 7.8–9.1)	11 (95% CI 10.4–12.5)	3	1
	D50 (%)			
	Time I	Time II	Δ	JND
Meal/craft room	89 (95% CI 87.7–90.9)	95 (95% CI 93.5–95.5)	5	5
Play room	86 (95% CI 84.5–88.1)	92 (95% CI 90.8–93.4)	6	5

JND denotes Just Noticeable Difference; and C50 denotes Clarity, D50 denotes Deutlichkeit.

with a raised voice, and a reduced perception of scraping and screeching sounds was associated with a 68% reduction in reporting a hoarse throat. Finally, Model III showed that a reduction in bodily perception of any sounds was associated with a large and significant reduction in the symptoms of stomach ache, headache, and increase in wellbeing, all in the range of 81–85%.

Room Acoustic Results

All measured rooms in the preschools had reverberation times in the order of 0.3–0.5 s before the intervention, and after the intervention the average reduction was in the order of 0.1 s. These measures are considered very low also before the intervention. The measurements showed the greatest reduction in the frequency range of 250–500 Hz for meal/craft room and in 250–500 and 2,000–4,000 Hz for play room, as seen in **Supplementary Figures S3A, S4**. The reduction in the frequency range 250–500 Hz was expected due to the configuration of suspended acoustic tiles with bass reinforcement. In one preschool the reverberation time increased after the intervention which is most likely due to how T20 is evaluated. The rooms in this preschool had broken decay curves which can occur when there are an uneven distribution of absorptive surfaces in one dimension. In this case it was most probably in the vertical plane due to the acoustic tiles in the ceiling (ISO 3382-2, 2008). The estimated standard deviation for T20 ranged from 0.001 to 0.009 s before the intervention for meal/craft room and from 0.001 to 0.008 s after. Values were in a similar range for play room. These values were obtained by using the described method in ISO 3382-1 on the averaged T20 results for each room type. The STI also improved slightly, increasing by around four to five percentage points, but as with the reverberation time the STI was already very good (more than 75%) before the interventions. (Larsson, 2011).

In addition to the reverberation time, an analysis of the rooms' frequency responses were performed. The measured impulse responses were averaged for each room and compared before and after the intervention. The general observation was that the room responses became more flat after the intervention. Hence the signal was less affected (coloured) by the room. The frequency responses in the frequency range 250–4,000 Hz were averaged by room type and the results are shown in **Supplementary Figures S5, S6**. The area under each curve was calculated using the "trapz" function in matlab in order to compare the dynamic of the responses. The integral was computed with equal spacing with steps of 1 Hz. The frequency responses were normalised to the minimum sound pressure level within the chosen frequency range for each room type. For further clarification see **Supplementary Material S2**—Calculations of area under the curve.

The average ratio of the areas under the curves between time I and time II for the two rooms are seen in **Table 6**. The area decreased at time II for both meal/craft room and play room.

Table 7 shows the C50 and D50 before and after the intervention. C50 and D50 showed an overall increase in the order of 3 dB and 5% which corresponds to 3 and 1 units of just noticeable difference (JND), respectively. However, the values before the interventions were C50 > 6 dB and D50 > 80% and it is not clear that a further increase of C50 and D50 in these rooms are beneficial or even perceivable. Centre time (Ts) was also evaluated and showed the same tendency as previously mentioned parameters. The measured Ts were in the range of 21–25 ms before and 15–17 ms after the interventions, hence the differences were then below JND and thereby most probably not perceivable. It can be noted that the values for Ts were considered to be very low also before the intervention meaning that the late energy in the impulse responses was attenuated.

DISCUSSION

The main finding of this study was a confirmed association between reduction of stationary noise levels after an acoustic intervention and a 30% reduction in children's perceptions of scraping and screeching sounds along with a similar reduction in self-reported stomach ache. In addition, when using children's perceptions of scraping and screeching sounds as predictor of the sound environmental change after the intervention, the change was associated with important oral communication outcomes, perception of teachers yelling and hoarse throat. Finally, when using the children's bodily sensation of any sound as a predictor of the sound environment change after the intervention, the change was associated with psychosomatic symptoms and wellbeing. The reduction of the sound levels seen in the intervention preschools was not found in the reference schools. As we unfortunately could not interview the children in the reference preschools, our conclusions are hampered by the lack of control for factors other than the acoustic interventions that may have affected the children's responses.

Given the very high sound equivalent and maximum levels that children were exposed to, as shown by the ample dosimeter measurements and stationary microphone measurements in the rooms, reductions of the indoor sound levels are highly needed. A reduction of the A-weighted equivalent noise level of 1–3 dB is not enough from a health perspective, but as it refers to a change of equivalent sound levels over one to 2 days per measurement point, it is clearly indicative. A change of up to 3 dB A-weighted equivalent level is also what can be expected with the fitting of absorbents to the ceiling.

The perceptual analyses, indicate that the interventions may to a lesser degree have affected perceived strength, but had a larger impact on perceptual qualities of the sound environment, that may be well less captured by the A-weighted equivalent sound level. The change in perceptual qualities could partly be explained by data showing that human neural processing of frequency develops at an earlier age than that of intensity (Sussman and Steinschneider, 2011), meaning that young children may be more attentive to changes in frequency characteristics than changes in intensity. It is also plausible that children more strongly direct their awareness toward (and hence are more likely to report on) sound characteristics that they perceived as most unpleasant. Scraping and screeching sounds were one unpleasant sound characteristics reported by children interviewed during the development of the INCH instrument (Dellve et al., 2013). Children typically described unease and bodily discomfort from sources such as the screeching sound of a swing or the scraping sounds of cutlery on plates. This observation was also confirmed in data obtained in German schools, where a relationship between binaural sharpness measured using a child's artificial head and children's reporting of sad reactions using INCH was indicated (Loh and Fels 2018). The amplification of high frequencies by small children's HRTF (Fels 2008) would enhance the perception of these types of high-frequency sounds, to a level that may be perceived as unpleasant. The intervention, which included fitting dampening cushions on chairs and absorbent tiling on the walls and ceilings, could, in combination with the already fitted acoustically soft table top, be of high importance in reducing the direct high-frequency sounds that reach the child's eardrum, however more studies are clearly needed to elucidate these matters.

The room acoustic interventions had the expected effect when studying the evaluated room acoustic parameters by room type. The measures were however considered to be good also before the interventions and thus the relative difference was small. The difference of C50 and D50 were close to or below JND in four out of six evaluated rooms and therefore, we cannot be certain that a change in acoustic quality was perceivable. Parameters such as C50 and D50 are greatly affected by the volume of the room due to more early reflections occurring because of the surfaces being closer to one another compared to a bigger room. It is not clear if these measures describe the perceived acoustic quality of small rooms from a child's perspective and further investigations are needed. Finally, when comparing the evaluated parameters per room, one of the preschools showed the opposite behaviour with increasing T20 and decreasing C50 and D50. This can be explained by an uneven distribution of absorptive surfaces as well as less diffusive elements in the horizontal plane.

Unexpectedly, the reduction in A-weighted noise levels after the intervention were not found to be associated to a lower rating of the perception of loud, strong sounds. One explanation for this could be that the loudness curves on which the A-weighting is based were derived from adults (ISO 226:2003, 2003) and as the HRTFs and ear canal of small children amplify higher frequencies differently compared to an adult (Fels, 2008), the loudness relation to frequency may not be similar for children and adults. If this is confirmed in future studies it would mean that that the A-weighted levels may not fully represent how young children perceive loudness.

A reduced perception of yelling sounds from other children would also have been plausible. However, children's yelling would mainly reach other children as direct sounds not well attenuated by the absorbent tiling placed on the ceiling or high up on the walls, where its main function would be to absorb reflected sounds. The lack of reduction of direct sounds was also shown by the lack of change in the dosimeter levels obtained for the staff and children before and after the intervention. An interesting finding was the tendency for the children to perceive the teacher as yelling or calling out less often after the intervention, a result also confirmed by a significant association in Model II. This suggests that the teachers experienced a higher acoustic speech comfort or were more aware of their noise-generating behaviour as a side-effect of the intervention, this could be of interest for future studies. The lack of an association between children's yelling and the intervention may also be interpreted as a non-confirmation of a reverse Lombard effect, as the acoustic interventions would be expected to make children less prone to raise their voices. It is possible that children's vocal effort would have benefited more if the absorbing tiles had been placed at child height, but this can currently only be speculated on, and needs to be furthered studied. Other studies investigating the effect of an acoustic intervention on vocal symptoms have focused on the teacher. For example alleviation of some vocal symptoms was reported among participating teachers and an improved perceived clarity and audibility of the teacher's voice was reported by school aged children after an acoustic interventions in a classroom (Pirilä et al., 2020).

Although the change of the traditional room acoustic parameters was small, the shape of the room response curves showed a tendency of being flatter after the interventions. The implication of frequency balance of low and high frequencies has previously been studied in other contexts e.g. "spectral balance" with emphasis on the perceptual qualities of low frequency noise from ventilation systems (Beranek 1989) and in music studios (International Telecommunication Union 2015), but not in this context. It is plausible however, that the frequency balance of the room response may be of importance for the perceived acoustic quality in preschools rooms, as small children seem to be more susceptible to high frequency contents in the sounds. On the other hand, a low frequency dominance may also impair children's language achievement as low frequency noise may mask frequencies important for speech communication (Pickett 1959). Though these studies have to the authors knowledge, not been performed within a child population. Taken together it seems as if a balanced or flattening of the room response curve would be advantageous also in a preschool environment, however this needs to be further studied.

The association between a lower perception of scraping, screeching sounds and a 70% reduction in children reporting a hoarse throat, as well as a similar reduction in being angry when hearing scraping and screeching sounds was seen in Model II. These findings could indicate an improved acoustic environment, where children have less need to raise their voice in order to be heard. This would be a very significant achievement, as using vocal behaviour to "be heard" was the main category reported by preschool teachers when describing how preschool noise affected children (Persson Waye et al., 2019). This category included more than 7,000 occurrences of words with the same meaning as "to be heard." The category in second place, which described children being distracted, included only 2,600 word counts.

The hypothesis of a reduction of scraping and screeching sounds being indicative of improved acoustic comfort is in line with the reduced occurrence of stomach ache found with the reduction of sound levels after the intervention. It is known from other areas of research that gastrointestinal organs are particularly sensitive to stress, especially in children (Ricour, 1989). Stomach ache, headache, and tiredness are reported as frequent causes of psychosomatic complaints in children (Silber and Pao, 2003), and a large study indicated that childhood stress may trigger psychosomatic and emotional symptoms (Vanaelst et al., 2012). The latter study collected parental reports on more than 4,000 children from eight European countries on headaches, stomach aches, sickness, and low emotional wellbeing, and related this to childhood adversity in a broad sense. The prevalence of frequent occurrence of at least one such psychosomatic and emotional symptom was broken down by age group, and it is interesting to note that the prevalence for 4-year-olds and 5-year-olds were 38.4 and 39.3%, respectively; remarkably similar to the percentages of 38 and 40% at times I and II in our study. It should be noted that our sample size was very small and hence prone to random errors, and that "sickness" was not included in our study. This again raises the need for further studies.

The findings from Model III showed a strong association between bodily perception and the psychosomatic symptoms of headache, stomach ache, and effects on wellbeing; a reduction of bodily perception was associated with around 80% reduction of symptoms. A previous qualitative study (Dellve et al., 2013) found that children perceive sounds and noise physically and within their body, and so it was important to include bodily perception in this study as a complement to auditory perception. Reducing psychosomatic symptoms would be a substantial achievement, as the prevalence of psychosomatic symptoms is greatly increasing among young children and adolescents (Luntamo et al., 2012; Vanaelst et al., 2012). Although noise is only one factor that may contribute to psychosomatic symptoms, it is a highly prevalent stressor for a large population of preschool-age children, both in Sweden and in other countries, and may therefore warrant concern.

Achieving an acoustic environment that reduces auditory and bodily discomfort is important not just from a wellbeing and health perspective but also for language acquisition and learning, as children tend to cope with aversive noise by ignoring or disregarding auditory inputs (Evans, 2006; Evans and Hygge, 2007). An unfortunate consequence of this is that important speech signals are also tuned out, which may result in impaired writing and reading abilities. Non-native language speaking children and children with language disorder and hearing impairments are at particular risk in such settings.

CONCLUSION

Achieving an acoustic environment that reduces auditory and bodily discomfort as well as supports language acquisition and learning is of major importance for children in the preschool environment. To reach this goal we need to know more on how small children perceive and react to sounds and noise in their environment. This study provided for the first time data on how children perceptually may perceive and react to a change of the sound environment and acoustic qualities resulting from an acoustic intervention. Furthermore, we were able to gather children's response using a questionnaire derived from children's own wording. It points to the importance of acoustical qualities that may not be included in today's standards of room acoustics and suggests that a revision of standards for preschools need to take as a point of departure small children's hearing, perception and reaction, paying attention to factors such as direct sounds from, i.e., friction between surfaces. Future studies should aim to include a larger sample of children, use a child perspective approach and perform a systematic evaluation of which room acoustics and sound qualities that are supportive of children's health and wellbeing.

Strengths and Limitations

An important limitation is the lack of interviews with the children from the reference preschools, where only measurements of noise levels were carried out. This makes our conclusions open to the influence of factors other than the acoustic interventions that may have affected the children's responses to the intervention at the preschools. The room acoustic measurements were only performed in three preschools to control that the intervention had the expected effect and we can therefore not make direct analyses of possible associations between the change of room acoustics and children response. A possible limitation for generalisability to today's situation could be that the study was performed between 2006 and 2009, and hence it is possible that the preschool environment has changed to the better. Unfortunately, there are little indications of a positive change of the sound environment as the Swedish National Agency for Education report that the number of children per group and the number of children per personnel is largely the same today, while the proportion of preschool teachers with a university degree has even decreased with about 3 percentage units. Furthermore, the follow up study performed among the preschool teachers in 2013–2014 indicated large problems with noise and similar reactions among the children (Persson Waye et al., 2019).

Another limitation is that Models II and III used self-report for the dependent and independent variables, hence including the possibility of information bias. However, this is currently unavoidable when investigating perceptual qualities from a child's perspective. Using only the measured values (A-weighted sound pressure levels) as independent variable may bias the results as the weighting is based on loudness estimations and derived from adult perceptions. However, as Model I, with the objective measures as independent variable, showed a relationship with children's perceptions of scraping screeching sounds, it gave a basis for the use of children's perceptions of these sound qualities as an independent variable in Model II. The INCH questionnaire has been validated and the similar reporting of prevalence of perceptions, reactions, symptoms, and wellbeing on the two occasions gives credibility to the instrument. A strength is that the questionnaire was constructed from focus group interviews and so the questions posed to the children were worded in their own "language."

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because ethics do not allow redistribution of data of the children, however acoustic data can be shared. Requests to access the datasets should be directed to kerstin.persson.waye@amm.gu.se.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Gothenburg (ref: 670-06). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

KPW aquired the funding, designed the study and supervised the work and carried out the main analyses of the intervention study. She drafted the article and revised. She also carried out part of the field work. JK reanalysed the room acoustic data and took part in the revision of the article.

FUNDING

The study was funded by Swedish national research funds FAS/Forte (Swedish research council for health working life and welfare), and 2008-0901 and FORMAS (Swedish Research Council for sustainable development) 2005-1153.

ACKNOWLEDGMENTS

The author would like to thank Agneta Agge for valuable help in performing the field studies; the Mölndal city employees who collaborated in this work; the children, parents, and staff at the preschools; and Assistant Professor Jeong Lim Kim and Assistant Professor Laith Hussain-Alkhateeb for their statistical guidance. The study was funded by FAS 2008-0901 and FORMAS 2005-1153.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2021.688836/full#supplementary-material

REFERENCES

AFS 2005: 16 (2005). Noise (In Swedish) the Swedish Work Environment Authority - Code of Statue. Solna. Available at: https://www.av.se/globalassets/filer/publikationer/foreskrifter/buller-foreskrifter-afs2005-16.pdf (Accessed February 12, 2020).

Astolfi, A., Puglisi, G. E., Murgia, S., Minelli, G., Pellerey, F., Prato, A., et al. (2019). Influence of Classroom Acoustics on Noise Disturbance and Well-Being for First Graders. *Front. Psychol.* 10, 2736. doi:10.3389/fpsyg.2019.02736

Beranek, L. L. (1989). Balanced Noise-criterion (NCB) Curves. *The J. Acoust. Soc. America* 86, 650–664. doi:10.1121/1.398243

Clark, C., and Paunovic, K. (2018). WHO Environmental Noise Guidelines for the European Region: a Systematic Review on Environmental Noise and Quality of Life, Wellbeing and Mental Health. *Ijerph* 15, 2400. doi:10.3390/ijerph15112400

Dellve, L., Samuelsson, L., and Persson Waye, K. (2013). Preschool Children's Experience and Understanding of Their Soundscape. *Qual. Res. Psychol.* 10, 1–13. doi:10.1080/14780887.2011.586099

Dockrell, J. E., and Shield, B. M. (2006). Acoustical Barriers in Classrooms: the Impact of Noise on Performance in the Classroom. *Br. Educ. Res. J.* 32, 509–525. doi:10.1080/01411920600635494

Evans, G., and Hygge, S. (2007). in *Noise and Cognitive Performance in Children and Adults*. Editors Luxon, L. M., and Prasher, D. (Chichester: Hoboken, NJ: Wiley).

Evans, G. W. (2006). Child Development and the Physical Environment. *Annu. Rev. Psychol.* 57, 423–451. doi:10.1146/annurev.psych.57.102904.190057

Eysel-Gosepath, K., Pape, H. G., Erren, T., Thinschmidt, M., Lehmacher, W., and Piekarski, C. (2010). Lärm in Kindertagesstätten. *HNO* 58, 1013–1020. doi:10.1007/s00106-010-2121-y

Fels, J. (2008). From Children to Adults: How Binaural Cues and Ear Canal Impedances Grow. Phd dissertation. Logos Verlag Berlin GmbH: RWTH Aache. doi:10.1145/1400468.1400522

Gokdogan, O., and Gokdogan, C. (2016). Determination of the Level of Noise in Nurseries and Pre-schools and the Teachers′ Level of Annoyance. *Noise Health* 18, 256–259. doi:10.4103/1463-1741.192475

IEC 60268-16 (2011). *Sound System Equipment – Part 16: Objective Rating of Speech Intelligibility by Speech Transmission index*. Geneva, Switzerland: International Electrotechnical Commission.

ISO 226 2003 (2003). *Acoustics — normal Equal-Loudness-Level Contours*. Stockholm, Sweden: ISO (The International Organization for Standardization.

ISO 3382-1 (2009). *Acoustics – Measurement of Room Acoustic Parameters – Part 1: Performance Spaces*. Stockholm, Sweden: ISO (The International Organization for Standardization).

ISO 3382-2 (2008). *Acoustics — Measurement of Room Acoustic Parameters — Part 2: Reverberation Time in Ordinary Rooms*. Stockholm, Sweden: ISO (The International Organization for Standardization.

International Telecommunication Union (2015). *Recommendation ITU-R BS. 1116-3: Methods for the Subjective Assessment of Small Impairments in Audio Systems*. Geneva, Switzerland: International Telecommunication Union.

Jonsdottir, V., Rantala, L., Oskarsson, G., and Sala, E. (2015). Effects of Pedagogical Ideology on the Perceived Loudness and Noise Levels in Preschools. *Noise Health* 17, 282–293. doi:10.4103/1463-1741.165044

Klatte, M., Hellbrück, J., Seidel, J., and Leistner, P. (2010). Effects of Classroom Acoustics on Performance and Well-Being in Elementary School Children: a Field Study. *Environ. Behav.* 42, 659–692. doi:10.1177/0013916509336813

Kuhl, P. K. (2010). Brain Mechanisms in Early Language Acquisition. *Neuron* 67, 713–727. doi:10.1016/j.neuron.2010.08.038

Lane, H., and Tranel, B. (1971). The Lombard Sign and the Role of Hearing in Speech. *J. Speech Hearing Res.* 14, 677–709. doi:10.1044/jshr.1404.677

Larsson, P. (2011). *God Ljudmiljö I Förskola - Beskrivning Av Rumsakustik Före Och Efter Åtgärdsprogram*. Gothenburg: Enheten för Arbet-soch miljömedicin, Avdelningen för Samhällsmedicin och Folkhälsa, Sahlgrenska Akademin vid Göteborgs universitet. Available at: http://hdl.handle.net/2077/35768 (Accessed February 12, 2020)

Lindstrom, F., Ohlsson, A.-C., Sjöholm, J., and Persson Waye, K. (20102010). Mean F0 Values Obtained through Standard Phrase Pronunciation Compared with Values Obtained from the normal Work Environment: a Study on Teacher and Child Voices Performed in a Preschool Environment. *J. Voice* 24 (3), 319–323. doi:10.1016/j.jvoice.2008.10.006

Loh, K., and Fels, J. (2018). *Binaural Psychoacoustic Parameters Describing Noise Perception of Children in Comparison to Adults*. Hersonissos, Crete, Greece: Euronoise

Luntamo, T., Sourander, A., Santalahti, P., Aromaa, M., and Helenius, H. (2012). Prevalence Changes of Pain, Sleep Problems and Fatigue Among 8-Year-Old Children: Years 1989, 1999, and 2005*. *J. Pediatr. Psychol.* 37, 307–318. doi:10.1093/jpepsy/jsr091

Maxwell, L. E., and Evans, G. W. (2000). The Effects of Noise on Pre-school Children's Pre-reading Skills. *J. Environ. Psychol.* 20, 91–97. doi:10.1006/jevp.1999.0144

Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2010). Combined Effects of Noise and Reverberation on Speech Recognition Performance of Normal-Hearing Children and Adults. *Ear & Hearing* 31, 336–344. doi:10.1097/AUD.0b013e3181d3d514

Nordtest Method NT ACOU 115 (2003). *Measurement of Occupational Noise Exposure of Workers: Part II: Engineering Method*, 19.

Persson Waye, K., Agge, A., Lindström, F., and Hult, M. (2011). *God Ljudmiljö I Förskola - Samband Mellan Ljudmiljö, Hälsa Och Välbefinnande Före Och Efter Åtgärdsprogram*. Gothenburg: Enheten för Arbetsoch miljömedicin, Avdelningen för Samhällsmedicin och Folkhälsa, Sahlgrenska Akademin vid Göteborgs universitet. Available at: http://hdl.handle.net/2077/35765 (Accessed February 12, 2020).

Persson Waye, K., Fredriksson, S., Hussain-Alkhateeb, L., Gustafsson, J., and van Kamp, I. (2019). Preschool Teachers' Perspective on How High Noise Levels at Preschool Affect Children's Behavior. *PLoS ONE* 14, e0214464. doi:10.1371/journal.pone.0214464

Persson Waye, K., Van Kamp, I., and Dellve, L. (2013). Validation of a Questionnaire Measuring Preschool Children's Reactions to and Coping with Noise in a Repeated Measurement Design. *BMJ Open* 3, e002408. doi:10.1136/bmjopen-2012-002408

Pickett, J. M. (1959). Low-Frequency Noise and Methods for Calculating Speech Intelligibility. *J. Acoust. Soc. America* 31 (9), 1259–1263. doi:10.1121/1.1907855

Pirilä, S., Jokitulppo, J., Niemitalo-Haapola, E., Yliherva, A., and Rantala, L. (2020). Teachers' and Children's Experiences after an Acoustic Intervention and a Noise-Controlling Workshop in Two Elementary Classrooms. *Folia Phoniatr Logop* 72, 454–463. doi:10.1159/000503231

Ricour, C. (1989). Stress-induced Disturbances of the Gastro-Intestinal Tract in Children. *Intensive Care Med.* 15, S32–S36. doi:10.1007/BF00260881

Silber, T. J., and Pao, M. (2003). Somatization Disorders in Children and Adolescents. *Pediatr. Rev.* 24, 255–264. doi:10.1542/pir.24-8-255

Sjödin, F., Kjellberg, A., Knutsson, A., Landström, U., and Lindberg, L. (2012). Noise Exposure and Auditory Effects on Preschool Personnel. *Noise Health* 14, 72. doi:10.4103/1463-1741.95135

SS 25268 2007+T1 (2017). *Acoustics – Sound Classification of Spaces in Buildings – Institutional Premises, Rooms for Education, Preschools and Leisure-Time Centres, Rooms for Office Work and Hotels*. Stockholm, Sweden: SIS – Swedish Standards Institute. . Available at: www.sis.se (Accessed January 15, 2020).

Sussman, E. S., and Steinschneider, M. (2011). Attention Modifies Sound Level Detection in Young Children. *Dev. Cogn. Neurosci.* 1, 351–360. doi:10.1016/j.dcn.2011.01.003

Tamburlini, G., Ehrenstein, O. S. V., and Bertollini, R. (2002). *Children's Health and Environment: A Review of Evidence*. Copenhagen/Rome: European Environment Agency.

Vanaelst, B., De Vriendt, T., Ahrens, W., Bammann, K., Hadjigeorgiou, C., Konstabel, K., et al. (2012). Prevalence of Psychosomatic and Emotional Symptoms in European School-Aged Children and its Relationship with Childhood Adversities: Results from the IDEFICS Study. *Eur. Child. Adolesc. Psychiatry* 21, 253–265. doi:10.1007/s00787-012-0258-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 09 July 2021
doi: 10.3389/fbuil.2021.687087

Check for updates

The Effect of Background Noise on a "Studying for an Exam" Task in an Open-Plan Study Environment: A Laboratory Study

Ella Braat-Eggen[1]*, *Jikke Reinten*[2], *Maarten Hornikx*[3] *and Armin Kohlrausch*[4]

[1]School of the Built Environment and Infrastructure, Avans University of Applied Sciences, Tilburg, Netherlands, [2]TNO, Soesterberg, Netherlands, [3]Department of the Built Environment, Eindhoven University of Technology, Eindhoven, Netherlands, [4]Human-Technology Interaction, Eindhoven University of Technology, Eindhoven, Netherlands

Students can be disturbed by background noise while working in an open-plan study environment. To improve the acoustic quality of open-plan study environments a study was done on the influence of different sound scenarios on students working on a typical student task, "studying for an exam". Three sound scenarios and a quiet reference sound scenario were developed, based on the sound environment of a real open-plan study environment, with a varying number of talkers in the background and different reverberation times of the study environment. Seventy students worked on a set of tasks simulating a "studying for an exam" task while being exposed to the sound scenarios. This task comprises a reading comprehension task with text memory by delayed answering questions about the text, with additional tasks being performed in the gap between studying the text and retrieving. These additional tasks are a mental arithmetic task and a logical reasoning task. Performance, self-estimated performance and disturbance of students were measured. No significant effect of the sound scenarios was found on performance of students working on the reading comprehension task with text memory and the mental arithmetic task. However, a significant effect of sound was found on performance of students working on the logical reasoning task. Furthermore, a significant effect of the sound scenarios was found on self-estimated performance and perceived disturbance for all tasks from which the reading comprehension task with text memory was the most disturbed task. It is argued that the absence of a detrimental sound effect on the performance of students working on a reading comprehension task with text memory is a result of focusing due to task engagement and task difficulty, both aspects working as a "shield against distraction".

Keywords: open-plan study environment, student task, background speech, task performance, noise disturbance, well-being, acoustic quality

OPEN ACCESS

Edited by:
Chiara Visentin,
University of Ferrara, Italy

Reviewed by:
Ellen Peng,
University of Wisconsin-Madison, United States
Raffaele Drgonetti,
University of Naples Federico II, Italy
Isabel Schiller,
RWTH Aachen University, Germany

***Correspondence:**
Ella Braat-Eggen
pe.braat-eggen@avans.nl

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *28 March 2021*
Accepted: *17 June 2021*
Published: *09 July 2021*

Citation:
Braat-Eggen E, Reinten J, Hornikx M and Kohlrausch A (2021) The Effect of Background Noise on a "Studying for an Exam" Task in an Open-Plan Study Environment: A Laboratory Study. Front. Built Environ. 7:687087. doi: 10.3389/fbuil.2021.687087

INTRODUCTION

Open-plan study environments (OPSEs) are becoming increasingly important in higher education. Not only the importance of their function but also the number of square meters is increasing (Montgomery, 2014; Beckers et al., 2015). The need for OPSEs is a result of changed visions on education and enables new ways of learning. In addition to education that is primarily aimed at knowledge transfer, education that focuses more on competencies is becoming increasingly important (Beckers et al., 2015; Koenen et al., 2015). This new type of education, in which skills and attitude of students are of great importance in addition to knowledge, have led to different work forms with corresponding assessment procedures (Koenen et al., 2015). Besides the well-known individual written and oral exams, the assessment of competences is often based on the outcome of individual or group assignments or projects (Koenen et al., 2015; Curry and Docherty, 2017). As a result of these educational changes, there is a need for workspaces where students can work on their assignments and projects, individually but also in groups. Accordingly, not only classrooms and lecture halls, but also OPSEs become part of buildings for higher education.

A survey on students tasks, perceived sound sources and noise disturbance among 496 students in five OPSEs showed that the tasks students perform in OPSEs are diverse, ranging from preparing for an individual exam to brainstorming for a group assignment (Braat-Eggen et al., 2017). Furthermore, the survey also showed that students are mostly bothered by noise when performing an individual complex cognitive task like studying for an exam, reading or writing. The variety of activities in the same OPSE implicates different demands on the acoustic environment. Disturbance can also occur because some students perform group tasks that will induce noise, while other students perform individual complex cognitive tasks.

Although the sound environment in OPSEs can be very disturbing, no recommendations or guidelines have so far been developed for the design of acoustically comfortable OPSEs. To do so, more knowledge is needed on tasks and the sound environment in an OPSE in relation to task performance and disturbance.

THEORETICAL BACKGROUND

Studying for an Exam in Higher Education

As showed in a study on noise in OPSEs, (Braat-Eggen et al., 2017) the most disturbed task students perform in an OPSE is "studying for exams". This task will be further investigated in this study. As far as we know, earlier research into the influence of different sound environments on a studying task has not yet been carried out in the context of performance and disturbance. Most studies on the influence of noise on cognitive performance were executed to find specific mechanisms responsible for distraction of a cognitive task. Therefore, these experiments are mostly performed on so-called "pure" cognitive tasks (Sörqvist, 2014) or sub-component cognitive abilities (Sörqvist, 2015), such as for instance short-term memory tasks (Haapakangas et al., 2011; Schlittmeier et al., 2011; Hughes, 2014) or tasks using retrieval from semantic memory (Jahncke, 2012; Jones et al., 2012). The use of experimental "sub-component ability" results may be complementary but not enough for understanding the effects of noise on a realistic complex cognitive task (Sörqvist, 2015). Therefore, in this research on OPSEs we will study the influence of noise on complex student tasks. It will be instrumental for developing recommendations for acoustically comfortable OPSEs.

Preparing for an examination is a typical student task and it is a very complex task. When students are learning for an exam, they have to analyse and understand the material. Moreover, they also have to make strategic choices and decide what to learn and to store in memory. Studies on participants performing self-regulated learning tasks are mostly performed in a quiet laboratory setting (Dunlosky and Ariel, 2011a; Dunlosky and Ariel, 2011b). In these studies, not only memorizing but also learning strategies are the subject of the research questions. In a recent study on self-regulated learning, the influence of noise as an environmental factor has been studied in relation to the strategic and metacognitive aspects of learning (Hanczakowski et al., 2018). The duration of the study time was related to the auditory distraction in the environment. The strategic choices of the participants were measured by how much time the participants had spent on various study items. It appeared that the duration of study time was not extended when the participants were disturbed by the noise during the study process, while it was expected that the participants would invest more study time when they were disturbed by the noise. Due to the lack of compensatory strategies, such as extending the study time, a decrease of performance was found. The researchers explained this as a distortion of time perception by auditory distraction (Hanczakowski et al., 2018).

Assessments in higher education are an essential part of a curriculum and evaluate the educational level of graduate students (Flores et al., 2015). There is a wide variety in ways to organize an exam. However, there are some basic characteristics of an exam in higher education. Exams at this educational level must include higher-order thinking skills and encourage conceptual understanding (Jensen et al., 2014). A model to describe different levels of cognitive skills has been developed by Bloom (Bloom and Krathwohl, 1956; Adams, 2015). His model describes six cognitive categories with increasing complexity: knowledge, comprehension, application, analysis, synthesis and evaluation. A revised version of his taxonomy changed the categories into more skill-based levels: remember, understand, apply, analyze, evaluate and create (Anderson and Kratwohl, 2001). In practice, it means that when students in higher education prepare for an exam they do not only have to remember and understand knowledge but also have to be able to apply, analyze, and evaluate that knowledge. "Creating"', the top of Bloom's pyramid, is the most complex cognitive skill and is often tested in (multidisciplinary) projects.

The Sound Environment

A study of five OPSEs (Braat-Eggen et al., 2017) showed various sound sources (e.g., unintelligible speech, walking sounds, noise of devices, telephones ringing) of which intelligible background speech was perceived as the most disturbing. Background noise and especially background speech has been proven to have a detrimental effect on cognitive performance (Szalma and Hancock, 2011; Klatte et al., 2013; Reinten et al., 2017). These results have been described by the duplex-mechanism account (Hughes, 2014). In this account, two ways of disruption have been distinguished; interference-by-process and attentional capture. The first mechanism, interference-by-process, arises if the processes needed to perform an intended task are similar to those needed to process background sound. For instance, the processes needed for a semantic task like reading a text will interfere with the unintended processing of background speech, which is a semantic task as well. The second mechanism of distraction is attentional capture, whereby sound causes disruption of cognitive performance when it removes the focus from the intended task. Specific attentional capture occurs when the content of the sound distracts you from the core task, like for instance hearing your own name (Conway, 2001). Another way of attentional capture is that a specific sound captures attention, due to the context in which it occurs (Hughes, 2014). For instance, the B within the sequence AAAAABAA will capture attention due to the deviation from the expected A (Hughes et al., 2005; Hughes et al., 2007). Auditory distraction can be overruled by cognitive control (Clark and Sörqvist, 2012). For instance, an increased task demand, a more difficult task or a greater engagement into the task can shield against distracting effects of noise on tasks, but if the task load is too heavy, it can also lead to abandonment (Engelmann et al., 2009; Halin et al., 2014a; Halin et al., 2014b; Hughes, 2014; Marsh et al., 2015). Furthermore, it should be acknowledged that even if students are able to shield against noise in terms of performance, they might require longer processing time, as has been shown in school aged children (age 6–7; 11–13) (Prodi et al., 2019; Schiller et al., 2020).

Generalization of the results of experimental studies on the influence of noise on task performance and disturbance into room acoustic requirements is difficult. A translation is only possible if the experimental sound environments are comparable with the real sound environment in which the task is expected to be performed. In a literature review on the influence of the indoor sound environment on human task performance (Reinten et al., 2017) it was found that only a limited number of studies made use of realistic variations of the room acoustic parameters in combination with realistic sound sources. The influence of room acoustic parameters is seldomly taken into account in experiments, and in many cases background speech consists of only one or two talkers which is an interesting disturbing sound environment (Keus van de Poll et al., 2014) but not the most representative setting for an OPSE.

Personal Factors

Different personal factors can influence the effect of noise on cognitive performance (Reinten et al., 2017). An important personal factor that can influence task performance and disturbance of people in noisy open environments is noise sensitivity (Haapakangas et al., 2014). In earlier studies on the influence of the sound environment of OPSEs on cognitive performance and disturbance, noise sensitivity was taken into account. In a field study on OPSEs, it was shown that students with a noise sensitivity score above the median score were more disturbed by noise than students with a noise sensitivity score below the median score (Braat-Eggen et al., 2017). In the experimental study on a collaboration task (together solving a problem) in an OPSE no influence of noise sensitivity was found (Braat-Eggen et al., 2019a), while in the experimental study on a writing task students with a noise sensitivity score above the median score showed to be more influenced by the sound environment resulting in a significantly lower writing performance in comparison to students less sensitive to noise (Braat-Eggen et al., 2019b). As some of the studies show an important influence of noise sensitivity of students on performance and disturbance in an OPSE, we will include noise sensitivity of students, measured by a well-tested questionnaire (Griefahn, 2008), as a personal factor also in this study.

The Aim of the Study

In this laboratory experiment, the influence of background speech on the performance and disturbance on a typical student task, "studying for an exam" in higher education, will be investigated by using a reading comprehension, logical reasoning, and mental arithmetic assignment. With regard to the importance of developing recommendations, this study will work with a variation in acoustical properties and different realistic sound sources in an acoustically simulated OPSE.

Based on the duplex-mechanism account, we hypothesize that a realistic sound environment with background speech will have a negative effect on performance and perceived disturbance while performing the "studying for an exam" task in an OPSE in comparison to a quiet environment. "Studying for an exam" has many sub-components as mentioned earlier, but based on the semantic elements within the task we expect that more intelligible background speech will reduce performance and will increase disturbance of students measured by a questionnaire (ISO/TS 15666, 2003). Also, the noise sensitivity of students is expected to affect how they perceive the disturbance of the background speech. We expect noise sensitive students to be more disturbed by the background speech and to perform less due to the background sound in comparison to less noise sensitive students.

MATERIALS AND METHODS

Design

To verify the hypotheses posed in Section *The aim of the study* a within-participants experimental design with repeated measurements was developed. The experiment included three tasks: a reading comprehension task, a logical reasoning task, and an arithmetic task, together representing a "studying for an exam" task. Four different sound scenarios with background speech were used in the experiment. Students had to perform each task four times, each time a different sound scenario was presented.

Participants

Seventy bachelor students from Avans University of Applied Sciences took part in the experiment. The results of four students were not included in the analysis. One of these students had severe hearing loss, the results of two other students were excluded due to computer problems during the test and the experiment of one student was interrupted by his mobile phone. All participating students were native Dutch speakers. The sixty-six students (24 female and 42 male) included in the analysis were between 17 and 30 years old (mean age = 20.2, SD = 2.7). As a reward for their participation, the students received an internet voucher or educational credits.

Research Settings

The experiments were conducted in a small two-person office (2.60 m × 2.25 m) with no windows, originally intended for audio processing. The walls were covered with acoustic absorbing material and the room was acoustically well insulated. During the experiment the participant was sitting at one desk while the researcher was sitting at the other desk, next to each other. The participant was working on a laptop with external sound card (ST Lab USB sound box) and was wearing headphones (Sennheiser HD 380 PRO) throughout the experiment.

Sound Conditions

To create realistic OPSE background sound scenarios, auralizations based on computed impulse responses were used. Therefore, a digital model of an existing OPSE at the Eindhoven University of Technology was constructed. The computational modeling and auralization was performed using the room acoustic modeling software Odeon (version 12.12). From this basic model two new models were developed, an sound absorbing model with a reverberation time of 0.6 s applying sound absorbing materials instead of the materials used in the real OPSE, and a reverberant model with a reverberation time of 2.4 s applying sound reflecting materials. These two models had also been used in the previous studies on the influence of background sound on student tasks (Braat-Eggen et al., 2019a; Braat-Eggen et al., 2019b).

Four sound scenarios were created for this experiment (**Table 1**), one quiet reference scenario and three scenarios with background speech. Not only the material properties of the OPSE but also the number of talkers in the OPSE were varied. The number of talkers in combination with the reverberation time in the modeled OPSEs resulted in sound scenarios with different levels of intelligibility of the background speech. In **Table 1** the four sound scenarios are described by the reverberation time, background sound level due to speech and the intelligibility of the background speech (Braat-Eggen et al., 2019b). The intelligibility is here based on the nearest speaker and is described by the estimated Speech Transmission Index (*STI*). *STI* is a dimensionless number between zero and one, where an excellent intelligibility results in an *STI* value of 1, and an *STI* value below 0.3 indicates almost unintelligible speech (Houtgast et al., 1980). The position of the talkers and their speech directions are described in **Figure 1**. More information about the modeling, materials, sound levels, and estimated *STI* values has been included in earlier research on the influence of background speech on a writing task (Braat-Eggen et al., 2019b) In this study the same OPSE models were used as in the current study. These models were used to research the influence of the sound environment of a (simulated) OPSE, varied by the number of background talkers and reverberation time, on performance and disturbance of students carrying out a writing task.

To create a realistic sound environment, recordings were made of students talking about their study, hobbies and work. Subsequently, the speech recordings were convolved with the binaural impulse responses using HRTFs (stereo effect) of the absorbing and reverberant model as calculated by Odeon. The quiet control sound condition consisted of a pink noise signal at 30 dB(A), which is equal to the background noise level in the existing, unoccupied OPSE (Braat-Eggen et al., 2019b). The sound pressure levels offered to the subjects by headphones were calibrated in accordance with the calculated sound pressure levels in the models (**Table 1**). The calibration was performed with a Head and Torso simulator (B&K 4128-C).

Measures

Task Performance

The typical student task "studying for an exam" was simulated by a series of assignments. The examination format chosen for this experiment was an individual written examination, a common format for examining knowledge in higher education (Curry and Docherty, 2017). One of the characteristics of this format is the time gap between the studying activity, that could take place in an OPSE, and the testing of the knowledge. To simulate the time gap in the experiment, after the study activity and before testing, two other assignments were introduced to the participants, a logical reasoning task and a mental arithmetic task. Performing these tasks not only simulates a time gap, but also what happens in real life: within the time span between studying for an exam and performing an exam, students are busy performing all kinds of tasks that take their focus away from the exam topic. The tasks which were chosen to fill in the time gap rely on cognitive skills that complement the study task in order to cover the cognitive skills described in Bloom's model. The combination of the three assignments used in the experiment represents five out of six levels of cognitive skills as described in Bloom's revised taxonomy (Anderson and Kratwohl, 2001):

- remembering: reading comprehension with delayed retrieval, mental arithmetic
- understanding: reading comprehension, logical reasoning, mental arithmetic
- applying: mental arithmetic
- analyzing: reading comprehension, logical reasoning, mental arithmetic
- evaluating: reading comprehension, logical reasoning

The highest level of cognitive tasks in Bloom's taxonomy, "creating", was not included in the assignments, to reduce the duration and complexity of the experiment. Each assignment was designed to represent the level of a beginning bachelor student. In this experiment the performance and disturbance of all three tasks, the studying task (reading comprehension task) and the tasks to simulate the time gap (the logical reasoning and mental arithmetic task) were analyzed.

Reading Comprehension With Delayed Answering

The "studying for an exam" task shows resemblance to a comprehensive reading test. At the start of the task, students were instructed to study an informative text, as if they were

TABLE 1 | Characteristics of the background sound scenarios.

Sound	Reverberation (time)	Background sound	Sound level L_{Aeq} background sound	Estimated *STI* values
Scenario				
A&3T	Absorbing (T_{30} = 0.6 s)	3 Talkers	41 dB(A)	0.62
A&14T	Absorbing (T_{30} = 0.6 s)	14 Talkers	54 dB(A)	0.38
R&14T	Reverberant (T_{30} = 2.3 s)	14 Talkers	64 dB(A)	0.18
Quiet	—	Pink noise	30 dB(A)	—

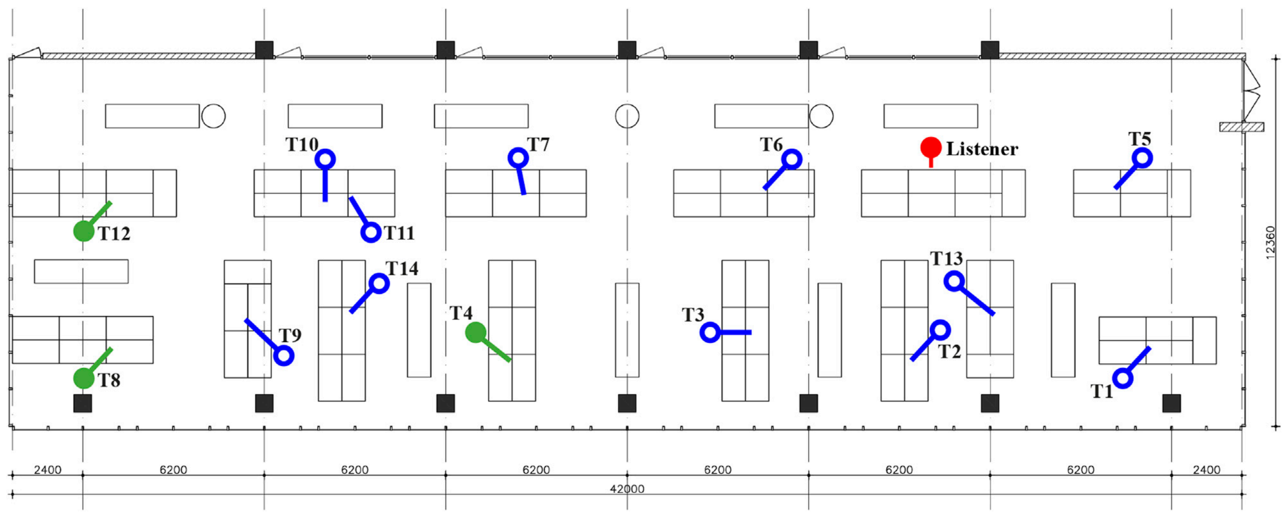


FIGURE 1 | Floor plan of the modeled OPSE with the positions of listener and three talkers (T4, T8, T12), and fourteen talkers (T1-T14), with lines indicating the listening and talking direction. Dimensions in mm (Braat-Eggen et al., 2019b).

preparing for an exam about the content of that text, that would be conducted later in the experiment. Four texts with the same length (mean = 645 words) and a similar level of complexity were selected. To this end, texts from 'The State Exams Dutch as a second Language (NT2)' were chosen. These texts are normally used for the national language proficiency exams for non-native adult speakers, who want to start a study at a Dutch University or want to work in the Netherlands. To study the influence of different background sound scenarios on a task in a repeated measurement design, it is very important to select four texts of the same difficulty level. Therefore, a pilot study was performed (n = 8) and from the analysis of the results the final four texts were selected.

The performance of the reading comprehension task was measured by the number of correct answers to the questions about the text, the exam. In total 10 multiple choice questions were formulated for each text. The students answered the questions after a time interval of 8 min. In these 8 min the students worked on two assignments, a logical reasoning task and a mental arithmetic task. These 'in-between' tasks were intended to simulate the time gap between studying and doing an exam.

Logical Reasoning

The logical reasoning task consisted of a set of so-called syllogisms. Students had to read two statements, subsequently they had to judge conclusions drawn from these two statements on validity. For example:

Statements:

- All mountains have rocks
- All countries have mountains

Conclusions:

1. All rocks have countries
2. All countries have rocks
3. Not all rocks have countries
4. No conclusion possible

A well-tested set of 40 syllogisms, developed by (Making Moves, 2019), was used. Each student had to solve ten syllogisms in all four different sound environments. The performance of the logical reasoning test was measured by the number of correct answers.

Mental Arithmetic

In the mental arithmetic test the students had to solve 18 calculations without the use of paper and pen or calculator. The calculations were examinations at a first-year bachelor educational level, in the Netherlands defined as level 3F (Citrus, 2018). The performance of the mental arithmetic test was measured by the number of correct answers.

Self-Estimated Performance and Perceived Disturbance

The self-estimated performance and perceived disturbance of tasks were measured by a questionnaire, on a 5-point scale, after each sound scenario (**Figure 2**). The questions were based on ISO/TS 15666 "Acoustics - Assessment of noise annoyance by means of social and social-acoustic surveys" and formulated in the Dutch language (ISO/TS 15666, 2003). Question 1, 3, and five measured noise disturbance when students were performing the three tests and question 2,4 and six measured the impact of noise on performance estimated by the students after performing the three tests.

1. Thinking about the last experiment, how much did noise bother, disturb or annoy you while studying the text: not at all - slightly—moderately—very—extremely?

2. Thinking about the last experiment, how much did the noise influence the number of correct answers on the questions about the text: not at all—slightly—moderately—very—extremely?

3. Thinking about the last experiment, how much did noise bother, disturb or annoy you while working on the logical reasoning statements: not at all—slightly—moderately—very—extremely?

4. Thinking about the last experiment, how much did the noise influence the number of correct answers on the logical reasoning statements: not at all—slightly—moderately—very—extremely?

5. Thinking about the last experiment, how much did noise bother, disturb or annoy you while working on the calculations: not at all—slightly—moderately—very—extremely?

6. Thinking about the last experiment, how much did the noise influence the number of correct answers on the calculations: not at all—slightly—moderately—very—extremely?

Noise Sensitivity

The noise sensitivity of the students was measured with the reduced version of the Noise Sensitivity Questionnaire (NoiSeQ-R), developed by Griefahn (Griefahn, 2008). The questionnaire was translated and offered in the Dutch language to the students. They had to indicate their agreement on twelve statements related to their sensitivity to noise. For each statement the level of agreement could be chosen on a 4-points scale: "disagree completely—slightly disagree—slightly agree—agree completely".

Procedure

The whole experiment took about 2 h and 30 min spread over two sessions (**Figure 2**). The first session started with an instruction by the experimental researcher, followed by a set of assignments to practice the type of questions and to get familiar with the procedure (**Figure 2**). After practicing, the first set of assignments was presented to the student while being exposed to one of the sound scenarios. After finishing the first set, a short break of 10 min was programmed before starting the second set of assignments. This set was presented to the student with another background sound scenario. This first session took about 80 min.

The second session took place on another day. The time interval between the two test sessions varied, the average was 7 days. During the second session each student worked on two new sets of assignments while being exposed to two different sound scenarios. Between the sets of assignments, a short break of 10 min was prescribed. At the end of the session the student had to fill in a questionnaire about noise sensitivity and personal factors like age, gender, and hearing. The second session took about 70 min.

Students worked individually on the experiment. All instructions about the assignments were displayed on the laptop and "start" and "stop" instructions were given orally through the headphone. The background sound conditions were offered through the headphones during both the study task and the assignments but not during answering the questions about the text.

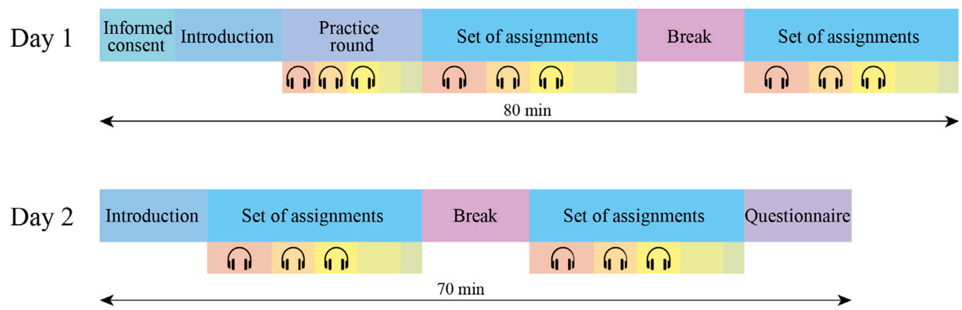
The set of assignments simulating the "studying for an exam" task started with reading and studying a text. The participating students were informed that they had to answer some questions about the text later in the experiment. The text was printed on paper and the use of pen and marker was allowed during their study activity. After 6 min, participants had to put the text, including all their notes, in a closed box. This task was followed by the logical reasoning task, assignments (syllogisms) were presented at the laptop screen. After 4 min the last task started, the mental arithmetic task. While working on the calculation exercises, making notes and using a calculator were forbidden. After 4 min this task was closed and the questions about the initial text were presented. Students had 4 min to answer the questions about the initial text. Finally, a questionnaire was presented about the perception of the background sound and the self-estimated influence of the sound scenario on performance. In total a set of tests (including the perception questionnaire) took 20 min (6 + 4 + 4 + 4 + 2 min), the practice set of tests took 13 min (3 + 3 + 3 + 2 + 2 min). An overview of order and duration of the assignments can be seen in **Figure 2**.

All tasks were announced on the laptop screen and after pushing the start button the time clock and assignments were started on the laptop. The elapsed time was shown on the screen, so the students knew how much time there was left to perform their task. The assignments were presented in the same sequence to all participants. The four sound scenarios were offered to the participants in a counter-balanced sequence.

Statistical Analysis

All statistical analyses were performed using SPSS 23.0. The influence of the background sound scenarios on the performance, self-estimated performance and perceived disturbance was analyzed by a single-factor repeated measures ANOVA. The significance of the differences between the means of the dependent variables due to the four sound scenarios was tested and a follow-up pairwise comparison to examine where the differences occurred was performed by using post-hoc t-tests with a Bonferroni correction.

The influence of noise sensitivity was studied after a median split was done to divide the subjects in two groups: a low noise sensitive group (below the median score) and a high noise sensitive group (above the median split). By using a factorial 4 (four sound scenarios) x 2 (low vs. high noise sensitivity) repeated



Assignment	practice [min]	test [min]
Studying text	3	6
Logical reasoning	3	4
Mental arithmetic	3	4
Questions text	2	4
Perception questionnaire	2	2

FIGURE 2 | The order and duration of a set of assignments in the experiment.

measures ANOVA, the influence of the noise sensitivity on performance, self-estimated performance and perceived disturbance was studied.

RESULTS

Impact of the Background Sound Scenario on Performance

Table 2 shows the influence of the different background sound scenarios on performance of students accomplishing a reading comprehension task, a logical reasoning task and a mental arithmetic task. The performance has been determined by the number of correctly answered questions for the assignments.

The analyses show that different sound scenarios do not have a significant effect on performance of a "reading comprehension" task ($p = 0.142$), and neither on the performance of a mental arithmetic task ($p = 0.934$). The analyses also show that different sound scenarios have a significant effect on performance of a logical reasoning task ($p = 0.013$). The background sound scenarios with speech lead to a decrease of performance. Follow-up t-tests with Bonferroni correction showed significant differences between the performance means for the quiet situation and the reverberant sound scenario with 14 talkers ($p = 0.008$). A 11% decrease in performance of the logical reasoning task is measured between the "reverberant 14 talkers" sound scenario and the "quiet" sound scenario. A performance reduction of an average of 7% is measured if all three sound scenarios are compared with the "quiet" sound scenario.

Impact of the Background Sound Scenario on Self-Estimated Performance

Figure 3 shows the influence of the different background sound scenarios on the self-estimated performance of students accomplishing the three tasks. The self-estimated performance was measured on a 5-point scale for each task. Scale value one indicated that students estimated their performance not at all to be influenced by background noise, while scale value five indicated that students estimated their performance to be extremely influenced by the background noise.

The analyses show the different sound scenarios to have a significant effect on the self-estimated performance of the students working on a reading comprehension task ($F(3,195) = 34.129$, $p < 0.0001$, $\eta_p^2 = 0.344$), a logical reasoning task ($F(3,189) = 38.468$, $p < 0.0001$, $\eta_p^2 = 0.379$), and a mental arithmetic task ($F(3,189) = 26.953$, $p < 0.0001$, $\eta_p^2 = 0.300$). The quiet condition was reported as the least influencing condition. Follow-up t-tests with Bonferroni adjustment for all tasks showed significant differences between the self-estimated performance means for the quiet condition and the three other sound scenarios ($p < 0.0001$).

Self-estimated performance of the mental arithmetic task seems the least influenced by the background sounds (**Figure 3**). A one-way repeated measures ANOVA shows that for each sound scenario the kind of task has no significant effect on self-estimated performance ($p > 0.05$).

Impact of the Background Sound Scenario on Perceived Disturbance

Figure 4 shows the influence of the different background sound scenarios on perceived disturbance of students working on the three tasks. The perceived disturbance has been measured on a 5-point scale for each task. Scale value one indicated that students felt not at all to be disturbed by background noise, while scale five indicated that students felt extremely disturbed by the background noise.

The analyses show different sound scenarios to have a significant effect on perceived disturbance of a study task (F

TABLE 2 | Mean number of correct answers as a performance measure of different tasks while exposed to different sound scenarios.

Task	Background sound scenario				F (3,192)	η_p^2
	Quiet condition	3 Talkers absorbing	14 Talkers absorbing	14 Talkers reverberant		
Reading comprehension	7.02	6.63	6.40	6.77	1.837	0.027
Logical reasoning	7.51	7.31	6.97	6.66	3.713[a]	0.055
Mental arithmetic	7.47	7.29	7.39	7.44	0.143	0.002

[a]*p < 0.05.*

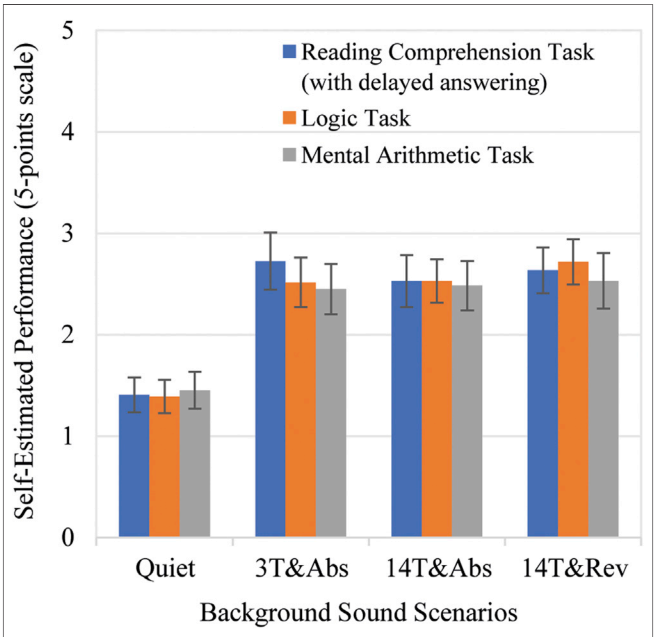


FIGURE 3 | Mean values and confidence intervals (95%) of the self-estimated performance of participants (*n* = 66) accomplishing three tasks with four different sound scenarios: Quiet, three Talkers and Absorbing (3T&Abs), 14 Talkers and Absorbing (14T&Abs), 14 Talkers and Reverberant (14T&Rev).

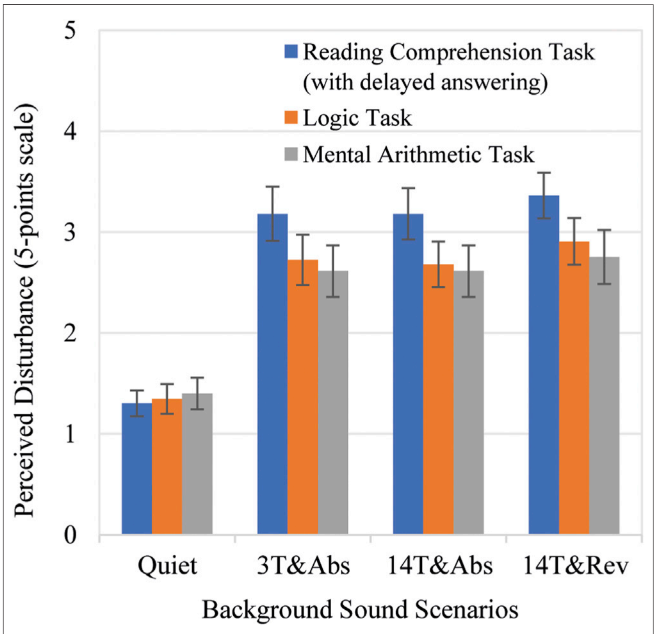


FIGURE 4 | Mean values and confidence intervals (95%) of the perceived disturbance of participants (*n* = 66) accomplishing three tasks with four different sound scenarios: Quiet, three Talkers and Absorbing (3T&Abs), 14 Talkers and Absorbing (14T&Abs), 14 Talkers and Reverberant (14T&Rev).

(3,195) = 94.280, $p < 0.0001$, $\eta_p^2 = 0.592$), a logical reasoning task (F (3,195) = 59.285, $p < 0.0001$, $\eta_p^2 = 0.477$) and a mental arithmetic task (F (3,192) = 44.976, $p < 0.0001$, $\eta_p^2 = 0.413$). The quiet condition was reported as the least disturbed sound condition. Follow-up t-tests with Bonferroni adjustment for all tasks showed significant differences between the perceived disturbance means for the quiet situation and all other sound scenarios ($p < 0.0001$).

Students rated the reading comprehension task with delayed answering as the most disturbed task due to the background sound scenarios (**Figure 4**). A one-way repeated measures ANOVA for all sound scenarios with speech (not the quiet scenario) shows that students are significantly more disturbed when performing a study task in comparison to the other tasks (3 Talkers Absorbing: F (2,130) = 13.389, $p < 0.0001$, $\eta_p^2 = 0.171$); 14 Talkers Absorbing: F (2,130) = 12.772, $p < 0.0001$, $\eta_p^2 = 0.164$); 14 Talkers Reverberant: F (2,130) = 11.353, $p < 0.0001$, $\eta_p^2 = 0.151$).

Impact of Noise Sensitivity of Participants on Task Performance and Disturbance

To verify the influence of noise sensitivity of participants on the three output measures, a general linear model with repeated measurements was used with sound scenarios as within-subject factor and noise sensitivity as between-subject factor. The participants were divided in two groups by a median noise sensitivity split. A low sound sensitivity group (mean = 2.51, *n* = 32 participants) was formed by participants with a noise sensitivity lower than the median (median = 2.83, scale1-4), and a high noise sensitivity group (mean = 3.21, *n* = 38 participants) was formed by participants with a noise sensitivity higher than the median.

No significant interaction effect was found for any of the independent variables (performance, self-estimated performance and perceived disturbance) for any of the tasks (reading comprehension, logical reasoning, mental arithmetic).

DISCUSSION

Impact of the Background Sound Scenarios on Performance

The analysis of the results (**Table 2**) showed no significant effect of the sound scenarios on the performance of students for the "reading comprehension" or "mental arithmetic" task. Only the performance of students working on the "logical reasoning" task was significantly impaired by the background sound scenarios with speech. Although for all tasks the quiet condition showed the highest student performance, only the "logical reasoning" task showed a significant reduction in performance due to background speech and reverberation. (**Table 1**).

"Studying for an exam" tasks in higher education have, for as far as we know, not been studied in an experimental setting until now. For comparison with previously conducted studies, experimental research into reading comprehension with delayed answers would be the best approach. A reading comprehension test with delayed answers by Martin *et al.* (Martin et al., 1998) indeed showed a similar procedure as the present study. The findings of this research showed a detrimental effect of unattended speech on comprehensive reading and the importance of semantic characteristics of speech. Also, a study of Oswald *et al.* (Oswald et al., 2000) on comprehensive reading showed that meaningful as well as meaningless speech decreased performance, although the procedure of this study was less comparable with the current study. Results of both studies are not in line with our results, as we could not establish significant effects of noise on performance. An essential difference between the previous studies and this study can be found in the characteristics of the sound environments. In the compared studies (Martin et al., 1998; Oswald et al., 2000), one voice with perfect intelligible speech was used as background noise, in contrast to the sound scenarios in the current study where a realistic OPSE sound environment was simulated with at least three voices. This might be an explanation for the differences between the results of the studies. Another important difference is the design of the experiments. In the current study, the comprehensive reading test with delayed answers has been presented as an exam, combined with several other tests. The importance and the difficulty of an exam might have affected the performance of the test.

Research on the influence of noise on a one-digit "mental arithmetic" task (Banbury and Berry, 1998) and on different "mental arithmetic" tasks (Caviola et al., 2021) showed a decrease of performance for noise with and without background speech. Also, a study of Jahncke (Jahncke, 2012) on a three-digit 'mental arithmetic' task showed a decrease of performance, although relatively low in comparison to other office tasks (less than 3%). Both studies showed that the performance in a mental arithmetic task was not determined by the intelligibility of the background speech. In the present study no significant effect of the sound scenarios on performance of the mental arithmetic task was found, and certainly no influence of the intelligibility of the background speech. The realistic three-digit calculation task of Jahncke showed a good similarity with the test and results of the present study. The small effects on performance are in line with the research of Jahncke (Jahncke, 2012) and in combination with the realistic sound scenarios used in this experiment, the effect size of the current study was probably too small to measure.

Impact of the Background Sound Scenario on Self-Estimated Performance and Perceived Disturbance

The subjective parameters, self-estimated performance and perceived disturbance (**Figures 3**, **4**) showed for all tasks to be significantly affected by background speech. Students expected the quiet sound scenario to have the least detrimental effect on their performance. We expected the most intelligible background sound scenario (3 talkers-absorbing) to be estimated as the most detrimental for self-estimated performance, but this was not supported by the results. The results of the self-estimated performance of the students was not in line with our hypothesis based on the 'interference of processes' theory of the DMAAD account (Hughes, 2014).

The analysis of the perceived disturbance of the participants during the different tasks showed major similarities with the self-estimated performance results. The least disturbance was experienced during the quiet sound environment, and the most intelligible sound scenario (3 talkers-absorbing) was not identified as the most disturbing. However, it is remarkable that when comparing the tasks among themselves, the participants were significantly more disturbed by the background noise when performing the task 'reading comprehension' compared to the performance of the other tasks (**Figure 4**). This is even more remarkable when one takes into account that the decrease in performance of the task 'reading comprehension' certainly did not the show the greatest decrease compared to the other tasks. A mean decrease of performance of students due to the background noise in comparison to the quiet environment was 5.9% for the "reading comprehension" task, 1.3% for the "mental arithmetic" task and 7.1% for the "logical reasoning" task. The major disturbance of the "reading comprehension" task with delayed answering is in accordance with the findings in a field study on five OPSEs (Braat-Eggen et al., 2017). In that research "studying for an exam" was identified by students as the most disturbed task by noise they perform in an OPSE.

"Studying for an exam" is a very important task for a student, as the odds for passing an examination depend on the quality of the studying phase. Therefore, it could be expected that the task engagement for "studying for an exam" is very high. Furthermore, an exam in higher education is a complex task that requires higher order thinking skills (Jensen et al., 2014), and therefore is a very difficult task. Both aspects, engagement and difficulty of a task, have shown to determine the amount of focusing on a task and will shield against distraction and a decrease of performance by the background noise (Engelmann et al., 2009; Halin et al., 2014a). In contrast, this shielding is not seen if we measure perceived disturbance. The perceived disturbance during the reading comprehension task by background noise was significantly higher than the perceived disturbance for both other tasks (**Figure 4**). This might also be the result of the difficulty and engagement of the task while an extra effort investment was needed of participants to perform the task which could lead to a feeling of disturbance. Schlittmeier *et al.* (Schlittmeier et al., 2008) call this the 'reactive effort enhancement', and this effect can lead to reduced performance differences and increased perceived disturbance differences (Kahneman, 1973; Schlittmeier et al., 2008). Also, a prolonged processing time as a result of the sound environment could lead to a feeling of disturbance, even if the performance is unaltered (Prodi et al., 2019; Schiller et al., 2020).

Impact of the Noise Sensitivity Performance and Disturbance

In this study no significant influence of the sound sensitivity of students was found on their performance and disturbance. This is in line with the findings in the experimental research on a collaboration task in OPSE's (Braat-Eggen et al., 2019a). On the other hand, in the field study on OPSEs (Braat-Eggen et al., 2017) and the experimental study on writing performance (Braat-Eggen et al., 2019b), noise sensitive students showed to be more disturbed by background sound than less noise sensitive students.

An explanation for the absence of a significant influence of noise sensitivity of students on performance and disturbance for a "studying for an exam" task could be the same as for the absence of significant sound effects on performance of students: decrease of importance of background noise due to task engagement and task difficulty. These aspects overrule the noise effect whereby noise sensitivity becomes less important.

Limitations of the Method

To study the influence of noise on a "studying for an exam" task, a repeated measurement design with four sound scenarios was used. This implicates that the "studying for an exam" task had to be tested four times. To simulate the studying task, a set of assignments was used that led to an extensive experiment with a long duration. In total, inclusive short breaks between sets of assignments and a practice set of tests, the experiment took 2 h and 30 min. Performing five times the set of tests could implicate fatigue, boredom and loss of concentration effects. The bias caused by these effects could only partly be removed by counterbalancing the sound conditions (Pan et al., 1994; Bergh and Vrana, 1998). Therefore, it was decided to split the experiments in two parts. The students had to perform a practice set of tests and two sets of assignments at the first day (approximately 80 min) and two sets of assignments on the second day (approximately 70 min). Splitting an experiment in two parts introduces possible sources of variation as well such as a spread in time-gap between the test days. However, a statistical comparison of the results of day 1 and day 2 did not show significant differences between the 2 days.

Repeated measurements can also implicate learning effects as a confounding factor. In this experiment we started with a practice set of tests to let the students get familiar with the assignments and the procedure, after all, significant learning effects occur mostly in the first tests (Collie et al., 2003). A learning effect was not expected for the reading comprehension' test; the texts and questions were very different. Syllogisms were used from a well-tested set of assignments and the mental algorithmic tests were diverse. A similar level of complexity of the tests is discussed in the method section.

Using a laboratory setting implies limitations in ecological validity of the sound environment. Although the modeling of the simulated sound environment is based on a real OPSE, which leads to a more realistic sound scenario than used in comparable research on the influence of noise on task performance and disturbance, the spaciousness is limited by the raytracing method used in Odeon, and also the use of headphones is limiting the spaciousness of the perception of the sound signal. Furthermore, not seeing the sources of the background noise (talkers) can contribute to a different perception of the sound field. The advantages of a laboratory study in giving the opportunity to study the influence of different parameters on performance and disturbance is obvious. In our view, this outweighs the disadvantages of a laboratory experiment.

CONCLUSION

In this study the complex task "studying for an exam" has been analyzed by a set of assignments. This typical student task was simulated by a comprehensive reading task with delayed answering (studying task), a mental arithmetic task, and a logical reasoning task, while being exposed to three sound scenarios and a quiet reference sound scenario. In our first hypothesis we expected that a sound environment with background speech would decrease performance and self-estimated performance and increase perceived disturbance of students working on a set of tasks simulating a "studying for an exam" task in an OPSE. This was not shown for the "reading comprehension" and "mental arithmetic" task performance. However, it was demonstrated for the "logic reasoning" task performance and also for self-estimated performance and perceived disturbance for all tasks.

Our second hypothesis claimed more intelligible background speech to have a negative influence on task performance of students and to find an increase of perceived disturbance of students. This hypothesis was not confirmed by the results. Also, no influence of noise sensitivity of students on performance and disturbance of students working on the study tasks was seen in this study.

The "reading comprehension" task with delayed answering showed the highest perceived disturbance in comparison with the other tasks, however, no significant decrease of performance was found due to the background sound scenarios. This might be the result of the difficulty and importance of the reading comprehension with delayed answering task. Both aspects, difficulty and importance, will lead to very high concentration levels for students, resulting in less influence of the background sound scenarios. On the other hand, mental stress and fatigue could be the consequence of prolonged high concentration and high disturbance levels. Therefore, background sound scenarios with background speech are not preferred for important cognitive tasks.

A minimal effect of the realistic simulated background sound scenarios on student performance for all tasks was shown. However, we observe significant effects of the sound scenarios on the subjective variables like self-estimated performance and perceived disturbance. This subjective negative perception of background noise will influence student's comfort. Therefore, it will be interesting to study the long-term impact of acoustically uncomfortable OPSEs in future work.

The translation of the experimental results to requirements for OPSEs is very difficult. All performance measures and all subjective measures of all tasks show the quiet situation to be preferred. A quiet OPSE is the best, this situation can be accomplished by separating different activities by creating activity zones. Strict behavioural rules are required in some of these zones, as no talking is allowed in silence zones.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

EB-E was responsible for the research design, conducted the experiment and wrote the manuscript. JR, MH, and AK gave feedback on the research design and contributed to the manuscript revision, read and approved the submitted version.

FUNDING

This work was supported by a Doctoral Grant for Teachers by the Netherlands Organization for Scientific Research (NWO) under Grant number 023.006.025.

ACKNOWLEDGMENTS

The authors would like to thank Marc Kalee for his contribution to the digital model of the open-plan study environment.

REFERENCES

Adams, N. E. (2015). Bloom's Taxonomy of Cognitive Learning Objectives. *J. Med. Libr. Assoc.* 103 (3), 152–153. doi:10.3163/1536-5050.103.3.010

Anderson, L. W., and Kratwohl, D. R. (2001). *"A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives"*. Boston: Allyn & Bacon.

Banbury, S., and Berry, D. C. (1998). Disruption of Office-Related Tasks by Speech and Office Noise. *Br. J. Psychol.* 89 (3), 499–517. doi:10.1111/j.2044-8295.1998.tb02699.x

Beckers, R., van der Voordt, T., and Dewulf, G. (2015). A Conceptual Framework to Identify Spatial Implications of New Ways of Learning in Higher Education. *Facilities* 33 (1/2), 2–19. doi:10.1108/f-02-2013-0013

Bergh, O. V., and Vrana, S. R. (1998). Repetition and Boredom in a Perceptual Fluency/Attributional Model of Affective Judgements. *Cogn. Emotions* 12, 533–553.

Bloom, B. S., and Krathwohl, D. R. (1956). *"Taxonomy of Educational Objectives: The Classification of Educational Goals," by a Committee of College and university Examiners*. New York: Longmans Green & Co.

Braat-Eggen, E., Poll, M. K. v. d., Hornikx, M., and Kohlrausch, A. (2019). Auditory Distraction in Open-Plan Study Environments: Effects of Background Speech and Reverberation Time on a Collaboration Task. *Appl. Acoust.* 154, 148–160. doi:10.1016/j.apacoust.2019.04.038

Braat-Eggen, E., Reinten, J., Hornikx, M., and Kohlrausch, A. (2019). The Influence of Background Speech on a Writing Task in an Open-Plan Study Environment. " *Building Environ.* 169, 106586. doi:10.1016/j.buildenv.2019.106586

Braat-Eggen, P. E., van Heijst, A., Hornikx, M., and Kohlrausch, A. (2017). Noise Disturbance in Open-Plan Study Environments: a Field Study on Noise Sources, Student Tasks and Room Acoustic Parameters. *Ergonomics* 60 (9), 1297–1314. doi:10.1080/00140139.2017.1306631

Caviola, A., Visentin, C., Borella, E., Mammarella, I., and Prodi, N. (2021). Out of the Noise: Effects of Sound Environment on Maths Performance in Middle-School Students. *J. Environ. Psychol.* 73, 101552. doi:10.1016/j.jenvp.2021.101552

Citrus, D. U. O. (2018). *Facet oefentoetsen en-examens*. College voor Toetsen en Examens. Available at: https://www.oefenen.facet.onl/facet/pages/oefen/start/.

Clark, C., and Sörqvist, P. (2012). A 3 Year Update on the Influence of Noise on Performance and Behavior. *Noise Health* 14, 292–296. doi:10.4103/1463-1741.104896

Collie, A., Maruff, P., Darby, D. G., and McStephen, M. (2003). The Effects of Practice on the Cognitive Test Performance of Neurologically normal Individuals Assessed at Brief Test-Retest Intervals. *J. Int. Neuropsychol. Soc.* 9 (3), 419–428. doi:10.1017/s1355617703930074

Conway, M. A. (2001). "Sensory–perceptual Episodic Memory and its Context: Autobiographical Memory. " *Philosophical Trans. R. Soc. Biol. Sci.* 356, 1413. doi:10.1098/rstb.2001.0940

Curry, L., and Docherty, M. (2017). Implementing Competency-Based Education. *Celt* 10, 61–74. doi:10.22329/celt.v10i0.4716

Dunlosky, J., and Ariel, R. (2011). Self-regulated Learning and the Allocation of Study Time. *Psychol. Learn. Motiv.* 54, 101–138.

Dunlosky, J., and Ariel, R. (2011). The Influence of Agenda-Based and Habitual Processes on Item Selection during Study. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 899–912. doi:10.1037/a0023064

Engelmann, J. B., Damaraju, E., Padmala, S., and Pessoa, L. (2009). Combined Effects of Attention and Motivation on Visual Task Performance: Transient and Sustained Motivational Effects. *Front. Hum. Neurosci. Cog.e Neurosc.e* 3, 1–17. doi:10.3389/neuro.09.004.2009

Flores, M. A., Veiga Simão, A. M., Barros, A., and Pereira, D. (2015). Perceptions of Effectiveness, Fairness and Feedback of Assessment Methods: a Study in Higher Education. *Stud. Higher Education* 40 (9), 1523–1534. doi:10.1080/03075079.2014.881348

Griefahn, B. (2008). Determination of Noise Sensitivity within an Internet Survey Using a Reduced Version of the Noise Sensitivity Questionnaire. *The J. Acoust. Soc. America* 123 (5), 3449. doi:10.1121/1.2934269

Haapakangas, A., Hongisto, V., Hyönä, J., Kokko, J., and Keränen, J. (2014). Effects of Unattended Speech on Performance and Subjective Distraction: The Role of Acoustic Design in Open-Plan Offices. *Appl. Acoust.* 86, 1–16. doi:10.1016/j.apacoust.2014.04.018

Haapakangas, A., Kankkunen, E., Hongisto, V., Virjonen, P., Oliva, D., and Keskinen, E. (2011). Effects of Five Speech Masking Sounds on Performance and Acoustic Satisfaction. Implications for Open-Plan Offices. *Acta Acustica united with Acustica* 97, 641–655. doi:10.3813/aaa.918444

Halin, N., Marsh, J. E., Haga, A., Holmgren, M., and Sörqvist, P. (2014). Effects of Speech on Proofreading: Can Task-Engagement Manipulations Shield against Distraction? *J. Exp. Psychol. Appl.* 20 (1), 69–80. doi:10.1037/xap0000002

Halin, N., Marsh, J. E., Hellman, A., Sörqvist, I., and Sorgvist, P. (2014). A Shield against Distraction. *J. Appl. Res. Mem. Cogn.* 3 (1), 31–36. doi:10.1016/j.jarmac.2014.01.003

Hanczakowski, M., Beaman, C. P., and Jones, D. M. (2018). Learning through Clamor: The Allocation and Perception of Study Time in Noise. *J. Exp. Psychol. Gen.* 147 (7), 1005–1022. doi:10.1037/xge0000449

Houtgast, T., Steeneken, H., and Plomp, R. (1980). Predicting Speech Intelligibility in Rooms from Modulation Transfer Function. I General Oom Acoustics. *Acustica* 46 (1), 60–72.

Hughes, R. W. (2014). Auditory Distraction: A Duplex-Mechanism Account. *PsyCh J.* 3, 30–41. doi:10.1002/pchj.44

Hughes, R. W., Vachon, F., and Jones, D. M. (2005). Auditory Attentional Capture during Serial Recall: Violations at Encoding of an Algorithm-Based Neural Model?. *J. Exp. Psychol. Learn. Mem. Cogn.* 31 (4), 736–749. doi:10.1037/0278-7393.31.4.736

Hughes, R. W., Vachon, F., and Jones, D. M. (2007). Disruption of Short-Term Memory by Changing and Deviant Sounds: Support for a Duplex-Mechanism Account of Auditory Distraction. *J. Exp. Psychol. Learn. Mem. Cogn.* 33 (6), 1050–1061. doi:10.1037/0278-7393.33.6.1050

ISO/TS 15666. (2003). *ISO/TS 15666 "Acoustics - Assessment of Noise Annoyance by Means of Social and Social-Acoustic Surveys"*.

Jahncke, H. (2012). Open-plan Office Noise: The Susceptibility and Suitability of Different Cognitive Tasks for Work in the Presence of Irrelevant Speech. *Noise Health* 14 (61), 315–320. doi:10.4103/1463-1741.104901

Jensen, J. L., McDaniel, M. A., Woodard, S. M., and Kummer, T. A. (2014). "Teaching to the Test or Testing to Teach: Exams Requiring Higher Order Thinking Skills Encourage Greater Conceptual Understanding. *Educ. Psychol. Rev.* 26 (2), 397–329. doi:10.1007/s10648-013-9248-9

Jones, D. M., Marsh, J. E., and Hughes, R. W. (2012). Retrieval from Memory: Vulnerable or Inviolable? *J. Exp. Psychol. Learn. Mem. Cogn.* 38 (4), 905–922. doi:10.1037/a0026781

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs: NJ: Prentice-Hall.

Keus van de Poll, M., Ljung, R., Odelius, J., and Sörqvist, P. (2014). Disruption of Writing by Background Speech: The Role of Speech Transmission index. *Appl. Acoust.* 81, 15–18. doi:10.1016/j.apacoust.2014.02.005

Klatte, M., Bergström, K., and Lachmann, T. (2013). Does Noise Affect Learning? A Short Review on Noise Effects on Cognitive Performance in Children. *Front. Psychol.* 4, 1–6. doi:10.3389/fpsyg.2013.00578

Koenen, A.-K., Dochy, F., and Berghmans, I. (2015). A Phenomenographic Analysis of the Implementation of Competence-Based Education in Higher Education. *Teach. Teach. Education* 50, 1–12. doi:10.1016/j.tate.2015.04.001

Making Moves, B. V. (2019). "Syllogismen Testen," Making Moves B.V. Amsterdam. Available at: https://www.assessment-training.com/.

Marsh, J. E., Sörqvist, P. R. W., and Hughes, R. W. (2015). Dynamic Cognitive Control of Irrelevant Sound: Increased Task Engagement Attenuates Semantic Auditory Distraction. *J. Exp. Psychol. Hum. Perception Perform.* 41, 1462–1474. doi:10.1037/xhp0000060

Martin, R. C., Wogalter, M. S., and Forlano, J. G. (1998). Reading Comprehension in the Presence of Unattended Speech and Music. *J. Mem. Lang.* 27 (4), 382–398. doi:10.1016/0749-596X(88)90063-0

Montgomery, S. E. (2014). Library Space Assessment: User Learning Behaviors in the Library. *The J. Acad. Librarianship* 40 (1), 70–75. doi:10.1016/j.acalib.2013.11.003

Oswald, C. J. P., Tremblay, S. D. M., and Jones, D. M. (2000). Disruption of Comprehension by the Meaning of Irrelevant Sound. *Memory* 8 (5), 345–350. doi:10.1080/09658210050117762

Pan, C. S., Shell, R. L., and Schleifer, L. M. (1994). Performance Variability as an Indicator of Fatigue and Boredom Effects in a VDT Data-entry Task. *Int. J. Human-Computer Interaction* 6, 37–45. doi:10.1080/10447319409526082

Prodi, N., Visentin, C., Borella, E., Mammarella, I. C., and Di Domenico, A. (2019). Noise, Age, and Gender Effects on Speech Intelligibility and Sentence Comprehension for 11-to 13-Year-Old Children in Real Classrooms. *Front. Psychol.* 10, 2166. doi:10.3389/fpsyg.2019.02166

Reinten, J., Braat-Eggen, P. E., Hornikx, M., Kort, H. S. M., and Kohlrausch, A. (2017). The Indoor Sound Environment and Human Task Performance: A Literature Review on the Role of Room Acoustics. *Building Environ.* 123, 315–332. doi:10.1016/j.buildenv.2017.07.005

Schlittmeier, S. J., Hellbrück, J., Thaden, R., and Vorländer, M. (2008). The Impact of Background Speech Varying in Intelligibility: Effects on Cognitive Performance and Perceived Disturbance. *Ergonomics* 51 (5), 719–736. doi:10.1080/00140130701745925

Schlittmeier, S. J., Weisz, N., and Bertrand, O. (2011). What Characterizes Changing-State Speech in Affecting Short-Term Memory? an EEG Study on the Irrelevant Sound Effect. *Psychophysiology* 48 (12), 1669–1680. doi:10.1111/j.1469-8986.2011.01263.x

Schiller, I. S., Morsomme, D., Kob, M., and Remacle, A. (2020). "Noise and a Speakers's Impaired Voice Quality Disrupt Spoken Language Processing in School-Aged Children: Evidence from Performance and Responde Time Measures. *J. Speech, Lamguage, Hearing Res.* 63, 2115–2131.

Sörqvist, P. (2014). On Interpretation and Task Selection in Studies on the Effects of Noise on Cognitive Performance. *Front. Psychol. Cogn.n* 5, 1–4.

Sörqvist, P. (2015). On Interpretation and Task Selection: the Sub-component Hypothesis of Cognitive Noise Effects. *Front. Psychol. Cog.n* 5, 1–4.

Szalma, J. L., and Hancock, P. A. (2011). Noise Effects on Human Performance: A Meta-Analytic Synthesis. *Psychol. Bull.* 137 (4), 682–707. doi:10.1037/a0023987

The Peak Performance Center (2019). Bloom's Taxonomy Revised. Available at: http://thepeakperformancecenter.com/educational-learning/thinking/blooms-taxonomy/blooms-taxonomy-revised/(Acessed 12 21, 2019).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 20 July 2021
doi: 10.3389/fbuil.2021.688016

How Reliable are 11- to 13-Year-Olds' Self-Ratings of Effort in Noisy Conditions?

Chiara Visentin * and Nicola Prodi*

Department of Engineering, University of Ferrara, Ferrara, Italy

Performing a task in noisy conditions is effortful. This is especially relevant for children in classrooms as the effort involved could impair their learning and academic achievements. Numerous studies have investigated how to use behavioral and physiological methods to measure effort, but limited data are available on how well school-aged children rate effort in their classrooms. This study examines whether and how self-ratings can be used to describe the effort children perceive while working in a noisy classroom. This is done by assessing the effect of listening condition on self-rated effort in a group of 182 children 11–13 years old. The children performed three tasks typical of daily classroom activities (speech perception, sentence comprehension, and mental calculation) in three listening conditions (quiet, traffic noise, and classroom noise). After completing each task, they rated their perceived task-related effort on a five-point scale. Their task accuracy and response times (RTs) were recorded (the latter as a behavioral measure of task-related effort). Participants scored higher (more effort) on their self-ratings in the noisy conditions than in quiet. Their self-ratings were also sensitive to the type of background noise, but only for the speech perception task, suggesting that children might not be fully aware of the disruptive effect of background noise. A repeated-measures correlation analysis was run to explore the possible relationship between the three study outcomes (accuracy, self-ratings, and RTs). Self-ratings correlated with accuracy (in all tasks) and with RTs (only in the speech perception task), suggesting that the relationship between different measures of listening effort might depend on the task. Overall, the present findings indicate that self-reports could be useful for measuring changes in school-aged children's perceived listening effort. More research is needed to better understand, and consequently manage, the individual factors that might affect children's self-ratings (e.g., motivation) and to devise an appropriate response format.

Keywords: children, classroom acoustics, listening effort, self-ratings, speech perception, comprehension, maths, noise

OPEN ACCESS

Edited by:
Hasim Altan,
Arkin University of Creative Arts and Design (ARUCAD), Cyprus

Reviewed by:
Harvey Dillon,
Macquarie University, Australia
Shabana Urooj,
Princess Nourah bint Abdulrahman University, Saudi Arabia

****Correspondence:***
Chiara Visentin
chiara.visentin@unife.it

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *30 March 2021*
Accepted: *09 July 2021*
Published: *20 July 2021*

Citation:
Visentin C and Prodi N (2021) How Reliable are 11- to 13-Year-Olds' Self-Ratings of Effort in Noisy Conditions?. Front. Built Environ. 7:688016. doi: 10.3389/fbuil.2021.688016

INTRODUCTION

Performing a listening task in adverse acoustic conditions demands a greater effort (Peelle, 2018). Speech signals can be degraded by a variety of factors, which can be categorized as listener-external (e.g., level and type of background noise, excessively long reverberation) or listener-internal (individual characteristics of auditory, linguistic, and cognitive processing) (Mattys et al., 2012; Lemke and Besser, 2016). Understanding speech in noise requires an explicit engagement of extra cognitive resources.

Effort has been defined as "the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task, with listening effort applying more specifically when tasks involve listening" (Pichora-Fuller et al., 2016). According to the Ease of Language Understanding model (ELU; Rönnberg et al., 2008; Rönnberg et al., 2013), listening to speech in ideal conditions is quick and easy, relying primarily on implicit processing. In unfavorable conditions, an explicit processing becomes necessary, posing greater cognitive demands, which the listener perceives as an increase in the effort involved. As explained in the Framework for Understanding Effortful Listening (FUEL; Pichora-Fuller et al., 2016), listening effort is also modulated by a listener's motivation, or "the resources or energy actually used by a listener to meet the cognitive demands" (Peelle, 2018). The stronger a listener's motivation, the more willing they will be to put effort into the task, regardless of its demands.

The concept of effort is especially relevant for children in classrooms, as greater cognitive demands related to listening could interfere with their ability to complete high-level cognitive tasks (e.g., comprehension; McGarrigle et al., 2019). Classrooms are usually far from optimal listening environments, with background noise levels that exceed the recommended normative values (Shield and Dockrell, 2008; Mealings, 2016) and with excessively long reverberation times. Research has also shown that elementary school children (from kindergarten up to grade 8) spend almost 90% of their time at school listening to speech in the presence of background noise (Crukley et al., 2011).

Previous studies found evidence of school-age children having to put more effort into listening tasks in background noise than in quiet, or when the level of background noise increased. This greater effort was revealed by slower response times (Lewis et al., 2016; McGarrigle et al., 2019; Prodi et al., 2019a; Prodi et al., 2019b; Picou et al., 2019), and by a larger task-evoked increase in pupil size (Steel et al., 2015; McGarrigle et al., 2017; Gómez-Merino et al., 2020). The results were not entirely consistent across the studies, however, presumably due to a different sensitivity of the listening effort measures (McGarrigle et al., 2017; McGarrigle et al., 2019), and to the children's difficulty in preferentially allocating their attention during dual tasks (Choi et al., 2008; Picou et al., 2019).

While numerous studies has investigated the use of behavioral and physiological measures of effort, limited data are available regarding the reliability of school-aged children's subjective ratings of effort in their classrooms. This seems rather odd because self-reports are the most direct and ecologically valid, non-invasive measures for tapping into the subjective experience of effortful listening. Being easy for participants to understand, and for experimenters to administer, self-reports classify as a potentially good candidate for gauging children's listening effort.

In the literature, self-reports were obtained mainly by means of study-specific questionnaires and rating scales (Oosthuizen et al., 2021a) developed specifically for a focused assessment on the conditions adopted in a given study (Moore and Picou, 2018). Children were variously asked to rate their perceived effort (von Lochow et al., 2018), ease of listening (Picou et al., 2019; Oosthuizen et al., 2021a), listening difficulty (Prodi et al., 2010), disturbance (Klatte et al., 2010), and perceived clarity (Gustafson et al., 2014). The results indicate that children report perceiving less effort: in quiet than in noisy conditions (von Lochow et al., 2018; Picou et al., 2019); in aided versus unaided conditions (Oosthuizen et al., 2021a); following digital noise reduction in hearing aids (Gustafson et al., 2014); and after increasing the signal-to-noise ratio by 4 dB in the presence of a four-talker babble-noise (Picou et al., 2019).

On the other hand, no significant differences emerged in children's self-ratings in noisy conditions when the type of background noise was changed (Klatte et al., 2010; von Lochow et al., 2018). Klatte et al. (2010) found no significant effect of the type of background noise (classroom noise, background speech) on first- and third-graders' disturbance ratings for a speech perception task. In the same listening conditions, adults reported finding classroom noise more disturbing than background speech. In the same study, when the children were administered a listening comprehension task, the results indicated no correlation between their ratings of perceived disturbance and their task performance. Von Lochow et al. (2018) examined how the number of competing speakers (one or four) influenced perceived effort in a passage comprehension task. There was no significant increase in the children's perceived effort when the number of speakers changed, despite a significant change in their accuracy.

Self-ratings rely on the assumption that listeners can accurately report the effort they experienced (Picou and Ricketts, 2018), but studies on adults indicate that self-ratings rarely correlate with behavioral or physiological measures of effort (Picou and Ricketts, 2018; Strand et al., 2018; Alhanbali et al., 2019; Lau et al., 2019; McGarrigle et al., 2020). Many variables possibly affecting different participants' performance and self-ratings of effort might obscure any correlation measured across participants, however. On the other hand, self-ratings of effort appear to correlate inversely with self-ratings of performance (Moore and Picou, 2018). Studies in which task difficulty was manipulated (e.g., presence/absence of background noise, a change of SNR) suggest that listeners might become aware of a change in the demands of a task and/or of their own performance, and use this impression as a substitute for judging the effort involved (McGarrigle et al., 2020).

Previous studies on school-age children generally showed no correlation between self-ratings of effort and task performance (von Lochow et al., 2018; Gustafson et al., 2014; Klatte et al., 2010). That said, Picou et al. (2019) examined the relationship between accuracy in a speech perception task, dual-task response time, and four questions related to effort (perceived performance, ease of listening, control, time) in a sample of 10- to 17-year-olds. The results suggested that changing the wording of the question and asking participants to assess attributes more readily-understandable than "listening effort" revealed significant associations between self-ratings and actual performance.

Very few studies have examined the relationship between self-ratings and objective measures of effort in children. Picou et al. (2019) found a significant association between dual-task response times and time perception ratings, but the correlation went in the opposite direction to the one hypothesized. Gustafson et al.

(2014) administered a speech perception task to a group of children (7–12 years old) with normal hearing to examine the impact of digital noise reductions in hearing aids. Response times were faster and clarity ratings were higher with the noise reduction algorithm, but there was no significant correlation between the two measures.

All in all, there seems to be a paucity of information available on the reliability of school-aged children's self-ratings of effort, and their correlation with objective measures. This information would be valuable because reliable subjective ratings of effort would facilitate the assessment of the sound environment in classrooms. A better understanding of how school-age children deal with the demands of listening in challenging acoustic conditions would enable us to promote the design of learning environments with better acoustics.

The purpose of the present study was twofold. The first aim was to explore the sensitivity of self-reports of perceived effort to changes in the demands of a task by taking action on the background noise. Three tasks were considered: 1) a speech perception task; and 2) two tasks highly representative of typical classroom activities, i.e., sentence comprehension and mental calculation. For all three tasks, we expected self-ratings of effort to be higher in noisy than in quiet conditions, reflecting the subjective perception of a greater effort being needed in more adverse listening conditions (von Lochow et al., 2018). The second aim was to examine how subjective measures of effort relate to task performance and response time (RT). In our experiments, RTs (acquired with a single-task paradigm) were used as a behavioral measure of effort. We also planned to see whether the pattern of correlations differed depending on the task.

The results of the present study will add to the current literature on listening effort in children, by:

1) Further exploring the relationship between self-ratings of effort and performance accuracy. This relationship has already been examined in adults, but few reports are currently available regarding children, and their findings are inconsistent (von Lochow et al., 2018; Picou et al., 2019);
2) Newly exploring the specific relationship between self-ratings and RTs in school-aged children.

MATERIALS AND METHODS

This study is part of a broader research project conceived to investigate the effects of noisy and reverberant classrooms on children's performance and perceived effort when engaged in linguistic and calculation tasks. The research also considered the potential role of mediating factors (age, gender, noise source, task difficulty, and room acoustics). The material and methods used in this study were consequently the same as those adopted in other, related studies (Prodi et al., 2019a; Caviola et al., 2021; Prodi et al., 2021; Prodi and Visentin, in press). The focus of the present study is very different, however, in that it investigates the effects of listening conditions and age on self-ratings of effort, whereas the above-mentioned studies only concerned behavioral measures of performance and listening effort.

Participants

A total of 182 children between the ages of 11 and 13 years took part in the study. They were from two schools in Ferrara (Italy), and the study involved three classes for each grade (6–8). Six children were excluded from the data analysis due to intellectual disabilities or a diagnosis of hearing impairment. Other children were also excluded because of their maths fluency assessment (n = 14) or their score in the reading comprehension task (n = 6). The final sample included 159 participants for the speech perception and sentence comprehension tasks, and 162 children for the mental calculation task.

The study was approved by the Ethics Committee at the University of Padova (Italy) and by the school management.

Reading Comprehension and Maths Fluency Tests

In addition to the experimental tasks, children were administered a standardized reading comprehension test [derived from Cornoldi et al. (2017)] and a standardized maths fluency test (Caviola et al., 2016). Both tests were administered collectively to the students in their classrooms, in quiet conditions, one week after the experimental tasks. These individual measures were used for data screening purposes: participants obtaining a standardized score lower than −2.5 (reading comprehension) and −3 (maths fluency) were excluded from our analysis. Their individual scores were included in the statistical model as a covariate, to control for the effects of comprehension abilities and maths fluency on children's perceived effort.

Listening Tasks

Speech Perception Task

The Italian version of the Matrix Sentence Test (Puglisi et al., 2015) was used to measure speech perception. The test consists of sentences with a fixed syntactic structure but no semantic predictability [e.g., *Chiara manda sette porte azzurre* (Chiara sends seven blue doors)], obtained from a base matrix of 50 words. After listening to a sentence, the children selected the words they had heard from the base-word matrix shown on a tablet. Sixteen trials (sentences) were presented in each listening condition. After the last sentence, the children rated their perceived effort in performing the task.

Sentence Comprehension Task

Sentence comprehension in the auditory modality was measured using the COMPRENDO test (Cecchetto et al., 2012), which consists of sentences of varying syntactic complexity [e.g., *La mamma sta inseguendo il bambino* (The mother is chasing the child)]. For each trial, children listened to the playback of a sentence. Then, at the audio offset, four images appeared on the tablet and they had to select the image that best matched the sentence content. Sixteen trials were presented in each listening condition. After the last sentence, the children rated their perceived effort.

Mental Calculation Task

The mental calculation task consisted in solving two-digit additions and subtractions, with or without borrowing and carrying procedures (Caviola et al., 2021). For each problem, the children listened to the playback of a female voice posing the problem. Then, they were asked to select the right answer from among the three options presented on the tablet. Twenty-eight trials (problems) were presented in each listening condition. After the last problem, the children rated their perceived effort.

Listening Conditions

Background Noises

Participants completed each task and provided self-ratings for the three background noise conditions (quiet, traffic noise, classroom noise).

In the quiet condition, no noise was played back, and children completed the tasks in the actual ambient noise of their classrooms, mainly consisting of noises coming from other classrooms and corridors. Traffic noise was obtained by obtaining recordings alongside a busy road (with cars and trucks passing by), then applying spectral filtering to correct for the sound insulation properties of a typical building façade. Classroom noise included sound events typical of a working classroom (e.g., pens rolling onto the floor, chairs scraping) which were mixed with a standard noise signal (ICRA noise; Dreschler et al., 2001). The ICRA noise was constructed from Italian sentences, processed to make them unintelligible but still retaining the long-term average spectrum and temporal envelope fluctuations of speech.

Test Environments

The experiments were carried out in two classrooms, one at each school. The two classrooms had a similar volume (152 and 155 m^3), size (7.3 x 7.0 x 3.1 and 8.3 x 6.0 x 3.1 m), and reverberation time (after one of the classrooms had been temporarily treated by installing sound-absorbing polyester fiber blankets).

During the tests, speech was presented from a Gras 44AB mouth simulator positioned in front of the teacher's desk (height: 1.50 m). The background noises were played back with a Look Line D303 omnidirectional source placed on the floor near a corner of the classroom. Audio playback, testing and data collection were managed with a laptop PC running a wireless test bench (Prodi et al., 2012).

For all conditions, the level of the speech signal was set to 63 dB(A), measured at 1 m in front of the speech source. The background noises (traffic noise, classroom noise) were played back at a level of 60 dB(A), measured as the spatial average of four receivers (positioned in the area where children were seated).

Acoustic Measurements

The reverberation time (T20), i.e., the time it takes, after a source of sound in an enclosure has stopped, for the sound to decrease by 60 dB, and the A-weighted equivalent sound pressure levels (LA,eq) were measured in the two classrooms in occupied conditions (Geneva: International Organization for Standardization, 2009). The measurements were obtained at four positions in the part of the classroom where children were seated, using an omnidirectional B&K4189 ½ inch microphone (height: 1.20 m). As the classrooms were small and the distance between the speech source and the listeners was short, equivalent listening conditions were ensured for the various seating positions. Differences between the acoustic parameters measured in the two classrooms were always below the minimum perceivable threshold, so the two classrooms could be considered equivalent in terms of acoustic perception. **Table 1** shows the listening conditions in the classrooms during the experiments; for more details, see Prodi et al. (2019a).

TABLE 1 | Listening conditions in the classrooms during the experiments, in terms of reverberation times ($T_{20,mid}$, averaged over the 0.5–2 kHz frequency bands) and sound pressure levels ($L_{A,eq}$). The reported values are the spatial averages across four positions in the audience and across repetitions over the classes.

Acoustic parameter	Listening condition		
	Quiet	Traffic noise	Classroom noise
$T_{20,mid}$ [s]	0.69	0.69	0.69
$L_{A,eq}$-speech dB(A)	60.0	60.0	60.0
$L_{A,eq}$-noise dB(A)	41.9	60.4	60.3
SNR dB	+18.1	−0.4	−0.3

Procedures

We used a within-subject study design, with all children performing each task in the three listening conditions. The order of the listening conditions was balanced across the classes in each school year. Children took part in the experiment as a whole class, and the tasks were administered collectively. The three tasks were completed in two sessions, one week apart. The children completed the mental calculation task in the first session, then the speech perception and sentence comprehension tasks in the second. Both sessions lasted about 1 h (including the acoustic measurements, repeated for each class after completing the experimental tasks). To avoid order and fatigue effects, the order of the speech perception and sentence comprehension tasks was counterbalanced across the classes in each school year. The children completed the standardized reading comprehension and maths fluency assessments in their classrooms nearly a week after the second experimental session.

For each task, the children were given instructions and practiced with three or four trials in quiet conditions. Then they were administered three tasks, one for each listening condition. The order of the listening conditions was balanced across the classes for each school year. During the tests, the background noise started approximately 1 s before the speech signal and ended simultaneously with it. In the quiet condition, an acoustic signal (a brief pure tone at 500 Hz) was played back 1 s before the spoken sentence. The next trial was automatically played back only after all participants had answered or reached the time limit (12, 15, or 20 s, depending on the task). Participants were instructed to pay attention to the task and to respond as accurately as possible.

Effort Ratings

Following the completion of each task in a given listening condition, the children were asked to report how much effortful the task had been (*"How much effort did doing this task require?"*). Their answers were given using on a categorical rating scale from one ("minimum effort") to five ("maximum effort"). The question and the scale were presented to the children on the tablet and the numbers they used for their answers were visible on the screen.

Data Analysis

Linear mixed-effects models (LMMs) were used for the statistical analysis, using the *R* software (R Core Team, 2017), and the *lme4* package (Bates et al., 2015). The outcome variable was the self-rated listening effort. The fixed effects included listening condition (quiet, traffic noise, classroom noise), age (categorical variable: 11, 12, 13 years), and their interaction. One model was set up for each task. Individual scores in the reading comprehension test were included as a covariate in the models for the speech perception and sentence comprehension tasks. Individual scores in the maths fluency test were included as a covariate in the models for the mental calculation task. The participant variable was included in all models as a random intercept. In LMMs, the fixed effects represent average trends in the data. Including participants as random effects in the model overcomes the problem of non-independence of the data, and accounts for the fact that each participant may react differently from the average trend. When analyzing RTs, for instance, this approach accounts for the fact that some participants respond more slowly than others.

The *p*-values and χ^2 values for the LMMs were obtained with the *afex* package (Singmann et al., 2021). The normality of the random effects and residuals was checked for each model to identify potential violations of statistical assumptions (Everitt and Hothorn, 2010). Post-hoc tests were run and standardized effect sizes (corresponding to Cohen's *d*) were calculated with the *emmeans* package (Lenth, 2020). In the case of multiple comparisons, the *p*-values were adjusted using the false discovery rate procedure (Benjamini and Hochberg, 1995).

Finally, a correlation analysis was run between self-ratings, task performance accuracy, and RTs. The correlations were examined using a repeated-measures correlation statistical technique, which examines the overall intra-individual association between two measures (Bakdash and Marusich, 2017). This method was chosen to account for the repeated-measures design of the experiment, yielding non-independent observations across listening conditions. The main advantages of this regression technique over standard approaches are: 1) the chance to analyze paired repeated measures without any averaging, and without violating independence assumptions (Bakdash and Marusich, 2017); and 2) the high statistical power, enabling within-subject associations between measures to be tested without any need for large samples of participants (McGarrigle et al., 2020). In the present study, the repeated-measures correlation was used to examine to what extent two measures (RT and self-ratings, or accuracy and self-ratings) show a corresponding variance as a function of changes in the within-subject factor (listening condition). The analysis was implemented using the *rmcorr* package in *R* (Bakdash and Marusich, 2017). The Bonferroni method was applied to adjust the *p*-values for multiple comparisons.

RESULTS

The findings are reported as the scores participants gave to the amount of effort they felt they had put into the tasks, on a scale from one to five (from less to more effort). **Figure 1** shows the average perceived effort ratings, by listening condition and age: it suggests that children found the tasks more effortful in background noise than in quiet conditions. The pattern varied, however, depending on the task. Age-related changes in perceived effort appeared to be quite small. A detailed statistical analysis of the effort ratings is reported in *Effort Ratings*.

Figure 2 shows the frequency distribution of the self-ratings by task and listening condition, over the five scores on the scale. The scores were generally low (from one to three). It was only in the mental calculation task, and in the classroom noise condition in particular, that the ratings were more evenly distributed over the whole scale.

Table 2 provides descriptive statistics on performance accuracy and RTs in the three tasks, by age and listening condition; these data are relevant to the correlation analyses. The RTs were defined as the time elapsing between the audio stimulus offset and the moment an answer was chosen. A detailed analysis of these results is reported elsewhere (Prodi et al., 2019a; Caviola et al., 2021).

Effort Ratings

Table 3 shows the statistical results for the three linear mixed-effects models (one for each task).

For the speech perception task, there was a significant main effect of listening condition ($\chi^2(2) = 63.83$, $p < 0.001$), with effort ratings higher for noisy than for quiet conditions (**Table 3**). The main effects of age ($p = 0.41$), reading comprehension score ($p = 0.84$), and the listening condition x age interaction ($p = 0.49$) were not significant. Pairwise comparisons indicated that perceived effort was significantly lower in quiet than in noisy conditions (quiet < traffic noise: $t = -4.25$, $p < 0.01$, $d = 0.51$, difference = 0.35; quiet < classroom noise: $t = -8.33$, $p < 0.01$, $d = 1.00$, difference = 0.67), and in traffic noise than in classroom noise (traffic noise < classroom noise: $t = -4.07$, $p < 0.01$, $d = 0.49$, difference = 0.33).

Regarding the sentence comprehension task, the analysis again revealed a significant main effect of listening condition ($\chi^2(2) = 35.33$, $p < 0.001$), with effort ratings higher for noisy than for quiet conditions (**Table 3**). The main effects of age ($p = 0.06$), reading comprehension score ($p = 0.36$), and the listening condition x age interaction ($p = 0.37$) were not significant. Pairwise comparisons indicated that perceived effort was significantly lower in quiet than in noisy conditions (quiet < traffic noise: $t = -4.47$, $p < 0.01$, $d = 0.54$, difference = 0.28; quiet < classroom noise: $t = -5.77$, $p < 0.01$, $d = 0.70$, difference = 0.37). There was no difference in the effort ratings between the two noisy conditions ($p = 0.55$).

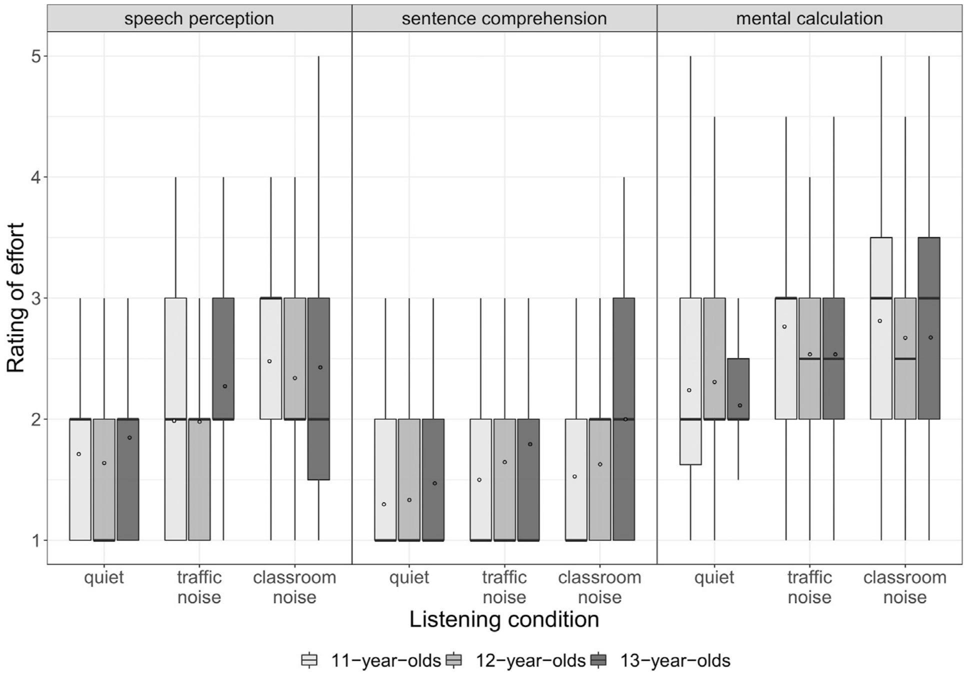


FIGURE 1 | Self-ratings of effort for each listening condition and age. In the boxplots the bottom and top boxes are the first and third quartiles of the data distributions; the central, bold line is the median, and the white circle is the mean; 99% of the data fall within the whiskers.

As for the mental calculation task, there was a significant main effect of listening condition in this case too ($\chi^2(2) = 44.42$, $p < 0.001$), with effort ratings higher for noisy than for quiet conditions (**Table 3**). The main effect of the maths fluency score was significant too ($\chi^2(1) = 9.89$, $p = 0.002$): examining the summary output (**Table 3**) showed that, when all other predictors were set to the reference level, an increase of one standard deviation in the maths fluency score coincided with an estimated 0.22 lower perceived effort. The main effect of age ($p = 0.63$) and the listening condition x age interaction ($p = 0.20$) were not significant. Pairwise comparisons indicated that perceived effort was rated significantly lower in quiet than in noisy conditions (quiet < traffic noise: $t = -5.28$, $p < 0.01$, $d = 0.59$, difference = 0.40; quiet < classroom noise: $t = -6.41$, $p < 0.01$, $d = 0.72$, difference = 0.49). There was no difference between the effort ratings in the two types of noise ($p = 0.76$).

Correlation Analysis

Repeated-measures correlation tests were run to explore the association between perceived effort, task performance accuracy and RTs, as a function of changes in the listening condition. The results showed that the correlation between the measures depended on the task.

In particular, for the speech perception task, there was an inverse relationship between perceived effort and performance accuracy, higher effort ratings being associated with a worse performance [$r = -0.47$; $p < 0.001$; 95%CI (−0.55–−0.38)]. There was a positive relationship instead between perceived effort and RTs, i.e., higher scores for effort were associated with longer RTs [$r = 0.16$; $p = 0.004$; 95%CI (0.03–0.27)].

No significant correlation emerged between effort ratings and RTs in the sentence comprehension task ($r = -0.006$, $p = 0.91$). No correlation analysis was run between effort ratings and accuracy due to the ceiling effect in task performance accuracy.

Finally, there was a significant relationship between effort ratings and performance accuracy in the mental calculation task [$r = -0.14$; $p = 0.011$; 95%CI (−0.026–−0.063)], but not between effort ratings and RTs ($r = 0.001$, $p = 0.98$).

DISCUSSION

The two aims of this study were: 1) to assess the effect of listening condition on children's self-ratings of the effort needed to perform a task; and 2) to investigate the relationship between children's effort ratings and their task performance accuracy, and a behavioral measure of effort (RTs). The two aims are discussed separately below.

Effects of Listening Condition on Effort Ratings

We used three tasks typical of daily classroom activities (speech perception, sentence comprehension, and mental calculation tasks) to examine the effect of listening condition in the classroom (quiet, traffic noise, classroom noise) on the effort that 11- to 13-year-olds reported it costing them to perform such tasks. Our analyses showed that it was more effortful to work in the two noisy conditions than in quiet, for all tasks. The children found classroom noise more disturbing than traffic noise, but only in the speech perception task.

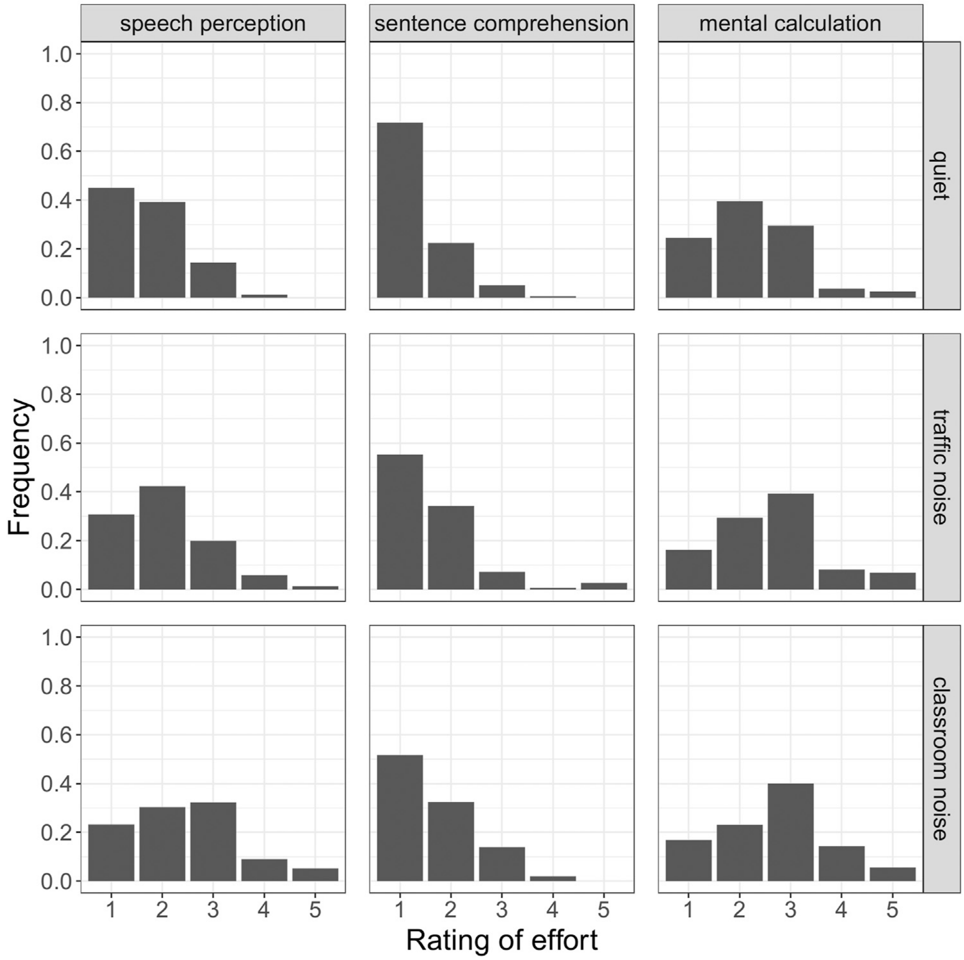


FIGURE 2 | Frequency distribution of the effort ratings, by task and listening condition.

TABLE 2 | Mean correct answers (in %) and response times (in ms) in each task, by listening condition. Standard deviations are given in parentheses.

Task	Correct answers (%)			Response times (ms)		
	Speech perception	**Sentence comprehension**	**Mental calculation**	**Speech perception**	**Sentence comprehension**	**Mental calculation**
Quiet	93.5 (7.4)	95.7 (5.9)	70.0 (18.1)	2,170 (361)	4,241 (789)	9,883 (2,930)
Traffic noise	91.9 (7.4)	95.3 (5.6)	68.5 (20.2)	2,176 (344)	4,316 (762)	9,867 (2,976)
Classroom noise	86.4 (10.8)	93.5 (7.2)	66.7 (19.7)	2,288 (394)	4,555 (903)	10,138 (2,847)

Our results are in line with previous research showing that school-age children found it more effortful to work in adverse (noisy) conditions than in quiet (von Lochow et al., 2018; Picou et al., 2019). These findings support the idea that performing a task in the presence of background noise (whatever its spectral characteristics or informational content) requires the allocation of more cognitive resources, and this is perceived by children as demanding a greater effort. It is worth noting that, when performance in quiet is already near-ceiling (as in our sentence comprehension task), adding background noise will increase perceived effort without affecting performance (Krueger et al., 2017).

Judging from our findings for noisy conditions, a difference in the effort required for different types of background noise only emerged in the speech perception task, in which case the children found classroom noise as more disturbing than traffic noise. The greater perceived effort required may be attributable more to the characteristics of the noise than to its level, since the SNR was the same in the two noisy conditions. Unlike the traffic noise (which had a stationary temporal envelope), the classroom noise used in

TABLE 3 | Linear mixed-effects models for predictors of effort ratings in the three tasks (Reference levels: Listening condition–quiet; Age–11 years).

	Speech perception		*Sentence comprehension*		*Mental calculation*	
Predictors	**Estimate (CI)**	***p***	**Estimate (CI)**	***p***	**Estimate (CI)**	***p***
(Intercept)	1.72 (1.51–1.94)	<0.001	1.30 (1.13–1.47)	<0.001	2.04 (1.74–2.34)	<0.001
Listening condition (quiet *vs* traffic noise)	0.28 (0.06–0.49)	0.013	0.21 (0.03–0.38)	0.02	0.56 (0.30–0.83)	<0.001
Listening condition (quiet *vs* classroom noise)	0.76 (0.54–0.97)	<0.001	0.23 (0.06–0.40)	0.008	0.61 (0.34–0.88)	<0.001
Age (11 *vs* 12)	–0.10 (–0.46 – 0.25)	0.57	0.01 (–0.28 – 0.29)	0.96	0.19 (–0.21 – 0.58)	0.36
Age (11 *vs* 13)	0.17 (–0.22 – 0.55)	0.40	0.17 (–0.13 – 0.48)	0.27	–0.11 (–0.50 – 0.28)	0.58
Reading comprehension/Maths fluency	–0.01 (–0.15 – 0.12)	0.84	–0.05 (–0.16 – 0.06)	0.37	–0.22 (–0.36 – 0.08)	0.002
Observations	464		459		479	
Marginal R^2/Conditional R^2	0.093/0.54		0.065/0.545		0.088/0.603	

our experiment had speech-like temporal fluctuations and included salient events. Given the mechanisms of auditory distraction (Hughes, 2014), the particular combination of the noise's changing state with its salient embedded events might have taxed the children's cognitive resources by competing with the speech material they needed to process, and by prompting them to become disengaged from the task. A greater use of cognitive resources would be needed to deal with the noise, thereby causing an increase in the experience of perceived effort. Our findings add to the limited literature on school-aged children's self-reported effort ratings in a speech perception task. Klatte et al. (2010) reported that a significant effect of the type of background noise on how adults rated the disturbance, but the same did not apply to children in first and third grade. Similarly, Prodi et al. (2010) found that task difficulty ratings by children 8–10 years old were less sensitive to listening condition than those of adults. On the other hand, Picou et al. (2019) showed that 10- to 17-year-olds' self-rated ease of listening was sensitive to changes in the SNR.

No significant difference in the effort ratings emerged between the two noisy conditions in the sentence comprehension task, while the RTs were significantly longer in classroom noise than in traffic noise (Prodi et al., 2019a; Caviola et al., 2021). This result is in line with reports from von Lochow et al. (2018), and Klatte et al. (2010). The former found no significant increase in the perceived effort to complete a passage comprehension task when the number of background speakers was increased from one to four. The latter found that children's ratings of perceived disturbance did not discriminate between the effects of classroom noise and background speech. Both studies concluded that, although children are aware of the disruption caused by background noise, they do not notice any change in the cognitive load of completing a task with different types of background noise. As concerns the sentence comprehension task used in the present study, an alternative explanation might lie in the fact that children's accuracy was at ceiling for all listening conditions. The presence of a four-option, forced-choice paradigm, and the contextual cues provided by the sentences might have made the task too easy for any changes in background noise to affect their accuracy or perceived effort (though it did influence their RTs).

The type of background noise also failed to affect the children's perceived effort in the mental calculation task. This was the only task in which there was a significant effect of individual proficiency (i.e., maths fluency score) on the self-ratings, however. Each child's maths fluency score significantly predicted their effort ratings, the children with higher scores perceiving the task as less effortful. For this task, the perceived effort seems to relate to the processing load involved: children more proficient in maths adopted more efficient strategies to complete the mental calculations, and this cost them less effort. It could also be that mastering mental calculation is associated with a more domain-general attentional control, which would be responsible for a better control over the distracting effect of noise. Unfortunately, no specific data were obtained in this study to test such a hypothesis.

Finally, it is worth noting that our children's effort ratings were rather low, mainly ranging from one to three (on a scale from one to five; **Figure 2**). Their tendency to use scores indicating a limited perceived effort might stem from subjective differences in effort "threshold" (McGarrigle et al., 2014), such that conditions rated as effortful by one sample of listeners might not be rated in the same way by another. Judging from the literature older adults tend to underestimate their perceived listening effort by comparison with young adults (Larsby et al., 2005); non-native listeners report less perceived effort than native listeners, despite the former having a worse performance (Visentin et al., 2018); and children 8–10 years old tend to rate their listening difficulty more favorably than adults (Prodi et al., 2010). We need to bear in mind that children's self-ratings could be affected by a social desirability response bias (King and Bruner, 2000); in other words, children may give researchers the answers they think researches would like to hear. For children's self-ratings to be usable, it is crucial to assess this issue, and try to control it experimentally. This will be the object of a specific study.

Alternatively, the children's reasons for using only the lower scores on the rating scale could have to do with their motivation towards the experiment and the tasks. According to the FUEL model (Pichora-Fuller et al., 2016), perceived effort depends not only on the demands of a task, but also on respondents' motivation, which mediates the cognitive demands of the task when listeners prioritize the allocation of their resources (Peelle, 2018). In the present case, the children were probably strongly motivated: the unconventional classroom experience involved in the listening tests was able to keep them engaged throughout the experiment. This was also noted by the experimenters, who could be considered as qualitative witnesses of the children's motivation. Unfortunately, no specific quantitative data were retrieved on this aspect, and our understanding of what motivates children to listen and engage in academic tasks in the classroom is rather limited (Rudner et al., 2018). More research is needed in this area, including motivation as a mediator in the analysis of self-rated effort.

Relations Between Effort Ratings, Accuracy, and RTs

A second goal of the present study was to examine within-subject associations between effort ratings, RTs, and performance accuracy in different tasks, as a function of the changes in listening condition.

An inverse within-subject association was found between effort ratings and accuracy (in the speech perception and mental calculation tasks), a better performance in the task being associated with lower effort ratings. This result is in line with previous findings referring mainly to speech reception tasks. For instance, Morimoto et al. (2004) found that a very high task performance (but still not at the ceiling) showed a strong negative correlation with the subjective judgement of listening difficulty and intelligibility–and self-rated listening difficulty was a more accurate indicator of performance than intelligibility. We did not find self-reported effort a more sensitive performance indicator than accuracy, but our results might be limited by the fact that our children only used the lower scores on the rating scale, so the variation in their effort ratings was limited.

A significant relationship between the two measures of effort considered here (self-rated effort and RT) only emerged for the speech reception task, but the two measures did not show a corresponding change with listening condition in the sentence comprehension and mental calculation tasks. These findings give the impression that the relationship between different measures of effort might depend on the type of task (e.g., characteristics, difficulty), potentially reinforcing the claim that they tap into different underlying cognitive dimensions (Alhanbali et al., 2019). For instance, recent research indicates a dissociation between ratings of "ease" or "effort" and behavioral measures of effort, in both adults (Lemke and Besser, 2016; Visentin et al., 2019) and children (Picou et al., 2019; Oosthuizen et al., 2021b). Studies exploring this relationship in school-age children are scarce, however, and future studies need to address this research gap.

Study Limitations and Future Directions

Several aspects of this study may limit the generalizability of our findings.

First, we assessed perceived effort based on a single question (*"How much effort did doing this task require?"*), which might be too difficult for the children to understand. According to Kahneman and Frederick (2002), people faced with a difficult question tend to answer a different, easier question. Formulating the question in a way that is easier for children to understand might generate results that are more reliable, and less biased toward positive responses (Picou et al., 2019).

Second, we used a 5-point rating scale for the self-ratings. It may be better to use a visual analogue scale to obtain a finer measure of the amount of effort perceived. This might give us more useful information on the construct assessed, and avoid systematic upward or downward bias deriving from the limited number of scores on the scale (Kuhlmann et al., 2017).

Third, none of the variables included in the statistical analysis (listening condition, age, individual proficiency in the baseline task) accounted for much of the substantial inter-individual variance in self-ratings of effort, as indicated by the conditional R-squared coefficients in **Table 3**. This suggests that other (intrinsic or extrinsic) factors might influence school-age children's effort ratings. For instance, the FUEL model indicates that listening effort depends both on the demands of a task and on the listener's motivation. The latter governs how hard listeners try to understand what is being said and governs how well their perceptual and cognitive abilities are used (Lidestam and Beskow, 2006). Future studies should include questions to assess children's motivation in order to shed light on how much they focus their attention on completing the task.

Another aspect of effort that future research could explore concerns confidence ratings, or how much guesswork respondents feel they have used in completing a task. This would give us an idea of their meta-cognitive monitoring abilities, i.e., the degree to which respondents are capable of monitoring whether they have understood the message correctly or not (Giovanelli et al., 2021). This aspect is especially relevant in the case of children working in classrooms in inadequate acoustic conditions, as an adequate meta-cognitive monitoring of their teacher's verbal communication would trigger compensatory strategies (e.g., the children would ask the teacher to repeat a sentence, or speak more slowly or louder) to help them cope with the adverse listening conditions.

CONCLUSION

This study investigated the effect of background noise conditions on self-ratings of listening effort in children aged 11–13 years completing three academic tasks (speech perception, sentence comprehension, and mental calculation). In all three tasks, the children's perceived effort was greater in background noise than in quiet conditions, but it was only in the speech perception task that the type of background noise (traffic versus classroom noise) influenced their effort ratings.

Our results indicate a significant within-subject association between children's effort ratings and their accuracy in all tasks. On the other hand, a significant link between their effort ratings and a behavioral measure of listening effort (response times) only emerged for the speech perception task.

Overall, the present findings go to show that self-ratings could be useful for measuring changes in school-aged children's perceived listening effort. More research is needed to clarify, and thus control the individual factors that influence children's effort ratings (such as motivation) and to devise an appropriate response format.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee at the University of Padova (Italy) and by the school managements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHORS CONTRIBUTIONS

CV and NP conceived and designed the study and took care of data collection. CV performed the statistical analysis and wrote the first draft of the paper, and both authors were actively involved in preparing the final draft.

FUNDING

This work was funded partly by the Emilia-Romagna Regional Authority's COMPRENDO project (Action POR FESR 2014-20–Axis 1 Action 1.2.2 (2019)), and partly by the BRIC 2019 project-topic ID 14 of the National Institute for Insurance against Accidents at Work (INAIL).

ACKNOWLEDGMENTS

The authors would like to thank Erika Borella and Irene Mammarella, who contributed advice, ideas and opinions on the selection and administration of the tasks, and Alberto Di Domenico and Sara Caviola, who provided the materials for the sentence comprehension and mental calculation tasks. The authors also thank the children and teachers who took part in the study. Finally, the authors acknowledge Hörtech GmbH, Oldenburg, Germany, for providing the speech recordings of the Matrix Sentence Test in the Italian language.

REFERENCES

Alhanbali, S., Dawes, P., Millman, R. E., and Munro, K. J. (2019). Measures of Listening Effort Are Multidimensional. *Ear and Hearing* 40 (5), 1084–1097. doi:10.1097/AUD.0000000000000697

Bakdash, J. Z., and Marusich, L. R. (2017). Repeated Measures Correlation. *Front. Psychol.* 8, 456. doi:10.3389/fpsyg.2017.00456

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Usinglme4. *J. Stat. Soft.* 67, 1–48. doi:10.18637/jss.v067.i01

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B (Methodological).* 57, 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Caviola, S., Gerotto, G., Lucangeli, D., and Mammarella, I. C. (2016). *Test AC-FL: Prove di fluenza per le abilità di calcolo [AC-FL Test: Maths fluency abilities test]*. Trento: Edizioni Centro Studi Erickson.

Caviola, S., Visentin, C., Borella, E., Mammarella, I., and Prodi, N. (2021). Out of the Noise: Effects of Sound Environment on Maths Performance in Middle-School Students. *J. Environ. Psychol.* 73, 101552. doi:10.1016/j.jenvp.2021.101552

Cecchetto, C., Di Domenico, A., Garraffa, M., and Papagno, C. (2012). *Comprendo. Batteria per la Comprensione di frasi negli adulti [Comprendo. Test battery for sentence comprehension in adults]*. Milano: Raffaello Cortina Editore.

Choi, S., Lotto, A., Lewis, D., Hoover, B., and Stelmachowicz, P. (2008). Attentional Modulation of Word Recognition by Children in a Dual-Task Paradigm. *J. Speech Lang. Hear. Res.* 51 (4), 1042–1054. doi:10.1044/1092-4388(2008/07610.1044/1092-4388(2008/076)

Cornoldi, C., Carretti, B., and Colpo, C. (2017). *Prove MT-Kit Scuola. Dalla valutazione degli apprendimenti di lettura e comprensione al potenziamento [MT Tests - School Kit. From the assessment of reading and comprehension to its improvement]*. Firenze: Organizzazioni Speciali Editore.

Crukley, J., Scollie, S., and Parsa, V. (2011). An Exploration of Non-quiet Listening at School. *J. Educ. Audiol.* 17 (1), 23–35.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., and Westermann, S. (2001). ICRA Noises: Artificial Noise Signals with Speech-like Spectral and Temporal Properties for Hearing Instrument Assessment: Ruidos ICRA: Señates de ruido artificial con espectro similar al habla y propiedades temporales para pruebas de instrumentos auditivos. *Int. J. Audiol.* 40, 148–157. doi:10.3109/00206090109073110

Everitt, B. S., and Hothorn, T. (2010). *A Handbook of Statistical Analyses Using R*. New York: Chapman & Hall.

Giovanelli, E., Valzolgher, C., Gessa, E., Todeschini, M., and Pavani, F. (2021). Unmasking the Difficulty of Listening to Talkers with Masks: Lessons from the COVID-19 Pandemic. *i-Perception.* 12 (2), 204166952199839. doi:10.1177/2041669521998393

Gómez-Merino, N., Gheller, F., Spicciarelli, G., and Trevisi, P. (2020). Pupillometry as a Measure for Listening Effort in Children: a Review. *Hearing, Balance Commun.* 18 (3), 152–158. doi:10.1080/21695717.2020.1807256

Gustafson, S., McCreery, R., Hoover, B., Kopun, J. G., and Stelmachowicz, P. (2014). Listening Effort and Perceived Clarity for normal-hearing Children with the Use of Digital Noise Reduction. *Ear Hear.* 35 (2), 183–194. doi:10.1097/01.aud.0000440715.85844.b8

Hughes, R. W. (2014). Auditory Distraction: a Duplex-Mechanism Account. *PsyCh J.* 3 (1), 30–41. doi:10.1002/pchj.44

Geneva: International Organization for Standardization (2009). *Acoustics – Measurement of Room Acoustic Parameters – Part 1: Performance Spaces*. Geneva, Switzerland, ISO3382–1.:2009.

Kahneman, D., and Frederick, S. (2002). "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," in *Heuristics and Biases: The Psychology of Intuitive Judgment*. Editors T. Gilovich, D. Griffin, and D. Kahneman (Cambridge: Cambridge University Press), 49–81. doi:10.1017/CBO9780511808098.004

King, M. F., and Bruner, G. C. (2000). Social Desirability Bias: a Neglected Aspect of Validity Testing. *Psychol. Mark.* 17 (2), 79–103. doi:10.1002/(SICI)1520-6793(200002)17:2<79::AID-MAR2>3.0.CO;2-0

Klatte, M., Lachmann, T., and Meis, M. (2010). Effects of Noise and Reverberation on Speech Perception and Listening Comprehension of Children and Adults in a Classroom-like Setting. *Noise Health* 12, 270–282. doi:10.4103/1463-1741.70506

Krueger, M., Schulte, M., Zokoll, M. A., Wagener, K. C., Meis, M., Brand, T., et al. (2017). Relation between Listening Effort and Speech Intelligibility in Noise. *Am. J. Audiol.* 26 (3S), 378–392. doi:10.1044/2017_AJA-16-0136

Kuhlmann, T., Dantlgraber, M., and Reips, U. D. (2017). Investigating Measurement Equivalence of Visual Analogue Scales and Likert-type Scales in Internet-Based Personality Questionnaires. *Behav. Res.* 49 (6), 2173–2181. doi:10.3758/s13428-016-0850-x

Larsby, B., Hällgren, M., Lyxell, B., and Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: effects of different noise backgrounds in normal-hearing and hearing-impaired subjects Desempeño cognitivo y percepción del esfuerzo en tareas de procesamiento del lenguaje: Efectos de las diferentes condiciones de fondo en sujetos normales e hipoacúsicos. *Int. J. Audiol.* 44 (3), 131–143. doi:10.1080/14992020500057244

Lau, M. K., Hicks, C., Kroll, T., and Zupancic, S. (2019). Effect of Auditory Task Type on Physiological and Subjective Measures of Listening Effort in Individuals with normal Hearing. *J. Speech Lang. Hear. Res.* 62 (5), 1549–1560. doi:10.1044/2018_JSLHR-H-17-0473

Lemke, U., and Besser, J. (2016). Cognitive Load and Listening Effort: Concepts and Age-Related Considerations. *Ear Hear.* 37, 77S–84S. doi:10.1097/AUD.0000000000000304

Lenth, R. V. (2020). Data From: Emmeans: Estimated Marginal Means, Aka Least-Squares Means. R package version 1.4.4. https://CRAN.R-project.org/package=emmeans.

Lewis, D., Schmid, K., O'Leary, S., Spalding, J., Heinrichs-Graham, E., and High, R. (2016). Effects of Noise on Speech Recognition and Listening Effort in Children with normal Hearing and Children with Mild Bilateral or Unilateral Hearing Loss. *J. Speech Lang. Hear. Res.* 59 (5), 1218–1232. doi:10.1044/2016_JSLHR-H-15-0207

Lidestam, B., and Beskow, J. (2006). Motivation and Appraisal in Perception of Poorly Specified Speech. *Scand. J. Psychol.* 47 (2), 93–101. doi:10.1111/j.1467-9450.2006.00496.x

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech Recognition in Adverse Conditions: A Review. *Lang. Cogn. Process.* 27, 953–978. doi:10.1080/01690965.2012.705006

McGarrigle, R., Dawes, P., Stewart, A. J., Kuchinsky, S. E., and Munro, K. J. (2017). Measuring Listening-Related Effort and Fatigue in School-Aged Children Using Pupillometry. *J. Exp. Child Psychol.* 161, 95–112. doi:10.1016/j.jecp.2017.04.006

McGarrigle, R., Gustafson, S. J., Hornsby, B. W. Y., and Bess, F. H. (2019). Behavioral Measures of Listening Effort in School-Age Children: Examining the Effects of Signal-To-Noise Ratio, Hearing Loss, and Amplification. *Ear Hear.* 40, 381–392. doi:10.1097/AUD.0000000000000623

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., et al. (2014). Listening Effort and Fatigue: What Exactly Are We Measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white Paper'. *Int. J. Audiol.* 53 (7), 433–445. doi:10.3109/14992027.2014.890296

McGarrigle, R., Rakusen, L., and Mattys, S. (2020). Effortful Listening under the Microscope: Examining Relations between Pupillometric and Subjective Markers of Effort and Tiredness from Listening. *Psychophysiology* 58 (1), e13703. doi:10.1111/psyp.13703

Mealings, K. (2016). "Classroom Acoustic Conditions: Understanding what Is Suitable through a Review of National and International Standards, Recommendations, and Live Classroom Measurements," in Proceedings of Acoustics 2016 – The Second Australasian Acoustical Societies Conference. Editors I. Hillock and D. Mee (Brisbane, Australia: The Australian Acoustical Society), 145–172.

Moore, T. M., and Picou, E. M. (2018). A Potential Bias in Subjective Ratings of Mental Effort. *J. Speech Lang. Hear. Res.* 61 (9), 2405–2421. doi:10.1044/2018_JSLHR-H-17-0451

Morimoto, M., Sato, H., and Kobayashi, M. (2004). Listening Difficulty as a Subjective Measure for Evaluation of Speech Transmission Performance in Public Spaces. *The J. Acoust. Soc. America* 116 (3), 1607–1613. doi:10.1121/1.1775276

Oosthuizen, I., Picou, E. M., Pottas, L., Myburgh, H. C., and Swanepoel, D. W. (2021b). Listening Effort in School-Age Children with normal Hearing Compared to Children with Limited Useable Hearing Unilaterally. *Am. J. Audiol.* 30, 309–324. doi:10.1044/2021_AJA-20-00082

Oosthuizen, I., Picou, E. M., Pottas, L., Myburgh, H. C., and Swanepoel, D. W. (2021a). Listening Effort in School-Aged Children with Limited Useable Hearing Unilaterally: Examining the Effects of a Personal, Digital Remote Microphone System and a Contralateral Routing of Signal System. *Trends Hearing* 25, 233121652098470. doi:10.1177/2331216520984700

Peelle, J. E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic challenge Are Reflected in Brain and Behavior. *Ear Hear.* 39 (2), 204–214. doi:10.1097/AUD.0000000000000494

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., et al. (2016). Hearing Impairment and Cognitive Energy: the Framework for Understanding Effortful Listening (FUEL). *Ear Hear.* 37, 5S–27S. doi:10.1097/AUD.0000000000000312

Picou, E. M., Bean, B., Marcrum, S. C., Ricketts, T. A., and Hornsby, B. W. Y. (2019). Moderate Reverberation Does Not Increase Subjective Fatigue, Subjective Listening Effort, or Behavioral Listening Effort in School-Aged Children. *Front. Psychol.* 10, 1749. doi:10.3389/fpsyg.2019.01749

Picou, E. M., and Ricketts, T. A. (2018). The Relationship between Speech Recognition, Behavioural Listening Effort, and Subjective Ratings. *Int. J. Audiol.* 57 (6), 457–467. doi:10.1080/14992027.2018.1431696

Prodi, N., and Visentin, C. (in press). A Slight Increase in Reverberation Time in the Classroom Affects Performance and Behavioral Listening Effort. Manuscript Accepted for Publication in Ear Hear. *Ear Hear.*

Prodi, N., Visentin, C., and Bellettini, C. (2012). "Listening Efficiency Testing," in Audio Engineering Society Conference: 45th International Conference: Applications of Time-Frequency Processing in Audio, Helsinki, Finland, March 1–4, 2012 (Audio Engineering Society).

Prodi, N., Visentin, C., Borella, E., Mammarella, I. C., and Di Domenico, A. (2019a). Noise, Age, and Gender Effects on Speech Intelligibility and Sentence Comprehension for 11- to 13-Year-Old Children in Real Classrooms. *Front. Psychol.* 10, 2166. doi:10.3389/fpsyg.2019.02166

Prodi, N., Visentin, C., Borella, E., Mammarella, I. C., and Di Domenico, A. (2021). Using Speech Comprehension to Qualify Communication in Classrooms: Influence of Listening Condition, Task Complexity and Students' Age and Linguistic Abilities. *Appl. Acoust.* 182, 108239. doi:10.1016/j.apacoust.2021.108239

Prodi, N., Visentin, C., and Farnetani, A. (2010). Intelligibility, Listening Difficulty and Listening Efficiency in Auralized Classrooms. *J. Acoust. Soc. America* 128, 172–181. doi:10.1121/1.3436563

Prodi, N., Visentin, C., Peretti, A., Griguolo, J., and Bartolucci, G. B. (2019b). Investigating Listening Effort in Classrooms for 5- to 7-Year-Old Children. *Lshss* 50 (2), 196–210. doi:10.1044/2018_LSHSS-18-0039

Puglisi, G. E., Warzybok, A., Hochmuth, S., Visentin, C., Astolfi, A., Prodi, N., et al. (2015). An Italian Matrix Sentence Test for the Evaluation of Speech Intelligibility in Noise. *Int. J. Audiol.* 54, 44–50. doi:10.3109/14992027.2015.1061709

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) Model: Theoretical, Empirical, and Clinical Advances. *Front. Syst. Neurosci.* 7, 31. doi:10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition Counts: a Working Memory System for Ease of Language Understanding (ELU). *Int. J. Audiol.* 47, S99–S105. doi:10.1080/14992020802301167

Rudner, M., Lyberg-Åhlander, V., Brännström, J., Nirme, J., Pichora-Fuller, M. K., and Sahlén, B. (2018). Listening Comprehension and Listening Effort in the Primary School Classroom. *Front. Psychol.* 9, 1193. doi:10.3389/fpsyg.2018.01193

Shield, B. M., and Dockrell, J. E. (2008). The Effects of Environmental and Classroom Noise on the Academic Attainments of Primary School Children. *J. Acoust. Soc. America* 123, 133–144. doi:10.1121/1.2812596

Singmann, H., Bolker, B., Westfall, J., Aust, F., and Ben-Shachar, M. S. (2021). Data From: Afex: Analysis of Factorial Experiments. R package version 0.28-1. https://CRAN.R-project.org/package=afex.

Steel, M. M., Papsin, B. C., and Gordon, K. A. (2015). Binaural Fusion and Listening Effort in Children Who Use Bilateral Cochlear Implants: a Psychoacoustic and Pupillometric Study. *PLOS ONE* 10 (2), e0117611. doi:10.1371/journal.pone.0117611

Strand, J. F., Brown, V. A., Merchant, M. B., Brown, H. E., and Smith, J. (2018). Measuring Listening Effort: Convergent Validity, Sensitivity, and Links with Cognitive and Personality Measures. *J. Speech Lang. Hear. Res.* 61 (6), 1463–1486. doi:10.1044/2018_JSLHR-H-17-0257

Visentin, C., Prodi, N., Cappelletti, F., Torresin, S., and Gasparella, A. (2019). Speech Intelligibility and Listening Effort in university Classrooms for Native and Non-native Italian Listeners. *Building Acoust.* 26 (4), 275–291. doi:10.1177/1351010X19882314

Visentin, C., Prodi, N., Cappelletti, F., Torresin, S., and Gasparella, A. (2018). Using Listening Effort Assessment in the Acoustical Design of Rooms for Speech. *Building Environ.* 136, 38–53. doi:10.1016/j.buildenv.2018.03.020

von Lochow, H., Lyberg-Åhlander, V., Sahlén, B., Kastberg, T., and Brännström, K. J. (2018). The Effect of Voice Quality and Competing Speakers in a Passage Comprehension Task: Perceived Effort in Relation to Cognitive Functioning and Performance in Children with normal Hearing. *Logop. Phoniatr. Vocology* 43 (1), 32–41. doi:10.1080/14015439.2017.1307446

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 26 July 2021
doi: 10.3389/fbuil.2021.689388

Contribution of Low-Level Acoustic and Higher-Level Lexical-Semantic Cues to Speech Recognition in Noise and Reverberation

Anna Warzybok[1]*, *Jan Rennies*[2] *and Birger Kollmeier*[1,2]

[1]*Medical Physics and Cluster of Excellence Hearing4all, Carl von Ossietzky Universität, Oldenburg, Germany,* [2]*Department for Hearing, Speech and Audio Technology, Fraunhofer Institute for Digital Media Technology IDMT and Cluster of Excellence Hearing4all, Oldenburg, Germany*

OPEN ACCESS

Edited by:
Arianna Astolfi,
Politecnico di Torino, Italy

Reviewed by:
Rosario Romano,
University of Naples Federico II, Italy
Christian Füllgrabe,
Loughborough University,
United Kingdom
Ellen Peng,
University of Wisconsin-Madison,
United States

***Correspondence:**
Anna Warzybok
a.warzybok@uol.de

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *31 March 2021*
Accepted: *21 June 2021*
Published: *26 July 2021*

Citation:
Warzybok A, Rennies J and Kollmeier B (2021) Contribution of Low-Level Acoustic and Higher-Level Lexical-Semantic Cues to Speech Recognition in Noise and Reverberation. Front. Built Environ. 7:689388. doi: 10.3389/fbuil.2021.689388

Masking noise and reverberation strongly influence speech intelligibility and decrease listening comfort. To optimize acoustics for ensuring a comfortable environment, it is crucial to understand the respective contribution of bottom-up signal-driven cues and top-down linguistic-semantic cues to speech recognition in noise and reverberation. Since the relevance of these cues differs across speech test materials and training status of the listeners, we investigate the influence of speech material type on speech recognition in noise, reverberation and combinations of noise and reverberation. We also examine the influence of training on the performance for a subset of measurement conditions. Speech recognition is measured with an open-set, everyday Plomp-type sentence test and compared to the recognition scores for a closed-set Matrix-type test consisting of syntactically fixed and semantically unpredictable sentences (c.f. data by Rennies et al., J. Acoust. Soc. America, 2014, 136, 2642–2653). While both tests yield approximately the same recognition threshold in noise in trained normal-hearing listeners, their performance may differ as a result of cognitive factors, i.e., the closed-set test is more sensitive to training effects while the open-set test is more affected by language familiarity. All experimental data were obtained at a fixed signal-to-noise ratio (SNR) and/or reverberation time set to obtain the desired speech transmission index (STI) values of 0.17, 0.30, and 0.43. respectively, thus linking the data to STI predictions as a measure of pure low-level acoustic effects. The results confirm the consistent difference between robustness to reverberation observed in the literature between the matrix type sentences and the Plomp-type sentences, especially for poor and medium speech intelligibility. The robustness of the closed-set matrix type sentences against reverberation disappeared when listeners had no a priori knowledge about the speech material (sentence structure and words used), thus demonstrating the influence of higher-level lexical-semantic cues in speech recognition. In addition, the consistent difference between reverberation- and noise-induced recognition scores of everyday sentences for medium and high STI conditions and the differences between Matrix-type and Plomp-type sentence scores clearly demonstrate the limited utility of the STI in predicting speech recognition in noise and reverberation.

Keywords: speech transmission index, speech in noise, reverberation, speech perception, matrix sentence test, everyday sentence test

INTRODUCTION

In realistic room scenarios speech intelligibility is mainly determined by background noise and reverberation. This has been confirmed by various studies that investigated the detrimental effects of background noise and reverberation in listeners with different hearing status. To model the combined effect of these factors for arbitrary situations, objective measures have been developed based on the concept of the Speech Transmission Index (Steeneken and Houtgast, 1980). Several studies confirmed a strong relationship between STI predictions and empirical speech recognition data as well as the detrimental influence of noise, reverberation, and hearing status of the listeners (e.g., Duquesnoy and Plomp, 1980; George et al., 2010). Rennies et al. (2014) employed a similar paradigm as George at al. (2010) by measuring speech recognition scores in normal-hearing listeners for different combinations of noise and reverberation that produced the same STI values. In contrast to the data of George et al. (2010) for their group of normal-hearing listeners, Rennies et al. (2014) reported significant discrepancies between the measured data and STI predictions. While the STI correctly accounted for the influence of noise on speech recognition, but the influence of reverberation was overestimated compared to the empirical data, i.e., the predicted detrimental effect of reverberation was larger than in the measured data. Rennies et al. (2014) speculated that the differences between these two studies may be caused by the fact that speech material differed in talker (male German speaker used by Rennies et al., 2014, vs. female Dutch talker used in the study of George et al., 2010). Furthermore, different types of speech material were used which could be crucial for the different outcomes. However, it remains unclear which of these differences is mainly responsible for the observed discrepancy.[1]

In the current study we therefore investigate if the observed discrepancies between the studies of George et al. (2010) and Rennies et al. (2014) are due to the type of speech material used in the experiments: Rennies et al. (2014) used a closed-set matrix-type sentence test consisting of semantically unpredictably and syntactically fixed sentences which are generated from a base matrix consisting of 50 words (10 names, 10 verbs, 10 numerals, 10 adjectives, and 10 nouns, Wagener et al., 1999). Before the actual measurements, listeners are always trained with at least two test lists of 20 sentences to get familiar with the speech material and account for training effect (Wagener et al., 1999; Kollmeier et al., 2015). In contrast, George et al. (2010) used open-set everyday sentence test VU98 (Versfeld et al., 2000) consisting of sentences with different syntax and vocabulary. Hence, the listeners were not aware of the sentence content and were not trained with the same material that was used for testing.

It is investigated here if a priori knowledge about the speech material (i.e., training to the speech material resulting in the familiarity with the limited set of 50 words and information about the fixed grammatical sentence structure) is the reason for the observed robustness of matrix tests in reverberant conditions. Furthermore, the effects of noise only, reverberation only, and combinations of noise and reverberation are systematically investigated using the method adapted from Rennies et al. (2014), i.e., speech recognition measurements at a fixed signal-to-noise ratio (SNR) and/or reverberation time set to obtain the desired STI values of 0.17, 0.30, and 0.43 but with a different type of speech material, namely the German everyday sentences test (so-called Kollmeier and Wesselkamp, 1997). This type of speech material is comparable to the Plomp-type speech material used by George et al. (2010). The GÖSA was recorded with the same speaker as the German matrix sentence test used by Rennies et al. (2014) which allows for excluding the potentially large effects of speaker on speech recognition (Hochmuth et al., 2015). This way a direct comparisons and examination of the effect of speech material type was possible without confounding effects of talker differences.

If the type of speech material is indeed the main factor responsible for different outcomes in the studies reported by George et al. (2010) and Rennies et al. (2014), the speech recognition scores measured here with the German everyday sentence test should agree with the findings of George et al. (2010) and the measured scores should be constant along iso-STI contours, i.e., for different combinations of SNR and reverberation time that produce the same STI.

Furthermore, we investigated the reasons for the high robustness of the German matrix sentence test against reverberation observed by Rennies et al. (2014). We hypothesize that the high recognition scores in strongly reverberant conditions as measured by Rennies et al. (2014) arise from the a priori knowledge about the speech material which is given in the training session. In order to test this hypothesis, two additional conditions with the matrix test in reverberation were included. The main difference to the study of Rennies et al. (2014) was that the listeners were not familiarized with the speech material prior actual measurements and by that they were not aware of the fixed grammatical structure of the sentences and the limited number of words (which effectively made the matrix test similar to an open-set speech test).

By comparing the experimental results with GÖSA and the untrained matrix test to the trained matrix test data collected by Rennies et al. (2014) and George et al. (2010), we can assess the influence of top-down processing (i.e., knowledge-driven cognitive processes utilizing lexical-semantic cues) for speech recognition in reverberant environments. Furthermore, the comparison of the measured data to predictions of the STI will provide an estimate of the signal-driven, low-level, bottom-up processing contribution in the conditions considered. Hence, an estimate of the role of cognitive

[1]While both the study of Rennies et al. (2014) and of George et al. (2010) employed normal-hearing listeners, the age range differed slightly (20–41 years vs. 26–57 years), i.e., the listeners group of George at al. (2010) included some older normal-hearing listeners. This may have contributed to differences between the studies. However, the mean speech recognition threshold and corresponding standard deviation observed by George et al. (2010) was comparable to the reference data of young normal-hearing listeners for the speech material used (VU98 corpus, Versfeld et al., 2000). Hence it can be assumed that differences in listener groups are not the main contributor to the systematic differences discussed by Rennies et al. (2014).

processes (with a focus on knowledge-driven cognitive processes utilizing lexical-semantic cues) in compensating for the detrimental effect of reverberation and background noise will become possible.

METHODS

Speech Transmission Index Calculations

The STI is based on the concept of the modulation transfer function (MTF), which describes the changes in the temporal modulation of the signal due to its transmission through a system. A simplified STI calculation method was adopted here from the study of Rennies et al. (2014). This method considers the calculation of the MTF as a product of the factor m_{rev}, which quantifies the convolutive distortion of the speech signal due to reverberation, and the factor m_{noise}, which characterizes the distortions of the speech signal due to the additive noise (IEC, 2003). If the room impulse response is approximated by an exponential decay, the factor m_{rev} can be then described as:

$$m_{rev}(F) = \left(1 + \left(\frac{2\pi F T_{60}}{13.8}\right)^2\right)^{-0.5}$$

where F is the modulation frequency in Hertz and T_{60} is the reverberation time in seconds (IEC, 2003).

For compatibility with the studies of George et al. (2010) and Rennies et al. (2014), the reverberation time was assumed to be frequency independent. Accordingly, all the room impulse responses (RIRs) used in this study were generated based on their broadband reverberation time. The factor m_{noise} is expressed as

$$m_{noise} = \left(1 + 10^{-SNR/10}\right)^{-1}$$

where SNR is signal-to-noise ratio in dB (IEC, 2003).

Since the long-term spectra of the speech material and the masking noise were similar, it was assumed that the SNR also is constant across frequencies.

Listeners

Fourteen normal-hearing listeners with a pure-tone threshold not exceeding 20 dB HL for octave frequencies between 125 Hz and 8 kHz participated in this study. They ranged in age from 21 to 27 years (mean age of 22.3 ± 2.2). None of them had previous experience with speech recognition measurements. All listeners were informed about the general purpose of the study, gave written informed consent, and were paid for their participation in the listening experiments. Ethical approval was obtained from the Research Ethical Committee of the Universität Oldenburg.

Speech Recognition Measurements

Set-Up

The Göttingen sentence test (Kollmeier and Wesselkamp, 1997) was used as a speech material in this study. It contains short, meaningful everyday sentences like the Plomp-type sentences (Plomp and Mimpen, 1979) and the HINT test (Nilsson et al., 1994), but employing word scoring and a numerical optimization procedure for homogenization across test items. The word corpus is rather large (1,194 words), the content of each sentence is unknown to the listener. Ten perceptually balanced lists of 20 sentences each are available. The lists were optimized for perceptual equivalence between lists, i.e., speech recognition scores do not depend on the test list used in the measurements.

TABLE 1 | Summary of the measurement settings including different combinations of signal-noise-ratio (SNR) and reverberation time (T60).

STI	Settings	Condition 1	Condition 2	Condition 3	Condition 4
0.17	SNR [dB]	−10	0	7	(∞)
	T60 [s]	(0)	4.45	7.71	9.38
0.30	SNR [dB]	−6	0	7	(∞)
	T60 [s]	(0)	1.63	3.25	4.06
0.43	SNR [dB]	−2	—	—	(∞)
	T60 [s]	(0)	—	—	2.03

In addition, two test lists of 20 sentences each from the German matrix sentence test (in Germany known as Oldenburg sentence test) were used to assess the influence of "a priori knowledge" about the speech material on speech recognition in reverberation. Speech recognition was measured without informing the listener about the structure of the test and with no training with the speech material. These sentences have a fixed grammatical structure and limited speech material of 50 words. Each word occurs in the test list exactly twice.

For the GÖSA, four different measurement conditions were included. In condition 1, only the influence of masking noise was considered. The noise had been generated by multiple, randomly time-shifted superpositions of sentences from the target talker and, hence, the long-term spectrum of the target material and the noise were very similar. Speech and noise were mixed at SNRs of −10, −6, and −2 dB to obtain the desired STI values of 0.17, 0.30, and 0.43, respectively. They were adapted from the study of Rennies et al. (2014) and corresponded to low, medium and high speech recognition scores. Conditions 2 and 3 included the combined influence of noise and reverberation. The signals were mixed at an SNR of 0 dB (condition 2) and 7 dB (condition 3), and the reverberation time was adapted to obtain the desired STI values (0.17 and 0.30, i.e., the same two lowest values as employed in condition 1). In condition 4, only reverberation was used as a detrimental factor. The reverberation times were chosen such that the same STI values were obtained as in the condition 1. The experimental settings are summarized in **Table 1**.

For the measurements with the German matrix test, two conditions with a reverberation time of 9.38 and 4.06 s were used resulting in the two lowest STI values used in this study (0.17 and 0.30, respectively). These conditions reflect the situations in which the robustness of the German matrix test was most prominent (Rennies et al., 2014). Pilot studies resulted in speech recognition scores for the reverberation time of 9.38 close to zero so that it can be assumed that presentation of one test list in this condition does not give sufficient possibility to get trained to the speech material. In other words, the second measurement with a reverberation time of 4.06 s can be considered as untrained measurement with no a priori knowledge about the sentence structure and linguistic content.

For all conditions containing reverberation, speech and noise (if applicable) were convolved with the desired RIR. To generate the RIR, white noise was multiplied with an exponential decay corresponding to the desired reverberation time. The length of the RIR was equal to the reverberation time. The same method of RIR generation was used by George et al. (2010) and Rennies et al. (2014) which makes it possible to define and vary the T60 in a systematic way.

The signals were calibrated to dB SPL using Brüel&Kjær instruments (artificial ear type 4153, microphone 4134, preamplifier 2669, and amplifier 2610). In the measurements, the speech level was fixed at 55 dB SPL and the level of the noise was varied to obtain the desired SNR. The masking noise was turned on 500 ms before and turned off 500 ms after sentence presentation. All signals were digitally preprocessed in MATLAB and the measurements were administered using the Oldenburg Measurement Application software (HörTech GmbH, Germany). The signals were run through an RME DIGI 96/8 PAD 24bit sound device and converted to analog signals (RME 4 ADI-8 Pro). The analog signals were then amplified by a TDT HB7 headphone amplifier and presented diotically through Sennheiser HD650 headphones in a sound attenuating booth (fulfilling the requirements of ANSI/ASA S3.1-1999, R2008).

Procedure

A constant stimulus-level method was used in all measurements. For each measurement condition, one test list of 20 sentences was used. The order of the measurement conditions with GÖSA was fully randomized. The two tests conducted with the German matrix test were presented between the 3rd and the 7th measurement with GÖSA, the exact order was randomized, but the most difficult condition with a reverberation time of 9.38 s was always presented before the condition with reverberation time of 4.03 s. This was done to exclude the possibility of training to the speech material. Overall, 12 different conditions were tested (10 with GÖSA, 2 with the German matrix test). The listeners' task was to repeat the understood words. The experimenter marked the correct responses. Word scoring was used meaning that each word in a sentence was judged separately as correct or incorrect. The percentage of correct responses was used as a measure of speech recognition.

Statistical Analysis

Speech recognition data were transformed using the rationalized arcsine transform (Studebaker, 1985) since recognition scores for the lowest and highest STI were close to 0 and 100%, respectively. The statistical tests were done on the transformed data. Non-parametric Friedman rank tests and Wilcoxon tests for pairwise comparisons were used since, in some of the conditions, the data were not normally distributed as indicated by Kolmogorov-Smirnov tests. If appropriate, the Wilcoxon test was used for post-hoc analysis to further explore the sources of significance—in this case the significance level of 0.05 was adjusted using Bonferroni corrections.

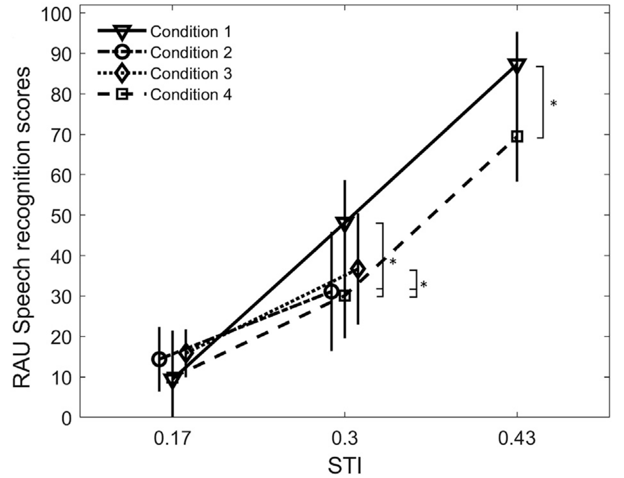


FIGURE 1 | Median RAU speech recognition scores for the GÖSA with corresponding interquartile ranges for different STIs and the measurement conditions listed in **Table 1**. Stars indicate statistically significant differences across measurements conditions.

RESULTS

Speech Recognition for Everyday Sentences

The median rationalized arcsine unit (RAU) scores with corresponding interquartile ranges for all conditions (listed in **Table 1**) measured with everyday sentences are shown in **Figure 1**. As expected, speech recognition scores were lowest for the STI of 0.17 and highest for the STI of 0.43. No statistically significant differences were found across the measurement conditions for the lowest STI [$\chi^2(3) = 3.51$, $p = 0.32$]. The median scores for the lowest STI averaged across all four conditions was 13.6% with an interquartile range of 9.1%. For the highest STI, the median score and corresponding interquartile range were 87.3 and 8.1% in condition 1, and 69.4 and 11.2% in condition 4, respectively. These differences were statistically significant [$Z = -3.3$, $p = 0.001$]. Statistically significant differences were found also across measurement conditions for the medium STI value [$\chi^2(3) = 28.11$, $p <$ 0.001], where median scores were 48.1, 31.1, 36.7, and 30.1% in conditions 1 to 4, respectively, and interquartile ranges varied from 7% (condition 1) to 14.7% (condition 2). Pairwise comparisons (with a significance level of 0.008) showed statistically significant differences in recognition score between condition 1 (noise only at an SNR of -6 dB) and condition 4 (reverberation only with T60 = 4.06 s; p = 0.001), condition 1 and condition 2 (SNR = 0 dB and T60 = 1.63 s, $p = 0.001$), condition 2 and condition 3 (SNR = 7 dB and T60 = 3.25 s, $p = 0.004$), and condition 3 and 4 ($p = 0.002$). Note that a consistent difference in speech scores across respective STI indicates that the measured effect of reverberation differs from the effect of noise on speech recognition even though no difference is predicted by the respective STI.

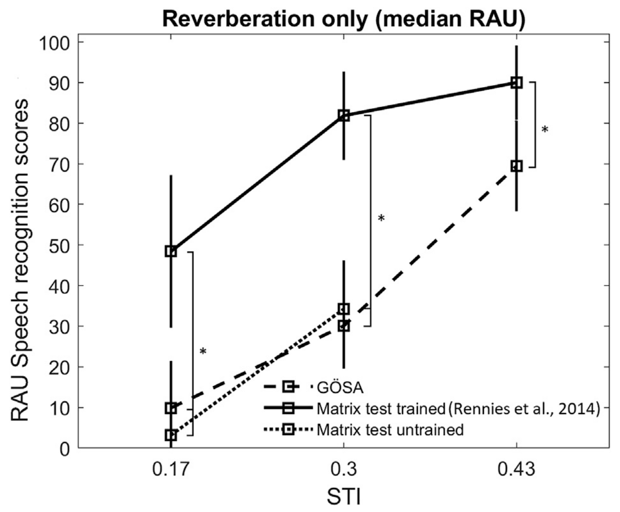


FIGURE 2 | Median RAU recognition scores and corresponding interquartile ranges for everyday sentence test (GÖSA), trained German matrix test (data from Rennies et al., 2014), and untrained German matrix test including only reverberation as detrimental factor. Stars indicate statistically significant differences across measurements conditions.

Influence of Speech Material on Speech Recognition in Noise and Reverberation

Median scores with corresponding interquartile ranges for measurements in reverberation (no noise) with GÖSA and untrained matrix sentences (present study) as well as trained matrix sentences (data from Rennies et al., 2014, transformed in RAU) are shown in **Figure 2**. Statistically significant differences across measurement conditions were found for each STI with $p < 0.001$. For the lowest and medium STI values, pairwise comparisons (with a significance level of 0.016) showed statistically higher scores for the trained matrix sentences than for the untrained matrix sentences as well as for the GÖSA (all comparisons with $p = 0.001$). No differences between GÖSA and untrained matrix sentences were found for both STIs (with $p = 0.041$ for both comparisons). Higher scores for trained matrix sentences than for the GÖSA were also confirmed for the highest STI ($p < 0.001$). The median scores of trained matrix sentences were 45.2 percentage points higher than for the untrained matrix sentences at an STI of 0.17, and 47.7 percentage points higher at an STI of 0.30. This indicates a strong effect of a priori knowledge (top-down processes) on speech recognition in reverberant conditions. The results of speech recognition in reverberation with untrained matrix sentences (no a priori knowledge about the speech material) are comparable with the outcomes of everyday sentence test (GÖSA).

Comparisons of the median recognition scores between trained matrix sentences and GÖSA for measurements in noise at different SNRs (no reverberation) are shown in **Figure 3**. The difference across the tests were significant for each STI. The median magnitude of this difference was 29.0% at an SNR of −10 dB (corresponding to an STI of 0.17, $p < 0.001$), 21.4% at an SNR of −6 dB (corresponding to an STI of 0.30, $p < 0.001$), and 5.1% at an SNR of −2 dB (corresponding to an STI of 0.43, $p = 0.001$). The relatively small difference at the highest STI probably resulted from ceiling effect observed at this SNR for both tests. The differences in speech recognition scores across the tests in noisy conditions were not expected since it is known from the literature that both tests have comparable reference speech recognition threshold in stationary noise (Kollmeier and Wesselkamp, 1997; Wagener et al., 1999; Brand et al., 2004; Warzybok et al., 2015). Possible reasons for this discrepancy will be elaborated in the discussion section.

DISCUSSION

The main aim of this research was to assess the role of different types of speech material on speech recognition in noise, in reverberation, and in combinations of noise and reverberation. The two speech material types used here, recorded with the same male talker, indicated significant differences in speech recognition even though the intelligibility should have been equal based on the STI predictions. The largest differences were observed in the conditions with reverberation as the only detrimental factor. In all reverberant conditions, the speech material of the German matrix test showed strong robustness (after training), i.e., the recognition scores were significantly higher than for the GÖSA. Moreover, comparing the outcomes of the measurements with the trained matrix test in noise and in reverberation, Rennies et al. (2014) found significantly higher speech recognition scores in reverberation (using the same white-noise RIRs as here) than in noise at STI values of 0.17 and 0.3, which is in disagreement with the STI predictions.

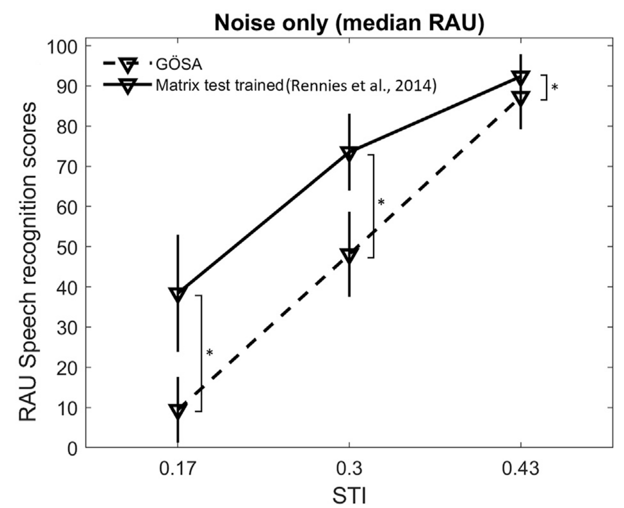


FIGURE 3 | Median RAU recognition scores and corresponding interquartile ranges for everyday sentence test (GÖSA) and trained German matrix test (data from Rennies et al., 2014) measured in stationary noise at different signal-to-noise ratios. Stars indicate statistically significant differences across measurements conditions.

The outcomes of the present study indicate that the a priori knowledge and training to the speech material has a substantial contribution to the robustness of the matrix test against reverberation. The knowledge of the sentence structure and familiarity with the limited speech material consisting of 50 words obtained within the training session resulted in much higher speech recognition scores than expected based on the STI predictions. However, the robustness of this type of speech material was not observed when the listeners were not trained prior the actual measurements. In the measurements with untrained matrix sentences, the recognition scores did not differ from the scores of the everyday sentence test. This shows the importance and contribution of the high-level top-down processes to speech recognition, which cannot be predicted by the STI since its calculations are based on the acoustic cues of the signals without consideration of top-down processes. Since each test list of the GÖSA contains unique sentences, i.e., the vocabulary differs across the lists, no training effect that would be comparable with that of the matrix-type sentences is expected for the everyday sentences. However, a strong contribution of higher-level top-down processes will be observed when the same test list is used for the second time in a short period. This is due to the context in the everyday sentences which makes these sentences easy to memorize and, in addition, enables the listener (to some degree) to guess the complete sentence from recognizing a single word. In comparison to first-time use of a meaningful sentence test, this would result in incorrectly high recognition scores. This is not the case for the matrix-type sentences, since due to their semantically unpredictable content, the sentences are difficult to memorize and there is no benefit available from sentence context.

Previous studies reported comparable speech recognition thresholds (SRTs), i.e., the SNRs corresponding to 50% speech recognition for GÖSA and the German matrix test (Kollmeier and Wesselkamp, 1997; Kollmeier et al., 2015). Hence, it was expected that the results of both tests would result in similar speech recognition scores in noise. However, the results showed higher scores for the matrix sentences than for the GÖSA. Rennies et al. (2014) reported that the good results observed in their study could be due to two extensively trained listeners participating in their experiments. To assess the impact of these listeners, we re-evaluated these data by excluding the two best listeners (corresponding to the two experienced listeners). However, the median speech recognition scores in different conditions only changed marginally (from 1 to maximally 5%) so that other reasons seem to be responsible for the good performance of the listeners in Rennies et al. (2014). Warzybok et al. (2015) and Brand et al. (2004) measured SRTs with a naïve group of normal-hearing listeners with the German matrix test and reported mean values of −6.7 and −6.8 dB, respectively. In the study of Rennies et al. (2014), the median speech recognition scores were 48% at an SNR of −10 dB and 84% at an SNR of −6 dB. This is considerably higher than the results obtained with naïve listeners by Brand et al. (2004) and Warzybok et al. (2015), supporting the assumption that listeners in the study of Rennies et al. (2014) were better than could be expected from a naïve listener panel. In contrast, Kollmeier and Wesselkamp (1997) reported a reference SRT of −6.2 dB for the GÖSA, which is in close agreement to the present data (median score of 48% at an SNR of −6 dB).

The measured data with the GÖSA can be also compared to the data from George et al. (2010) who adaptively measured SRTs with everyday sentences (VU98 corpus, Versfeld et al., 2000) in noise, in reverberation, and in combinations of noise and reverberation. George et al. (2010) assessed their listeners using sentence scoring. Because sentence scoring produces lower recognition rates than word scoring, the present data were re-calculated using sentence scoring in order to be directly comparable. Sentence-scored speech intelligibility was achieved by scoring a sentence as correct only if all the words of a sentence were repeated correctly. If the listener misunderstood one or more words, the answer was scored as incorrect. Then the number of correctly understood sentences was divided by the number of sentences presented to the listener (for GÖSA $N = 20$) and % correct responses were obtained. The re-calculation was possible since all the listener answers were digitally stored. Re-calculation was done for all conditions corresponding to STI values of 0.30 and 0.43 (the lowest STI was excluded from the comparisons since it resulted already in very low recognition scores for word scoring). For an STI of 0.30, sentence-scored medium recognition scores decreased similarly across measurement conditions and were on average 19.4% lower than scores obtained with word scoring. The median scores were 22.5% in condition 1 (SNR = −6 dB), 2.5% in condition 2 (SNR = 0 dB, T60 = 1.63 s) and 4 (T60 = 4.06 s), and 10.0% in condition 3 (SNR = 7 dB, T60 = 7.71 s). For an STI of 0.43, the median scores with sentence scoring were 90% in condition 1 (SNR = -6 dB) and 62.5% in condition 2 (T60 = 2.03 s). George et al. (2010) reported 50% speech recognition at an SNR of −3.9 dB when only noise was considered as a detrimental factor. For measurements in reverberation, 50% speech recognition was measured for T60 of 2.03 s. Considering the results of the present study, the 50% threshold in noise using sentence scoring can be estimated to be at about −4.4 dB (by interpolation) which is in line with the threshold measured by George et al. (2010) for the Dutch everyday sentence test. Sentence recognition scores in reverberation only at T60 of 2.02 s were 12.5% higher for GÖSA than the VU98 corpus, however, the significance of this difference remains unclear.

George at al. (2010) found a good correlation between the speech recognition data and STI predictions. They showed that the STI can account for the influence of noise, reverberation, and combination of both. In the study reported here, the recognition scores in reverberation only (condition 4) were significantly lower than in noise (condition 1) although the calculated STI was the same for both conditions. Hence, the detrimental influence of reverberation on speech recognition (using the same type of white-noise RIRs as used by George et al., 2010) was found to be greater than it was predicted by the STI. This effect occurred for both, sentence and word scoring methods, so that the scoring method does not seem to be the underlying reason for the observed discrepancies.

Apart from the scoring method, these two studies differ also in other aspects including the talker (female Dutch talker vs. male German talker) or speaking style (more informal for the Dutch speech material). However, the influence and interaction of these factors on the observed differences and mainly on the relative susceptibility to reverberation is unclear and could be a subject of future studies. Furthermore, the STI calculation differs slightly between the two studies. George et al. (2010) used a modified STI version (Houtgast et al., 1980) including 18 modulation frequency bands instead of the classic 14 (used in the present study). They argued that the classical STI underestimates the adverse effect of reverberation on speech intelligibility when informal, conversational speech is concerned. Systematic investigation and comparisons of STI predictions with different number of modulation frequency bands and for different types of speech material could be investigated in future studies.

CONCLUSION

In summary, it was shown that the difference between robustness to reverberation observed in the study of Rennies et al. (2014) and of George et al. (2010) may be attributed to the speech material type (closed-set matrix type sentences with high familiarity/training effect vs. unfamiliar, short meaningful Plomp-type sentences) because our listeners basically exhibited the same difference for comparable speech materials. The impact of the speech material type seems to be stronger at low and medium STI values, corresponding to poor and medium speech intelligibility than for high STI resulting in very good speech intelligibility and being limited by a ceiling effect. The robustness of the closed-set matrix type sentences against reverberation disappeared when listeners had no a priori knowledge about the speech material (sentence structure and words used).

This provides some evidence about the relative importance of high-level, top-down processing strategies in difficult reverberation situations. It remains unclear if the same applies for situations with interfering noise without reverberation. Further studies are needed with a direct comparison within the same subjects to assess the importance of bottom-up and top-down processing across different acoustic conditions.

Nevertheless, the consistent difference between reverberation- and noise-induced recognition scores of everyday sentences for medium and high STI conditions and the differences between Matrix-type and Plomp-type sentence scores clearly demonstrate the limited utility of the STI for predicting speech recognition in conditions with varying susceptibility to noise and/or reverberation.

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethical Committee of the Universität Oldenburg. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AW contributed to the study design, data collection, statistical analyses, and paper draft. JR was involved in study design, data analysis and manuscript preparation. BK contributed to the study design and the manuscript.

FUNDING

The project was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 352015383—SFB 1330 A1 and A5 and research project "Model-based multilingual rehabilitative audiology" (No. 32543918).

ACKNOWLEDGMENTS

The authors thank Henning Schepker for the valuable comments and contribution to the experimental design.

REFERENCES

ANSI/ASA S3.1-1999 (R2008). *Maximum permissible ambient noise levels for audiometric test rooms.* Washington, USA: American National Standards Institute.

Brand, T., Wittkop, T., Wagener, K., and Kollmeier, B. (2004). *Vergleich von Oldenburger Satztest und Freiburger Wörtertest als geschlossenen Versionen (Comparison of Oldenburg Sentence test and Freiburger word test in the closed-set response format).* Leipzig, Germany: Proceedings of 7th Congress of the German Society of Audiology.

Duquesnoy, A. J., and Plomp, R. (1980). Effect of reverberation and noise on the intelligibility of sentences in cases of presbyacusis. *The J. Acoust. Soc. America* 68, 537–544. doi:10.1121/1.384767

George, E. L. J., Goverts, S. T., Festen, J. M., and Houtgast, T. (2010). Measuring the effects of reverberation and noise on sentence intelligibility for hearing-impaired listeners. *J. Speech Lang. Hear. Res.* 53, 1429–1439. doi:10.1044/1092-4388(2010/09-0197)

Hochmuth, S., Jürgens, T., Brand, T., and Kollmeier, B. (2015). Talker- and language-specific effects on speech intelligibility in noise assessed with bilingual talkers: Which language is more robust against noise and reverberation? *Int. J. Audiol.* 54, 23–34. doi:10.3109/14992027.2015.1088174

Houtgast, T., Steeneken, H. J. M., and Plomp, R. (1980). Predicting speech intelligibility in rooms from the modulation transfer function: I. General room acoustics. *Acustica* 46, 59–72.

IEC (2003). "Sound System Equipment - Part 16: Objective rating of speech intelligibility by speech transmission index," in *International Standard IEC 60268-16* (Geneva, Switzerland: International Electrotechnical Commission)).

Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M., Uslar, V. N., Brand, T., et al. (2015). The multilingual matrix test: principles, applications and comparison across languages – a review. *Int. J. Audiol.* 54 (Suppl. 2), 3–16. doi:10.3109/14992027.2015.1020971

Kollmeier, B., and Wesselkamp, M. (1997). Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *J. Acoust. Soc. America* 102, 2412–2421. doi:10.1121/1.419624

Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J. Acoust. Soc. America* 95, 1085–1099. doi:10.1121/1.408469

Plomp, R., and Mimpen, A. M. (1979). Speech-reception threshold for sentences as a function of age and noise level. *J. Acoust. Soc. America* 66, 1333–1342. doi:10.1121/1.383554

Rennies, J., Schepker, H., Holube, I., and Kollmeier, B. (2014). Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *J. Acoust. Soc. America* 136 (5), 2642–2653. doi:10.1121/1.4897398

Steeneken, H. J. M., and Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *J. Acoust. Soc. America* 67, 318–326. doi:10.1121/1.384464

Studebaker, G. A. (1985). A "Rationalized" Arcsine Transform. *J. Speech Lang. Hear. Res.* 28, 455–462. doi:10.1044/jshr.2803.455

Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *J. Acoust. Soc. America* 107, 1671–1684. doi:10.1121/1.428451

Wagener, K., Brand, T., and Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil III: Evaluation des Oldenburger Satztests (Development and evaluation of a German speech intelligibility test. Part III: Evaluation of the Oldenburg sentence test). *Z. Audiol.* 38, 86–95.

Warzybok, A., Brand, T., Wagener, K. C., and Kollmeier, B. (2015). How much does language proficiency by non-native listeners influence speech audiometric tests in noise?. *Int. J. Audiol.* 54 (Suppl. 2), 88–99. doi:10.3109/14992027.2015.1063715

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 14 September 2021
doi: 10.3389/fbuil.2021.688484

Remote Working in the COVID-19 Pandemic: Results From a Questionnaire on the Perceived Noise Annoyance

Giuseppina Emma Puglisi, Sonja Di Blasio, Louena Shtrepi and Arianna Astolfi*

Department of Energy, Politecnico di Torino, Turin, Italy

Noisiness in the working environment was largely proved to have effects on the working activity and performance. To limit the spreading of the COVID-19 pandemic in the first wave between March and May 2020, Italian workers had massively started performing remote working. Insights on the subjective perception of noise annoyance under the remote working settings were thus necessary. Workers from a university and from several large and small Italian companies, resulting in 1,934 participants overall, answered to a questionnaire on the perception of noise annoyance in the remote working environment. A total of 57% of the responding workers stated to be sensitive to noise. The questionnaire was delivered online; data were recorded anonymously and then aggregated for statistical analyses. Results show that 55% of the workers perform their activity in an isolated room of the home environment, 43% in a shared room (e.g., kitchen, living room), and 2% in an outdoor space, with the majority of workers (57%) performing activity without other people in the environment. Among the noise sources investigated, 25% of workers recognize the noise generated by people (e.g., talking, moving, calling, listening to music) as the main source of disturbance. The negative consequences of noise annoyance during the remote working hours are mainly related to a loss of concentration and to a difficulty in relaxing. Furthermore, workers reported to get easily irritated by noise generated from the neighborhoods or from the housemates as it tends to distract from finishing a task.

Keywords: well-being, noise annoyance, office acoustics, remote working, noise sensitivity

OPEN ACCESS

Edited by:
Bin Yang,
Xi'an University of Architecture and Technology, China

Reviewed by:
Yongzheng Yao,
China University of Mining and Technology, Beijing, China
Simone Torresin,
University of Trento, Italy

****Correspondence:***
Giuseppina Emma Puglisi
giuseppina.puglisi@polito.it

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *30 March 2021*
Accepted: *30 July 2021*
Published: *14 September 2021*

Citation:
Puglisi GE, Di Blasio S, Shtrepi L and Astolfi A (2021) Remote Working in the COVID-19 Pandemic: Results From a Questionnaire on the Perceived Noise Annoyance. Front. Built Environ. 7:688484. doi: 10.3389/fbuil.2021.688484

INTRODUCTION

During the COVID-19 pandemic, Italian workers have been pushed to perform remote working to limit the increasing number of infections from the virus, especially within the first wave in March-May 2020, with a progressively growing portion of population working from home. According to a recent report from the Italian National Institute of Statistics (ISTAT-Istituto Nazionale di Statistica, 2020), remote working was engaged by 8% of microenterprises (i.e., with 3-9 workers), 19% of small-size enterprises (i.e., 10–49 workers), 50% of medium-size enterprises (i.e., 50–249 workers), 77% of large-size enterprises (i.e., more than 250 workers). Such a change in the working organization has thus brought a great number of people to live and work in the same location, that is, at home. Recent studies highlighted the quantity of positive aspects related to remote working, such as improved working performance, cutting of traveling costs, saving of time, and increasing of employee satisfaction (Barbuto et al., 2020; Thulin et al., 2020). From an acoustic point of view, an investigation approach based on indoor soundscape (Torresin et al., 2020a) revealed a number of positive effects on remote working and people's well-being. As an example, Torresin et al. (2021) reviewed a number of studies that proved the significant association between perceived soundscapes rated as positive and 1) a faster recovery from stress, 2) better self-reported health conditions, and 3) higher self-reported well-being. Furthermore, they found a positive and significant association between perceived well-being and comfortable-rated soundscapes from people who worked from home, particularly when a little component of "content" (i.e., so-defined as an empty or full of content continuum) was present in the environment. However, there are several detrimental effects on the perceived well-being related to the remote working condition. Several authors have shown the increase of stress, discomfort, and anxiety during the remote working, especially because of a continuous usage of technology (Tarafdar et al., 2010; Salanova et al., 2013; Molino et al., 2019) and of an increased sedentary behavior associated with longer sitting and screen time (McDowell et al., 2020). Furthermore, working from home involves different job routines alternated to family needs: frequent changes in the working process, and various cognitive tasks in turn, were proved to lead to additional negative effects such as a sense of frustration and feeling guilt (Spagnoli et al., 2020). In this framework, the features of a built environment may play a critical role in health. Amerio et al. (2020) reviewed several studies that proved the onset of mood lability, depressive and anxiety symptoms, alcohol abuse, irritability, in subjects who have experienced COVID-19 quarantine. In their study based on the administration of online surveys, they observed that a poor quality view from the inside to the outside of a house was associated with moderate and severe depressive symptoms, and that this combination of factors generated a significant loss in working performance. Andargie et al. (2021) observed in a study on Canadian remote workers that the problem of insulation from airborne sounds and impact noises in buildings was the major cause of annoyance during working hours.

Together with personal issues, effects on the working activity may be due to recurrent environmental conditions; however, the extent to which they act in the remote working setting is largely unexplored. Before the pandemic, in fact, studies have mainly focused on understanding how noise from neighbors could annoy dwellers, revealing that noise from neighboring flats is the second most relevant source of noise (32% of answers) when staying at home (WHO-World Health Organization, 2007) right after traffic (38% of answers). Therefore, it is reasonable to investigate on the effect of noises in the living environment once people are asked to stay at home also for working.

When working from a specific workplace, it is fundamental to recognize the noise sources that mainly cause disturbance and degrade productivity (D'Orazio et al., 2019). In shared and open-plan offices the noise that is generated from colleagues who converse, laugh, or talk at the phone (i.e., the so-called irrelevant speech) was found to be one of the main causes of annoyance and reduced productivity, and of growing of symptoms related to mental health and well-being (Di Blasio et al., 2019). In addition, having a positive acoustic environment when working has consequences on the perceived comfort and well-being, in fact noisy offices may bring to frequent headaches (Pejtersen et al., 2006; Kaarlela-Tuomaala et al., 2009), loss of concentration (Banbury and Berry, 2005; Pejtersen et al., 2006; Kaarlela-Tuomaala et al., 2009) and motivation (Jahncke et al., 2011), general sense of stress (Evans and Johnson, 2000; WHO-World Health Organization, 2000). Under such conditions, speech intelligibility, and so the ability of understanding words with a reduced involvement of cognitive resources, is strongly challenged too (Colle and Welsh, 1976; Hongisto et al., 2007; Haapakangas et al., 2014; Schlittmeier and Liebl, 2015; D'Orazio et al., 2018). As a remote working environment typically hosts a whole family or house-mates group, irrelevant speech can thus be considered one of the main causes of noise annoyance as it happens in proper workplaces. Nevertheless, understanding which are other recurrent and annoying noise sources may help practitioners in contributing to the acoustic design of homes also to support remote working premises.

Following an article by the authors that focused on the effects of irrelevant speech noise in offices of different sizes (Di Blasio et al., 2019), the aim of this work is to extend outcomes to the environments where remote working is performed as its practice is getting more and more common. The approach adopted in the present study may only partially represent the actual perception of indoor soundscapes during the remote working situation, as it does not account for the possible positive features related to the environment. However, the authors aimed at investigating the annoyance generated by noise from the very beginning. In particular, the aim of the present study is twofold: 1) to investigate the effects of noise on the perceived annoyance, productivity, mental health, and well-being; 2) to assess the relationship between noise annoyance, subjective and environmental characteristics.

METHODOLOGY

Participants

Workers were recruited *via* online questionnaire in May 2020. The overall response rate corresponded to 20%, which is in agreement with other works that used the same method of online questionnaires administration (Nulty, 2008). Owing to incomplete answers, a final number of 1,934 respondents were considered who belonged to one university (n = 1,104), one large size company (n = 731), 25 research and development units related to university (n = 59), five research centers (n = 10), and 19 small-size companies (n = 30).

The respondents' characterization with the description of the information related to the city where the remote working was experienced, age, gender, professional sectors, and number of people in the working environment and in the overall living environment too is provided in **Table 1**.

Concerning the remote working environment, respondents were grouped into five clusters based on the most recurrent typology of room/space where they were placed, that is, in a separate environment (e.g., a room where the worker could isolate herself or himself; 54.6% of the answers), in a shared environment

TABLE 1 | Main characteristics of the total sample (*n* = 1934) given overall.

Background information		Overall	
		n	%
Gender	Female	1,127	58
	Male	807	42
Age range	18–25	74	4
	26–35	509	26
	26–50	534	28
	51–65	783	40
	65+	34	2
Nationality	Italian	1889	98
	Other	45	2
Remote working city	Northern Italy	1,560	81
	Central Italy	122	6
	Southern Italy	228	12
	Other	24	1
Professional sector	Technical	381	20
	Engineering	490	25
	Management	164	8
	Administration	445	23
	Creative, design and architecture	123	6
	Other	80	4
	Teaching	40	2
	Researcher	97	5
	Sales and public affairs	48	2
	Teaching and researcher	37	2
	Services	29	1
Number of people in the remote working environment (yourself excluded)	0	1,100	57
	1–2	743	38
	3–4	85	4
	5+	6	0
Number of people in the overall living environment (yourself excluded)	0	284	15
	1–2	1,061	55
	3–4	547	28
	5+	42	2

(e.g., an open space, the kitchen, the living room; 43.1% of the answers), outdoor (e.g, a balcony, a terrace, a garden; 0.8% of the answers), other environments (1.1% of the answers), a mix of the above (when they reported to change repeatedly the workstation in the house; 0.4% of the answers).

The location of the city in which the company/research center/university were settled was asked as there can be sociocultural differences whether it is in the North or in the South of the country (Carboni and Russu, 2018). In the present work, however, the location of the city where the remote working activity was performed was considered, as sociocultural factors could also be mixed with geographical premises that led, for instance, to keep windows opened due to weather conditions that brought to an increased perception of noise from the outside rather than from the inside.

Questionnaire

The questionnaire was designed according to a previous work (Di Blasio et al., 2019). It was delivered online, and the participation was given on a voluntary base: an email was sent with a link that directed to the compilation of a Google Form approved by the head of the human resources of each company. The email was accompanied by an invitation letter with a brief explanation to inform the workers about the aim of the study, the confidential treatment of their personal data, and the anonymity of the answers.

An overall number of 22 questions with close answer were included in the questionnaire, which was delivered both in Italian and in English to reach the majority of workers possible. The compilation time was estimated in approximately 2 minutes, as questions were designed to be very easy to be read and the possible answers were either organized on a 5-point scale or on a single-choice selection.

Seven questions were related to the respondents' background about gender, age, nationality, company name, location of the city of the company, city of remote working, professional sector, and two more questions were asked with regards to the number of people in the working environment and in the overall living environment—excluding the respondent herself or himself—as already reported in

TABLE 2 | Questionnaire layout.

Topic	ID	Question	Scale	Label(s)
Annoyance	Q1	How much does noise annoy you during your smart working activity	5	Not at all (1) - Extremely (5)
Mental health well-being (feelings and symptoms)	Q2	What is the main feeling (or symptom) related to noise during your remote working activity	Single choice	• Stress • Negative feeling such as feeling displeased • Negative feelings toward other housemates • Loss of concentration • Anger • Loss of motivation • Headache • Tiredness and overstrain • None • Other
		How much do you agree with the following statements?		
Work productivity	Q3	Noise often interrupts me during my smart working activity	5	Strongly disagree (1)—Strongly agree (5)
	Q4	Noise does not allow me to work as much as I would like during my remote working activity		
	Q5	Noise significantly reduces my work performance during my remote working activity		
Mental health well-being (interpersonal relationships)	Q6	Noise during my remote working activity compromises the harmony at home	5	Strongly disagree (1)—Strongly agree (5)
Occupants' behavior (personal strategies)	Q7	What is the main strategy that you use to reduce the annoyance resulting from noise during your remote working activity?	Single choice	• Take a break • Change work task • Headphones with music • Noise cancelling headphones • Open the window • Close the window • Change room • Close the room door • Plan the return to office • Ask people to reduce their voice volume • None • Other
		How much do you agree with the following statements?		
Occupants' sensitivity and reaction to noise	Q8	I am sensitive to noise	5	Strongly disagree (1)—Strongly agree (5)
	Q9	I find it hard to relax in a place that is noisy		
	Q10	I get mad at people who make noise that keeps me from falling asleep or getting work done		
	Q11	I get annoyed when my neighbors are noisy		
	Q12	I get used to most noises without much difficulty		
Noise source perception	Q13	What is the main source of noise present during your remote working activity?	Single choice	• Technological noise (household appliances, systems, television, tablets) • Traffic (vehicular, rail, air) • Sirens (ambulances and firefighters) • Anthropic noise generated by children under the age of 5 years, anthropic noise generated by children aged 6–13 years • Anthropic noise generated by adults (conversations, video calls, physical activity, music) • Noise from own pets • Noise of nature (chirping of birds, noise from neighborhood animals, wind, water) • Neighborhood noise (trampling, shouting, loading, and unloading of goods, music)

Table 1. Then, 12 questions that were specifically oriented to assess the relationship between noise sensitivity, annoyance, well-being, and work productivity were given and are summarized in **Table 2**. A last question in **Table 2**, i.e., Q13, was added to ask for the main perceived source of noise during the remote working activity.

TABLE 3 | Mean and mode values of the answers given by nonsensitive (*n* = =358) to noise respondents on noise annoyance (Q1), work productivity (Q3, Q4, Q5), and mental health and well-being (Q6) for different typologies of remote working environments, considering. Two-tailed *p*-values of significance for the differences between across environments are reported according to the Kruskal Wallis (KW) Test. Specific significant differences (*p*-value < 0.10) between environments are given as mean values in italics as a result of the application of the Mann Whitney U Test.

ID	Separate		Shared		Outdoor		Other		Mix of environments		KW *p*-value
	(*n* = =198)		(*n* = =151)		(*n* = =3)		(*n* = =6)		(*n* = =0)		
	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode	
Noise annoyance											
Q1	*1.69*	2.00	*1.83*	2.00	1.67	2.00	2.50	1.00	-	-	0.09
Work productivity											
Q3	*1.76*	1.00	*2.00*	2.00	1.67	2.00	2.50	1.00	-	-	0.06
Q4	1.71	1.00	1.83	2.00	1.67	2.00	2.33	1.00	-	-	0.31
Q5	1.66	1.00	1.82	1.00	1.00	1.00	2.00	1.00	-	-	0.13
Mental health and well-being (interpersonal relationships)											
Q6	1.72	1.00	1.77	1.00	1.67	1.00	2.00	1.00	-	-	0.68

Each question, the Likert ranking based on the 5-point scale, and the list of items in the single-choice option were all appropriately defined based on previous studies. Particularly, the article by Di Blasio et al. (2019) and its references were used as main baseline to design the presented questionnaire. Furthermore, specific questions on noise sensitivity were added based on Senese et al. (2012) who validated the Italian version of the "Weinstein Noise Sensitivity Scale (WNSS)."

Answers related to mental health and well-being (Q2) were grouped as suggested in Di Blasio et al. (2019). In particular, clusters turned into: loss of concentration, mental illness (stress), emotional and social feelings (negative feeling such as feeling displeased, negative feelings toward other housemates, anger, loss of motivation), physical symptoms (headache, tiredness, and overstrain), none, others.

Answers related to personal strategies due to occupants' behavior (Q7) were grouped, again as suggested in Di Blasio et al. (2019). In particular, clusters turned into: use of technological tools (headphones with music, noise cancelling headphones), use of adaptive behaviors (take a break, change work task, open the window, close the window, change room, close the room door, plan the return to office), asking people to reduce their voice, none, other, mix of the above. As in the category "other" several answers frequently recurred, the following clusters were added, that is, working in different time, using earplugs, listening to music.

Statistical Analyses

Statistical analyses were performed using SPSS (IBM Statistics20, IBM, Armonk, NY, United States). Nonparametric tests were applied to the database that contained data measured with ordinal and nominal scales, according to Sigel and Castellan (1988). Answers given to the questions related to the perceived noise annoyance, work productivity, mental health, and well-being, were separated per each type of environment were respondents performed their remote working activity were considered (i.e., separate room, shared room, outdoor space, other, mix of environments). Then, comparisons were performed first with the Kruskal Wallis (KW) test, which fits the comparison of results for more than two groups of observation, and then with the Mann-Whitney U Test (MWU), which is used to compare two groups of independent observations. The KW test was also applied to the dataset to investigate how noise annoyance is related to different age ranges, professional sectors, and number of people in an office. The relationship between noise annoyance and gender was instead investigated using the MWU test.

According to question Q8 (see **Table 2**) related to the respondents' perceived noise sensitivity, answers were both considered overall and also grouped into two clusters. In such a second case, if an answer was given with rating 1 or 2, respondents were considered "non-sensitive to noise," vice versa if an answer was given with ratings 3–5, respondents were considered "sensitive to noise." The dichotomization of the sample was applied to understand the extent to which the perception of being sensitive to noise could affect the answers related to the influence of noise on annoyance, productivity, mental health, and well-being (see **Tables 3, 5**). Although an overall noise sensitivity index could be obtained as an average among items (Senesi et al., 2012; Aletta et al., 2018), this work only focused on considering the way in which a participant self-perceived of being sensitive to noise. The other items from Q9 to Q12, in fact, were related to features of reaction to noise and were not accounted for an average. In general, this approach in the data analysis is based on past studies that have highlighted the need of including the sensitivity to noise as a factor (Stansfeld et al., 2021). Furthermore, it allowed exploring the answers given to the perceived noise annoyance (Q1) considering the sample's general characteristics of gender, age range, professional sector, and city of remote working (see **Tables 7, 8**), and the number of people that the respondents shared the working or living environment with (see **Tables 10** and **11**). Based on the aforementioned dichotomization criterion, among the 1934 respondents 1,576 and 358 reported to be sensitive and nonsensitive to noise, respectively.

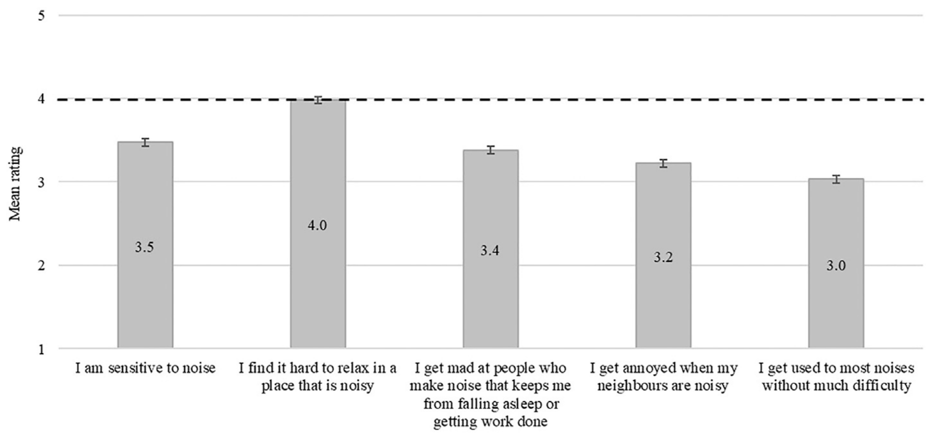


FIGURE 1 | Mean values of the sample's noise sensitivity with respect to specific questions. Error bars represent the 95% confidence interval (scores from 1, strong disagreement, to 5, strong agreement). The dashed black line corresponds to the mode value obtained for each question (i.e., rating 4).

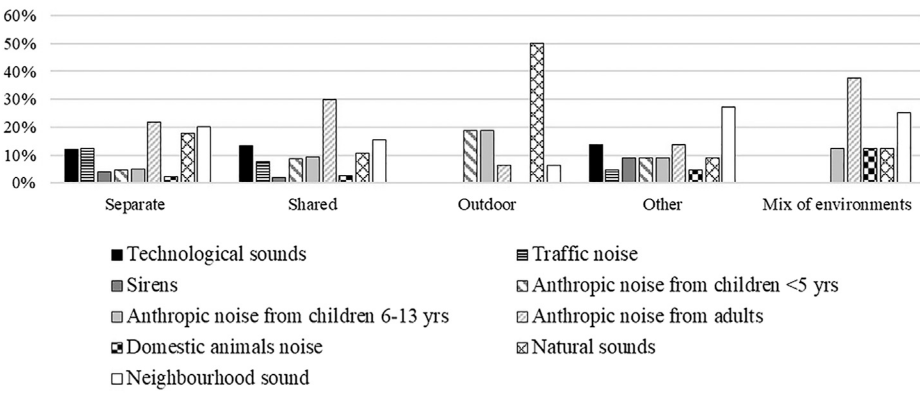


FIGURE 2 | Percentages of main sound source perceived during the remote working hours across all 1,934 respondents, divided per environment where the remote working activity takes place.

RESULTS

Figure 1 shows the mean values, with errors whiskers in terms of 95% confidence interval, that workers gave to questions Q8–Q12 related to occupants' sensitivity, i.e., the general reaction of occupants to noise. As rating 1 was labeled as "strongly disagree" and rating 5 as "strongly agree," for questions Q8–Q11 the higher the mean values, the higher can be considered the annoyance due to noise. As the mode values for all these questions consisted in rating 4, it means that respondents result to be significantly annoyed by noise (Q8) and to find hard relaxing in a noisy environment (Q9). At the same time, as far as Q12 is concerned, the mode value consisting in rating four means that respondents get used to most noises without difficulties although they perceive of being annoyed by them.

As far as the types of noise that respondents were immersed in, the questionnaire aimed at understanding which was the main noise source during the remote working hours. **Figure 2** reports the percentages of the answers acquired, which are homogeneously distributed across options based on the remote working environment in which the activity takes place and slightly well represents a correspondence between environment and noise (i.e., the "natural sounds" are reported to be the main source of noise by those who work from outdoor spaces). As separate or shared rooms are the principal environments for remote working, it is interesting that the "anthropic noise generated by adults" is there reported to be the main noise source (22 and 30% in separate and shared rooms, respectively). This outcome, in fact, allows considering such a source as the main one to be controlled.

Effects of Noise on the Perceived Annoyance, Productivity, Mental Health, and Well-Being

Table 4 shows the results of the extent to which respondents were subjected to noise annoyance (Q1) and the way they perceived an

TABLE 4 | Mean and mode values of the answers given by sensitive (n = 1,576) to noise respondents on noise annoyance (Q1), work productivity (Q3, Q4, Q5), and mental health and well-being (Q6) for different typologies of remote working environments, considering. Two-tailed p-values of significance for the differences between across environments are reported according to the Kruskal Wallis (KW) Test. Specific significant differences (p-value < 0.05) between environments are given as mean values in italics as a result of the application of the Mann Whitney U Test.

ID	Separate		Shared		Outdoor		Other		Mix of environments		KW p-value
	(n = 857)		(n = 682)		(n = 13)		(n = 16)		(n = 8)		
	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode	
Noise annoyance											
Q1	*2.19*	2.00	*2.46*	2.00	2.23	2.00	2.25	1.00	2.25	2.00	0.00
Work productivity											
Q3	*2.33*	2.00	*2.75*	3.00	2.31	2.00	2.44	1.00	2.63	2.00	0.00
Q4	*2.33*	2.00	*2.69*	2.00	2.77	2.00	2.25	1.00	2.38	3.00	0.00
Q5	*2.28*	2.00	*2.57*	2.00	2.31	2.00	2.06	1.00	2.00	1.00	0.00
Mental health and well-being (interpersonal relationships)											
Q6	*2.18*	1.00	*2.49*	2.00	2.23	1.00	2.25	2.00	2.25	1.00	0.00

TABLE 5 | Mean and mode values of the answers on noise annoyance (Q1), work productivity (Q3, Q4, Q5), and mental health and well-being (Q6) for different typologies of remote working environments. Two-tailed p-values of significance for the differences between across environments are reported according to the Kruskal Wallis (KW) Test. Specific significant differences (p-value < 0.05) between environments are given as mean values in italics as a result of the application of the Mann Whitney U Test.

ID	Separate		Shared		Outdoor		Other		Mix of environments		KW p-value
	(n = 1,055)		(n = 833)		(n = 16)		(n = 22)		(n = 8)		
	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode	Mean	Mode	
Noise annoyance											
Q1	*2.09*	2.00	*2.34*	2.00	2.13	2.00	2.32	1.00	2.25	2.00	0.00
Work productivity											
Q3	*2.22*	2.00	*2.61*	2.00	2.19	2.00	2.45	1.00	2.63	2.00	0.00
Q4	*2.22*	1.00	*2.53*	2.00	2.56	2.00	2.27	1.00	2.38	3.00	0.00
Q5	*2.16*	1.00	*2.43*	2.00	2.06	2.00	2.05	1.00	2.00	1.00	0.00
Mental health and well-being (interpersonal relationships)											
Q6	*2.10*	1.00	*2.36*	2.00	2.13	1.00	2.18	2.00	2.25	1.00	0.00

effect of noise on work productivity (Q3, Q4, and Q5), mental health, and well-being (Q6). Results are given as mean and mode values divided per each environment in which respondents perform their remote working activity. To understand if answers were significantly different among environments, the KW test was first applied overall and, whenever it resulted statistically significant (p-value < 0.05), followed by the MWU test across couples to have specific insights.

All the investigated aspects (Q1, Q3, Q4, Q5, and Q6) resulted to have statistically significant differences whenever respondents performed their remote working activity in a separate or in a shared room of the living environment. Mean values were higher in the case of shared room in the living environment, resulting in a higher degree of perceived noise annoyance, of reduction of work productivity, and of reduction of harmony in the interpersonal relationships at home.

As far as the clustering of the sample is concerned, the answers were also divided per subjective sensitivity to noise (Q8) and the results are reported in **Tables 3, 5** for nonsensitive and sensitive respondents, respectively. Although the trend is having higher mean values, thus a higher degree of perceived noise annoyance and reduction of productivity and well-being when the remote working activity is performed in a shared environment rather than in a separate one, only sensitive subjects revealed that such a difference is statistically significant (p-value < 0.05). Answers from nonsensitive respondents were statistically significant with a p-value < 0.10 only in the case of perceived noise annoyance and feeling of need to interrupt the working activity under noisy conditions.

Noise Annoyance, Subjective and Environmental Characteristics

As the main aspect under investigation was related to the perceived noise annoyance during the remote working activity in the COVID-19 pandemic, this section reports the answers to

TABLE 6 | Mean and standard deviation (St.dev.) values of the answers on noise annoyance (Q1) related to gender, age range, professional sector, and remote working city for different types of environment of the overall sample (*n* = 1934). Two-tailed *p*-values of significance of the differences according to the Mann Whitney U (MWU) or Kruskal Wallis (KW) Test are reported. Statistically significant differences with *p*-values < 0.05 are reported in bold.

–	Separate			Shared			Outdoor			Other			Mix of environments		
	Mean		**St.dev.**	**Mean**		**St.dev.**	**Mean**		**St.dev.**	**Mean**		**St.dev.**	**Mean**		**St.dev.**
Gender															
Female	2.05		0.92	2.35		0.94	1.90		1.20	2.44		1.42	2.00		1.00
n		665			440			10			9			3	
Male	2.16		0.93	2.34		0.93	2.50		1.05	2.23		1.48	2.40		0.55
n		390			393			6			13			5	
MWU *p*-value		0.05			0.85			0.22			0.70			0.57	
Age range															
18–25	2.42		0.78	2.05		0.69	2.00		1.41	-		-	-		-
n		52			20			2			-			-	
26–35	2.33		0.97	2.38		0.87	2.00		1.41	1.67		0.58	2.50		0.71
n		296			204			4			3			2	
36–50	2.18		0.98	2.57		0.99	2.67		1.21	2.67		1.37	2.20		0.84
n		268			249			6			6			5	
51–65	1.83		0.80	2.20		0.92	1.50		0.58	2.31		1.60	2.00		0.00
n		412			353			4			13			1	
65+	1.93		1.14	1.86		0.69	-		-	-		-	-		-
n		27			7			-			-			-	
KW *p*-value		**0.00**			**0.00**			0.36			0.55			0.81	
Professional sector															
Technical	1.84		0.82	2.25		0.96	1.80		0.45	2.33		1.75	2.50		0.71
n		201			167			5			6			2	
Engineering	2.39		0.95	2.48		0.92	2.50		1.73	2.67		2.08	-		-
n		307			176			4			3			-	
Management	1.80		0.84	2.13		0.92	1.00		0.00	2.50		1.52	-		-
n		85			72			1			6			-	
Administration	1.83		0.78	2.31		0.91	2.00		1.41	2.33		1.53	2.50		0.71
n		192			246			2			3			2	
Creative, design and architecture	2.47		1.01	2.67		1.04	2.00		1.41	-		-	2.50		0.71
n		74			45			2			-			2	
Other	2.00		0.93	2.48		0.79	-		-	-		-	-		-
n		57			23			-			-			-	
Teaching	2.27		0.98	2.29		1.05	-		-	2.00		0.00	-		-
n		22			17			-			1			-	
Researcher	2.51		0.95	2.63		0.93	-		-	2.00		0.00	1.50		0.71
n		51			43			-			1			2	
Sales and public affairs	1.79		0.92	2.05		0.76	-		-	-		-	-		-
n		28			20			-			-			-	
Teaching and researcher	2.14		1.04	2.45		0.82	3.00		1.41	1.50		0.71	-		-
n		22			11			2			2			-	
Services	1.88		0.72	1.62		0.65	-		-	-		-	-		-
n		16			13			-			-			-	
KW *p*-value		**0.00**			**0.00**			0.71			0.99			0.44	
Remote working city															
Northern Italy	2.14		0.92	2.35		0.92	2.20		1.15	2.17		1.34	2.33		0.52
n		865			656			15			18			6	
Central Italy	1.85		0.91	2.02		0.84	-		-	4.00		0.00	-		-
n		65			56			-			1			-	

(Continued on following page)

TABLE 6 | (*Continued*) Mean and standard deviation (St.dev.) values of the answers on noise annoyance (Q1) related to gender, age range, professional sector, and remote working city for different types of environment of the overall sample (*n* = 1934). Two-tailed *p*-values of significance of the differences according to the Mann Whitney U (MWU) or Kruskal Wallis (KW) Test are reported. Statistically significant differences with *p*-values < 0.05 are reported in bold.

–	Separate		Shared		Outdoor		Other		Mix of environments	
	Mean	St.dev.	Mean	St.dev.	Mean	St.dev.	Mean	St.dev.	Mean	St.dev.
Southern Italy	1.86	0.96	2.47	1.07	-	-	3.00	2.83	3.00	0.00
n	111		114		-		2		1	
Other	2.21	0.89	2.29	0.49	1.00	0.00	2.00	0.00	1.00	0.00
n	14		7		1		1		1	
KW *p*-value	**0.00**		0.05		0.21		0.68		0.16	

the specific question on noise annoyance (Q1) in relationship to different characteristics of the sample.

Noise Annoyance and Subjective Characteristics

Mean values and standard deviations of the noise annoyance scores were first analyzed according to the subjective characteristics of gender, age range, professional sector, and city of remote working activity. **Table 6** reports the results for the overall sample (*n* = 1934), always divided based on the environment where the remote working activity was mainly performed. Applying the proper statistical tests of KW and/or MWU, the main outcomes can be summarized as follows:

- The factor gender does not bring to statistically significant different results on the perceived noise annoyance;
- The factor age range brings to statistically significant differences on the perceived noise annoyance whether the remote working activity was performed in a separate or in a shared room of the living environment. In particular, respondents who worked in a separate room of the living environment revealed to be annoyed by noise to a greater extent if they were younger. Statistically significant differences were found for the age group 18–25 with respect to 36–50, 51–65, and 65+; then for the age group 26–35 with respect to 36–50, 51–65, and 65+; then for the age group 36–50 with respect to 51–65. A different trend was found for respondents who performed their remote working activity in a shared room of the environment, with the maximum perception of noise annoyance for the 36–50 age group. In such a case, statistically significant differences were found for the age group 18–25 with respect to 36–50; then for the age group 26–35 with respect to 36–50 and 51–65; then for the age group 36–50 with respect to 51–65 and 65+;
- The factor professional sector brings again to statistically significant differences on the perceived noise annoyance whether the remote working activity was performed in a separate or in a shared room of the living environment. In the first case, i.e., separate room, researchers exhibited the highest mean value of noise annoyance. As far as the differences are concerned, people working in the technical sector gave different answers with respect to people working in management, administration, creative/design/architecture sectors; people working in the engineering sector gave different answers with respect to people working in management, administration, creative/design/architecture; people working in the management sector gave different answers with respect to people administration. In the second case, i.e., shared room, workers in the creative, design, and architecture field reported to be the most annoyed by noise. Then, specifically, people working in the technical sector gave different answers with respect to people working in management; people working in the engineering sector gave different answers with respect to people working in management and administration;

TABLE 7 | Mean and standard deviation (St.dev.) values of the answers given by nonsensitive to noise respondents ($n = 358$) on noise annoyance (Q1) related to gender, age range, professional sector, and remote working city for different types of environment. Two-tailed *p*-values of significance of the differences according to the Mann Whitney U (MWU) or Kruskal Wallis (KW) Test are reported. Statistically significant differences with *p*-values < 0.05 are reported in bold.

	Separate		**Shared**		**Outdoor**		**Other**		**Mix of environments**	
	Mean	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**
Gender										
Female	1.65	0.65	1.89	0.65	1.50	0.71	3.25	1.26	-	-
n	133		87		2		4		-	
Male	1.75	0.71	1.77	0.68	2.00	-	1.00	0.00	-	-
n	65		64		1		2		-	
MWU *p*-value	0.38		0.24		0.48		0.06		-	
Age range										
18–25	1.93	0.62	1.80	0.45	-	-	-	-	-	-
n	14		5		-		-		-	
26–35	1.77	0.67	1.95	0.71	-	-	1.00	-	-	-
n	52		41		-		1		-	
36–50	1.77	0.75	1.96	0.70	2.00	0.00	3.00	0.00	-	-
n	48		46		2		2		-	
51–65	1.57	0.61	1.66	0.61	1.00	-	2.67	2.08	-	-
n	79		58		1		3		-	
65+	1.20	0.45	2.00	-	-	-	-	-	-	-
n	5		1		-		-		-	
KW *p*-value	0.06		0.15		0.16		0.40		-	
Professional sector										
Technical	1.56	0.55	1.81	0.75	2.00	-	2.67	2.08	-	-
n	41		31		1		3		-	
Engineering	1.86	0.64	1.85	0.49	2.00	-	1.00	-	-	-
n	56		39		1		1		-	
Management	1.48	0.60	1.82	0.81	1.00	-	3.00	0.00	-	-
n	21		17		1		2		-	
Administration	1.56	0.56	1.74	0.67	-	-	-	-	-	-
n	36		34		-		-		-	
Creative, design and architecture	2.17	0.98	2.00	0.82	-	-	-	-	-	-
n	6		4		-		-		-	
Other	1.61	0.70	2.13	0.64	-	-	-	-	-	-
n	18		8		-		-		-	
Teaching	1.67	0.58	1.67	-	-	-	-	-	-	-
n	3		6		-		-		-	
Researcher	2.00	1.10	2.60	-	-	-	-	-	-	-
n	6		5		-		-		-	
Sales and public affairs	1.40	0.55	1.67	-	-	-	-	-	-	-
n	5		3		-		-		-	
Teaching and researcher	2.25	1.50	2.00	-	-	-	-	-	-	-
n	4		1		-		-		-	
Services	2.00	0.00	1.33	-	-	-	-	-	-	-
n	2		3		-		-		-	
KW *p*-value	0.20		0.40		0.37		0.40		-	
Remote working city										
Northern Italy	1.75	0.67	1.90	0.63	2.00	0.00	2.00	1.15	-	-
n	162		118		2		4		-	
Central Italy	1.64	0.63	1.40	0.52	-	-	-	-	-	-
n	14		10		-		-		-	
Southern Italy	1.21	0.42	1.64	0.79	-	-	5.00	-	-	-
n	19		22		-		1		--	
Other	1.67	1.15	3.00	-	1.00	-	2.00	-	-	-
n	3		1		1		1		-	
KW *p*-value	**0.01**		**0.01**		0.16		0.32		-	

people working in the management sector gave different answers with respect to people administration and creative/design/architecture;

- The factor remote working city brings to statistically significant differences on the perceived noise annoyance whether the remote working activity was performed in a

TABLE 8 | Mean and standard deviation (St.dev.) values of the answers given by sensitive to noise respondents (n = 1,576) on noise annoyance (Q1) related to gender, age range, professional sector, and remote working city for different types of environment. Two-tailed p-values of significance of the differences according to the Mann Whitney U (MWU) or Kruskal Wallis (KW) Test are reported. Statistically significant differences with p-values < 0.05 are reported in bold.

		Separate		Shared		Outdoor		Other		Mix of environments	
		Mean	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**
Gender											
	Female	2.15	0.95	2.47	0.97	2.00	1.31	1.80	1.30	2.00	1.00
	n	532		353		8		5		3	
	Male	2.25	0.95	2.45	0.93	2.60	1.14	2.45	1.51	2.40	0.55
	n	325		329		5		11		5	
	MWU *p*-value	0.15		0.79		0.25		0.28		0.51	
Age range											
	18–25	2.61	0.75	2.13	0.74	2.00	1.41	-	-	-	-
	n	38		15		2		-		-	
	26–35	2.45	0.98	2.48	0.87	2.00	1.41	2.00	0.00	2.50	0.71
	n	244		163		4		2		2	
	36–50	2.27	1.00	2.70	0.99	3.00	1.41	2.50	1.73	2.20	0.84
	n	220		203		4		4		5	
	51–65	1.90	0.82	2.30	0.94	1.67	0.58	2.20	1.55	2.00	-
	n	333		295		3		10		1	
	65+	2.09	1.19	1.83	0.75	-	-	-	-	-	-
	n	22		6		-		-		-	
	KW *p*-value	**0.00**		**0.00**		0.49		0.79		0.81	
Professional sector											
	Technical	1.91	0.86	2.35	0.98	1.75	0.50	2.00	1.73	2.50	0.71
	n	160		136		4		3		2	
	Engineering	2.51	0.96	2.66	0.93	2.67	2.08	3.50	2.12		
	n	251		137		3		2			
	Management	1.91	0.89	2.22	0.94	1.00	-	2.25	1.89		
	n	64		55		1		4			
	Administration	1.89	0.81	2.41	0.91	3.00	-	2.33	1.53	2.50	0.71
	n	156		212		1		3		2	
	Creative, design and architecture	2.50	1.01	2.73	1.05	2.00	1.41	-	-	2.50	0.71
	n	68		41		2		-		2	
	Other	2.18	0.97	2.67	0.82	-	-	-	-	-	-
	n	39		15		-		-		-	
	Teaching	2.37	1.01	2.64	1.12	-	-	2.00	-	-	-
	n	19		11		-		1		-	
	Researcher	2.58	0.92	2.63	0.94	-	-	2.00	-	1.50	0.71
	n	45		38		-		1		2	
	Sales and public affairs	1.87	0.97	2.12	0.78	-	-	-	-	-	-
	n	23		17		-		-		-	
	Teaching and researcher	2.11	0.96	2.50	0.85	3.00	1.41	1.50	0.71	-	-
	n	18		10		2		2		-	
	Services	1.86	0.77	1.70	0.68	-	-	-	-	-	-
	n	14		10		-		-		-	
	KW *p*-value	**0.00**		**0.00**		0.59		0.86		0.44	
Remote working city											
	Northern Italy	2.23	0.95	2.45	0.94	2.23	1.24	2.21	1.42	2.33	0.52
	n	703		538		13		14		6	
	Central Italy	1.90	0.96	2.15	0.84	-	-	4.00	-	-	-
	n	51		46		-		1		-	
	Southern Italy	2.00	0.98	2.67	1.04	-	-	1.00	-	3.00	-
	n	92		92		-		1		1	
	Other	2.36	0.81	2.17	0.41	-	-	-	-	1.00	-
	n	11		6		-		-		1	
KW *p*-value		**0.00**		**0.03**		NA		0.29		0.16	

TABLE 9 | Mean and standard deviation (St.dev.) values of the answers on noise annoyance (Q1) related to the number of people in the remote working environment and in the overall living environment, considering the overall sample (n = 1934). Two-tailed p-values of significance of the differences according to the Kruskal Wallis (KW) Test are reported. Statistically significant differences with p-values < 0.05 are reported in bold.

		Separate		Shared		Outdoor		Other		Mix of environments	
		Mean	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**
Number of people in the remote working environment (yourself excluded)	0	2.04	0.92	1.94	0.83	1.86	1.07	1.83	1.11	1.67	0.58
	n	843		235		7		12		3	
	1–2	2.29	0.92	2.44	0.89	2.29	1.38	2.33	1.75	2.60	0.55
	n	190		535		7		6		5	
	3–4	2.63	0.96	3.08	1.10	2.50	0.71	3.33	0.58	-	-
	n	19		61		2		3		-	
	5+	1.67	0.58	2.00	0.00	-	-	5.00	0.00	-	-
	n	3		2		-		1		-	
	KW p-value	**0.00**		**0.00**		0.51		0.11		0.07	
Number of people in the overall living environment (yourself excluded)	0	1.97	0.95	1.96	0.87	1.00	0.00	1.50	0.53	2.00	0.00
	n	176		94		2		10		2	
	1–2	2.07	0.89	2.31	0.91	2.17	0.98	2.33	2.31	2.00	1.00
	n	563		486		6		3		3	
	3–4	2.19	0.95	2.54	0.96	2.38	1.30	3.38	1.41	2.67	0.58
	n	294		234		8		8		3	
	5+	2.59	1.14	2.74	0.87	-	-	2.00	0.00	-	-
	n	22		19		-		1		-	
	KW p-value	**0.01**		**0.00**		0.18		0.07		0.42	

separate or in a shared room of the living environment. For people working in a separate room, being in northern Italy brought to a greater higher degree of perceived noise annoyance with respect to central and southern Italy respondents. For people working in shared rooms, statistically significant differences were found for northern vs. central Italy and for central vs. southern Italy.

Sensitive and Nonsensitive to Noise Subjects

Considering the clustered sample, mean values and standard deviations of the noise annoyance scores were then analyzed for respondents who revealed to be nonsensitive (n = 358, **Table 7**) or sensitive (n = 1,576, **Table 8**) to noise. In both the cases, the factor gender did not bring to statistically significant differences, when the remote working activity was performed neither in a separate nor in a shared room of the living environment.

Then, in the case of nonsensitive to noise respondents (n = 358, **Table 7**), the smaller number of cases considered did not lead to statistically significant differences on perceived noise annoyance for the factors age range and professional sector too. The factor remote working city showed differences for northern and southern Italy, and for northern and central and southern Italy whether the respondent used to work in a separate or shared room of the living environment, respectively.

In the case of sensitive to noise respondents (n = 1,576, **Table 8**), the main outcomes can be summarized as follows:

- The same results in terms of significant differences outlined for the overall sample were found for the factor age range. In particular, statistically significant differences on the perceived noise annoyance were found whether the remote working activity was performed in a separate or a shared room of the living environment;
- The factor professional sector revealed some slight differences. Overall, in the case of people working in a separate room, researchers exhibited the highest mean value of noise annoyance. Then, people working in the technical sector gave different answers with respect to people working in engineering, creative/design/architecture, teaching, and research sectors; people working in the engineering sector gave different answers with respect to people working in management, administration, other, sales and public affairs, services; people working in the management sector gave different answers with respect to people working in creative/design/architecture, research; people working in the administration sector gave different answers with respect to people working in creative/design/architecture, teaching, research; people working in the creative/design/architecture sector gave different answers with respect to people working in sales and public affairs, services; people working in other sectors gave different answers with respect to people working in research; people working in the research sector gave different answers with respect to people working in sales and public affairs, teaching and research, services. For people working in a shared room, workers in the creative, design, and architecture field reported to be the most annoyed by noise. Then, people working in the technical sector gave different answers with respect to people working in engineering and services; people working in engineering gave different answers with respect to people working in management, administration, sales and public affairs, services; people working in the management sector gave different answers with respect to people working in creative/design/

TABLE 10 | Mean and standard deviation (St.dev.) values of the answers given by nonsensitive to noise respondents (*n* = 358) on noise annoyance (Q1) related to the number of people in the remote working environment and in the overall living environment. Two-tailed *p*-values of significance of the differences according to the Kruskal Wallis (KW) Test are reported. Statistically significant differences with *p*-values < 0.05 are reported in bold.

		Separate		**Shared**		**Outdoor**		**Other**		**Mix of environments**	
		Mean	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**
Number of people in the remote working environment (yourself excluded)	0	1.66	0.64	1.69	0.62	1.00	-	1.50	0.71	-	-
	n	160		48		1		2		-	
	1–2	1.71	0.67	1.86	0.62	2.00	-	1.00	-	-	-
	n	35		93		1		1		-	
	3–4	3.00	1.00	2.30	1.06	2.00	-	3.00	0.00	-	-
	n	3		10		1		2		-	
	5+	-	-	-	-	-	-	5.00	-	-	-
	n	-		-		-		1		-	
	KW *p*-value	**0.04**		0.10		0.37		0.20		-	
Number of people in the overall living environment (yourself excluded)	0	1.58	0.71	1.88	0.62	1.00	-	1.00	-	-	-
	n	48		16		1		1		-	
	1–2	1.67	0.64	1.78	0.64	2.00	-	-	-	-	-
	n	100		88		1		-		-	
	3–4	1.80	0.68	1.86	0.68	2.00	-	2.80	1.48	-	-
	n	46		42		1		5		-	
	5+	2.00	0.82	2.40	1.14	-	-	-	-	-	-
	n	4		5		-		-		-	
	KW *p*-value	0.24		0.50		0.37		0.23		-	

TABLE 11 | Mean and standard deviation (St.dev.) values of the answers given by sensitive to noise respondents (*n* = 1,576) on noise annoyance (Q1) related to the number of people in the remote working environment and in the overall living environment. Two-tailed *p*-values of significance of the differences according to the Kruskal Wallis (KW) Test are reported. Statistically significant differences with *p*-values < 0.05 are reported in bold.

		Separate		**Shared**		**Outdoor**		**Other**		**Mix of environments**	
		Mean	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**	**Mean**	**St.dev.**
Number of people in the remote working environment (yourself excluded)	0	2.13	0.95	2.01	0.86	2.00	1.10	1.90	1.20	1.67	0.58
	n	683		187		6		10		3	
	1–2	2.43	0.92	2.56	0.89	2.33	1.51	2.60	1.82	2.60	0.55
	n	155		442		6		5		5	
	3–4	2.56	0.96	3.24	1.05	3.00	-	4.00	-	-	-
	n	16		51		1		1		-	
	5+	1.67	0.58	2.00	0.00	-	-	-	-	-	-
	n	3		2		-		-		-	
	KW *p*-value	**0.00**		**111111110.00**		0.58		0.44		0.07	
Number of people in the overall living environment (yourself excluded)	0	2.11	0.99	1.97	0.91	1.00	-	1.56	0.53	2.00	0.00
	n	128		78		1		9		2	
	1–2	2.15	0.92	2.42	0.92	2.20	1.10	2.33	2.31	2.00	1.00
	n	463		398		5		3		3	
	3–4	2.26	0.97	2.69	0.95	2.43	1.40	4.33	0.58	2.67	0.58
	n	248		192		7		3		3	
	5+	2.72	1.18	2.86	0.77	-	-	2.00	-	-	-
	n	18		14		-		1		-	
	KW *p*-value	0.07		0.00		0.44		0.10		0.42	

architecture, research; people working in the services sector gave different answers with respect to people working in technical, engineering, administration, creative/design/architecture, other, research;

- The factor remote working city brings once more to statistically significant differences on the perceived noise annoyance whether the remote working activity was performed in a separate or in a shared room of the living environment. Indeed, the same results in terms of significant differences outlined for the overall sample were found.

Noise Annoyance and Number of People in the Environment

As a second goal of the data analysis, mean values and standard deviations of the noise annoyance scores were then analyzed according to the number of people who were present either in the

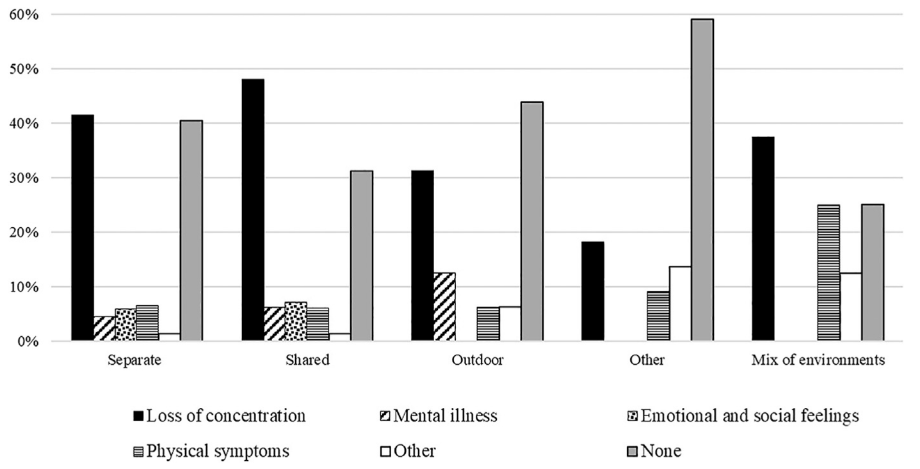


FIGURE 3 | Percentages of the subjective ratings of the effects of noise in the different remote working environments on mental health and well-being (i.e., feelings and symptoms).

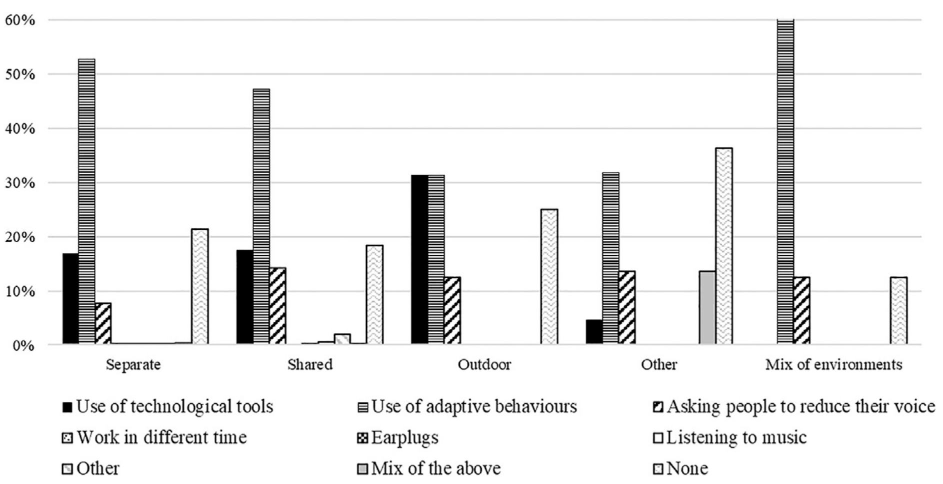


FIGURE 4 | Percentages of the subjective ratings of the effects of noise in the different remote working environments on occupants' behavior (i.e., strategies to cope with noise).

same room of remote working or in the overall living environment. **Table 9** reports the results for the overall sample. Overall, the greater was the number of people in the working environment or in the general living environment, the greater was the perceived noise annoyance.

Considering the factor of number of people in the remote working environment, statistically significant differences in the case of separate room for 0 and 1–2, 3–4 groups, and in the case of shared room for 0 and 1–2, 3–4 groups, and for 1–2 and 3–4 groups. As far as the number of people in the overall living environment is concerned, statistically significant differences in the case of separate room for 0 and 3–4, 5 + groups, and for 1–2 and 5 + groups, and in the case of shared room for 0 and 1–2, 3–4, 5 + groups, and for 1–2 and 3–4, 5 + groups.

Sensitive and Nonsensitive to Noise Subjects

Similarly, also for the number of people in the environments the sample was split based on the respondents' sensitivity to noise.

In the case of nonsensitive to noise respondents (n = 358, **Table 10**), consider the factor of number of people in the remote working environment, statistically significant differences in the case of separate room for 0 and 3–4 groups, and 1–2 and 3–4 groups. As far as the number of people in the overall living environment is concerned, no statistically significant differences among groups were found.

In the case of sensitive to noise respondents (n = 1,576, **Table 11**), consider the factor of number of people in the remote working environment, statistically significant differences in the case of separate room for 0 and 1–2 groups, and 1–2 and 3–4 groups, and in the case of shared room for 0 and 1–2, 3–4 groups. As far as the number of people in the overall living environment is concerned, statistically significant differences in the case of shared room for 0 and 5 + groups, and for 1–2 and 5 + groups.

Occupants' Behavior

Questions Q2 and Q7 were designed to understand the perceived mental health well-being, in terms of symptoms generated by noise annoyance, and the actions that respondents were willing to make to reduce annoyance, respectively.

Figure 3 shows that, overall, remote workers either do not complain about noise and report of not feeling symptoms (39.9% of the respondents) or complain about feeling a loss of concentration (35.3%). Symptoms related to mental illness (i.e., stress) and emotional/social feelings (i.e., negative feeling such as feeling displeased, negative feelings toward other housemates, anger, loss of motivation) are reported by only workers who perform remote working activity in a separate or shared room (less than 10% of respondents anyway) and, in the first case, in outdoor spaces (12.5% of respondents).

Figure 4 reports the main strategy used to reduce the annoyance resulting from noise during remote working activity. Less than 23% of the respondents does not adopt any strategy to this aim, and on average the 14 and 12% of respondents either use technological tools to mask or cancel noise (e.g., wearing headphones) or ask people to reduce their voice, respectively. The majority of respondents (47.6%) reported to use adaptive behaviors to actively solve the problem of noise annoyance in the working hours. In particular, the main strategies adopted were related to take a break, change room/ environment, switch between working tasks, or interact with the environment by opening/closing the windows to change the environment's soundscape.

DISCUSSION

Several studies carried out before and during the pandemic period of COVID-19 revealed a significant decrease in outdoor measured noise levels, as strong measures such as the "stay at home" strategy were taken by the National Governments to limit the spreading of the infection. Aletta et al. (2020a) measured a decrease by about 65% of the use of vehicles in a study on noise mapping in Rome (IT) during the pandemic period. Bartalucci et al. (2020), on a weekly basis, assessed a reduction up to 10 dB in terms of L_{den} during a long-term monitoring of traffic noise in Monza (IT). An average decrease in L_{den} by 5 dB was reported by Hornberg et al. (2021), who reviewed a number of studies related to noise level decreases during and after the pandemic period. In a study by the Soundscape and noise observatory of Greater Lyon (Acoucité, 2020), 21 monitoring stations spread in five French cities measured outdoor noise before and during the lockdown period due to the pandemic, revealing a reduction in L_{den} by up 6 dB in the weekdays and up to 9 dB in the weekends. At the same time, this latter study also investigated on the dwellers' perception of the sound environment during the lockdown: they gave more positive attributes to their impression, such as "calm," "pleasant," "peaceful," and thus revealed a link between the soundscape perception and composition. As before the pandemic period traffic noise was typically predominant in the cities, during the lockdown the hierarchy of sound sources has been reversed allowing for more anthropic sounds to be heard (Aletta et al., 2020b; Sakagami, 2020). Furthermore, Manzano et al. (2021) report a significant shift from human-generated (e.g., traffic) and anthropic sourcers to animal and natural sources. Şentop Dümen and Şaher (2020) stressed the negative effect of anthropic sounds as they observed, from the outcomes of an online survey administered in Turkey to 1,053 subjects, that annoyance from noise generated by neighbors did not change significantly before and during the pandemic period, whereas annoyance from noise generated by dwellers significicantly increased. These studies corroborate at several levels the findings of the present study, as the categories of sources related to "anthropic noise," "natural sounds," and "neighborhood sounds" were the most perceived at home during the remote working hours. On the contrary, "traffic noise" and "sirens" were reported to be the predominantly perceived noise source by less than 15% of respondents on average.

As the soundscape of remote working environments, which correspond to living environments from the spreading of COVID-19 pandemic, has profoundly changed in the last year, it is necessary to account for the new main sound sources perceived to provide an effective home design. In light of this, noise control should not be the only approach, but should be integrated with a perceptual and multisensory perspective as suggested by Torresin et al. (2020b), also considering participatory design practices that account for the dweller/worker premises to enable the complex building–user interrelations.

Perceived Noise Annoyance, Productivity, Mental Health, and Well-Being

To the authors' knowledge, only few studies investigated the problem of noise annoyance under a remote working setting. However, it is possible to make a comparison with the outcomes from studies related to noise annoyance in offices as in both settings the predominantly perceived noise source was found to be of anthropic nature.

Overall, remote workers were most annoyed by noise (Q1) when they performed their working activity in a shared space of the house than in a separate environment. This outcome was confirmed also splitting the sample based on their sensitivity to noise (Q8), and corroborated studies in which noise annoyance was assessed as higher in shared and larger offices than in smaller ones (Danielsson, 2005; Di Blasio et al., 2019).

As far as the perception of productivity is concerned, workers who performed activity in shared environments of the house reported a higher sense of loss. In summary, they perceived more that noise interrupts them during the remote working activity, does not allow them to work as much as they would, and reduces their working performance. Again, this outcome confirms past works by Di Blasio et al. (2019) and is similar to what Kaarlela-Tuomaala et al. (2009) found in relation to the perceived feeling of wasting time and loose productivity when changing workspace from a private to an open-plan office.

In relation to mental health and well-being perception, the questionnaire distinguished queries on the base of issues related to feelings and symptoms (Q2) and interpersonal relationships (Q6). As far as symptoms are concerned, despite that about an average 40% of respondents reported to have no experience of feelings/symptoms related to noise during the remote working hours, the main consequence of noise indicated by approximately 35% of respondents, on average, was a loss of concentration. This outcome is in agreement with other studies that highlighted a significant increase in deconcentration during the working hours (Banbury and Berry, 2005; Kaarlela-Tuomaala et al., 2009; Di Blasio et al., 2019) and getting even worse as the office size increased itself (Pejtersen et al., 2006). As far as the interpersonal relationships are concerned, workers from shared environments in the house reported a sense of compromising the harmony at home to a statistically significant greater extent than workers from separate environments. These findings are in agreement with Brennan et al. (2002) and Di Blasio et al. (2019), who reported difficulties and less satisfaction in co-worker relationships more when they worked in open-plan and large offices than in shared, private, and small ones.

Noise Annoyance, Subjective and Environmental Characteristics

No differences in noise annoyance were found with respect to gender, as both male and female respondents gave, on average, not statistically significantly different ratings to the question on noise annoyance (Q1) under the different remote working settings (i.e., separate, shared, outdoor, other, mix of environments). This result is coherent with the findings of Di Blasio et al. (2019) who did not assess any statistically significant difference in gender when analyzing noise annoyance in shared rooms, to which the remote working settings of separate or shared environments at home can be compared.

As far as the analysis of age-range in relation to noise annoyance is concerned, working in a separate environment of the house was related to a higher degree of perceived noise annoyance for younger respondents. Working in a shared environment, instead, was related to the highest higher degree of perceived noise annoyance in the 36–50 years of age range. These outcomes seem to be in contradiction to other studies, in which a dependency of annoyance and age was assessed with elder workers being more annoyed by noise than younger (Pierrette et al., 2015; Di Blasio et al., 2019). However, the difference in the results can be due to the different sizes of the participant samples considered in the studies and also to the fact that the working environment is not exactly the same, so some comparisons can be done but not all the results can be matched between situations. Indeed, the outcome of the present study related to noise annoyance for workers in shared environments of the house, which revealed a greater annoyance in subjects of 36–50 years, corresponds to a situation in which the number of people and even of children in the whole home is higher, thus noise annoyance can depend on other psychological aspects too (e.g., the need of answering to the other premises and a major request to switch between cognitive and practical tasks). Furthermore, this outcome is in agreement with Van Gerven et al. (2009) who found middle-aged subjects—peaking around 45 years of age—to be the most annoyed by noise in a transversal study across all the lifespan, regardless of the noise exposure level and of the individually perceived noise sensitivity.

The professional sector to which respondents belonged to revealed differences, again, when workers performed the remote activity in separate or shared environments of the house. In particular, "researchers" and workers in the field of "creative, design and architecture" were mostly annoyed by noise in separate and shared environments, respectively. Overall, it is not possible to establish a comparison with other studies, as different work categories were either used or group sizes were available. Therefore, future works should establish more similar categories related to the professional sector, maybe introducing a clustering related to the predominant cognitive task carried out. As an example, anthropic noise was found to annoy workers performing mathematical tasks by Logie and Baddeley (1987). Associating, for instance, engineers and technicians with such a cognitive task could help in finding more evident trends.

In relation to the location of the cities where remote working was performed, respondents from northern Italy were most annoyed by noise if they worked from a separate environment of the house. On the opposite, when considering a shared environment of the house, workers were most annoyed by noise if they were in southern Italy. This outcome needs to be deepened, maybe performing a repeated assessment *via* questionnaire in different periods of the year to understand if weather issues influence this answer or if such perception is recurrent based on location.

Last, considering the number of people in the working environment and in the overall living environment, the same significant trend was found, as expected: the more people were present either in the environment or in the whole house, the more respondents were annoyed by noise.

Occupants' Behavior

It is worth giving an insight into the outcomes related to the potential involvement of occupants to increase well-being and reduce noise annoyance under remote working settings. To this aim, Q7 was designed to understand whether a worker had an inclination to activate personal strategies to reduce noise annoyance. Less than one-fifth of the respondents do not adopt any strategy. The 14% of respondents use technological tools to mask or cancel noise (e.g., wearing headphones), whereas the 12% of respondents actively ask other mates to reduce their voice level to keep high focus on the working task. The majority of responses was interestingly concentrated on "use adaptive behaviors" to find the most adequate soundscape to perform working (e.g., change room, switch between working tasks, interact with the environment to change its soundscape); therefore, remote workers are interested in being active part of the occupant–environment relationship to solve a problem and increase the sense of well-being in it. A similar result was obtained by Di Blasio et al. (2019) who found that workers from shared offices, which are almost comparable for size and occupation to the shared environments of the house, reported to adopt active strategies to reduce noise annoyance.

STRENGTHS AND LIMITATIONS OF THE STUDY

A strength related to the present work is its capability of giving insights on a real condition in which a great portion of workers is asked to perform nowadays. Literature has focused on the understanding of the perception of the physical office environment in terms of air quality, thermal, visual, and acoustic comfort, even providing practitioners and researchers with important information on how to optimize the offices' design. This emergency, however, gives a good opportunity to ameliorate the design of indoor environments too, both in terms of sound insulation of buildings (Andargie et al., 2021) and in terms of indoor soundscapes, to support both living and working premises. In light of this latter aspect, there are very recent studies such as the one of Torresin et al. (2020a) that define some initial discussions that can be integrated with the present outcomes to build a perceptual, multisensory, and well-being-oriented design paradigm.

A main drawback of the study is related to the adoption of a negative connotation of noise as a criterion at the base of the questionnaire. The reason for this was related to the willing of extending outcomes from a similar study and to making the obtained ones as comparable as possible with the literature available. In the next future, however, a shift in the paradigm to an approach oriented on the indoor soundscape assessment will be foreseen, in order to account for the positive effects of the sound environment on the working activity as well as on perceived well-being. Then, another limitation of the present work relates to the little possibility of comparing its outcomes with other similar studies, as the unique condition related to the COVID-19 pandemic has taken workers to change their everyday life in a fast way and like never in history. Therefore, the main comparisons of the present study concern outcomes from investigations on offices and thus differences can still be found, or some results are difficult to be explained in depth. To this aim, it would be interesting to perform further investigations applying the same methodology that relies on online questionnaires provision to 1) increase the database of responses, 2) corroborate obtained outcomes, and 3) understand possible further changes in the occupants' subjective perception and behavior during a remote working setting that follows the emergency of one of the first pandemic periods in March-May 2020.

CONCLUSION

The aim of the present study was twofold: 1) to investigate on the effects of noise on the perceived annoyance, productivity, mental health, and well-being; and 2) to assess the relationship between noise annoyance, subjective and environmental characteristics. To this aim an online questionnaire has been administered to more than 1,934 people.

Although some of the outcomes of the present work could be expected, it is worth putting in light some aspects that should be taken into account to the aim of supporting a holistic design of home environments that are, nowadays, no longer only living but also working spaces. Indeed, as remote working seems to persist in time, results will contribute to understand the extent to which working from home can be supported by the indoor soundscape.

First, noise annoyance affects work productivity, mental health, and well-being not only in office settings but also in remote working settings, that is, when workers perform their activity from home. In particular, sharing a room—regardless of its dimensions—brings to a higher degree of perceived noise annoyance with respect to working from a separate environment in the house. Having a positive soundscape at home is thus a growing need to support several premises in one's everyday life.

Second, subjective characteristics must be taken into account when investigating the extent to which noise annoys the working activity from home. Different outcomes, in fact, were found in relation to the location of the city of remote working, as well as in relation to the age of the respondents. Further studies should better categorize respondents based on the typology of the performed working tasks (e.g., linguistic/humanistic, mathematical, technical) rather than on their specific professional sector to have a more robust clustering of the acquired data.

Third, a design approach—or practical suggestions—introducing proper spaces to be used during the remote working hours is necessary. This can be done, where possible, designing separate rooms in the house to this aim. However, when this is not possible, it would be worth integrating specific sound shields to give a greater separation of the workstation from the rest of the shared environment.

Fourth, occupant's behavior and attitude should be considered to define the ability of a built environment typically used for living, to support the intense and prolonged working activity too. This study highlighted the active behavior that workers adopt to ameliorate the soundscape of their remote working environment. The abovementioned brand-new design approach should then be supported through an integrated participatory practice that actively engages workers.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical approval was not provided for this study on human participants because the material used in this study, consisting in the online administration of an anonymous questionnaire, was approved by the Ethic Committee of Politecnico di Torino in January 2019 for a former investigation performed and published by the same authors. In the context of the spreading of the COVID-19 pandemic in May 2020, the authors used such material and received an approval from the management area of each company involved for the questionnaires' administration. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

GEP, SDB, LS, and AA contributed in the conceptualization of the study, designed and developed the questionnaire, defined the formal analysis; SDB managed the diffusion of the questionnaire and collected data; GEP applied the formal analysis and wrote the manuscript; LS and AA reviewed the article drafts; AA supervised the research activity. All the authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

The kind collaboration of the management area of each company and the participation of the workers have made this work possible.

REFERENCES

Acoucité (2020). *Soundscape and Noise Observatory of Greater Lyon, Lockdown during COVID-19 Pandemic: Impact on Sound Environment*. Summary Report June 11th.

Aletta, F., Brinchi, S., Carrese, S., Gemma, A., Guattari, C., Mannini, L., et al. (2020a). Analysing Urban Traffic Volumes and Mapping Noise Emissions in Rome (Italy) in the Context of Containment Measures for the COVID-19 Disease. *Noise Mapp.* 7, 114–122. doi:10.1515/noise-2020-0010

Aletta, F., Oberman, T., Mitchell, A., Tong, H., and Kang, J. (2020b). Assessing the Changing Urban Sound Environment during the COVID-19 Lockdown Period Using Short-Term Acoustic Measurements. *Noise Mapp.* 7, 123–134. doi:10.1515/noise-2020-0011

Aletta, F., Van Renterghem, T., and Botteldooren, D. (2018). Influence of Personal Factors on Sound Perception and Overall Experience in Urban green Areas. A Case Study of a Cycling Path Highly Exposed to Road Traffic Noise. *Ijerph* 15 (1118), 1–17. doi:10.3390/ijerph15061118

Amerio, A., Brambilla, A., Morganti, A., Aguglia, A., Bianchi, D., Santi, F., et al. (2020). COVID-19 Lockdown: Housing Built Environment's Effects on Mental Health. *Ijerph* 17 (5973), 1–10. doi:10.3390/ijerph17165973

Andargie, M. S., Touchie, M., and O'Brien, W. (2021). Case Study: A Survey of Perceived Noise in Canadian Multi-Unit Residential Buildings to Study Long-Term Implications for Widespread Teleworking. *Building Acoust.*, 1351010X2199374. in press. doi:10.1177/1351010X21993742

Banbury, S., and Berry, D. (2005). Office Noise and Employee Concentration: Identifying Causes of Disruption and Potential Improvements. *Ergonomics* 48, 25–37. doi:10.1080/00140130412331311390

Barbuto, A., Gilliland, A., Peebles, R., Rossi, N., and Shrout, T. (2020). Telecommuting: Smarter Workplaces. Available at: http://hdl.handle.net/1811/91648 (accessed 17 June 2020)

Bartalucci, C., Borchi, F., and Carfagni, M. (2020). Noise Monitoring in Monza (Italy) during COVID-19 Pandemic by Means of the Smart Network of Sensors Developed in the LIFE MONZA Project. *Noise Mapp.* 7, 199–211. doi:10.1515/noise-2020-0017

Brennan, A., Chugh, J. S., and Kline, T. (2002). Traditional versus Open Office Design. *Environ. Behav.* 34, 279–299. doi:10.1177/0013916502034003001

Carboni, O. A., and Russu, P. (2018). Measuring and Forecasting Regional Environmental and Economic Efficiency in Italy. *Appl. Econ.* 50, 335–353. doi:10.1080/00036846.2017.1313954

Colle, H. A., and Welsh, A. (1976). Acoustic Masking in Primary Memory. *J. Verbal Learn. Verbal Behav.* 15, 17–31. doi:10.1016/S0022-5371(76)90003-7

Danielsson, C. B. (2005). *Office Environment, Health & Job Satisfaction. An Explorative Study of Design's Influence*. Stockholm, Sweden: School of ArchitectureKTH Royal Institute of Technology.

Di Blasio, S., Shtrepi, L., Puglisi, G., and Astolfi, A. (2019). A Cross-Sectional Survey on the Impact of Irrelevant Speech Noise on Annoyance, Mental Health and Well-Being, Performance and Occupants' Behavior in Shared and Open-Plan Offices. *Ijerph* 16 (280), 1–17. doi:10.3390/ijerph16020280

D'Orazio, D., Rossi, E., de Salvio, D., and Garai, M. (2019). "Measuring and Identifying Background Noises in Offices during Work Hours," in Proceedings of the 23rd International Congress on Acoustics: Integrating the 4th EAA Euroregio, 9. Aachen: Aachen, September.

D'Orazio, D., Rossi, E., and Garai, M. (2018). Comparison of Different *In Situ* Measurements Techniques of Intelligibility in an Open-Plan Office. *Building Acoust.* 25 (2), 111–122. doi:10.1177/1351010X18776431

Evans, G. W., and Johnson, D. (2000). Stress and Open-Office Noise. *J. Appl. Psychol.* 85, 779–783. doi:10.1037/0021-9010.85.5.779

Haapakangas, A., Hongisto, V., Hyönä, J., Kokko, J., and Keränen, J. (2014). Effects of Unattended Speech on Performance and Subjective Distraction: the Role of Acoustic Design in Open-Plan Offices. *Appl. Acoust.* 86, 1–16. doi:10.1016/j.apacoust.2014.04.018

Hongisto, V., Virjonen, P., and Keränen, J. (2007). "Determination of Acoustic Conditions in Open Offices and Suggestions for Acoustic Classification," in Proceedings of the 19th International Congress on Acoustics. Madrid: September 2, 7.

Hornberg, J., Haselhoff, T., Lawrence, B. T., Fischer, J. L., Ahmed, S., Gruehn, D., et al. (2021). Impact of the COVID-19 Lockdown Measures on Noise Levels in Urban Areas-A Pre/during Comparison of Long-Term Sound Pressure Measurements in the Ruhr Area, Germany. *Ijerph* 18 (4653), 1–16. doi:10.3390/ijerph18094653

ISTAT-Istituto Nazionale di Statistica (2020). Situazione e prospettive delle imprese nell'emergenza sanitaria COVID-19. Available at: https://www.istat.it/it/files//2020/12/REPORT-COVID-IMPRESE-DICEMBRE.pdf (accessed 3 February 2021)

Jahncke, H., Hygge, S., Halin, N., Green, A. M., and Dimberg, K. (2011). Open-plan Office Noise: Cognitive Performance and Restoration. *J. Environ. Psychol.* 31, 373–382. doi:10.1016/j.jenvp.2011.07.002

Kaarlela-Tuomaala, A., Helenius, R., Keskinen, E., and Hongisto, V. (2009). Effects of Acoustic Environment on Work in Private Office Rooms and Open-Plan Offices - Longitudinal Study during Relocation. *Ergonomics* 52, 1423–1444. doi:10.1080/00140130903154579

Logie, R. H., and Baddeley, A. D. (1987). Cognitive Processes in Counting. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 310–326. doi:10.1037/0278-7393.13.2.310

Manzano, J. V., Pastor, J. A. A., Quesada, R. G., Aletta, F., Oberman, T., Mitchell, A., et al. (2021). The "sound of Silence" in Granada during the COVID-19 Lockdown. *Noise Mapp.* 8 (1), 16–31. doi:10.1515/noise-2021-0002

McDowell, C. P., Herring, M. P., Lansing, J., Brower, C., and Meyer, J. D. (2020). Working from home and Job Loss Due to the COVID-19 Pandemic Are Associated with Greater Time in Sedentary Behaviors. *Front. Public Health* 8, 597619. doi:10.3389/fpubh.2020.597619

Molino, M., Cortese, C., and Ghislieri, C. (2019). Unsustainable Working Conditions: the Association of Destructive Leadership, Use of Technology, and Workload with Workaholism and Exhaustion. *Sustainability* 11, 1–14. doi:10.3390/su11020446

Nulty, D. D. (2008). The Adequacy of Response Rates to Online and Paper Surveys: What Can Be Done? *Assess. Eval. Higher Edu.* 33, 301–314. doi:10.1080/02602930701293231

Pejtersen, J., Allermann, L., Kristensen, T. S., and Poulsen, O. M. (2006). Indoor Climate, Psychosocial Work Environment and Symptoms in Open-Plan Offices. *Indoor Air* 16, 392–401. doi:10.1111/j.1600-0668.2006.00444.x

Pierrette, M., Parizet, E., Chevret, P., and Chatillon, J. (2015). Noise Effect on comfort in Open-Space Offices: Development of an Assessment Questionnaire. *Ergonomics* 58, 96–106. doi:10.1080/00140139.2014.961972

Sakagami, K. (2020). How Did the 'state of Emergency' Declaration in Japan Due to the COVID-19 Pandemic Affect the Acoustic Environment in a rather Quiet Residential Area? *Ucloe* 2 (1), 06. doi:10.14324/111.444/ucloe.000009

Salanova, M., Llorens, S., and Cifre, E. (2013). The Dark Side of Technologies: Technostress Among Users of Information and Communication Technologies. *Int. J. Psychol.* 48, 422–436. doi:10.1080/00207594.2012.680460

Schlittmeier, S. J., and Liebl, A. (2015). The Effects of Intelligible Irrelevant Background Speech in Offices - Cognitive Disturbance, Annoyance, and Solutions. *Facilities* 33, 61–75. doi:10.1108/F-05-2013-0036

Senese, V. P., Ruotolo, F., Ruggiero, G., and Iachini, T. (2012). The Italian Version of the Weinstein Noise Sensitivity Scale. *Eur. J. Psychol. Assess.* 28 (2), 118–124. doi:10.1027/1015-5759/a000099

Şentop Dümen, A., and Şaher, K. (2020). Noise Annoyance during Covid19 Lockdown: A Research of Public Opinion before and during the Pandemic. *J. Acoust. Soc. Am.* 148 (6), 3489–3496. doi:10.1121/10.0002667

Sigel, S., and Castellan, N. J. (1988). *Non Parametric Statistics for the Behavioral Sciences*. 2nd ed. New York, NY, USA: McGraw-Hill, 116–126. 184–194, ISBN-13: 978-0070573574, ISBN-10: 0070573573.

Spagnoli, P., Molino, M., Molinaro, D., Giancaspro, M. L., Manuti, A., and Ghislieri, C. (2020). Workaholism and Technostress during the COVID-19 Emergency: the Crucial Role of the Leaders on Remote Working. *Front. Psychol.* 11, 620310. doi:10.3389/fpsyg.2020.620310

Stansfeld, S., Clark, C., Smuk, M., Gallacher, J., and Babisch, W. (2021). Road Traffic Noise, Noise Sensitivity, Noise Annoyance, Psychological and Physical Health and Mortality. *Environ. Health* 20 (1), 1–24. doi:10.1186/s12940-021-00720-3

Tarafdar, M., Tu, Q., and Ragu-Nathan, T. S. (2010). Impact of Technostress on End-User Satisfaction and Performance. *J. Manage. Inf. Syst.* 27, 303–334. doi:10.2753/MIS0742-1222270311

Thulin, E., Vilhelmson, B., and Johansson, M. (2020). New Telework, Time Pressure, and Time Use Control in Everyday Life. *Sustainability* 11, 3067. doi:10.3390/su11113067

Torresin, S., Albatici, R., Aletta, F., Babich, F., Oberman, T., Siboni, S., et al. (2020a). Indoor Soundscape Assessment: A Principal Components Model of Acoustic Perception in Residential Buildings. *Building Environ.* 182 (107152), 1–16. doi:10.1016/j.buildenv.2020.107152

Torresin, S., Albatici, R., Aletta, F., Babich, F., Oberman, T., Stawinoga, A. E., et al. (2021). Indoor Soundscapes at home during the COVID-19 Lockdown in London - Part I: Associations between the Perception of the Acoustic Environment, Occupantś Activity and Well-Being. *Appl. Acoust.*, 183. Applied Acoustics, 108305. in press. doi:10.1016/j.apacoust.2021.108305

Torresin, S., Aletta, F., Babich, F., Bourdeau, E., Harvie-Clark, J., Kang, J., et al. (2020b). Acoustics for Supportive and Healthy Buildings: Emerging Themes on Indoor Soundscape Research. *Sustainability* 12, 6054. doi:10.3390/su12156054

Van Gerven, P. W. M., Vos, H., Van Boxtel, M. P. J., Janssen, S. A., and Miedema, H. M. E. (2009). Annoyance from Environmental Noise across the Lifespan. *The J. Acoust. Soc. America* 126 (1), 187–194. doi:10.1121/1.3147510

WHO-World Health Organization (2000). Mental Health and Work: Impact, Issues and Good Practices. Available at: https://apps.who.int/iris/bitstream/handle/10665/42346/WHO_MSD_MPS_00.2.pdf?sequence=12&isAllowed=y (accessed 20 March 2021).

WHO-World Health Organization (2007). Large Analysis and Review of European Housing and Health Status. Available at: https://www.euro.who.int/__data/assets/pdf_file/0007/107476/lares_result.pdf (accessed 3 February 2021).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 03 November 2021
doi: 10.3389/fbuil.2021.688395

Higher Sound Levels in K-12 Classrooms Correlate to Lower Math Achievement Scores

Laura C. Brill[1,2] *and Lily M. Wang*[1]*

[1]*Durham School of Architectural Engineering and Construction, University of Nebraska–Lincoln, Omaha, NE, United States,*
[2]*Threshold Acoustics, Chicago, IL, United States*

Sound levels from occupied classrooms have been gathered from 220 classrooms across four grade levels (3, 5, 8 and 11) over six school days each and processed with k-means clustering into speech and non-speech clusters. Three metrics describing the classroom acoustics, including the average daily A-weighted equivalent level for non-speech, the average daily difference between the A-weighted equivalent levels for speech and non-speech (a signal to noise ratio), and the mid-frequency averaged reverberation time, were analyzed against classroom-aggregated standardized reading and math achievement test scores, while controlling for classroom demographics including socioeconomic status. Interactions between the metrics and demographics were also tested. A statistically significant relationship was found between the average daily non-speech levels in classrooms and math test scores; higher daily non-speech levels were correlated with lower math test scores ($p < 0.05$). No statistically significant main effects of acoustic metrics were found on reading achievement. There were some significant differences and an interaction found between grades, but these may be due to uneven sample distributions as there were fewer grade 8 and 11 classrooms measured. Children learn in occupied classrooms, and the findings from this investigation based on data from occupied conditions suggest that designing for lower unoccupied sound levels can lead to occupied environments that are conducive to better student learning outcomes.

Keywords: classroom acoustics, sound levels, noise, reverberation time, math achievement, signal to noise ratio, classrooms, children

OPEN ACCESS

Edited by:
Chiara Visentin,
University of Ferrara, Italy

Reviewed by:
Jessica Massonnie,
University College London,
United Kingdom
Nicola Prodi,
University of Ferrara, Italy

Correspondence:
Lily M. Wang
lilywang@unl.edu

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *30 March 2021*
Accepted: *15 October 2021*
Published: *03 November 2021*

Citation:
Brill LC and Wang LM (2021) Higher Sound Levels in K-12 Classrooms Correlate to Lower Math Achievement Scores. Front. Built Environ. 7:688395. doi: 10.3389/fbuil.2021.688395

1 INTRODUCTION

Acoustic conditions in K-12 classrooms affect the clarity and ease of verbal communication and consequently are expected to impact learning, language development, and development of cognitive skills in children (Leibold, 2017). As reviewed in this section, previous work has demonstrated how conditions with higher noise levels and/or excessive reverberation are related to worse performance by primary and secondary school students on speech intelligibility, reading or listening comprehension, short-term memory, and assorted reasoning tasks. Poor acoustic conditions have also been shown to lead to increased response times and greater listening effort. Fewer studies, though, have correlated *in situ* classroom acoustic conditions with student achievement on a large scale. This paper presents the results of such an investigation where acoustic metrics compiled over multiple school days from 220 K-12 classrooms are correlated with classroom-aggregated student achievement scores in math and reading taken at the end of the school year, controlling for student demographics such as socioeconomic status.

Earlier studies on classroom acoustics focused on determining preferable conditions and criteria for optimizing speech communication, often by measuring speech intelligibility in terms of the percent of words, phrases, or sentences recognized correctly (Picard and Bradley, 2001; Yang and Bradley, 2009; Wróblewski et al., 2012). Research conducted by Bistafa and Bradley (2000) suggested that ideal maximum classroom background noise levels are 25 dB below the voice level from 1 m away from the talker whereas acceptable classroom background noise levels are 20 dB below the voice level under the same conditions. From combining ideal maximum background noise levels and recommended reverberation times, they suggested a minimum signal-to-noise (SNR) of 15 dB for classrooms. Later Bradley and Sato (2008) revisited these conclusions and suggested that a 15 dB SNR might not be sufficient for younger students who need a higher level of speech intelligibility. Neuman et al. (2010) confirmed that younger children require higher SNR to perform at the same levels as those who are older.

Based on the accumulated body of knowledge particularly around the desired minimum SNR, the ANSI S12.60 standard gives guidance that the greatest 1-h average A-weighted background noise level measured in an unoccupied classroom with mechanical systems on should not exceed 35 dBA for a single mode mechanical system or 37 dBA for multiple mode mechanical systems with "multiple stages of cooling or heating, multiple or variable fan speeds, or ventilation only modes". Additionally, the reverberation times at the mid-frequency octave bands of 500, 1,000, and 2,000 Hz should not exceed 0.6 s for classrooms smaller than 283 m^3 (10,000 ft^3). Previous research has demonstrated that the reverberation time recommendations are more easily met in classrooms than the unoccupied background noise level guidelines (Knecht et al., 2002; Shield and Dockrell, 2004; Nelson et al., 2007; Astolfi and Pellerey, 2008; Ronsse and Wang, 2010, 2013; Shield et al., 2015).

An underlying assumption has been that improving speech intelligibility leads to improved student learning and achievement; however, few studies prior to the ANSI standard's introduction in 2002 showed a direct link between classroom acoustics and student learning outcomes. Bronzaft and McCarthy (1975) and Bronzaft (1981) are two early studies that showed statistically significant lower results of annual reading achievement tests in classrooms more heavily exposed to noise from passing trains. Investigations since the publication of ANSI S12.60 have provided more evidence that poor classroom acoustic conditions correlate to worse performance on tasks that require more comprehension than the recognition of words, phrases, or sentences. Studies have investigated children's reading or listening comprehension performance, in which pupils demonstrate their understanding of meaning from cues (Klatte et al., 2010b; Valente et al., 2012; Klatte et al., 2013; Lewis et al., 2014; Klatte et al., 2017; Rudner et al., 2018; Connolly et al., 2019; Prodi et al., 2019). Some of these gathered and compared results from both speech recognition and speech comprehension tasks. For example, Klatte et al. (2010b) found that the performance of first and third graders on listening comprehension tasks was worse than on speech perception tests when exposed to background speech. Valente et al. (2012) also found that increasing background noise or reverberation resulted in worse performance on comprehension tasks but had minimal effect on sentence recognition tasks.

As found in adults (Kryter, 1985; Jones and Broadbent, 1998; Tiller et al., 2010; Lee et al., 2017), higher noise levels and/or excessive reverberation have been related to decreased performance also by children on various other tasks including short-term memory (Klatte et al., 2010a), basic math (Ljung et al., 2009; Caviola et al., 2021), and categorization or validation tasks (Meinhardt-Injac et al., 2015). In some of these studies, the students' response times were captured and shown to be longer under worse acoustic conditions (Meinhardt-Injac et al., 2015; Puglisi et al., 2018; Connolly et al., 2019; Prodi et al., 2019). Furthermore, the relation between louder and/or more reverberant conditions and task performance is usually more strongly negative for younger students compared to older students or adults (ANSI, 2010; Klatte et al., 2010b; Neuman et al., 2010; Valente et al., 2012; Wróblewski et al., 2012; Klatte et al., 2013; Meinhardt-Injac et al., 2015; Prodi et al., 2019; Caviola et al., 2021), although a few studies have reported stronger effects on older students in their samples (Shield and Dockrell, 2008; Connolly et al., 2019). Negative effects are expected to be more pronounced for pupils with hearing impairments (McCreery et al., 2019) or for persons communicating in a non-native-language (Nelson et al., 2005; Cooke and Lecumberri, 2012). Peng and Wang found that adult speech comprehension performance was significantly worse (Peng and Wang, 2016) and listening effort significantly greater (Peng and Wang, 2019) for non-native English listeners compared to native English listeners when the background noise levels were above 48 dBA or the reverberation times were greater than 0.6 s.

The vast majority of studies reviewed above were conducted under controlled conditions during which subjects were asked to complete tasks over a short period of time (typically less than 1 h) while listening to auralizations presented via headphones in labs or in rooms with noise added via loudspeakers. Only a few studies have investigated student learning outcomes by considering standardized student achievement test scores. In the multi-national RANCH project, Stansfeld et al. (2005) found that exposure of schools to higher aircraft noise levels correlated with lower reading comprehension scores for students aged 9–10 years. The study controlled for student socioeconomic status (SES) in the statistical models. Math test scores were not analyzed in the investigation, though.

Rather than at school-level, classroom-level analyses of standardized test results for literacy, math, and science at grades 2 and 6 were reported by Shield and Dockrell (2008), due to external and internal noise sources found commonly at primary schools. Besides corroborating effects of external road traffic noise, they found statistically significant relationships between grade 2 math scores and grade 6 English scores with background noise levels in occupied and unoccupied classrooms;

TABLE 1 | Number of sampled classrooms, sorted by grade level and school district.

	District A (15 Schools)	District B (13 Schools)	District C (6 Schools)	District D (2 Schools)	District E (4 Schools)	Total (40 Schools)
3rd Grade	25	21	15	5	8	74
5th Grade	20	24	14	5	7	70
8th Grade	15	8	6	0	3	32
11th Grade	20	12	10	0	2	44
# Classrooms	80	65	45	10	20	220

higher noise levels correlated with lower test scores. Many of the relationships lost statistical significance when SES factors were included, though. This may be due to the fact that the number of classrooms for which internal sound levels were available was not large (n = 16 for occupied, n = 14 for unoccupied).

Ronsse and Wang (2010) investigated 58 classrooms across 14 elementary schools within a school district located in Council Bluffs, Iowa, United States, and found that higher unoccupied background noise levels correlated with lower classroom-aggregated student achievement scores in reading, while controlling for SES. They analyzed a second set of measurements from another school district near Omaha, Nebraska, United States, surveying grade 3 (n = 34) and grade 5 classrooms (n = 33) at 14 schools. The findings were similar in that higher unoccupied background noise levels correlated with lower student achievement scores in reading and language subject areas, but the relationship lost significance when controlling for SES (Ronsse and Wang, 2013). In both of those studies, no statistically significant results were found with math scores, nor were any sound levels measured in occupied classrooms.

This paper presents analyses of standardized achievement test results in the math and reading areas across a larger number of classrooms (n = 220) from five different school districts in Iowa and Nebraska. Both primary and secondary school classrooms have been surveyed, specifically at grades 3, 5, 8 and 11. Sound levels were logged in the occupied classrooms over six complete school days, three times seasonally (fall, winter, spring) throughout an academic year. The logged levels have been processed into metrics that describe the classroom acoustic conditions, such as when speech was occurring, when it was not, and the experienced SNR. Reverberation times have also been calculated from impulse response measurements made in the unoccupied classrooms. Details on the assorted calculated metrics may be found in Wang and Brill (2021). Herein, results from statistically analyzing relationships between the classroom-aggregated acoustic metrics and student achievement data, while controlling for SES and other student demographics, are presented to understand better how classroom acoustic conditions relate to student achievement.

2 MATERIALS AND METHODS

In-situ indoor environmental measurements capturing information about acoustics, lighting, thermal comfort, and indoor air quality were conducted in 220 K-12 classrooms, 110 of which were measured during the 2015–2016 academic year and another 110 during 2016–2017. The sample was composed of 3rd, 5th, 8th, and 11th grade classrooms in 40 schools from five school districts in Iowa and Nebraska (**Table 1**). These classrooms represent third and fifth grade homeroom classrooms where both math and language arts are taught and subject specific eighth and eleventh grade classrooms to align with the achievement data collected. This paper isolates the acoustic measurements and achievement data; more details on the complete set of indoor environmental measurements may be found in Kuhlenengel et al. (2017) and Kabirikopaei et al. (2019).

The measured classrooms ranged in volume from 101 to 331 m^3, with a mean volume of 201 m^3 and standard deviation of 32.4 m^3. Classrooms were measured with 22 student occupants on average, ranging from 11 to 32 with a standard deviation of 2.7 pupils. Seven of the classrooms were in portable buildings; none were open plan designs. Classrooms were furnished, and their surface materials were typically gypsum board or concrete-masonry unit walls, thin carpet on floors, acoustical tiles on ceilings, and at least one exterior window.

Equivalent sound levels were measured with two BSWA 309 Type 2/Class 2 sound levels meters. The levels were recorded every 10 seconds with an integration period of 10 seconds. The two sound level meters were placed in locations representative of the teaching position (i.e., in the front of the classroom) and the farthest listening position. The meter at the teaching position was at work plane height (80 cm) enclosed in an open-air wire container along with other equipment. The second meter was attached to the ceiling above the farthest listening position to minimize its distraction to students in class. All meters were placed away from noise-making equipment like projector fans or ventilation outlets/inlets and operated on external battery packs. Meters were deployed in the classroom before school started and then collected the next day after school dismissal, capturing approximately 36 h of measurements. The logging measurements were repeated three times during one academic year in an attempt to capture seasonal differences resulting in measurements of sound levels over six school days. Meters were placed in the same locations for all three sets of measurements to ensure comparability.

Impulse responses were measured in each classroom under unoccupied conditions using the software EASERA, a Larson Davis 831 sound level meter, and an omnidirectional Larson Davis dodecahedron loudspeaker. The loudspeaker was positioned in the front of the classroom where an instructor would typically lead class, at least 1 m away from reflective

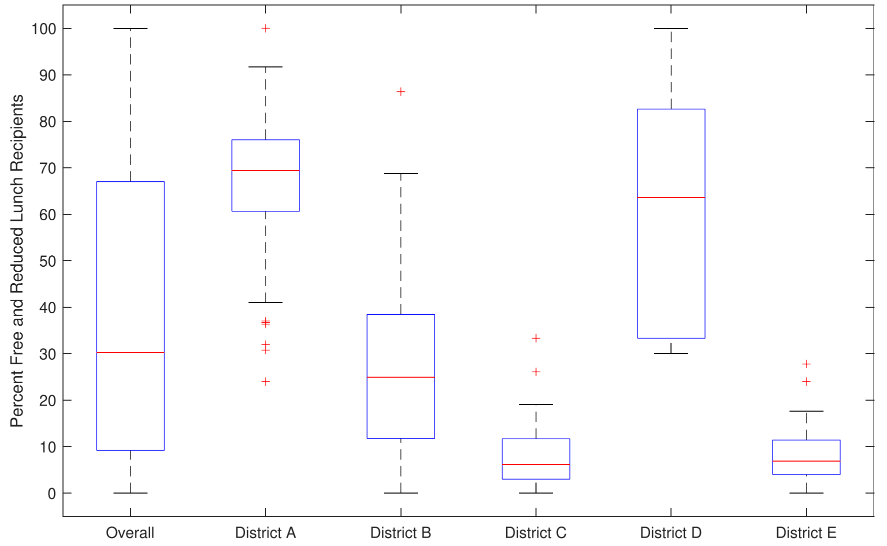


FIGURE 1 | Box plot of the percent of students in the measured classrooms (*n* = 216) receiving free or reduced-price lunches, shown by district.

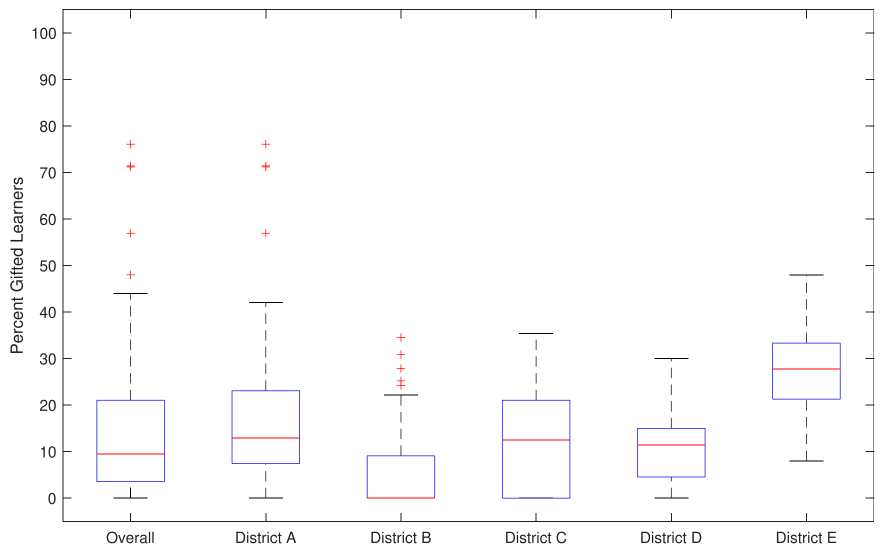


FIGURE 2 | Box plot of the percent of students in the measured classrooms (*n* = 216) designated as gifted learners, shown by district.

surfaces. Two receiver positions were used: one at a seated student's ear height in the middle of the classroom, and another at a seated student's ear height at the farthest listening position. The swept sine method in EASERA was used to acquire the impulse response, with sweeps that were at least 1.2 s long and eight repetitions; EASERA then calculated assorted room acoustic metrics, such as the reverberation time (T20) in each octave band, following ISO 3382-2 (ISO, 2008).

For each measured classroom, the school districts provided the following demographic information aggregated at the classroom level: the percent of students in each classroom who 1) received free or reduced-price lunches (referred to as %FRL), 2) were designated as gifted learners (referred to as %Gifted), and 3) were designated as special education learners (referred to as %SPED). The first of these is commonly used as an indicator of socioeconomic status, which has been shown to have significant relation to student achievement, while a higher percentage of the latter two student categories in a classroom is likely to also impact test scores. Consequently these three demographic variables are controlled for in the statistical analyses. School districts in the United States are required to report the number of gifted pupils and the number of special education learners, but the specific definitions of these categories are often left up to the districts to decide. In the school districts that participated in this study, gifted students were typically defined as performing in the top 5% of their grade, while

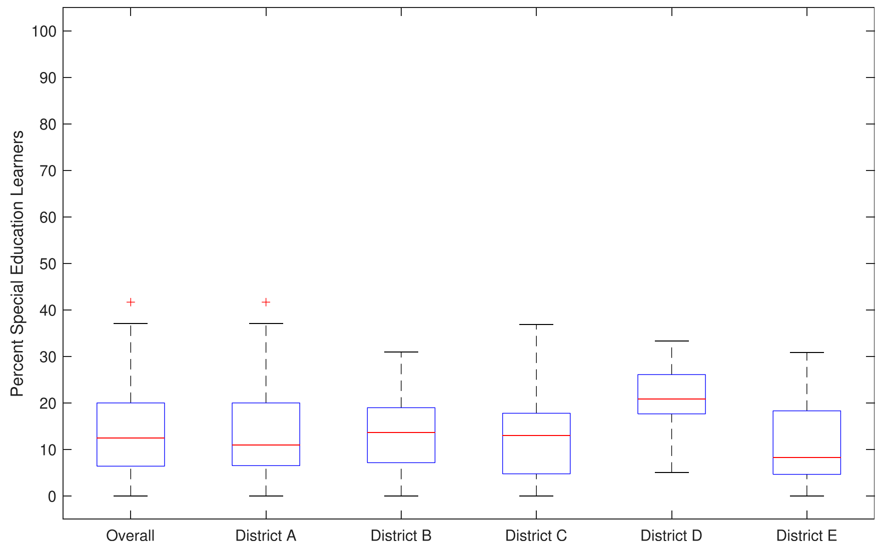


FIGURE 3 | Box plot of the percent of students in the measured classrooms (n = 216) designated as special education learners, shown by district.

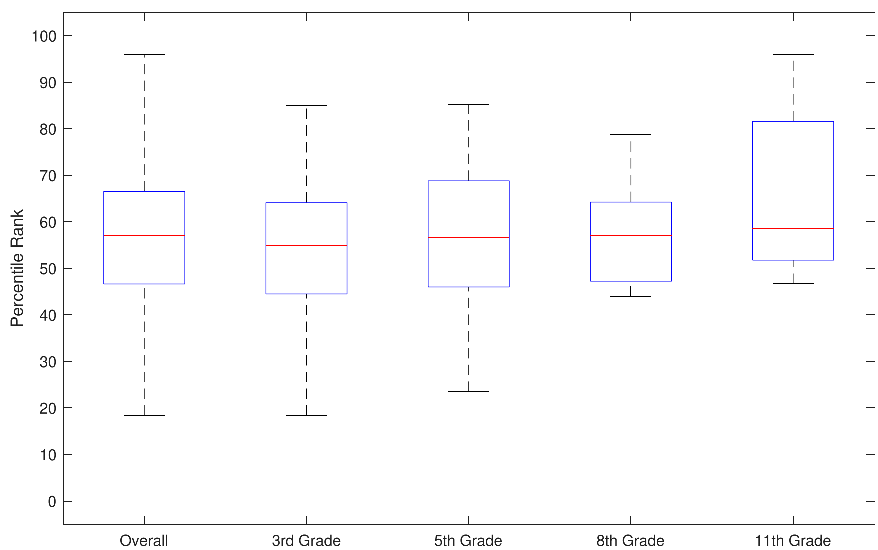


FIGURE 4 | Box plot of the math achievement scores in terms of percentile ranks, averaged for each classroom (n = 178), shown by grade.

special education students were defined as those whose learning abilities are discrepant from their peers, often falling in identified categories outlined in the United States Individuals with Disabilities Education Act (US Department of Education, 2015). **Figures 1–3** present box-plots of the classroom demographic values across the analyzed sample, where the median, 25th, and 75th percentiles are marked by the box, and the whiskers extend to the minimum and maximum data points.

Students in each classroom completed state-wide achievement tests typically in April each year. For this study, achievement was quantified by the results from this state-wide standardized testing [either the Nebraska State Accountability (NeSA) assessment or the Iowa Test of Basic Skills (ITBS)] and not by assessments designed and administered by the researchers. These assessments measure proficiency in fundamental subject areas including math and reading, as compared to state and national standards, and are typically administered towards the end of the academic year in the classrooms in which students receive instruction. The school districts provided results on math achievement and reading achievement in terms of a classroom-level aggregate national percentile rank for each classroom. Districts scored the tests, converted the raw scores to standard scores based on state standards, and then converted the standard scores to a national percentile rank. **Figures 4**, **5** show box-plots of math achievement scores and reading achievement scores, respectively,

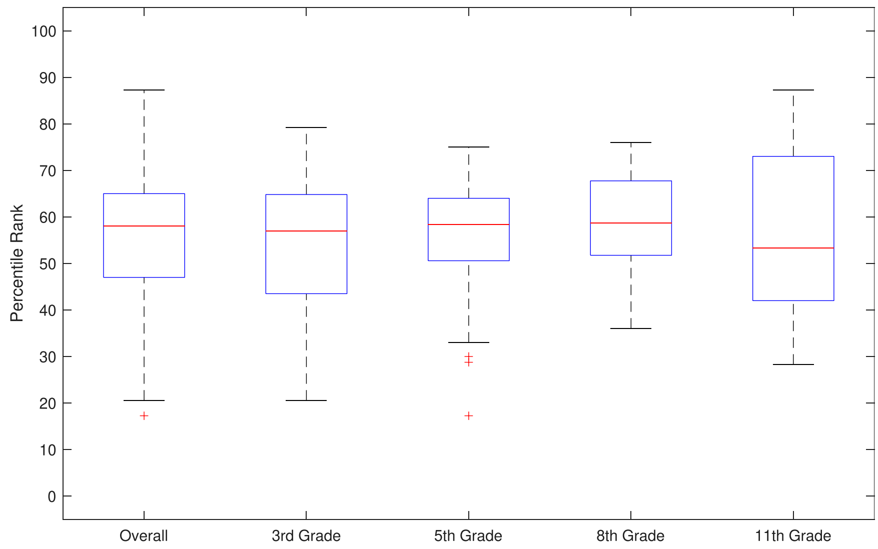


FIGURE 5 | Box plot of the reading achievement scores in terms of percentile ranks, averaged for each classroom (n = 180), shown by grade.

by grade. The total sample size for the statistical analyses on math achievement was n = 178 and the total sample size for the statistical analyses on reading achievement was n = 180. The samples include all third and fifth grade classrooms as both subjects were taught in the same room, as well as the specific eighth and eleventh grade classrooms where math or English classes were held.

This project was reviewed by the University of Nebraska—Lincoln's Institutional Review Board, which determined that individual informed consent was not required as data were provided and analyzed at a classroom-aggregated level with no personally identifiable information.

2.1 Data Analysis

Four classrooms out of the 220 measured were not included in the analyzed sample set. Two of the high school classrooms that had been identified by our school district partners as math classrooms before the start of the school year were not included because they were reassigned to science classrooms which did not correspond to the assessment subjects. Analysis of Mahalanobis distances was used to identify other outliers in the data set, resulting in the exclusion of two other classrooms that were dedicated to special education learners.

Each sound level meter reported A, B, C, and Z-weighted equivalent levels at an interval of every 10 s, in addition to equivalent octave band levels with center frequencies ranging from 32 Hz to 8 kHz. Because the focus of this investigation is on sound levels experienced during the school day, only sound level data recorded during published academic hours for each school were used in the following analyses. An energy-average of the data from the two sound level meters within each classroom was taken at every time interval across the school day, and the energy-averaged data were then used to calculate assorted acoustic metrics for each school day. As with any data set and project of this size, there were occasional missing data, equipment malfunctions, and operator mistakes. Importing routines were programmed to create a log of missing files and missing data, as well as to flag possibly spurious data to be removed from subsequent averaging. Of the 216 classrooms, 83% had data logged over six complete school days on both sound level meters in the classroom, while 15% had missing data on one meter impacting one to 2 days, 0.5% had missing data on one meter impacting three to 4 days, and none had missing data on one meter impacting five to 6 days. In all cases, there were at least data logged on one meter over the six school days.

K-means clustering is an unsupervised statistical learning technique that partitions data into K number of clusters by minimizing the distance between observations within a cluster while maximizing the distance between the clusters (Alpaydin, 2020). For this study, k-means clustering was performed on the nine-dimensional octave band equivalent levels for each observation to provide more information for the partitioning. $K = 2$ was chosen to separate the recorded sound levels into two categories; **Figure 6** graphs box plots of the two clusters, from which it is clear that one represents observations containing high levels across speech frequencies while the other does not. Wang and Brill (2021) provides more detail on the k-means clustering application to the logged data and how the clustered groups more accurately estimate speech levels and non-speech levels in the occupied classrooms than other metrics previously presented in the literature, such as from applying Gaussian mixture modeling or from daily equivalent and statistical levels. These clustered groups were then utilized to calculate the various metrics utilized in the statistical analysis.

Assorted acoustic metrics were calculated to assess the acoustic conditions of the classrooms in this investigation, including equivalent and percentile levels across a full occupied day, equivalent and percentile levels for the speech cluster and the non-speech cluster over the school day, the percent of time that speech or non-speech levels exceeded

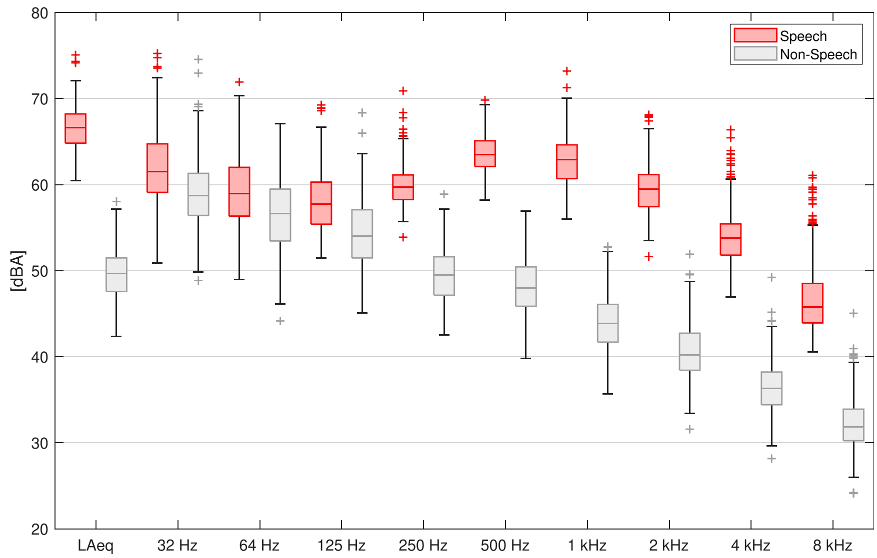


FIGURE 6 | Box plots of the spectra across the two data clusters obtained from k-means clustering, demonstrating that one cluster includes those data with higher levels in speech frequencies while the other does not.

certain values in a day, assorted metrics over octave bands or frequency ranges, and more (Wang and Brill, 2021). Many of the metrics are strongly correlated with correlation coefficients above 0.8 and should consequently not be included together in a statistical model. Preliminary studies led the research team to use three main acoustic metrics in the statistical model: 1) a quantifier of the daily non-speech levels which serves as an estimate of the occupied ambient noise levels, taken to be the A-weighted equivalent sound level of the daily non-speech data (L_{AeqN}), averaged over the six measured school days for each classroom; 2) a quantifier of the daily SNR between the speech and non-speech levels in the classroom, taken to be the daily difference between the A-weighted equivalent sound level of the daily speech data (L_{AeqS}) and of the daily non-speech data (L_{AeqN}) which will be labeled as "SNR" for the remainder of this paper, averaged over the six measured school days; and 3) a quantifier for room reverberance, taken to be the unoccupied mid-frequency reverberation time ($T20_m$) averaged across the 500 Hz, 1 kHz, and 2 kHz octave bands. These three metrics align somewhat with the ANSI S12.60 classroom acoustics standard (ANSI, 2010), as the standard sets guidelines for unoccupied background noise levels and the reverberation times in mid-frequency octave bands, in the hopes of achieving an acceptable SNR of at least 15 dB in occupied conditions, as reported in this paper. Notably, the daily averages of a classroom's speech levels, non-speech levels, and SNR were not found to vary greatly across the six measured school days measured, with average standard deviations of less than 2 dBA, and 3 dBA respectively (Wang and Brill, 2021).

2.2 Statistical Analysis

Descriptive statistics for the demographic variables, acoustic metrics, and student test outcomes are shown in **Table 2**. All of the variables follow a normal distribution except for the demographic ones. Histograms and other analyses of the measured L_{AeqS}, L_{AeqN}, and SNR are provided in Wang and Brill (2021). Pearson's correlation coefficients between the acoustic metrics to be used as predictors in the regression model are provided in **Table 3**. In all statistical analyses presented in this paper, a statistically significant finding is one in which the p-value was less than 0.05. As expected, the average daily SNR significantly correlates to the average daily L_{AeqS} and L_{AeqN} levels, with correlation coefficients $R = 0.46$ and $R = -0.66$ respectively; this is understandable as the daily calculation of SNR is taken as the difference between the other two's daily values. Note that between SNR and L_{AeqN}, the correlation coefficient is negative and larger in magnitude than with L_{AeqS}; as the average daily non-speech levels in classrooms increase, the average daily SNR that students experience decreases. The reverberation time

TABLE 2 | Descriptive statistics of the variables in this investigation. These include classroom demographics: the percent of students receiving free or reduced-price lunch (%FRL), the percent of gifted students (%Gifted), the percent of special education students (%SPED); acoustic metrics: the A-weighted equivalent levels of the speech (L_{AeqS}) and non-speech clusters (L_{AeqN}), the SNR taken as their daily difference, and the mid-frequency averaged reverberation time ($T20_m$); and the test score outcomes in math and reading, given in terms of percentile ranks.

	Mean	Std dev	Min	Max
%FRL	37.3	29.6	0	100
%Gifted	13.6	13.7	0	76.1
%SPED	13.7	9.0	0	41.7
L_{AeqS} (dBA)	66.2	2.41	60.3	74.1
L_{AeqN} (dBA)	49.3	2.90	42.0	57.6
SNR (dBA)	16.9	3.09	9.68	27.1
$T20_m$ (s)	0.47	0.11	0.29	0.84
Math	56.7	14.7	18.3	96.0
Reading	55.9	13.6	17.3	87.3

TABLE 3 | Correlations between the input acoustic variables in this investigation: L_{AeqS}, L_{AeqN}, SNR, $T20_m$.

	L_{AeqS}	L_{AeqN}	SNR	$T20_m$
L_{AeqS}	1	–	–	–
L_{AeqN}	0.35[a]	1	–	–
SNR	0.46[a]	−0.66[a]	1	–
$T20_m$	0.13	0.14[b]	−0.03	1

[a]$p < 0.01$.
[b]$p < 0.05$.

TABLE 4 | Results from the multivariate regression model (*N* = 178) with acoustic metrics as predictors, classroom demographic variables as covariates, and math test scores as outcomes. Grade results shown are against grade 3.

	Estimate B	Standard error	β
%FRL	−0.26[a]	0.03	−0.52
%Gifted	0.58[a]	0.05	0.54
%SPED	−0.31[a]	0.09	−0.19
G5 v G3	8.01	10.35	0.25
G8 v G3	−3.39	10.52	−0.08
G11 v G3	18.63[b]	7.28	0.49
L_{AeqN}	−0.87[b]	0.35	−0.17
SNR	−0.42	0.30	−0.09
$T20_m$	−0.22	7.65	−0.00
SNR × (G5 v G3)	−0.64	0.55	−0.36
SNR × (G8 v G3)	0.22	0.67	0.09
SNR × (G11 v G3)	−1.35[a]	0.41	−0.58

[a]$p < 0.01$.
[b]$p < 0.05$.

TABLE 5 | Results from the multivariate regression model (*N* = 180) with acoustic metrics as predictors, classroom demographic variables as covariates, and reading test scores as outcomes. Grade results shown are against grade 3.

	Estimate B	Standard error	β
%FRL	−0.17[a]	0.03	−0.37
%Gifted	0.61[a]	0.07	0.62
%SPED	−0.29[a]	0.09	−0.19
G5 v G3	0.75	9.90	0.03
G8 v G3	−35.98[b]	15.71	−0.94
G11 v G3	−10.90	20.08	−0.31
L_{AeqN}	−0.26	0.32	−0.06
SNR	−0.49	0.41	−0.11
$T20_m$	11.45	8.23	0.09
SNR × (G5 v G3)	−0.24	0.56	−0.15
SNR × (G8 v G3)	1.88[b]	0.81	0.85
SNR × (G11 v G3)	−0.18	1.24	−0.08

[a]$p < 0.01$.
[b]$p < 0.05$.

$T20_m$ is only significantly correlated to L_{AeqN} with a relatively low $R = 0.14$; note that the range of $T20_m$ in the sample was 0.29–0.84 s (**Table 2**), though, with the majority of rooms meeting ANSI S12.60 guidelines (Wang and Brill, 2021).

The software R 4.0.2 with the Lavaan package version 0.6-7 (Rosseel, 2012) was used to conduct the multivariate regression analyses of acoustic metrics on math and reading scores, while controlling for classroom demographics. Outcome residuals for math and reading scores were allowed to covary, necessitating a multivariate model. The three demographic descriptors (%FRL, %Gifted, and %SPED) as well as grade level were used as covariates. Non-independence of classrooms within schools was accounted for by applying robust cluster standard errors. First, interactions were explored by considering each demographic variable one at a time in separate sub-models. In the sub-models, demographic variables were permitted to moderate association of an acoustic metric and its effect on math or reading scores. Statistically significant interactions from the sub-models were then retained in the full model.

3 RESULTS

When exploring which interactions with demographic variables should be retained in the full model, four group differences were found to be statistically significant from using the Wald test: 1) L_{AeqN} and grade level on math scores, 2) SNR and grade level on math scores, 3) SNR and grade level on reading scores, and 4) SNR and the percent of students receiving free or reduced-price lunch on math scores. These interactions were then probed in the full multivariate regression model. The SNR and %FRL interaction was not retained in the full model as it did not reach statistical significance. Final regression results on the math and reading scores are shown respectively in **Tables 4**, **5**.

As expected, the classroom demographic variables had statistically significant relationships with the math and reading test outcomes. Higher %FRL and higher %SPED values correlated with lower test scores, while higher %Gifted correlated with higher test scores. Controlling for these, the results indicate only one statistically significant main effect between L_{AeqN} and math test scores; higher daily non-speech levels in a classroom correlated with lower math test scores. No other main effects between acoustic predictors and math or reading test outcomes reached statistical significance.

The entries in **Tables 4**, **5** pertaining to grade compare a higher grade's results against those from grade 3. For math scores, only grade 11 indicates a significant difference from grade 3; **Figure 4** illustrates that the distribution of grade 11 math scores extended higher and not as low in range as grade 3 math scores. This difference in distribution likely plays a role in the statistically significant interaction between SNR and grade 11 on math scores. As grade 11 did not have many low test scores, possibly due to the sample including less grade 11 classrooms, the authors suggest that this significant interaction between SNR and grade 11 is likely not indicative of a true relationship. Similarly for reading scores, grade 8 shows a statistically different result from grade 3, as well as an interaction with SNR. This result is again likely due to the distribution of grade 8 reading scores being quite different from that of grade 3 (**Figure 5**), which could be due to the lower number of grade 8 classrooms in the sample.

The R^2 values associated with the regression results presented above are 0.644 for the math scores and 0.536 for the reading scores. When running the model without acoustic predictors but with all other demographic variables, the R^2 values are 0.618 for the math results and 0.506 for the reading results. A comparison between these two models (without and with acoustic variables) indicates that the overall models are different at a statistically significant level ($p < 0.01$), as well as specifically for the math achievement prediction ($\chi^2 = 20.47$, $df = 6$, $p < 0.01$) and the reading achievement prediction ($\chi^2 = 13.04$, $df = 6$, $p < 0.05$). Adding in the acoustic variables did result in a model that accounted for more variance in the results at a statistically significant level.

4 DISCUSSION

A statistically significant relationship between the average daily non-speech levels in occupied classrooms and math test scores has been found from multivariate regression analysis based on data from 178 classrooms across four grade levels (3, 5, 8 and 11), with controls for classroom demographics. Higher daily non-speech levels are correlated with lower math test scores. This is the first time to the authors' knowledge that a significant relationship between noise levels in classrooms and math achievement scores has been reported. Previous investigations involving standardized test scores have instead found significant relationship between higher noise levels and lower reading achievement scores. One of those studies did not report on any math scores (Stansfeld et al., 2005). The others (Shield and Dockrell, 2008; Ronsse and Wang, 2010, 2013) included only primary school students, had much smaller sample sizes (less than 70 classrooms) than the current paper, and found that some relationships lost significance when SES factors were included.

There has been other evidence in the literature, reporting effects of noise on children's performance of math tasks. Ljung et al. (2009) ran tests that included basic math and math reasoning tasks on 187 12 or 13 year old pupils under different noise conditions within a classroom, and found that the road traffic noise condition did impair performance on the math task compared to the other noise conditions. Meinhardt-Injac et al. (2015) asked 21 second-graders and 25 sixth-graders to complete tasks including validations of math problem, while listening to different noise conditions over headphones; younger pupils did worse on the math validation task when exposed to irrelevant speech but not to classroom noise without speech. More recently, Caviola et al. (2021) reviewed the different skills and cognitive components related to math performance, when reporting on their study wherein 162 11–13 year olds were asked to complete a variety of math tasks under different noise conditions. Their results show that the younger pupils did perform worse when exposed to classroom noise than under quiet or traffic noise conditions, although as the task difficulty increased, the effect faded. While these previous studies have presented performance on short-term math tasks, rather than on standardized math tests that may be more indicative of math learning outcomes, they do support the finding in this paper of a relationship between non-speech levels in occupied classrooms and math achievement.

In the presented regression analyses, L_{AeqN} is the metric that accounts for the most variance in the math test scores; L_{AeqN} and SNR are significantly correlated (**Table 3**) with $R = -0.66$, so in these models L_{AeqN} is accounting for most of the variance to which SNR may also have contributed. An interpretation of this is as follows: lower non-speech levels in occupied classrooms correlate with higher standardized math test scores. Those lower non-speech levels also significantly correlate with higher SNR conditions, which has been an overall goal of classroom acoustic design standards like ANSI S12.60 (ANSI, 2010). Consequently, designing classrooms for lower unoccupied noise levels that lead to lower non-speech levels in occupied classrooms and higher SNR in classrooms is recommended.

As **Table 3** shows, there is a statistically significant correlation between the speech and non-speech levels whereby higher speech levels are correlated with higher non-speech levels ($r = 0.35$, $p < 0.01$). Linear regression analysis finds the relationship to show 0.29 dBA increase in speech levels for every 1 dBA increase in non-speech levels, but there is a lot of variance in speech levels that are not accounted for by the non-speech levels, due possibly for example to talker variability or vocal strength (Wang and Brill, 2021). Other recent studies have reported Lombard effects measured at the talker ranging from a +0.51 to +0.72 dBA increase in speech levels for every +1 dBA increase in noise levels (Sato and Bradley, 2008; Bottalico and Astolfi, 2012; Sarantopoulos et al., 2014). In applying any of these Lombard effect slopes, increasing noise levels results in lower SNR because speech levels increase in less than a one to one ratio.

The results in the presented analyses are interpreted to represent the chronic or accumulated effects of noise. Noise levels were not measured in the test rooms at the time students were taking these assessments, so it is not possible for this study to base any interpretations on the acute effect of noise. That does not mean that acute effects do not exist. The large number of classrooms was intentionally chosen to distill the chronic effects rather than the acute effects.

Grade has been used as a proxy for student age in this paper, and other studies have shown more strongly negative relationships between acoustic conditions and task performance for younger students compared to older students (Klatte et al., 2010b; Neuman et al., 2010; Valente et al., 2012; Wróblewski et al., 2012; Meinhardt-Injac et al., 2015; Prodi et al., 2019; Caviola et al., 2021). That conclusion cannot be made based on the regression models presented here. More investigations that span the grades covered in this investigation and both math and reading achievement test scores are needed.

5 CONCLUSION AND FUTURE WORK

Regression models have been run relating acoustic data gathered from 216 classrooms across four grade levels (3, 5, 8 and 11) with classroom-aggregated standardized math and reading test scores, while controlling for classroom demographics including socioeconomic status. A statistically significant relationship was found between the average daily non-speech levels in classrooms and math test scores; higher daily non-speech levels were correlated with lower math test scores ($p < 0.05$). No statistically significant main effects of acoustic metrics were found on reading achievement. There were some significant differences and an interaction found between grades, but the authors believe that they are due to uneven sample distributions across grade as there were fewer grade 8 and 11 classrooms measured.

One limitation to the current investigation is that other classroom demographics were not available for inclusion, such as the percent of students in each classroom with hearing impairments or those learning in a non-native language, so it is unclear how other demographics may relate to the results. Further investigations that include such demographics is recommended. Also, this investigation has primarily been a correlational study; thus, one should interpret the results of the presented multivariate regression models with caution. Causation cannot be assumed without further investigations in which acoustic conditions are deliberately changed and the effect on student achievement scores assessed while controlling for other factors that can impact test scores. Future studies should work with school districts to test changes or manipulations aimed at lowering non-speech levels in classrooms to see if improved student test scores are achieved; refer for example to Bronzaft (1981) and Massonnié et al. (2020). Another idea for future study is to consider studying achievement at the level of individual students, rather than aggregated at the classroom level. The 220 classrooms measured in this study represent rooms in which more than 7,000 students learned. Investigating how an individual's exposure to sound levels throughout their school day and in other indoor spaces they occupy (home, recreational facilities, etc.) is related to their learning outcomes, while controlling for that person's demographics, may be difficult but worthwhile.

This is the first investigation to the authors' knowledge that has shown a significant relationship between non-speech levels in occupied classrooms and math achievement scores. Unlike previous studies, no significance was found with reading scores, but the current study differs from earlier ones in that both primary and secondary classrooms were included and three classroom demographics (percent of students receiving free or reduced-price lunches, percent gifted, and percent special education) were used as covariates. Furthermore, the results are based on detailed sound level data logged across six school days per classroom over an academic year, thereby more effectively capturing occupied acoustic conditions experienced by students in the classrooms. The logged data were separated using k-means clustering in nine dimensions into one group representing when speech occurs and another when speech does not. The daily averages of a classroom's speech levels, non-speech levels, and SNR did not vary greatly across the 6 days measured across three seasons in a school year, with standard deviations of less than 3 dBA typically. So while K-12 classrooms are complex environments in which different teaching modalities are used, ranging from single instructor to individual work to small group activities (Shield and Dockrell, 2004), the daily values of acoustic metrics were not found to vary greatly in this study for a specific classroom occupied by a consistent instructor. Having found a statistically significant correlation between the average daily occupied non-speech levels with math achievement is a step forward towards better evidence-based classroom acoustics design.

Indeed, better evidence-based design of classrooms as a whole requires that acoustic conditions be considered in balance with other indoor environmental conditions, such as indoor air quality, thermal, and lighting. How do the relationships presented here with acoustics vary when other measured environmental metrics are included? Researchers are looking into this, and additional work along those lines is recommended so that the school design community can prioritize evidence-based design aimed at benefiting human well-being and performance.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Nebraska—Lincoln Institutional Review Board. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

LW conceived and designed the overall project, secured funding, and oversaw data collection and analyses. LB managed the data collection, led work on data analyses, and proposed the application of k-means clustering to the logged sound level data. Both authors were involved with drafting and revising the manuscript.

FUNDING

This research was supported by the United States Environmental Protection Agency Grant Number R835633. Open access publication fees were covered by the University of Nebraska—Lincoln Durham School of Architectural Engineering and Construction.

ACKNOWLEDGMENTS

The authors gratefully acknowledge all members of the research team who worked on the project, listed at the University of Nebraska—Lincoln Healthy Schools website (https://engineering.unl.edu/healthy-schools), with special thanks to Michael Kuhlenengel, Kieren Smith McCord, Jayden Nord, and Dr. James Bovaird. The authors would also like to thank the school districts who partnered with us on this project, and Dr. Ralph Muehleisen who served as a technical advisor in the area of acoustics on the Project Advisory Committee.

REFERENCES

Alpaydin, E. (2020). *Introduction to Machine Learning*. Cambridge, MA: The MIT Press.

ANSI (2010). *American National Standard Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools, Part 1: Permanent Schools. S12.60*. New York: American National Standards Institute.

Astolfi, A., and Pellerey, F. (2008). Subjective and Objective Assessment of Acoustical and Overall Environmental Quality in Secondary School Classrooms. *J. Acoust. Soc. Am.* 123, 163–173. doi:10.1121/1.2816563

Bistafa, S. R., and Bradley, J. S. (2000). Reverberation Time and Maximum Background-Noise Level for Classrooms from a Comparative Study of Speech Intelligibility Metrics. *J. Acoust. Soc. Am.* 107, 861–875. doi:10.1121/1.428268

Bottalico, P., and Astolfi, A. (2012). Investigations into Vocal Doses and Parameters Pertaining to Primary School Teachers in Classrooms. *J. Acoust. Soc. Am.* 131, 2817–2827. doi:10.1121/1.3689549

Bradley, J. S., and Sato, H. (2008). The Intelligibility of Speech in Elementary School Classrooms. *J. Acoust. Soc. Am.* 123, 2078–2086. doi:10.1121/1.2839285

Bronzaft, A. L., and McCarthy, D. P. (1975). The Effect of Elevated Train Noise on reading Ability. *Environ. Behav.* 7, 517–528. doi:10.1177/001391657500700406

Bronzaft, A. L. (1981). The Effect of a Noise Abatement Program on reading Ability. *J. Environ. Psychol.* 1, 215–222. doi:10.1016/S0272-4944(81)80040-0

Caviola, S., Visentin, C., Borella, E., Mammarella, I., and Prodi, N. (2021). Out of the Noise: Effects of Sound Environment on Maths Performance in Middle-School Students. *J. Environ. Psychol.* 73, 101552. doi:10.1016/j.jenvp.2021.101552

Connolly, D., Dockrell, J., Shield, B., Conetta, R., Mydlarz, C., and Cox, T. (2019). The Effects of Classroom Noise on the reading Comprehension of Adolescents. *J. Acoust. Soc. Am.* 145, 372–381. doi:10.1121/1.5087126

Cooke, M., and Lecumberri, M. L. G. (2012). The Intelligibility of Lombard Speech for Non-native Listeners. *J. Acoust. Soc. Am.* 132, 1120–1129. doi:10.1121/1.4732062

ISO (2008). *Acoustics – Measurement of Room Acoustic Parameters – Part 2: Reverberation Time in Ordinary Rooms*, 3382-2. Geneva: International Standards Organization.

Jones, D. M., and Broadbent, D. E. (1998). "Chapter 24: Human Performance and Noise," in *Handbook of Acoustical Measurements and Noise Control*. Editor C. Harris (Melville, NY: Acoustical Society of America), 24.

Kabirikopaei, A., Kuhlenengel, M., Arthur, A., Bovaird, J., Lau, J., Waters, C., et al. (2019). "The Effects of Indoor Environmental Factors on Students' Academic Achievement," in Proceedings of the 2019 Architectural Engineering Institute Conference, Tysons, VA, April 3–6, 2019 (State College, PA: American Society of Civil Engineers), 257–264. doi:10.1061/9780784482261.030

Klatte, M., Bergström, K., and Lachmann, T. (2013). Does Noise Affect Learning? A Short Review on Noise Effects on Cognitive Performance in Children. *Front. Psychol.* 4, 578. doi:10.3389/fpsyg.2013.00578

Klatte, M., Hellbrück, J., Seidel, J., and Leistner, P. (2010a). Effects of Classroom Acoustics on Performance and Well-Being in Elementary School Children: a Field Study. *Environ. Behav.* 42, 659–692. doi:10.1177/0013916509336813

Klatte, M., Lachmann, T., and Meis, M. (2010b). Effects of Noise and Reverberation on Speech Perception and Listening Comprehension of Children and Adults in a Classroom-like Setting. *Noise Health* 12, 270–282. doi:10.4103/1463-1741.70506

Klatte, M., Spilski, J., Mayerl, J., Möhler, U., Lachmann, T., and Bergström, K. (2017). Effects of Aircraft Noise on reading and Quality of Life in Primary School Children in Germany: Results from the NORAH Study. *Environ. Behav.* 49, 390–424. doi:10.1177/0013916516642580

Knecht, H. A., Nelson, P. B., Whitelaw, G. M., and Feth, L. L. (2002). Background Noise Levels and Reverberation Times in Unoccupied Classrooms. *Am. J. Audiol.* 11, 65–71. doi:10.1044/1059-0889(2002/009)

Kryter, K. D. (1985). "Chapter 9: Mental and Psychomotor Task Performance in Noise," in *The Effects of Noise on Man* (New York: Academic Press), 545–585.

Kuhlenengel, M. B., Brill, L. C., Deng, S., Lester, H. F., Bovaird, J., Lau, J., et al. (2017). "An Investigation of School Environmental Effects on Student Achievement," in Proceedings of the 2017 Architectural Engineering Institute Conference, Oklahoma, April 11–13, 2017 (Oklahoma City, OK: American Society of Civil Engineers), 927–939. doi:10.1061/9780784480502.078

Lee, J., Francis, J. M., and Wang, L. M. (2017). How Tonality and Loudness of Noise Relate to Annoyance and Task Performance. *Noise Cont. Engng. J.* 65, 71–82. doi:10.3397/1/376427

Leibold, L. J. (2017). Speech Perception in Complex Acoustic Environments: Developmental Effects. *J. Speech Lang. Hear. Res.* 60, 3001–3008. doi:10.1044/2017_jslhr-h-17-0070

Lewis, D. E., Manninen, C. M., Valente, D. L., and Smith, N. A. (2014). Children's Understanding of Instructions Presented in Noise and Reverberation. *Am. J. Audiol.* 23, 326–336. doi:10.1044/2014_aja-14-0020

Ljung, R., Sorqvist, P., and Hygge, S. (2009). Effects of Road Traffic Noise and Irrelevant Speech on Children′s reading and Mathematical Performance. *Noise Health* 11, 194–198. doi:10.4103/1463-1741.56212

Massonnié, J., Frasseto, P., Mareschal, D., and Kirkham, N. Z. (2020). Scientific Collaboration with Educators: Practical Insights from an in-Class Noise-Reduction Intervention. *Mind Brain Educ.* 14, 303–316. doi:10.1111/mbe.12240

McCreery, R. W., Walker, E. A., Spratford, M., Lewis, D., and Brennan, M. (2019). Auditory, Cognitive, and Linguistic Factors Predict Speech Recognition in Adverse Listening Conditions for Children with Hearing Loss. *Front. Neurosci.* 13, 1093. doi:10.3389/fnins.2019.01093

Meinhardt-Injac, B., Schlittmeier, S., Klatte, M., Otto, A., Persike, M., and Imhof, M. (2015). Auditory Distraction by Meaningless Irrelevant Speech: a Developmental Study. *Appl. Cognit. Psychol.* 29, 217–225. doi:10.1002/acp.3098

Nelson, E. L., Smaldino, J., Erler, S., and Garstecki, D. (2007). Background Noise Levels and Reverberation Times in Old and New Elementary School Classrooms. *J. Ed. Aud.* 14, 16–22.

Nelson, P., Kohnert, K., Sabur, S., and Shaw, D. (2005). Classroom Noise and Children Learning through a Second Language. *Lang. Speech Hear. Serv. Sch.* 36, 219–229. doi:10.1044/0161-1461(2005/022)

Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2010). Combined Effects of Noise and Reverberation on Speech Recognition Performance of normal-hearing Children and Adults. *Ear. Hear.* 31, 336–344. doi:10.1097/aud.0b013e3181d3d514

Peng, Z. E., and Wang, L. M. (2016). Effects of Noise, Reverberation and Foreign Accent on Native and Non-native Listeners' Performance of English Speech Comprehension. *J. Acoust. Soc. Am.* 139, 2772–2783. doi:10.1121/1.4948564

Peng, Z. E., and Wang, L. M. (2019). Listening Effort by Native and Nonnative Listeners Due to Noise, Reverberation, and Talker Foreign Accent during English Speech Perception. *J. Speech Lang. Hear. Res.* 62, 1068–1081. doi:10.1044/2018_jslhr-h-17-0423

Picard, M., and Bradley, J. S. (2001). Revisiting Speech Interference in Classrooms: Revisando la interferencia en el habla dentro del salón de clases. *Int. J. Audiol.* 40, 221–244. doi:10.3109/00206090109073117

Prodi, N., Visentin, C., Borella, E., Mammarella, I. C., and Di Domenico, A. (2019). Noise, Age, and Gender Effects on Speech Intelligibility and Sentence Comprehension for 11- to 13-Year-Old Children in Real Classrooms. *Front. Psychol.* 10, 2166. doi:10.3389/fpsyg.2019.02166

Puglisi, G. E., Prato, A., Sacco, T., and Astolfi, A. (2018). Influence of Classroom Acoustics on the reading Speed: A Case Study on Italian Second-Graders. *J. Acoust. Soc. Am.* 144, EL144–EL149. doi:10.1121/1.5051050

Ronsse, L. M., and Wang, L. M. (2010). Effects of Noise from Building Mechanical Systems on Elementary School Student Achievement. *ASHRAE Trans.* 116, 347–354.

Ronsse, L. M., and Wang, L. M. (2013). Relationships between Unoccupied Classroom Acoustical Conditions and Elementary Student Achievement Measured in Eastern Nebraska. *J. Acoust. Soc. Am.* 133, 1480–1495. doi:10.1121/1.4789356

Rosseel, Y. (2012). Lavaan: AnRPackage for Structural Equation Modeling. *J. Stat. Soft.* 48, 1–36. doi:10.18637/jss.v048.i02

Rudner, M., Lyberg-Åhlander, V., Brännström, J., Nirme, J., Pichora-Fuller, M. K., and Sahlén, B. (2018). Listening Comprehension and Listening Effort in the Primary School Classroom. *Front. Psychol.* 9, 1193. doi:10.3389/fpsyg.2018.01193

Sarantopoulos, G., Lykoudis, S., and Kassomenos, P. (2014). Noise Levels in Primary Schools of Medium Sized City in Greece. *Sci. Total Environ.* 482-483, 493–500. doi:10.1016/j.scitotenv.2013.09.010

Sato, H., and Bradley, J. S. (2008). Evaluation of Acoustical Conditions for Speech Communication in Working Elementary School Classrooms. *J. Acoust. Soc. Am.* 123, 2064–2077. doi:10.1121/1.2839283

Shield, B., Conetta, R., Dockrell, J., Connolly, D., Cox, T., and Mydlarz, C. (2015). A Survey of Acoustic Conditions and Noise Levels in Secondary School Classrooms in England. *J. Acoust. Soc. Am.* 137, 177–188. doi:10.1121/1.4904528

Shield, B., and Dockrell, J. E. (2004). External and Internal Noise Surveys of London Primary Schools. *J. Acoust. Soc. Am.* 115, 730–738. doi:10.1121/1.1635837

Shield, B. M., and Dockrell, J. E. (2008). The Effects of Environmental and Classroom Noise on the Academic Attainments of Primary School Children. *J. Acoust. Soc. Am.* 123, 133–144. doi:10.1121/1.2812596

Stansfeld, S., Berglund, B., Clark, C., Lopez-Barrio, I., Fischer, P., Öhrström, E., et al. (2005). Aircraft and Road Traffic Noise and Children's Cognition and Health: a Cross-National Study. *Lancet* 365, 1942–1949. doi:10.1016/s0140-6736(05)66660-3

Tiller, D. K., Musser, A., Wang, L. M., and Radik, M. J. (2010). Combined Effects of Noise and Temperature on Human comfort and Performance. *ASHRAE Trans.* 116, 522–540.

US Department of Education (2015). Individuals with Disabilities Education Act. Available at: https://sites.ed.gov/idea/ (Accessed March 29, 2021).

Valente, D. L., Plevinsky, H. M., Franco, J. M., Heinrichs-Graham, E. C., and Lewis, D. E. (2012). Experimental Investigation of the Effects of the Acoustical Conditions in a Simulated Classroom on Speech Recognition and Learning in Children. *J. Acoust. Soc. Am.* 131, 232–246. doi:10.1121/1.3662059

Wang, L. M., and Brill, L. C. (2021). Speech and Noise Levels Measured in Occupied K-12 Classrooms. *J. Acoust. Soc. Am.* 150, 864–877. doi:10.1121/10.0005815

Wróblewski, M., Lewis, D. E., Valente, D. L., and Stelmachowicz, P. G. (2012). Effects of Reverberation on Speech Recognition in Stationary and Modulated Noise by School-Aged Children and Young Adults. *Ear Hear* 33, 731–744. doi:10.1097/aud.0b013e31825aecad

Yang, W., and Bradley, J. S. (2009). Effects of Room Acoustics on the Intelligibility of Speech in Classrooms for Young Children. *J. Acoust. Soc. Am.* 125, 922–933. doi:10.1121/1.3058900

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Built Environment

ORIGINAL RESEARCH
published: 08 February 2022
doi: 10.3389/fbuil.2022.688847

Check for updates

Toward Child-Appropriate Acoustic Measurement Methods in Primary Schools and Daycare Centers

Karin Loh[1]*, *Manuj Yadav*[1,2], *Kerstin Persson Waye*[3], *Maria Klatte*[4] *and Janina Fels*[1]

[1]Institute for Hearing Technology and Acoustics, RWTH Aachen University, Aachen, Germany, [2]Sydney School of Architecture, Design and Planning, University of Sydney, Sydney, NSW, Australia, [3]Occupational and Environmental Medicine, School of Public Health and Community Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, [4]Cognitive and Developmental Psychology, University of Kaiserslautern, Kaiserslautern, Germany

Children spend a considerable amount of time in educational institutions, where they are constantly exposed to noisy sound environments, which has detrimental effects on children's health and cognitive development. Extensive room acoustics measurements and long-term *in-situ* measurements in such institutions are scarce and are generally conducted using omnidirectional microphones. This study provides preliminary results of room acoustics in unoccupied conditions and *in-situ* noise measurements during occupancy, in classrooms and playrooms in Germany using an omnidirectional microphone, an adult HATS (head and torso simulator), and a child HATS. The results indicate that room acoustics of most of the sampled rooms need improvement (mid-frequency reverberation time, T_{30} (s) = 0.6 (0.3–1.1) and clarity index, C_{50} (dB) = 6.1 (1.6–10.4); speech transmission index (STI) = 0.7 (0.6–0.8); mean values and range); the sound pressure level (SPL) during activities was around 66 dB (A-weighted equivalent level SPL) in both classrooms and playrooms using omnidirectional measurements, which is somewhat lower than similar measurements in other countries that varied in measurement periods; psychoacoustics parameters relating to sound fluctuation (fluctuation strength and roughness) show variation with increasing room volumes; and that there may be some benefit in considering child HATS for *in-situ* noise measurements. While the validity of these results in relation to children's perceptual evaluation (using questionnaires, etc.) is subject to future investigations, the results highlight some of the nuances in the choice of transducers in measurements with children and potential benefits of psychoacoustic parameters in complementing the SPL-based parameters in more comprehensively characterizing the noise environments in educational institutions.

Keywords: noise assessment, educational buildings, children's hearing, binaural acoustics, classroom acoustics

OPEN ACCESS

Edited by:
Arianna Astolfi,
Politecnico di Torino, Italy

Reviewed by:
Valtteri Hongisto,
Turku University of Applied Sciences, Finland
Lori Leibold,
Boys Town National Research Hospital, United States

****Correspondence:***
Karin Loh
karin.loh@akustik.rwth-aachen.de

Specialty section:
This article was submitted to Indoor Environment, a section of the journal Frontiers in Built Environment

Received: *31 March 2021*
Accepted: *03 January 2022*
Published: *08 February 2022*

Citation:
Loh K, Yadav M, Persson Waye K, Klatte M and Fels J (2022) Toward Child-Appropriate Acoustic Measurement Methods in Primary Schools and Daycare Centers. Front. Built Environ. 8:688847. doi: 10.3389/fbuil.2022.688847

1 INTRODUCTION

Noise and unfavorable room acoustics in educational institutions, such as primary schools and preschools, is a well-known problem. While several studies have reported results of room acoustics measurements (in both occupied and unoccupied rooms), long-term noise measurements are scarce, especially in daycare settings. Furthermore, for characterizing noise and room acoustic measurements in such institutions, two possibilities include using omnidirectional and binaural

TABLE 1 | Summary of noise and room acoustic parameters in previous studies. Reported values are mean and the range (in brackets) unless indicated otherwise.

Study	Summary of conditions	$L_{A,eq}$ (dB)	Percentiles (dB)	Reverberation time (s)	*STI*, C_{50} (dB), U_{50} (dB)
Sala and Rantala, 2016 (others' studies)	Elementary school classrooms	42–100	L_{A90}: 40–61 (L_{A95})	0.7 (0.2–1.27)	Classrooms: *STI* = 0.68 (0.44–0.81), schools for 6-year-olds: *STI* = 0.77 (0.59–0.92)
	Preschool classrooms	60–85	L_{A90}: 39–47	0.55 (0.41–0.85)	*STI* = 0.74 (0.65–0.81)
Sala and Rantala, 2016 (own study)	40 schools, children ages 7–12; 19 students per classroom on average	69 (57–89)	L_{A10}: 68 (57–77), L_{A50}: 55 (42–64), L_{A90}: 42 (29–51)	0.55 (0.41–0.85)	*STI* = 0.75 (0.65–0.81)
Astolfi et al., 2019b ("good" acoustics)	20 classrooms with an average of 18 children each, predominantly 6–7 years old	60–75	—	(0.5–0.8)	2.9 dB ≤ C_{50} ≤ 7.6 and −0.8 dB ≤ U_{50} ≤ 4.0
Astolfi et al., 2019b ('bad' acoustics)	Same as above	62–72	—	0.5 < value <0.8	−2.2 dB ≤ C_{50} ≤ 2.7 and −2.6 dB ≤ U_{50} ≤ 0.9
Persson Waye and Karlberg, (2021)	Dosimeter results from seven preschools in Sweden for 56 children aged 4–5 years old before acoustic intervention	85	L_{A5}: 90 L_{A25}: 87	0.3–0.5	8 ≤ C_{50} ≤ 10
Persson Waye and Karlberg, (2021)	After acoustic intervention	83	L_{A5}: 91 L_{A25}: 86	Lower by 0.1 s on average to above	11 ≤ C_{50} ≤ 13
Wang and Brill, (2021)	220 K-12 classrooms. Values averaged over two SLM measurements and over 6 days per room. Average of 22 students per room (SD: 2.7)	Speech: 65 (SD: 2.5), noise: 47 (SD: 3.5)	—	(0.2–1.1)	−2.0 dB ≤ C_{50} ≤ 14.4
Södersten et al. (2002)	Binaural recordings in 10 daycare centers for teachers of 1–6 year old children	76 (73–78)	—	—	—
McAllister et al. (2009)	Binaural recordings of 10 children (5 years old each) in three daycare centers	83 (82–84)	—	—	—

transducers. The former allows a range of measurements including standardized ones (Bradley et al., 1999; American National Standards Institute, 2002; Building Bulletin 93, 2015; Deutsches Institut für Normung, 2016; Astolfi et al., 2019a), while the latter generally incorporated as microphones near human (who may or may not have freedom of movement) ears or within ear canals of head and torso simulator (HATS). Binaural transducers allow measurements that can represent some of the effects of head, shoulders, and outer ear processing for static listeners (i.e., without head movements). Adult HATS in binaural measurement procedures are relatively common in research settings including in classrooms (CRs) for children (e.g., Shinn-Cunningham et al., 2005; Peng et al., 2012), and there is at least one example of a head and shoulder simulator (Fels et al., 2004; Prodi et al., 2007) that has been qualified to closely represent children aged approximately 3–6 years (hereinafter, referred to as children/child HATS). This study reports on room acoustics measurements in unoccupied conditions (furnished rooms) and long-term noise measurements during typical hours of occupancy, in several primary schools and daycare centers, using both omnidirectional and binaural transducers. This includes investigating the extent to which relevant acoustic and psychoacoustic parameters vary across the educational settings and between the transducers, that is, omnidirectional, adult and child HATSs, with the latter two providing first-order representations of teachers and students' perception, respectively. These considerations may be important for characterizing the noise environment and room acoustics and determining appropriate measurement methods in a variety of educational institutions for children.

As listed in **Table 1**, which refers to measurements in CRs mainly in primary schools and daycare centers across various countries, the sound pressure levels (SPLs, in decibel) in such educational institutions are considerable. There are also considerable variations between studies due to factors such as the number of children present (Sala and Rantala, 2016); the age groups, with daycare centers generally reporting higher levels than primary school CRs (Picard and Bradley, 2001); activities involved; room acoustics due to excessively low or high reverberation times (RTs) (Astolfi et al., 2019b); measurement methods including duration (Sala and Rantala, 2016; Wang and

Brill, 2021), transducer type and locations, for example, omnidirectional vs binaural recordings vs dosimeters (last two rows in **Table 1**), with microphones in front of ears (Södersten et al., 2002; McAllister et al., 2009); and pedagogical aspects. Representing a wide range of such factors, Sala and Rantala (2016) summarized SPLs from studies conducted in Finland, Germany, Sweden, United Kingdom, and United States over several years, reporting a range of SPLs from $L_{A,eq}$ (A-weighted equivalent energy SPL) = 42–100 dB in schools and $L_{A,eq}$ = 60–85 dB in preschools measured for periods ranging from 2 min up to five working days. Their own investigations included $L_{A,eq}$ as well as several percentile levels including L_{A90} representing the background noise in occupied CRs, L_{A10} representing the higher levels, and L_{A50} representing the median level. In terms of room acoustics, Sala and Rantala (2016) reported RTs, speech transmission index (*STI*) values (values ≥0.85 considered adequate for a wide range of hearing and learning conditions for children), and mean background noise level in unoccupied CRs as 34.5 dB (27–44 dB). Astolfi et al. (2019b) used a consistent measurement setup across CRs in Italy, which were classified either as rooms with “good” or “bad” acoustics according to the occupied rooms’ RT ($T_{20,occ}$), as listed in **Table 1**. They also reported clarity index (C_{50} in dB, ratio between the energy arriving in the first 50 ms and the remaining energy) and the ratio of useful to detrimental energy values (U_{50} in dB) to express speech intelligibility, which were highly correlated with $T_{20,occ}$. For the CRs with “good” acoustics, reported values for C_{50} and U_{50} were mostly within the range of optimum values, with $C_{50} \geq 3$ dB considered good and $U_{50} \geq 1$ dB considered optimal; and the CRs with “bad” acoustics had corresponding values outside this optimal range. Persson Waye and Karlberg (2021) reported results from a study in Sweden in unoccupied furnished rooms, before and after an acoustic intervention. Wang and Brill (2021) reported estimated noise and speech levels from measurements in the United States CRs, along with RT and C_{50} values in unoccupied rooms.

The studies mentioned above have typically used omnidirectional microphones at fixed locations and/or single-channel noise dosimeters to measure the sound environment in educational institutions. Binaural recordings of children and teachers moving freely within CRs have been performed in at least two studies (last two rows in **Table 1**) where microphones were placed in front of both ears of teachers and children in preschool CRs in Sweden, and values reported are power averages of left and right ear values. For typical daily activities in CRs, these values represent a closer representation of hearing levels for both teachers and children. The almost 6 dB difference in the mean $L_{A,eq}$ values in these two studies using similar measurement methods was partly attributed by the authors to the differences in heights and distances between the teachers and children. The values reported in Södersten et al. (2002) and McAllister et al. (2009) do not include contributions due to self-speech of the participants wearing binaural microphones, which, besides other measurement factors, may partly account for slightly lower values compared to the dosimeter values reported in Persson Waye and Karlberg (2021) which presumably include contributions due to the participant’s own speech. In the latter, significant differences were found between children and personnel amounting to 6–8 dB.

Regardless of the measurement method, the high SPLs reported in **Table 1** can have detrimental effect on adults’ well-being at work (e.g., Åhlander et al., 2011) and on children’s behavior and development (for reviews, Shield and Dockrell, 2003, 2008; Klatte et al., 2013). Unfavorable room acoustics, such as long RTs relative to the room volume, characterized as being outside the 0.5–0.8 s optimum range (in occupied rooms) in Astolfi et al. (2019b), have been shown to lower performance in phoneme identification in adults and children (Neuman and Hochberg, 1983; where RT = 0.6 s was detrimental compared to RT = 0.4 s or no reverberation), impairment in primary school children’s speech perception and listening comprehension (Klatte et al., 2010a), short-term memory (Klatte et al., 2010b), and negative effects on performance, well-being, and social climate at school (Klatte et al., 2010c). Since children spend a considerable amount of time in these educational institutions, noise assessment and control are crucial toward providing optimal learning and development environments. Noise assessment and subsequent control would benefit from long-term measurements using a consistent method for more reliable parameter values compared to previous studies, which include a range of measurement periods (see Sala and Rantala, 2016, for a review).

While measurements using omnidirectional transducers have several advantages, binaural measurements are a closer representation of hearing conditions. Binaural transducers placed near human ears, as in Södersten et al. (2002) and McAllister et al. (2009), perhaps represent one possibility, with its own set of logistical issues. HATSs have limitations in terms of fixed location, and generic head-related transfer function (HRTF); the latter characterizes the frequency-dependent amplifications in the signals when measured at the ear canal entrance (Møller et al., 1995). However, the advantages of HATSs include a potentially more robust and repeatable setup compared to putting transducers on humans, with a major limitation being the use of additional equipment that may not be as readily available as individual microphones. Another overhead includes additional binaural analyses due to processing two channels instead of one in general and the potential use of computational expensive binaural models such as those for binaural loudness (Moore and Glasberg, 2007). Yet, to avoid intrusive methods involving humans (especially children), HATSs represent a rather convenient middle ground for noise measurements in CRs, which can be used to augment information provided by standard methods using omnidirectional microphones.

In terms of HATS sizes, children have smaller ears, head, and shoulder sizes than adults, and arguably a HATS representing adult morphology may not represent those of children. Hence, differences in anthropometric sizes between adults and children need to be considered to represent children’s perspectives more appropriately. Indeed, different adult HATSs can also have different HRTFs, but for the sake of brevity, this is not explored further here, and instead the focus is on comparisons between a selected adult and child HATS. Fels et al. (2004)

reported more amplification in the higher frequency bands (starting from 4 to 5 kHz) for children vs adults HRTFs. Differences were also observed for different directions in the horizontal plane and the median plane (Fels and Vorländer, 2009). More gain in higher frequencies in children's HRTFs might explain the higher sensitivity to high-frequency sounds reported for children (Persson Waye and Karlberg, 2021). With this in mind, it may be expected that differences between the transducers might also be observable in certain room acoustic and noise parameters, such as RT and SPL, when analyzed on a band-by-band basis, especially in the higher frequency bands.

To further characterize the behavioral effects of spectral and temporal (and spatial) aspects of human sound perception, psychoacoustic models and associated parameters (hitherto based largely on adults' perception) are, at least in principle, better suited than level-based parameters. Psychoacoustic loudness is perhaps the most common (for both stationary and time-varying sounds; ISO 532-1). However, other psychoacoustic parameters such as sharpness, roughness, fluctuation strength, etc., have been useful in investigations of several subjective attributes of various sound environments for adults. However, the use of psychoacoustic parameters in CR studies has been very limited, and it is unclear whether there is any benefit in considering psychoacoustic models based on adults' perception to characterize children's perception; psychoacoustic models specifically for children, and adults' models adapted for children are possible too, but not the focus here. Yet, the scope of existing psychoacoustic parameters has the potential to complement and even go beyond investigations that are possible with SPL-based parameters. This includes, but is not limited to, exploring the higher sensitivity of children to high-frequency sounds compared to adults. This is possible by comparing, for instance, SPL of lower vs higher-frequency octave bands with psychoacoustic sharpness (S) and whether there is a benefit in using one approach over another. Additionally, one may expect higher sharpness values based on measurement with child HATS compared to adult HATS due to higher amplification in higher frequencies for children's HRTFs compared to adults' HRTFs (Fels et al., 2004). Similarly, to explore the effect of fluctuations in the sound environment on human perception, it is possible to compare the performance of SPL-based parameters that quantify the level fluctuations above the ambient SPL (e.g., L_{A10}—L_{A90}, etc.) and psychoacoustic parameters fluctuation strength (FS) and roughness (R). These psychoacoustic parameters characterize human perception to slower (FS) and faster (R) amplitude fluctuations and have been shown to be related to annoyance due to air-conditioning, and auditory distraction due to many sounds including speech in office simulations Schlittmeier et al. (2012), respectively.

This study has two main aims:

1. Providing pilot results of room acoustics (unoccupied furnished rooms) and long-term measurements during occupancy in several primary schools and daycare centers using an adult and a child HATS, representing a teacher and a child in CRs, respectively, along with an omnidirectional microphone that is used in most previous studies.
2. Studying the relationship between relevant acoustic and psychoacoustic parameters that characterize the sound environment in CRs, based on the measurements in aim 1 with regard to the differences that might be introduced by transducers including differences in anthropometric sizes. The results may be beneficial for future studies where the subjective perceptions of children and adults in CRs are characterized in relation to one or more types of these transducers.

Both the aims are steps toward determining measurement and analysis methodologies that best characterize the subjective perception of children and of adults in CRs, which should be linked in further studies to children's responses on subjective perception in child-appropriate questionnaires (e.g., Persson Waye et al., 2013).

2 METHODS

2.1 Educational Institutions Measured

Acoustic measurements were conducted in ten educational buildings in Germany (Aachen) including four primary schools and six daycare centers. In total, N = 8 CRs and N = 10 playrooms (PRs) were measured. An overview of all selected rooms is given in **Table 2** including information on connected rooms, acoustical treatments, room dimensions, and A-weighted ambient background noise levels outside occupied hours. In terms of PRs in daycare centers, most of them are directly connected with one or two smaller connected rooms (for example, an extra eating room or an extra sleeping room), which are presented as additional volumes in **Table 2**. The doors to these rooms are seldom closed to enable continuous supervision by the educators. Therefore, the room volumes of these smaller rooms were added to the overall room volumes, and they were also considered while evaluating the room acoustic measurements (**Section 2.2** and **Section 4.1**). Furnishings corresponded to the purpose of the educational institutions and remained unchanged for the acoustic measurements. None of the rooms had mechanical ventilation systems, and ventilation was mostly managed through windows that were closed for the room acoustics measurements and were open during some of the *in-situ* measurements. However, the opening of windows during *in-situ* measurements was not controlled in this study.

In the CRs, on average 22 children (f: 50.0%) and 1 adult (mostly female), while on the PRs, in average 15 children (f: 53.5%) and 2 adults (f: 91.3%) were present during noise measurements. The adults were teachers of the corresponding groups of children. Children in the primary schools were between 6–10 years old and more than 50% of adults were in the age group between 31–50 years. In the daycare centers, children were between 3–6 years old and more than 50% of the adults were between 21–40 years old.

All involved adults gave signed informed consent, and in the case of children, all parents signed the informed consent for the participation of their children. The procedure was approved by the Medical Ethics Committee at the RWTH Aachen University, Germany (EK 321/16 and EK 218/18).

2.2 Standardized Room Acoustic Measurements

2.2.1 Measurement Procedure

Room acoustic measurements were conducted in unoccupied furnished rooms according to ISO 3382-2 (International

TABLE 2 | Overview of selected classrooms (CRs) and playrooms (PRs). Background noise measurements conducted in unoccupied rooms.

Room	Connected rooms	Acoustic treatment	Area (m^2)	Height (m)	Additional volumes (area (m^2) * height (m))	Total volume (m^3)	Background noise level (dBA)	Average number of people present			
								Children		Adults	
								f	m	f	m
CR01	No	Yes	60.9	3.2		194.9	26.4	7	12	2	—
CR02	No	Yes	71.1	3.0		213.3	25.7	0	17	1	—
CR03	No	No	69.3	3.2		224.6	21.9	13	8	2	—
CR04	No	No	69.3	3.2		224.6	28.0	13	15	1	—
CR05	No	No	66.8	3.6		240.5	32.0	9	13	2	—
CR06	Sloped ceiling	No	56.7	3.4	56.7 × 1.9/2	246.5	25.2	14	9	1	—
CR07	No	Yes	82.7	3.0		248.2	25.6	24	0	1	—
CR08	Sloped ceiling	Yes	61.3	3.0	61.3 × 2.3/2	254.4	25.1	7	13	1	—
PR09	No	No	38.4	3.1		117.0	27.3	7	9	2	—
PR10	Yes	No	32.2	2.8	(2.8 + 16.2 + 8.3) × 2.8	164.8	22.4	10	6	2	1
PR11	Yes	No	54.9	2.8	7.7 × 2.8	172.8	30.0	9	10	3	—
PR12	Yes	No	58.7	3.2	5.3 × 3.2 + 3.1 × 2.0	209.7	25.7	8	6	2	—
PR13	Yes	No	45.5	3.7	12.0 × 3.7	210.5	34.0	4	2	2	—
PR14	Yes	No	49.5	2.7	23.8 × 2.7	194.3	22.2	7	10	3	—
PR15	No	No	44.2	3.0		132.6	29.7	7	7	2	—
PR16	Yes	No	44.1	2.7	(16.2 + 9.4) × 2.7	187.5	23.2	7	5	2	—
PR17	Yes	No	72.4	2.8	19.7 × 2.8	254.2	25.3	10	7	1	1
PR18	Yes	No	49.6	4.0	20.4 × 4.0	280.1	21.9	8	5	2	—

Organization for Standardization, 2008) at precision level with two source and six receiver positions. As the sound source, the Institute of Technical Acoustics (ITA)'s 3-way omnidirectional dodecahedron loudspeaker was used. Simultaneous measurements were executed with the ITA adult HATS (Schmitz, 1995) equipped with Schoeps CCM2H microphones, ITA child HATS (Fels et al., 2004) equipped with Sennheiser KE4 microphones, and a ½" diffuse field omnidirectional microphone (B&K Type 4134) as a reference. Positions were chosen according to ISO 3382-2 with as little overlap as possible without removing the furnishings inside the rooms. All receivers were positioned to represent standing situations since the chosen positions were quite far from the tables and chairs. It further represents reasonably the behavior of teachers and educators in the room, who are standing most of the time. The reference microphone was positioned at the height of 1.2 m, the ear axis of the adult HATS was adjusted to 1.5 m, and of the child HATS to 1.0 m height. The measurement signal was an exponential sweep with a duration of 5,944 s, and it was repeated five times per position.

Furthermore, for all six receiver positions, the ambient equivalent A-weighted background noise level over 30 s ($BNL_{A,eq,30s}$) was measured according to ISO 9568 (International Organization for Standardization, 1993) using the reference microphone (½" diffuse field microphone B&K Type 4134).

2.2.2 Data Processing and Analysis

Room acoustic parameters T_{20}, T_{30}, *EDT* (early decay time; to potentially represent subjective "reverberance" (Bradley, 2011)), C_{50} (clarity index), D_{50} (definition), and T_S (center time) were computed according to ISO 3382-2 (International Organization for Standardization, 2008) and ISO 3382-1 (International Organization for Standardization, 2008) for a frequency range of 125 Hz–16 kHz octave bands center frequencies. The A-weighted background noise level over 30 s was evaluated according to ISO 9568 (International Organization for Standardization, 1993) for a frequency range of 31.5 Hz–16 kHz. The *STI* was calculated using the indirect method following IEC 60268-16 (International Electrotechnical Commission, 2012), which computed the STI using the measured impulse response neglecting effects from masking and background noise. Hereby, MATLAB and the ITA toolbox (Berzborn et al., 2017) were used. Since some of the rooms measured were connected with smaller volumes (**Table 2**), the degree of non-linearity in the reverberant energy decay of the measured impulse responses was examined using the method in Annex B of ISO 3382-2. All measurement positions where the degree of curvature of the decay (comparing T_{30} and T_{20}) for the reference microphone exceeded the 10% threshold, signifying substantial deviation from linearity, were removed from further room acoustic analyses (**Table 3**) (International Organization for Standardization, 2008). Subsequently, only T_{30} values are reported. Results including T_{20} values are provided in **Supplementary Table S2**.

To approximate binaural versions of the standard room acoustic parameters, two approaches were considered. Firstly, the computed parameters from the left and right ear were averaged ($\frac{^{X}Left + ^{X}Right}{2}$). This method is indicated in the following with "A-HATS$_{Av}$/C-HATS$_{Av}$", representing the values from the adult and child HATSs, respectively. Secondly, the value from the prominent ear was chosen. In this work, it is assumed to be the higher value out of the left and right ear values. The idea here is that the prominent ear represents the conservative approximation of a binaural model, except for *STI* in which the higher value or the "better-ear" *STI*

TABLE 3 | Room acoustic parameters from omnidirectional microphone for classrooms (CRs) and playrooms (PRs).

Room	N_{Pos}	T_{30} (s)		EDT (s)		C_{50} (dB)		D_{50} (%)		T_S (s)		STI
		BB	Mid	BB	Mid	BB	Mid	BB	Mid	BB	Mid	Mean
CR01	11	0.50	0.49	0.42	0.44	7.8	6.7	83.2	81.8	0.03	0.03	0.78
CR02	11	0.59	0.60	0.53	0.58	5.0	4.0	74.3	70.8	0.04	0.04	0.72
CR03	12	0.97	1.05	0.88	1.02	1.6	0.3	58.1	51.9	0.06	0.07	0.61
CR04	10	0.97	1.10	0.90	1.06	2.0	0.5	58.6	52.7	0.06	0.07	0.61
CR05	10	0.54	0.59	0.42	0.46	7.5	6.3	83.0	80.1	0.03	0.03	0.77
CR06	11	0.54	0.62	0.52	0.63	5.6	3.9	76.5	70.4	0.03	0.04	0.71
CR07	10	0.61	0.58	0.54	0.58	4.5	3.7	72.0	69.2	0.04	0.04	0.71
CR08	8	0.51	0.56	0.49	0.54	6.6	5.5	80.3	77.4	0.03	0.03	0.74
PR09	12	0.47	0.50	0.44	0.50	6.5	4.7	78.5	74.0	0.03	0.04	0.75
PR10	6	0.42	0.41	0.37	0.36	8.8	8.5	84.8	86.5	0.03	0.02	0.81
PR11	10	0.35	0.34	0.31	0.32	10.4	9.5	89.7	88.5	0.02	0.02	0.82
PR12	8	0.47	0.49	0.46	0.48	6.4	5.0	78.1	75.2	0.03	0.04	0.76
PR13	11	0.56	0.58	0.48	0.54	6.6	4.9	78.3	74.3	0.03	0.04	0.75
PR14	10	0.45	0.42	0.44	0.45	7.3	5.9	78.5	77.4	0.03	0.03	0.79
PR15	10	0.48	0.49	0.42	0.45	7.6	6.3	82.5	80.2	0.03	0.03	0.78
PR16	4	0.47	0.47	0.44	0.44	7.0	6.7	78.3	78.4	0.03	0.03	0.79
PR17	12	0.53	0.46	0.48	0.46	6.7	6.4	78.9	79.8	0.03	0.03	0.76
PR18	12	0.68	0.78	0.62	0.73	4.6	2.8	69.9	64.9	0.04	0.05	0.69

Note. *BB = broadband average over octave bands with center frequency range 125 Hz–16 kHz, mid = average over octave bands with 500 Hz, 1 kHz, and 2 kHz center frequencies.*

(signifying better signal-to-noise ratio) was used as it has been shown to perform well in relation to a binaural *STI* model (van Wijngaarden and Drullman, 2008). This method is referred to with "A-$HATS_{Prom}$/ C-$HATS_{Prom}$" for the adult and child HATSs, respectively.

All the room acoustic parameters were examined according to the factors "connected rooms" (single room vs coupled rooms vs rooms with sloping ceilings), "acoustic treatment" (with vs without), "room type" (CRs vs PRs), and "measurement method" (omnidirectional microphone vs A-$HATS_{Av}$ vs C-$HATS_{Av}$) with respect to 125 Hz–16 kHz octave bands center frequencies. One-way ANOVAs were carried out for the room acoustic parameters (as listed in **Table 3** with a single value for each room) to examine possible differences between the measurement methods (e.g., between omnidirectional microphone vs A-$HATS_{Av}$, etc.).

2.3 *In-situ* Noise Measurements

2.3.1 Measurement Procedure

The *in-situ* noise measurements were conducted during the daily activities of children in CRs and PRs. The same equipment as stated for the room acoustic measurements in unoccupied rooms was used to execute the *in-situ* measurements. All three receivers were positioned together in the center (less than 30 cm to each other, cf. **Figure 1**) of the main room of activity so that people were able to move around them. While this potentially introduces acoustic shadowing and interference issues for the transducers, the location of the transducers was due to logistical concerns including ensuring that the measurement equipment did not adversely interfere with the usual behavior of the adults and children (e.g., by attracting too much children's attention). The positioning of measurement equipment was discussed forehand with the teachers, and the study and the equipment were explained to the children 1 day before the measurements started in each educational institution.

In-situ measurements were conducted over 2 days per CR and PR during normal daily activities. On the first day, all dummy heads were positioned to represent a standing position (ear axis of the adult HATS at 1.5 m, ear axis of the child HATS at 1.0 m, and the omnidirectional microphone at 1.2 m). On the second day, they were positioned to represent a sitting position (ear axis of the adult HATS at 1.2 m, ear axis of the child HATS at 0.8 m, and the omnidirectional microphone at 1.2 m) or to represent a playing height (ear axis of the child HATS at 0.5 m), respectively, according to the dominant scenario of each educational institution. Only periods with children present in the room were considered, where the sound pressure level exceeded 35 dBZ (Z-weighted). In other words, a cutoff sound pressure level of 35 dBZ was used to distinguish between children's presence and absence in the rooms. This cutoff was based on inspecting several samples within the recordings when children were not present. This resulted in up to 6 hours of recordings on average per room, which were used for further analyses.

2.3.2 Data Processing and Analysis

Three types of noise parameters were taken into consideration. Firstly, parameters based on A-weighted SPL were computed, including $L_{A,eq}$ and percentiles (L_{A10} and L_{A90}). Secondly, to consider sound fluctuations over time in rooms above the background noise (generally signified using L_{A90}), level-based fluctuation parameters typically used in areas with multi-talker speech, for example, open-plan offices (Yadav et al., 2021), were calculated. These included noise climate $NCl = L_{A10} - L_{A90}$ (Kryter, 2013), noise pollution level $NPL = L_{A,eq} + (L_{A10} - L_{A90})$ (Ayr et al., 2003), and $M_{A,eq} = L_{A,eq} - L_{A90}$ (Lenne et al., 2020). Thirdly, a set of psychoacoustic parameters was computed using the ArtemiS SUITE 11.0 by HEAD Acoustics (Herzogenrath, Germany): loudness N for time-varying sounds, with the unit "sone" following ISO 532-1 (International Organization for Standardization, 2017); sharpness S with the unit "acum" according to DIN 45692 (Deutsches Institut für Normung, 2009), roughness R (unit: "asper"); and fluctuation strength FS (unit: "vacil") according to the Hearing Model by Sottek (1993). To address time dependent effects over

FIGURE 1 | Example of the centered positioning of the measurement transducers during an *in-situ* measurement in a classroom.

the entire measurement period to an extent, the parameters were calculated for 15-min frames and then averaged over all frames, that is, the whole measurement period.

To further understand the relation between the sharpness and A-weighted SPL with respect to low and high frequencies (below and above 1 kHz), a low–high-frequency ratio of $L_{A,eq}$ was calculated as follows: $L_{A,eq(L,v,H)} = \frac{mean(L_{A,eq[31.5-1k HZ]})}{mean(L_{A,eq[12.5-16k HZ]})}$. For all noise parameters, the binaural parameters were calculated using the average and prominent ear methods described in **Section 2.2.2**. For the A-weighted SPL parameters, level summation was computed using the left and right ear SPL values instead of the averaging method (further also indicated as A-HATS$_{Av}$/C-HATS$_{Av}$). The A-weighted SPL $L_{A,eq}$ was examined according to the factors "connected rooms" (single room vs coupled rooms vs rooms with sloping ceilings), "acoustic treatment" (with vs without), "room type" (CRs vs PRs), "measurement method" (omnidirectional microphone vs A-HATS$_{Av}$ vs C-HATS$_{Av}$) differences between $L_{Z,eq}$, $L_{A,eq}$, L_{A10}, and L_{A90} with respect to the one-third octave bands with center frequencies between 31.5 Hz and 16 kHz. One-way ANOVAs were carried out for all noise parameters to examine possible differences between the measurement methods (omnidirectional microphone vs A-HATS$_{Av}$ vs C-HATS$_{Av}$ vs A-HATS$_{Prom}$ vs C-HATS$_{Prom}$).

3 RESULTS

3.1 Room Acoustics (Unoccupied, Furnished Rooms)

Table 3 lists the results of the room acoustic measurements (using the omnidirectional microphone) of all rooms averaged over all positions that met the decay curvature criteria within ISO 3382–2 (i.e., those with decay curvature under 10%; Section 2.2.2), and **Table 4** presents a summary of the details provided in **Table 3**. Results averaged over all measured positions are provided in **Supplementary Table S1**. In four cases (CR08, PR10, PR12, and PR16; **Table 3**), ≥4 positions had to be discarded. In further analyses, PR16 was excluded due to especially a low number of measurement positions that met the curvature criteria.

The room acoustics of two CRs (CR03 and CR04; **Table 3**) were noticeably different from the other six (averaged T_{30} = 0.55 s and averaged STI = 0.74 in these CRs). Possible explanation could be combined effect of the large room volumes with flat ceilings compared to other larger volumes like C06, C07, and the absence of acoustic treatment in these rooms; although C05 still has comparable room volume and room acoustics to C06 and C07, it has no acoustic treatment and has a flat ceiling similar to C03 and C04.

Figure 2 presents the relationship between *STI* and other room acoustic parameters. **Figure 3** presents the mean values of room acoustic parameters over 125 Hz–16 kHz octave band center frequencies, grouped according to room type.

All results of the room acoustic measurements using the adult and child HATSs including both evaluation methods (averaging and prominent-ear) can be found in the supplementary material (**Supplementary Table S3–S6**) including results on the inter-aural correlation coefficient (IACC) though it is not further discussed in this work (**Supplementary Table S7** and **Supplementary Table S8**).

Figure 4 presents the mean value of room acoustic parameters over 125 Hz–16 kHz octave band center frequencies, grouped according to the three measurement methods: using the omnidirectional microphone (Ref), adult and child HATSs with the averaging method, that is A-HATS$_{Av}$ and C-HATS$_{Av}$, respectively. Results for the prominent-ear method can be found in **Supplementary Figure S1**. For the mid-frequency octave bands (500 Hz, 1

TABLE 4 | Summary of room acoustic parameters in **Table 3**. Reported values are mean and the range (in brackets).

Room acoustic parameters		All	Classrooms (CRs)	Playrooms (PRs)
T_{30} (s)	BB	0.57 (0.35–0.97)	0.65 (0.50–0.97)	0.49 (0.35–0.68)
	Mid	0.60 (0.34–1.10)	0.70 (0.49–1.10)	0.49 (0.34–0.78)
EDT (s)	BB	0.52 (0.31–0.90)	0.59 (0.42–0.90)	0.45 (0.31–0.62)
	Mid	0.57 (0.32–1.06)	0.66 (0.44–1.06)	0.47 (0.32–0.73)
C_{50} (dB)	BB	6.1 (1.6–10.4)	5.1 (1.6–7.8)	7.2 (4.6–10.4)
	Mid	5.0 (0.3–9.5)	3.9 (0.3–6.7)	6.1 (2.8–9.5)
D_{50} (%)	BB	76.5 (58.1–89.7)	73.2 (58.1–83.2)	79.7 (69.9–89.7)
	Mid	73.6 (51.9–88.5)	69.3 (51.9–81.8)	77.9 (64.9–88.5)
T_S (s)	BB	0.04 (0.02–0.06)	0.04 (0.03–0.06)	0.03 (0.02–0.04)
	Mid	0.04 (0.02–0.07)	0.04 (0.03–0.07)	0.03 (0.002–0.05)
STI	Mean	0.74 (0.61–0.82)	0.71 (0.61–0.78)	0.77 (0.69–0.82)

Note. *BB = broadband average over octave bands with center frequency range 125 Hz–16 kHz, mid = average over octave bands with 500 Hz, 1 kHz, and 2 kHz center frequencies.*

and 2 kHz center frequencies), values were very similar for all the room acoustics parameters across the measurement methods. Beyond 2 kHz, some deviations can be seen, which can broadly be attributed to the anthropomorphic features (i.e., HRTFs) of the binaural transducers becoming important for smaller wavelengths. In this regard, the values for the adult HATS varied more in comparison to the other measurement methods. Furthermore, for the parameters C_{50}, D_{50}, and T_S, which are all ratios of early sound energy to late/reverberant energy, the values for the child and adult HATSs are similar but deviated from the omnidirectional microphone for the 4 kHz band, and the parameter values for the adult HATS exhibited a distinct deviation in comparison to the corresponding values for the child HATS and the omnidirectional microphone, which have similar values, for the 8 kHz octave band. These deviations, starting from the 4 kHz octave band, can also be observed in the HRTF magnitude response of child HATS in comparison to adult HATS as presented in Fels et al. (2004). However, the statistical analyses revealed no significant difference between the five measurement methods (Ref vs A-HATS$_{Av}$ vs C-HATS$_{Av}$ vs A-HATS$_{Prom}$ vs C-HATS$_{Prom}$; **Supplementary Table S9**).

3.2 *In-situ* Acoustics (Occupied Rooms)

Table 5 and **Table 6** present the noise parameters calculated from the omnidirectional microphone recordings for each room, which are summarized in **Table 7**.

Differences between CRs in primary schools and PRs in daycare centers are mainly observable in loudness (average N = 11.9 sone (9.3–15.3 sone) vs 10.8 sone (7.5–13.9 sone); **Table 7**) and in sharpness (average S = 1.5 acum (1.38–1.63 acum) vs 1.46 acum (1.40–1.51 acum), and in the percentiles (average N_5 = 25.1 vs 22.9 sone). An increase in loudness is understandably related to increasing $L_{A,eq}$, with an R^2 of 0.85. Loudness N_{mean} (as y) is predicted from $L_{A,eq}$ (as x) with the equation y = −61.90x + 1.12. Almost no relationship between increasing sharpness and increasing high-frequency content in the *in-situ* sound was found (R^2 of 0.00 for Ref, A-HATS$_{Av}$, and A-HATS$_{Prom}$; R^2 of 0.02 for C-HATS$_{Av}$; R^2 of 0.04 for C-HATS$_{Prom}$).

Figure 5 shows the SPL variation according to the different measurement methods, room conditions, and SPL-based parameters. Results of the level-based and the *in-situ* psychoacoustic parameters from the other measurement methods using HATSs including both evaluation methods (averaging and prominent-ear) for each room can be found in **Supplementary Tables S10–S17**.

3.2.1 Variation in Noise Parameters Across the Measurement Methods

Figure 6 shows the difference between the mean values for several parameters across the measurement methods (Ref vs A-HATS$_{Av}$ vs C-HATS$_{Av}$ and Ref vs A-HATS$_{Prom}$ vs C-HATS$_{Prom}$). Statistical analyses revealed differences in all noise parameters, except for the two level-based sound fluctuation parameters *NCl* and $M_{A,eq}$ (**Supplementary Table S18**). For the level-based parameters, the post-hoc analyses yielded significant differences for all comparisons across measurement methods for the parameter *NPL*, with significant differences between the Ref and HATS values for $L_{A,eq}$ using the averaging method and significant difference between the HATSs only for $L_{A,eq}$ using the prominent ear method. In terms of the low–high-frequency ratio of $L_{A,eq}$, the post-hoc analyses revealed significant differences between the Ref and child HATS values while the Ref and adult HATS values were not significant for both averaging and prominent ear method. Differences between HATSs in $L_{A,eq(L.v.H)}$ for both evaluation methods were significant.

For the psychoacoustic parameters, the post-hoc analyses showed no significant differences for loudness (N_{mean} and N_5) between both HATSs using the averaging method and the omnidirectional microphone, while differences were significant in terms of using the prominent-ear method. However, no differences were found between the adult and child HATSs. Considering N_{90}, differences between omnidirectional microphone and adult HATS were significant as well between the two HATSs. For sharpness (S_{mean}), all results from the HATS were significantly different to the omnidirectional microphone and within each other (A-HATS$_{Av}$ vs C-HATS$_{Av}$ and A-HATS$_{Prom}$ vs C-HATS$_{Prom}$). However, for S_{90}, no differences between the HATSs with both evaluation methods were observed. For roughness (R_{mean}, R_5 and R_{90}), differences were observed between the omnidirectional microphone and the HATSs using the

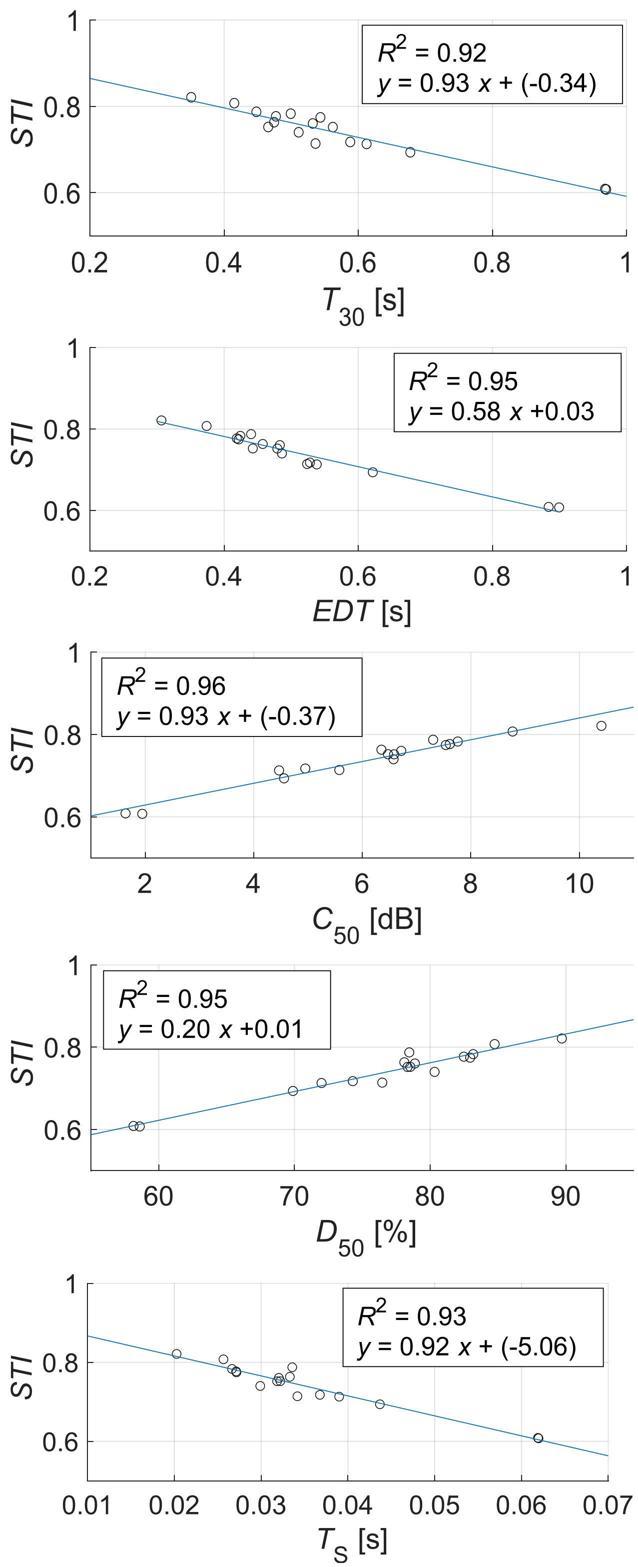


FIGURE 2 | Linear regression models between speech transmission index and other room acoustics parameters.

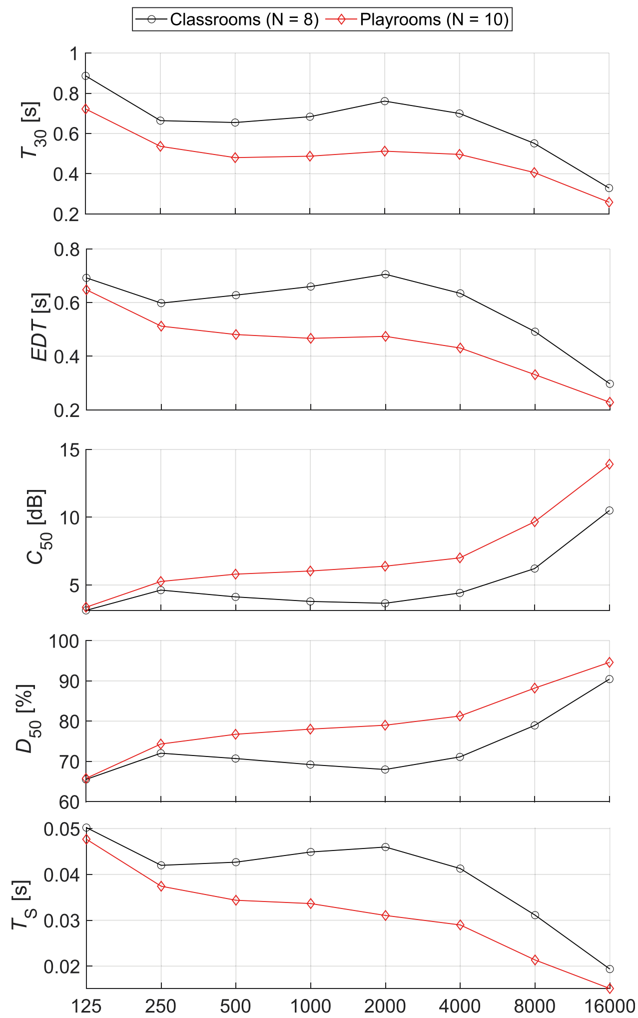


FIGURE 3 | Room acoustic parameters over 125 Hz–16 kHz octave band frequencies for classrooms and playrooms.

averaging method, while differences between the HATSs and both evaluation methods were not significant. For fluctuation strength, all measurement methods were significantly different from each other for FS_{90}, while FS_{mean} was only significantly different for omnidirectional microphone vs C-HATS$_{Av}$; and for FS_5, the only significant difference was between the omnidirectional microphone and the HATSs using the averaging evaluation method.

3.2.2 Correlation Between Noise Parameters Differences and Room Volume

Supplementary Table S19 summarizes the results of the correlation analyses between the noise parameters and the room volume according to the different measurement methods. Significant correlations were only yielded for R_5, FS_{mean}, and FS_5. In R_5, only the adult and child HATSs' measurement results using the averaging method were

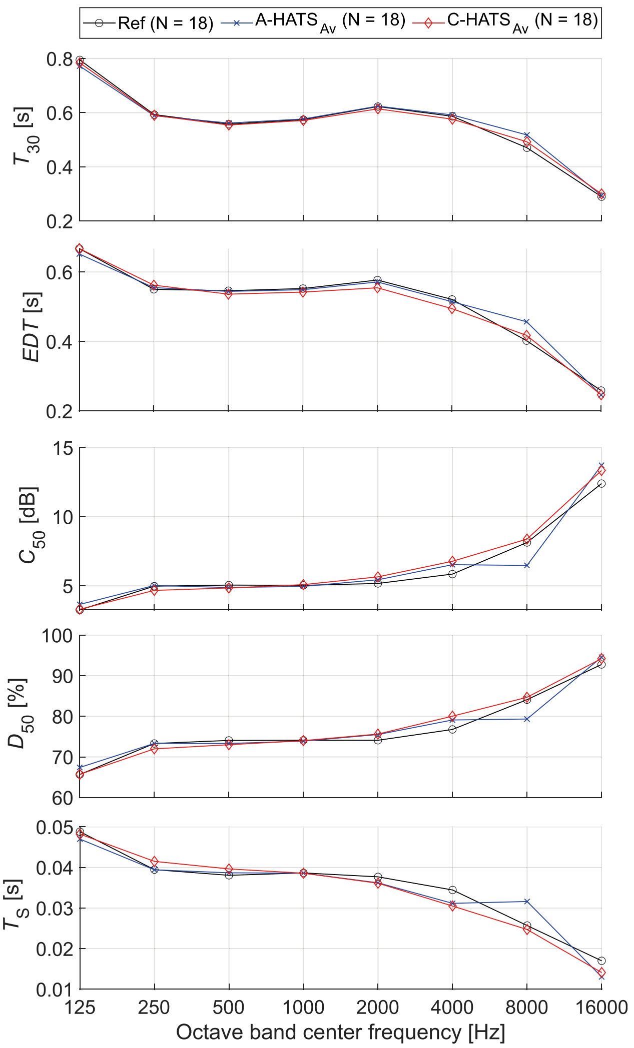


FIGURE 4 | Room acoustic parameters according to the different measurement methods: omnidirectional microphone (Ref), adult and child HATSs with the averaging method (A-HATS$_{Av}$/C-HATS$_{Av}$).

correlated with the room volume, while the FS_{mean} showed a correlation above 0.5 for all HATS methods. All correlations with these methods were negatively correlated.

4 DISCUSSION

4.2 Evaluation of Room Acoustics

Results from this study (as listed in **Table 4**) add to the insights from previous studies across various countries (Klatte et al., 2013; Sala and Rantala, 2016; Astolfi et al., 2019b; Persson Waye and Karlberg, 2021; Wang and Brill, 2021). In the following, the mid-frequency octave bands (500 Hz, 1 kHz, 2 kHz; mid in **Table 3** and **Table 4**) values will be discussed (except for *STI*), which were used in most previous studies in **Table 1**. In general, primary school CRs measured in Germany had higher RTs (T_{30}) and marginally lower *STI* values than CRs in Finland (Sala and Rantala, 2016), but are within the range of corresponding values reported in previous studies summarized in Sala and Rantala from CR in several countries. Current sample of PRs had similar RT values compared to preschools in Sweden in a recent study (Persson Waye and Karlberg, 2021; **Table 1**), albeit with the largest RT reported there being 0.5 s compared to 0.7 s in the current sample. Overall, PRs had better room acoustic properties than CRs, with lower *T30*, *EDT*, and T_S, and higher C_{50}, D_{50}, and *STI* values. Rooms with acoustic treatment understandably yielded better room acoustic values. PRs with coupled volumes had better room acoustic properties than single-volume PRs. However, this needs to be interpreted with caution since additional volumes, besides involving more complicated energy decays and subsequent analyses (**Section 2.2.2** and **Table 3**), also provide extra sound absorption, which would disappear if the doors connecting these volumes were shut. Yet, the current findings are closer in representing the daily operation in PRs, where these doors are typically kept open.

If the room acoustic classification from the work by Astolfi et al. (2019b) is used, which is based on criteria for both T_{20} (T_{30} values used for the current sample) and C_{50} values, although for occupied rooms, all CRs except CR03 and CR04 would be classified as having "good acoustics," while all PRs except PR10, 11, 14, and 18 would be classified as having "good acoustics." It is likely, however, that some of these values may meet the criteria for "good acoustics" as per Astolfi et al. (2019b) if the measurements were conducted during occupation, as was shown for university CRs, especially with mostly reflective surfaces (Choi, 2016). *STI* (unoccupied) is recommended to be at least ≥0.80 and ≥0.85 for educational institutions for children without and with hearing, cognition, and/or behavioral issues, respectively (Finnish Standards Association, 2004). For the current sample, except for PR10 and PR11, none of the CRs or PRs meet the recommended *STI* values for even children without hearing and/or learning difficulties. *STI*, overall, had a strong linear relationship with RT (and other room acoustic parameters), similar to previous studies (Leccese et al., 2018; although not in Sala and Rantala, 2016), which can be used to estimate global *STI* values based on the simpler way to calculate RT values. While this implies that RT could perhaps be used as a primary indicator to represent the room acoustics in CRs, more studies with larger sample sizes are needed to determine the strength of the relationship between *RT* and *STI*, and their relationship with subjective impressions. U_{50} values, which have been used in several studies of CR acoustics, were not calculated for the current sample due to the measurement issues. The relevance of the good/bad acoustics in the current sample can further be explored based on subjective impressions of children in these rooms, as was done in Astolfi et al. (2019b), which is proposed for a future study. Nevertheless, based on room acoustics measurements alone, with higher RTs in some rooms and lower values for intelligibility than recommended, most CRs measured in the current study are likely to be not optimal for learning purposes (however, these rooms in occupied conditions might be better suited though it is not examined in this study) and may affect cognitive and behavioral development of children (Klatte et al., 2010a; 2010b; 2010c), and especially for children

TABLE 5 | *In-situ* sound pressure level parameters from omnidirectional microphone.

Room	$L_{A,eq}$ (dB)	$L_{A,eq(L.v.H.)}$	L_{A10} (dB)	L_{A90} (dB)	*NCI* (dB)	*NPL* (dB)	$M_{A,eq}$ (dB)
CR01	65.9	0.82	69.0	45.9	23.2	70.0	20.1
CR02	64.2	0.85	64.4	37.5	26.9	67.5	26.6
CR03	67.9	0.80	71.5	49.9	21.7	71.5	18.0
CR04	65.8	0.84	69.1	38.9	30.3	71.8	26.9
CR05	64.3	0.86	67.4	45.8	21.6	67.1	18.5
CR06	65.8	0.85	69.3	47.7	21.6	68.8	18.1
CR07	64.4	0.83	66.9	40.4	26.5	68.3	24.0
CR08	67.3	0.83	70.5	45.3	25.2	71.0	21.9
PR09	67.1	0.84	70.3	49.5	20.8	70.1	17.6
PR10	64.0	0.85	66.8	42.6	24.2	67.9	21.4
PR11	67.4	0.80	70.4	49.1	21.4	71.3	18.4
PR12	67.1	0.85	70.1	50.0	20.1	70.2	17.1
PR13	61.8	0.80	63.9	38.1	25.8	65.3	23.7
PR14	68.8	0.82	71.6	51.4	20.2	72.0	17.4
PR15	63.4	0.80	66.0	39.6	26.3	68.5	23.8
PR16	64.1	0.83	66.2	41.7	24.4	67.5	22.4
PR17	64.4	0.84	67.2	42.2	24.9	68.0	22.2
PR18	67.0	0.83	69.8	48.1	21.7	70.1	18.9

Note. *CR = classroom (primary school), PR = playrooms (day care center).*

TABLE 6 | *In-situ* psychoacoustic parameters from omnidirectional microphone.

Room	N_{mean}	N_5 (sone)	N_{90}	S_{mean}	S_5 (acum)	S_{90}	R_{mean}	R_5 (asper)	R_{90}	FS_{mean}	FS_5 (vacil)	FS_{90}
CR01	12.1	24.5	4.5	1.57	2.13	1.27	0.03	0.05	0.01	0.09	0.19	0.04
CR02	9.3	20.8	2.7	1.38	1.88	0.98	0.03	0.10	0.01	0.05	0.14	0.01
CR03	15.3	29.4	6.3	1.55	2.00	1.28	0.03	0.05	0.02	0.08	0.16	0.04
CR04	10.5	26.3	2.9	1.48	1.94	1.16	0.02	0.05	0.01	0.06	0.13	0.02
CR05	11.7	23.5	5.0	1.54	2.03	1.23	0.03	0.05	0.01	0.07	0.15	0.03
CR06	13.0	26.4	5.4	1.62	2.13	1.31	0.03	0.05	0.01	0.09	0.19	0.04
CR07	10.1	22.2	3.4	1.57	2.06	1.24	0.02	0.05	0.01	0.07	0.18	0.01
CR08	13.3	27.7	4.6	1.63	2.13	1.32	0.03	0.05	0.01	0.09	0.19	0.04
PR09	13.1	26.1	5.8	1.47	1.93	1.18	0.03	0.05	0.02	0.10	0.21	0.04
PR10	9.1	20.3	3.0	1.49	2.03	1.16	0.02	0.05	0.01	0.10	0.22	0.04
PR11	12.5	25.5	5.1	1.51	1.98	1.24	0.03	0.05	0.01	0.11	0.24	0.05
PR12	13.3	25.9	6.0	1.47	1.90	1.19	0.03	0.05	0.02	0.10	0.19	0.05
PR13	7.5	17.5	2.3	1.46	2.06	1.10	0.02	0.05	0.01	0.08	0.19	0.03
PR14	13.9	27.4	6.3	1.42	1.85	1.15	0.03	0.05	0.02	0.10	0.21	0.04
PR15	8.2	19.0	2.5	1.48	2.04	1.14	0.02	0.05	0.01	0.09	0.20	0.03
PR16	8.5	19.4	2.9	1.48	2.03	1.14	0.02	0.05	0.01	0.10	0.20	0.04
PR17	9.6	22.1	3.3	1.40	1.88	1.09	0.02	0.05	0.01	0.10	0.21	0.03
PR18	12.5	25.5	5.2	1.45	1.91	1.17	0.03	0.05	0.01	0.08	0.17	0.04

Note. *CR = classroom (primary school), PR = playroom (daycare center).*

with hearing loss and/or learning difficulties (Crandell and Smaldino, 2000). The CRs that have "bad acoustics" (CR03, 04; **Table 2**) did not have any acoustic treatment, although the same was true of some CRs that could be classified as with "good acoustics" as per Astolfi et al. (2019b). None of the PRs had any substantial acoustic treatment but some of these may benefit from the extra absorption due to open doors to connecting volumes. With this in mind, **Figure 3** shows that there is plenty of scope to improve the room acoustic conditions in the CRs and the PRs (if adjoining volumes are not included), with basic acoustic treatment to manage the excessive reverberation in the higher frequency while improving/maintaining appropriately high speech intelligibility (which could be estimated using **Figure 2**). Given that high-frequency sound absorption is relatively straightforward to accomplish (e.g., ceiling tiles, carpet, and/or wall absorption), and may partly be provided by the occupants, the results here are at least encouraging in terms of providing some impetus and guidance for solving the issues related to bad/insufficient room acoustics in CRs.

4.2 Evaluation of the Level-Based Parameters

The results in **Table 7** add to the previous *in-situ* measurements in primary schools and daycare centers and introduce some level-based sound fluctuation parameters previously used in other

TABLE 7 | Summary of noise parameters. Reported values are mean and the range (in brackets).

Noise parameters		All	Classrooms	Playrooms
$L_{A,eq}$ (dB)	Ref	65.6 (61.8–68.8)	65.7 (64.2–67.9)	65.5 (61.8–68.8)
	A-HATS$_{Av}$	73.0 (70.7–76.1)	72.3 (70.9–74.5)	73.6 (70.7–76.1)
	C-HATS$_{Av}$	72.0 (69.5–75.5)	70.8 (69.5–73.4)	72.9 (70.7–75.5)
	A-HATS$_{Prom}$	69.8 (66.7–76.7)	69.6 (66.7–76.7)	70.0 (68.0 74.5)
	C-HATS$_{Prom}$	67.8 (65.0–71.4)	66.9 (65.0–71.4)	68.5 (66.4–71.2)
$L_{A,eq(L.v.H)}$	Ref	0.83 (0.80–0.86)	0.84 (0.76–0.86)	0.82 (0.80–0.85)
	A-HATS$_{Av}$	0.84 (0.81–0.86)	0.84 (0.83–0.86)	0.83 (0.81–0.86)
	C-HATS$_{Av}$	0.80 (0.78–0.84)	0.80 (0.78–0.84)	0.80 (0.78–0.83)
	A-HATS$_{Prom}$	0.82 (0.80–0.85)	0.83 (0.81–0.85)	0.82 (0.80–0.84)
	C-HATS$_{Prom}$	0.78 (0.76–0.82)	0.79 (0.76–0.82)	0.78 (0.76–0.82)
L_{A10} (dB)	Ref	68.4 (63.9–71.6)	68.5 (64.4–71.52)	68.2 (63.9–71.6)
L_{A90} (dB)	Ref	44.6 (37.5–51.4)	43.9 (37.5–49.9)	45.2 (38.1–51.4)
NCI (dB)	Ref	23.8 (20.1–30.3)	24.6 (21.6–30.3)	23.0 (20.1–26.3)
NPL (dB)	Ref	69.3 (65.3–72.0)	69.5 (67.5–71.8)	69.1 (65.3–72.0)
$M_{A,eq}$ (dB)	Ref	21.0 (17.1–26.9)	21.8 (18.0–26.9)	20.3 (17.1–23.8)
N_{mean} (sone)	Ref	11.4 (7.5–15.3)	11.9 (9.3–15.3)	10.8 (7.5–13.9)
N_5 (sone)	Ref	24.0 (17.5–29.4)	25.1 (20.8–29.4)	22.9 (17.5–27.4)
N_{90} (sone)	Ref	4.3 (2.3–6.3)	4.4 (2.7–6.3)	4.2 (2.3–6.3)
S_{mean} (acum)	Ref	1.50 (1.38–1.63)	1.54 (1.38–1.63)	1.46 (1.40–1.51)
S_5 (acum)	Ref	2.00 (1.85–2.13)	2.04 (1.88–2.13)	1.96 (1.85–2.06)
S_{90} (acum)	Ref	1.19 (0.98–1.32)	1.22 (0.98–1.32)	1.15 (1.09–1.24)
R_{mean} (asper)	Ref	0.03 (0.02–0.03)	0.03 (0.02–0.03)	0.03 (0.02–0.03)
R_5 (asper)	Ref	0.05 (0.05–0.10)	0.06 (0.05–0.10)	0.05 (0.05–0.05)
R_{90} (asper)	Ref	0.01 (0.01–0.02)	0.01 (0.01–0.02)	0.01 (0.01–0.02)
FS_{mean} (vacil)	Ref	0.09 (0.05–0.11)	0.08 (0.05–0.09)	0.10 (0.08–0.11)
FS_5 (vacil)	Ref	0.18 (0.13–0.24)	0.17 (0.13–0.19)	0.20 (0.17–0.24)
FS_{90} (vacil)	Ref	0.03 (0.01–0.05)	0.03 (0.01–0.04)	0.04 (0.03–0.05)

Note. *Ref = omnidirectional microphone; A-HATS$_{Av}$/C-HATS$_{Av}$ = adult and child HATS, with the averaging method; A-HATS$_{Prom}$/C-HATS$_{Prom}$ = adult and child HATS, with the prominent-ear method.*

fields with multi-talker speech environments like open-plan offices (Yadav et al., 2021). For omnidirectional microphones, the mean $L_{A,eq}$ values in CRs and PRs were almost the same (~66 dB), with a relatively wider range of values in the latter. These values are within the range of values reported in previous studies, but the range of values in the current sample generally has lower upper limits, that is, the CRs and PRs with higher $L_{A,eq}$ values had lower $L_{A,eq}$ values compared to previous studies. This includes $L_{A,eq}$ values reported from omnidirectional measurements in CRs of Italy (Astolfi et al., 2019a; 2019b), Finland (Sala and Rantala, 2016), and United States (for the speech levels, Wang and Brill, 2021), where the measurement devices were placed at fixed locations.

Omnidirectional $L_{A,eq}$ values in the current study were around 20 dB lower than values reported in studies wherein children wore dosimeters (Persson Waye and Karlberg, 2021), and in McAllister et al. (2009) where children in daycare centers had a microphone placed near each ear. Compared to McAllister et al. (2009), where children were free to move around, the HATSs in the current study had fixed locations and with microphones at the entrance of the ear canal instead. This, combined with the overall quieter CRs in the current sample, may partly explain the lower mean binaural $L_{A,eq}$ values in the current sample of PRs for the adult and child HATSs of around 13 dB, and 14.5 dB, respectively, for the prominent ear values. The mean $L_{A,eq}$ values calculated using the adult HATS were higher than those for the child HATS, which is opposite to what was reported in McAllister et al. (2009), where they compared values of similar measurement methods using binaural measurements for adults (Södersten et al., 2002) and children. McAllister et al. (2009) had partly attributed their results to children being the primary noise sources, which is also relevant for the current sample. Hence, the counterintuitive finding of higher $L_{A,eq}$ for the adult compared to child HATS in the current study, which is most likely due to the particular transducer placement, is suggested as a question for future research. At the very least, this comparison highlights the issues in the selection of transducers for child-appropriate *in-situ* studies where the location of the transducers is fixed.

Moreover, $L_{A,eq}$ values calculated using omnidirectional microphone were significantly different from HATS, and the adult HATS was at least significantly different from child HATS for the prominent ear condition (**Section 3.2** and **Supplementary Table S18**). Based on the octave-band spectra in **Figure 5**, differences between the adult and child HATSs (SPLs calculated using level summation for the left and right ear values) and the omnidirectional microphone are largely linked to the 6 dB introduced by the level summation till around 1 kHz, followed by a more complicated trend till 16 kHz. There can be many contributing factors here, including the peak (around 4 kHz) and notches (around 8 and 10 kHz) in the magnitude response of the adult HATS HRTFs. In terms of the prominent ear values for the adult and child HATSs, no differences to the omnidirectional microphone are observed up to 2 kHz. Differences around 4 kHz and beyond 8 kHz are again observable as in the room acoustic parameters, which is in line with the work by Fels et al. (2004) and Fels and Vorländer (2009),

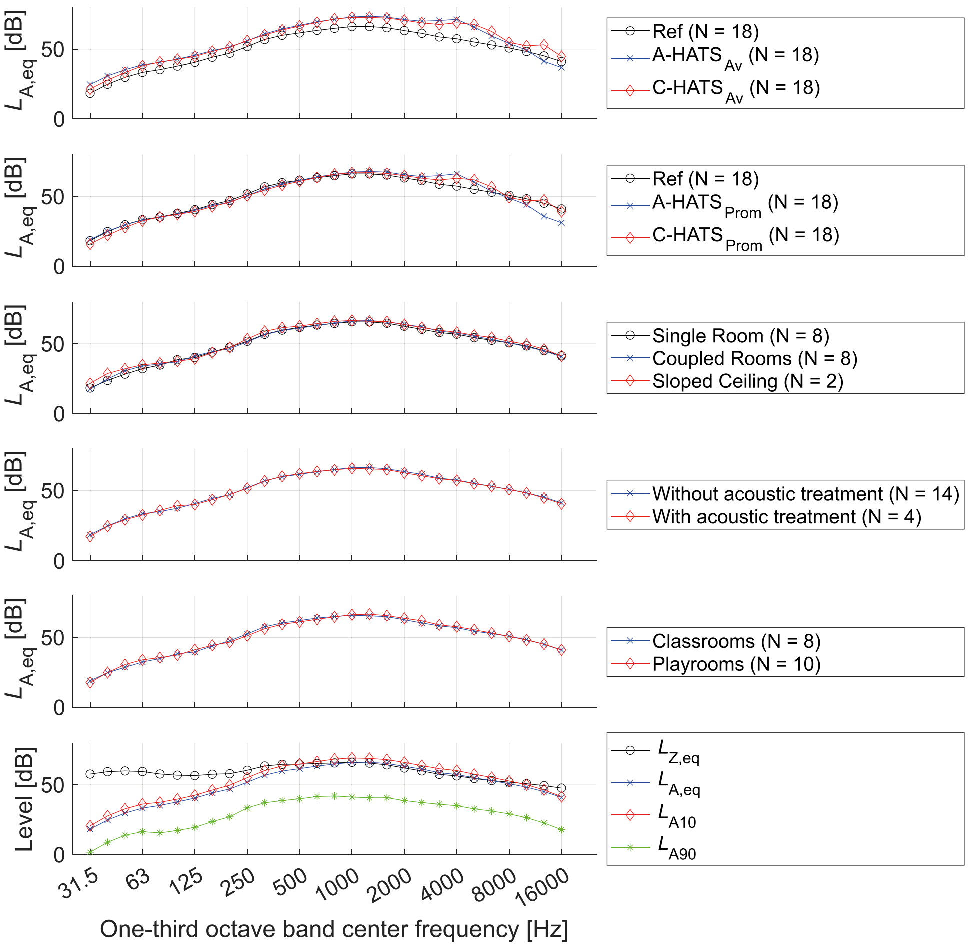


FIGURE 5 | One-third octave band sound pressure levels (A-weighted) for the measurement methods and various room properties. Here, Ref = omnidirectional microphone, A-HATS$_{Av}$/C-HATS$_{Av}$ = adult/child HATS (level summation), and A-HATS$_{Prom}$/C-HATS$_{Prom}$ = adult/child HATS (prominent-ear method).

who explain these effects with the anthropometric differences between adults and children; however, there are no noticeable differences between the different room conditions (coupled rooms, acoustic treatment, and room types). Since SPL above 8 kHz diverged between the adult and child HATSs, it can be assumed that the fine structure of the ear played a role in the evaluation and that the $L_{A,eq}$ could be sensitive to differences introduced by the anthropometric sizes of the ear. Altogether, the spectral variation in the SPL between the different transducers points toward some benefit in using a child HATS over an adult HATS and/or an omnidirectional microphone for $L_{A,eq}$ values.

For the percentile levels, the L_{A10} and L_{A90} values in the current study were similar to the ones reported in Sala and Rantala (2016) and were within the range of values of previous studies summarized in Sala and Rantala (2016). These percentile levels were, however, not significantly different between the adult and child HATSs. For the level-based fluctuation parameters, while only *NPL* showed significant differences across the measurement methods (different transducers), the usefulness of such parameters will depend on whether they are able to explain children's perception of the noise environment in CRs and PRs, which is suggested as a question for future studies.

4.3 Variation in Psychoacoustic Parameters

The use of psychoacoustic parameters is not common in CR acoustics literature. One of the aims of this study was to present results that may be useful for future studies, including ones that compare the performance and potential benefits of level-based and psychoacoustic parameters in relation to children's perception of their acoustic environments. Hence, the discussion here will be limited to a preliminary comparison between the level-based and psychoacoustic parameters.

Fluctuation strength (*FS*) and roughness characterize amplitude modulations up to 20 Hz, and between 15–300 Hz, respectively. Roughness has been shown to be related to noise annoyance due to faster sound fluctuations, and *FS* has been shown to be related to auditory distraction (Schlittmeier et al., 2012). In the current study, the main finding related to these parameters is that they both show a significant decrease in values (R_5, FS_{mean}, and FS_5) with increasing room volumes for the HATSs (**Section 3.2.2**), out of all the parameters tested (including level-based parameters). This finding

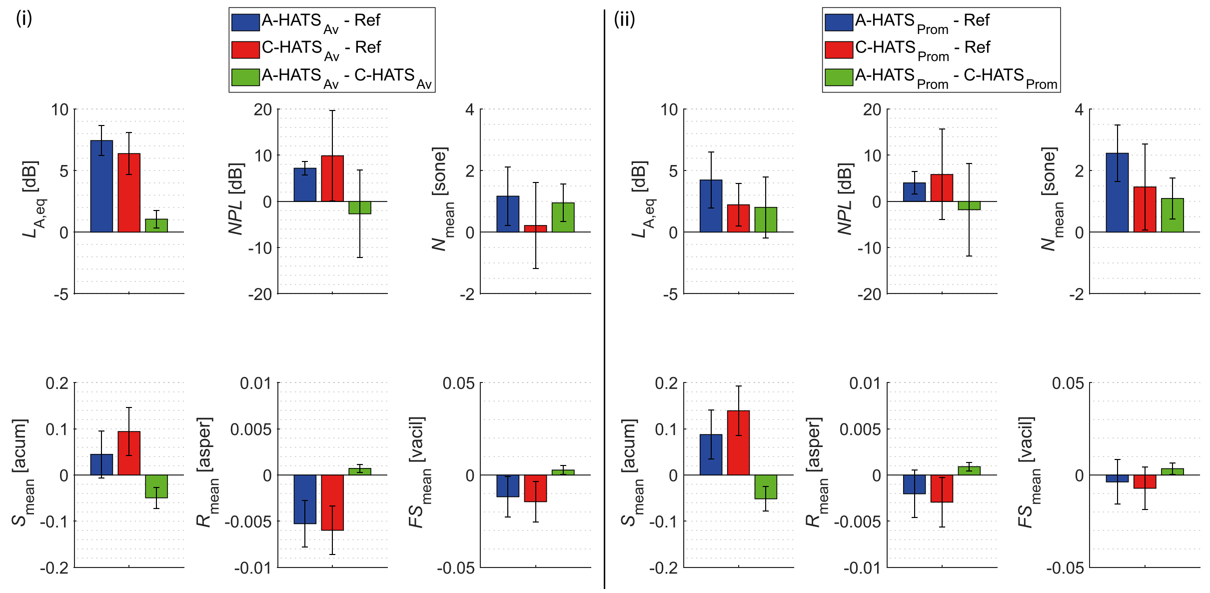


FIGURE 6 | Each subplot shows differences in parameter (*y*-axis) values between three measurement methods. Each bar represents the difference in mean values as per the legend and its 95% confidence interval. Two groups of differences are presented: (i) differences between omnidirectional microphone (Ref), adult and child HATS with the averaging method (A-HATS$_{Av}$/C-HATS$_{Av}$) and (ii) differences between Ref, and adult and child HATS with the prominent-ear method (A-HATS$_{Prom}$/C-HATS$_{Prom}$).

broadly suggests decreasing amplitude fluctuations with increasing room volumes, which is expected, presumably if the number of children is not increasing disproportionately. However, this needs future studies where the number of children in a room is considered more systematically than the current study.

Mean sharpness values had significant differences between all the transducers. In terms of HATSs, this is consistent with previous findings (Fels et al., 2004) and the expectations of the current study, with higher values of high-frequency content measured with child HATS compared to adult HATS in accordance with anthropometric differences, as was discussed in **Section 4.2**. This observation is supported by the findings regarding the low–high-frequency ratio of $L_{A,eq}$, though significant differences were only found between the omnidirectional transducer and the child HATS and between the HATSs. However, these results need validation in terms of children's perception in future studies, where it would be possible to comment on whether the use of HATSs, which introduce measurement and analysis overheads, is sufficiently justified over the more traditional use of omnidirectional transducers, which further allow relatively convenient measurement setups.

Overall, the current results show very limited benefit in considering computationally expensive loudness and sharpness over the more traditional and easier to calculate level-based parameters, although there may be some benefit in considering fluctuation strength and roughness, which showed variation with increasing room values for the HATSs.

4.4 Limitations

The measurement methods in this study have several limitations that are generally related to logistical difficulties in conducting measurements with children and/or at educational institutions. For the room acoustic measurements, some of the doors to adjoining rooms were not closed since they are typically left open during daily activities. This was discussed in detail in **Section 2.2** and **Section 2.3**, including the method to address nonlinear decays. Room acoustics measurements were not conducted during occupancy, which limits the characterization of rooms to unoccupied conditions only, and further limits comparisons with previous studies.

In terms of the *in-situ* measurements, a major limitation was the fixed locations of the transducers, which was to avoid too many disruptions to the normal activities of children. The close-by positioning of the transducers can lead to shadowing effects that were not examined in detail within this study, which is acknowledged as a limitation. Furthermore, the head directions of the HATSs were static and not changed during the measurement durations, so the effects introduced by head movements could not be analyzed within this study. These issues could be improved in a future study wherein several measurement locations including using a combination of HATS locations and measurements using microphones placed near children's ears, etc., as in McAllister et al. (2009).

Moreover, the number of children in the rooms fluctuated during the day, which could not be monitored. Hence, detailed analyses using the number of children as a factor were not possible and are recommended for future studies. Additionally, the type of noises and associated activities were not specifically analyzed in this study. In further studies, the measurement method could be chosen to allow studying the impact of noise sources (e.g., speech-based and impact sounds). Finally, since only one example of adult and child HATSs was used, results are limited to these HATS. It is likely that using a different HATSs will lead to different results to an extent, which can be considered in a separate study.

CONCLUSION

This study presents pilot results from acoustic assessments in German primary schools and daycare centers with a focus on long-term measurements using an adult and a child HATS, and an omnidirectional microphone, besides room acoustics. Main conclusions are as follows:

1. The room acoustics in both CRs and PRs in Germany has a lot of scope for improvements to meet guidelines for "good acoustics" outlined in recent studies. The current findings point toward the use of typical room acoustics treatment for high-frequency sound absorption to control the RTs while ensuring high speech intelligibility.
2. Based on omnidirectional measurements, long-term $L_{A,eq}$ values in CRs and PRs are very similar (~66 dB), which are in general lower and with a smaller range than similar measurements in other countries in Europe and the United States . Similar trends are reported for percentile levels (L_{10}, L_{90}).
3. There are some indications that psychoacoustic parameters (especially fluctuation strength and roughness) may be beneficial in complementing SPL-based parameters.
4. Overall, while the findings here suggest some benefit in considering child HATS over adult HATS and/or omnidirectional measurements in terms of characterizing *in-situ* noise measurements, especially in the higher frequencies where anthropomorphic details are important, these findings are specific to the current measurement method with fixed locations for the transducers including orientation for the HATSs. More research that considers various transducer locations and orientations is necessary to validate these findings.
5. The findings here, especially for points 3 and 4 above, are considered preliminary and need further studies that characterize children's perception in relation to the wide range of parameters studied here that included room acoustics, level-based, and psychoacoustic parameters.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee at the RWTH Aachen University, Germany. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

The concept of this study and manuscript was conceived and designed by JF, KP, MY, and KL. Data collection and processing were carried out by KL. Results analysis was conducted by KL and MY with support by MK in terms of statistical analysis. Results interpretations and literature search were driven by MY and KL in support of all contributing authors. KL took the lead in writing the manuscript and was supported by the other authors. All authors provided critical feedback and helped shaping the research, analysis, and manuscript. JF supervised the project and provided all necessary resources.

FUNDING

This work was funded through the HEAD-Genuit-Stiftung (project ID: P-16/17-W), the European Union's Horizon 2020 research and innovation program under the grant agreement no. 874724 (Equal-Life), and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number: 401278266. Furthermore, the project was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—project number: 368482240/GRK2416.

ACKNOWLEDGMENTS

Special thanks to the student assistants Simon Klein and Julian Burger for assisting acoustic measurements and data entry. Many thanks to Florian Pausch for helpful discussions and advice. Finally, the authors would like to thank all children, parents, educators, and administration from the involved educational institutions for participating in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2022.688847/full#supplementary-material

REFERENCES

Åhlander, V. L., Rydell, R., and Löfqvist, A. (2011). Speaker's Comfort in Teaching Environments: Voice Problems in Swedish Teaching Staff. *J. Voice* 25, 430–440. doi:10.1016/j.jvoice.2009.12.006

American National Standards Institute (2002). *ANSI-S12.60-2002 - Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools*. Melville, NY: Inc..

Astolfi, A., Parati, L., D'Orazio, D., and Garai, M. (2019a). "The New Italian Standard UNI 11532 on Acoustics for Schools," in Proceedings of the 23rd International Congress on Acoustics, Aachen, Germany, 7004–7011.

Astolfi, A., Puglisi, G. E., Murgia, S., Minelli, G., Pellerey, F., Prato, A., et al. (2019b). Influence of Classroom Acoustics on Noise Disturbance and Well-Being for First Graders. *Front. Psychol.* 10, 2736. doi:10.3389/fpsyg.2019.02736

Ayr, U., Cirillo, E., Fato, I., and Martellotta, F. (2003). A New Approach to Assessing the Performance of Noise Indices in Buildings. *Appl. Acoust.* 64, 129–145. doi:10.1016/S0003-682X(02)00075-0

Berzborn, M., Bomhardt, R., Klein, J. C., Richter, J.-G., and Vorländer, M. (2017). "The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing," in *Fortschritte der Akustik - DAGA 2017 (Kiel: Deutsche Gesellschaft für Akustik e.V. (DEGA))*. Available at: https://publications.rwth-aachen.de/record/687308.

Bradley, J. S., Reich, R., and Norcross, S. G. (1999). A Just Noticeable Difference in C 50 for Speech. *Appl. Acoust.* 58, 99–108.

Bradley, J. S. (2011). Review of Objective Room Acoustics Measures and Future Needs. *Appl. Acoust.* 72, 713–720. doi:10.1016/j.apacoust.2011.04.004

Building Bulletin 93 (2015). *Acoustic Design of Schools: Performance Standards*. London.

Choi, Y.-J. (2016). Effect of Occupancy on Acoustical Conditions in university Classrooms. *Appl. Acoust.* 114, 36–43. doi:10.1016/j.apacoust.2016.07.010

Crandell, C. C., and Smaldino, J. J. (2000). Classroom Acoustics for Children with Normal Hearing and with Hearing Impairment. *Lshss* 31, 362–370. doi:10.1044/0161-1461.3104.362

Deutsches Institut für Normung, e. V. (2016). *DIN 18041 - Acoustic Quality in Rooms - Specifications and Instructions for the Room Acoustic Design*.

Deutsches Institut für Normung, e. V. (2009). *DIN 45692 - Measurement Technique for the Simulation of the Auditory Sensation of Sharpness*.

Fels, J., Buthmann, P., and Vorländer, M. (2004). Head-related Transfer Functions of Children. *Acta Acustica united with Acustica* 90, 918–927.

Fels, J., and Vorländer, M. (2009). Anthropometric Parameters Influencing Head-Related Transfer Functions. *Acta Acustica united with Acustica* 95, 331–342. doi:10.3813/aaa.918156

Finnish Standards Association (2004). SFS 5907 - Acoustic Classification of Spaces of Buildings.

International Electrotechnical Commission (2012). IEC 60268-16 - Sound System Equipment - Part 16: Objective Rating of Speech Intelligibility by Speech Transmission index.

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

International Organization for Standardization (2017). ISO 532-1 - Acoustics - Methods for Calculating Loudness - Part 1: Zwicker Method.

International Organization for Standardization (1993). ISO 9568 - Cinematography - Background Acoustic Noise Levels in Theatre, Review Rooms and Dubbing Rooms.

Klatte, M., Bergström, K., and Lachmann, T. (2013). Does Noise Affect Learning? A Short Review on Noise Effects on Cognitive Performance in Children. *Front. Psychol.* 4, 578. doi:10.3389/fpsyg.2013.00578

Klatte, M., Hellbrück, J., Seidel, J., and Leistner, P. (2010a). Effects of Classroom Acoustics on Performance and Well-Being in Elementary School Children: A Field Study. *Environ. Behav.* 42, 659–692. doi:10.1177/0013916509336813

Klatte, M., Lachmann, T., and Meis, M. (2010b). Effects of Noise and Reverberation on Speech Perception and Listening Comprehension of Children and Adults in a Classroom-like Setting. *Noise Health* 12, 270. doi:10.4103/1463-1741.70506

Klatte, M., Lachmann, T., Schlittmeier, S., and Hellbrück, J. (2010c). The Irrelevant Sound Effect in Short-Term Memory: Is There Developmental Change? *Eur. J. Cogn. Psychol.* 22, 1168–1191. doi:10.1080/09541440903378250

Kryter, K. D. (2013). *The Effects of Noise on Man*. Elsevier.

Leccese, F., Rocca, M., and Salvadori, G. (2018). Fast Estimation of Speech Transmission Index Using the Reverberation Time: Comparison between Predictive Equations for Educational Rooms of Different Sizes. *Appl. Acoust.* 140, 143–149. doi:10.1016/j.apacoust.2018.05.019

Lenne, L., Chevret, P., and Marchand, J. (2020). Long-term Effects of the Use of a Sound Masking System in Open-Plan Offices: A Field Study. *Appl. Acoust.* 158, 107049. doi:10.1016/j.apacoust.2019.107049

McAllister, A. M., Granqvist, S., Sjölander, P., and Sundberg, J. (2009). Child Voice and Noise: A Pilot Study of Noise in Day Cares and the Effects on 10 Children's Voice Quality According to Perceptual Evaluation. *J. Voice* 23, 587–593. doi:10.1016/j.jvoice.2007.10.017

Møller, H., Sørensen, M. F., Hammershøi, D., and Jensen, C. B. (1995). Head-related Transfer Functions of Human Subjects. *J. Audio Eng. Soc.* 43, 300–321.

Moore, B. C. J., and Glasberg, B. R. (2007). Modeling Binaural Loudness. *The J. Acoust. Soc. America* 121, 1604–1612. doi:10.1121/1.2431331

Neuman, A. C., and Hochberg, I. (1983). Children's Perception of Speech in Reverberation. *J. Acoust. Soc. America* 73, 2145–2149. doi:10.1121/1.389538

Peng, Z., Lau, S.-K., Wang, L. M., Browne, S. D., and Roy, K. P. (2013). Binaural Room Impulse Response Database Acquired from a Variable Acoustics Classroom. *Proc. Mtgs. Acoust.* 15, 015003. doi:10.1121/1.4793567

Persson Waye, K., and Karlberg, J. (2021). Sound Quality Characteristics of Importance for Preschool Children's Perception and Wellbeing after an Acoustic Intervention. *Front. Built Environ.* 7, 91. doi:10.3389/fbuil.2021.688836

Persson Waye, K., van Kamp, I., and Dellve, L. (2013). Validation of a Questionnaire Measuring Preschool Children's Reactions to and Coping with Noise in a Repeated Measurement Design. *BMJ open* 3, e002408. doi:10.1136/bmjopen-2012-002408

Picard, M., and Bradley, J. S. (2001). Revisiting Speech Interference in Classrooms: Revisando la interferencia en el habla dentro del salón de clases. *Int. J. Audiol.* 40, 221–244.

Prodi, N., Farnetani, A., Yiluya, S., and Fels, J. (2007). "Investigating Classroom Acoustics by Means of Advanced Reproduction Techniques," in *Audio Engineering Society Convention*, 122.

Sala, E., and Rantala, L. (2016). Acoustics and Activity Noise in School Classrooms in Finland. *Appl. Acoust.* 114, 252–259. doi:10.1016/j.apacoust.2016.08.009

Schlittmeier, S. J., Weißgerber, T., Kerber, S., Fastl, H., and Hellbrück, J. (2012). Algorithmic Modeling of the Irrelevant Sound Effect (ISE) by the Hearing Sensation Fluctuation Strength. *Atten Percept Psychophys* 74, 194–203. doi:10.3758/s13414-011-0230-7

Schmitz, A. (1995). A New Digital Artificial Head Measuring System. *Acta Acustica united with Acustica* 81, 416–420.

Shield, B. M., and Dockrell, J. E. (2008). The Effects of Environmental and Classroom Noise on the Academic Attainments of Primary School Children. *J. Acoust. Soc. America* 123, 133–144. doi:10.1121/1.2812596

Shield, B. M., and Dockrell, J. E. (2003). The Effects of Noise on Children at School: A Review. *Building Acoust.* 10, 97–116. doi:10.1260/135101003768965960

Shinn-Cunningham, B. G., Kopco, N., and Martin, T. J. (2005). Localizing Nearby Sound Sources in a Classroom: Binaural Room Impulse Responses. *J. Acoust. Soc. America* 117, 3100–3115. doi:10.1121/1.1872572

Södersten, M., Granqvist, S., Hammarberg, B., and Szabo, A. (2002). Vocal Behavior and Vocal Loading Factors for Preschool Teachers at Work Studied with Binaural DAT Recordings. *J. Voice* 16, 356–371. doi:10.1016/S0892-1997(02)00107-8

Sottek, R. (1993). *Models for Signal Processing in the Human Auditory System*. [dissertation]. Aachen, Germany: RWTH Aachen University.

van Wijngaarden, S. J., and Drullman, R. (2008). Binaural Intelligibility Prediction Based on the Speech Transmission index. *J. Acoust. Soc. America* 123, 4514–4523. doi:10.1121/1.2905245

Wang, L. M., and Brill, L. C. (2021). Speech and Noise Levels Measured in Occupied K-12 Classrooms. *J. Acoust. Soc. America* 150, 864–877. doi:10.1121/10.0005815

Yadav, M., Cabrera, D., Kim, J., Fels, J., and de Dear, R. (2021). Sound in Occupied Open-Plan Offices: Objective Metrics with a Review of Historical Perspectives. *Appl. Acoust.* 177, 107943. doi:10.1016/j.apacoust.2021.107943

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership