# SOUNDSCAPE ASSESSMENT

EDITED BY: Östen Axelsson, Catherine Guastavino and Sarah R. Payne
PUBLISHED IN: Frontiers in Psychology

**frontiers** Research Topics

## 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: researchtopics@frontiersin.org

# SOUNDSCAPE ASSESSMENT

Topic Editors:
**Östen Axelsson,** Stockholm University, Sweden
**Catherine Guastavino,** McGill University, Canada
**Sarah R. Payne,** Heriot-Watt University, United Kingdom

# Table of Contents

# Editorial: Soundscape Assessment

Östen Axelsson[1]*, Catherine Guastavino[2] and Sarah R. Payne[3]

[1] Gösta Ekman Laboratory, Department of Psychology, Stockholm University, Stockholm, Sweden, [2] School of Information Studies and Centre for Interdisciplinary Research in Music Media and Technology, McGill University, Montreal, QC, Canada, [3] The Urban Institute, Heriot-Watt University, Edinburgh, United Kingdom

**Editorial on the Research Topic**

**Soundscape Assessment**

## SOUNDSCAPE RESEARCH, 50 YEARS

In 2019, soundscape research celebrates 50 years as a scientific field. In 1969, the first scientific article using the term "soundscape" was published in the inaugurate issue of the premier scientific journal for environmental psychology, Environment and Behavior (Southworth, 1969). The author was Michael Southworth, a PhD student in city planning at MIT in Boston, today Professor Emeritus of Urban Design at UC-Berkley. The article was based on his Master's Thesis in city planning, which he completed at MIT in 1967.

In the 1970s and 1980s, soundscape was largely associated with the Canadian composer R. Murray Schafer at Simon Fraser University in Vancouver. In 1972, Schafer begun the World Soundscape Project by a detailed study of the soundscape of Vancouver.

The topic gained international momentum after being introduced to the wider community of noise and health researchers at the International Congresses on Acoustics in Seattle 1998. In 2008, the International Organization for Standardization (ISO) formed a working group on the topic to develop the ISO 12913 series, which is the first international standard in this field.

Part 1 of ISO 12913 was published in August 2014. It defined the term of "soundscape" as "acoustic environment as perceived or experienced and/or understood by a person or people, in context" (ISO, 2014). It also provides a conceptual framework, distinguishing the acoustic environment, as a physical phenomenon, from the soundscape, as a perceptual construct. Part 2 identifies data collection and reporting requirements (ISO, 2018), while Part 3 will identify data analysis aspects. The present Research Topic was initiated to support the development of the ISO 12913 series by investigating methods for soundscape assessment.

## DIVERSITY OF APPROACHES

A wide range of methods and subjects are covered in this Research Topic, indicating that soundscape assessment should be approached from a holistic, multisensory perspective to capture outcomes that extend well-beyond auditory judgments. Contributions encompass theoretical and practical approaches, highlighting their complimentary, and informative role in furthering soundscape assessments.

Two contributions investigated the relationship between public space usage and soundscape. Bild et al. used behavioral mapping and questionnaires to investigate how social interaction influences soundscape assessment. Meng et al. investigated how music in a public space influences crowd density and walking patterns. In these instances, behavioral observations identified variations in soundscape assessments by groups of people.

Two contributions took a cognitive approach by using free sorting tasks conducted by individual participants. Bones et al. developed a sound taxonomy to investigate how people categorize environmental sounds. Aletta et al. identified holistic acoustic properties through sorting of spectrograms, instead of soundscape recordings.

Several contributions focus on refining or improving existing soundscape assessment methods, particularly semantic scales that are often used in questionnaire studies. Welch et al. developed semantic scales based on creative writing as a way of exploring a wider range of soundscape descriptors. Payne and Guastavino investigated the validity of the Perceived Restorativeness Soundscape Scale (PRSS) using psycholinguistic analysis. van den Bosch et al. developed a theoretic framework to provide the underpinnings of existing results in soundscape research and a rational for current assessment models.

Contributions also highlight the importance of person-related and contextual factors in soundscape assessment, which have previously received limited attention. Sun et al. investigated how individual differences in the ability to process audio-visual information (named *audio-visual aptitude*) influences the interaction between landscape and soundscape appraisal. Benfield et al. studied how attitudes to motorized recreation in national parks and to its regulation may moderate the effect of the sound of motorized recreation on scenic evaluation. Memoli et al. discovered that deviation from the expected flight path influences noise annoyance from incoming, landing aircrafts.

## COLLABORATION BUILT ON DIVERSITY

The diversity of methods among the contributions reflects the two-way interaction between theory and practice, revisiting the traditional dichotomy between deductive and inductive approaches. While theories may guide concrete soundscape interventions, the complexity of real-world applications inform and enrich theories and models. The methods include field and laboratory studies, as well as qualitative and quantitative methods, suggesting that no single method may capture all the different facets of a soundscape.

With this diversity of soundscape assessment methods, integration of the various approaches, and comparability across results is increasingly difficult. Soundscape researchers

consider different approaches depending on the object of study (e.g., individual sensory experience, group behavior, acoustic properties, or invested meaning). The diversity of research methods used suggests that soundscape, both as an object of study and as a field of research, is still under development. Consequently, standardization efforts should focus on identifying or developing a reference method for enhanced comparability among studies, as opposed to a single soundscape method. This observation reflects the outcome of ISO/TS 12913-2 that recommends multiple assessment methods (ISO, 2018), as a single method could not be agreed upon.

The increasing interest in contextual and person-related factors is worth noting, and reflects the ISO definition (ISO, 2014). Contextual and person-related factors extend beyond auditory judgment, providing a more holistic representation of the soundscape. A focus on context also enables the inclusion of applied research, around practical soundscape interventions, alongside more fundamental research that advances theory development.

Another recurring challenge in soundscape research is the main sources of variance: individual variation among participants providing the soundscape assessments, and variation among soundscapes. Essentially, soundscapes result from a variety of sound sources, in varied contexts. Frequently, researchers investigate only one kind of place, such as parks or plazas. This provides an in-depth understanding of issues related to these particular contexts, but limits the generalizability to other places, where sound sources and contextual factors will vary. Recognizing the range of contexts and investigating the transferability of research findings from one context to another provides further directions for soundscape assessment research.

While it is important to relate critically to existing methods and results, and to examine their validity, it is also important to seek a common ground and common objectives for a research field to develop further. Consequently, it is important to encourage continued and deepened international and interdisciplinary collaboration.

## AUTHOR CONTRIBUTIONS

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

## REFERENCES

ISO (2014). *ISO 12913-1:2014 Acoustics — Soundscape — Part 1: Definition and Conceptual Framework*. Geneva: International Organization for Standardization.

ISO (2018). *ISO/TS 12913-2:2018 Acoustics — Soundscape — Part 2: Data Collection and Reporting Requirements*. Geneva: International Organization for Standardization.

Southworth, M. (1969). The sonic environment of cities. *Environ. Beha*v. 1, 49–70.

Check for
updates

# Dimensions Underlying the Perceived Similarity of Acoustic Environments

Francesco Aletta[1,2]*, Östen Axelsson[3] and Jian Kang[1]*

[1] Acoustics Group, School of Architecture, University of Sheffield, Sheffield, United Kingdom, [2] WAVES Research Group, Department of Information Technology, Ghent University, Ghent, Belgium, [3] Gösta Ekman Laboratory, Department of Psychology, Stockholm University, Stockholm, Sweden

Scientific research on how people perceive or experience and/or understand the acoustic environment as a whole (i.e., soundscape) is still in development. In order to predict how people would perceive an acoustic environment, it is central to identify its underlying acoustic properties. This was the purpose of the present study. Three successive experiments were conducted. With the aid of 30 university students, the first experiment mapped the underlying dimensions of perceived similarity among 50 acoustic environments, using a visual sorting task of their spectrograms. Three dimensions were identified: (1) Distinguishable–Indistinguishable sound sources, (2) Background–Foreground sounds, and (3) Intrusive–Smooth sound sources. The second experiment was aimed to validate the results from Experiment 1 by a listening experiment. However, a majority of the 10 expert listeners involved in Experiment 2 used a qualitatively different approach than the 30 university students in Experiment 1. A third experiment was conducted in which 10 more expert listeners performed the same task as per Experiment 2, with spliced audio signals. Nevertheless, Experiment 3 provided a statistically significantly worse result than Experiment 2. These results suggest that information about the meaning of the recorded sounds could be retrieved in the spectrograms, and that the meaning of the sounds may be captured with the aid of holistic features of the acoustic environment, but such features are still unexplored and further in-depth research is needed in this field.

Keywords: soundscape, perceived similarity, acoustic environment, PCA, listening experiment

## INTRODUCTION

One of the first definitions of 'soundscape' was given in the Handbook for Acoustic Ecology (first published in 1978) – "An environment of sound (or sonic environment) with emphasis on the way it is perceived and understood by the individual, or by a society" (Truax, 1978). The concept has attracted interest from various scientific and social disciplines: acoustics, psychology, sociology, urban planning, ecology, and more. Due to its strong interdisciplinary appeal it is a field of wide experimentation. The literature in the field is growing, proposing both theoretical models and practical approaches (Schulte-Fortkamp and Dubois, 2006; Cain et al., 2009, 2013; Axelsson et al., 2010; Davies, 2013; Schulte-Fortkamp and Kang, 2013). In 2008 the International Organization for Standardization (ISO) created a new working group with the mission to develop the first International Standard on soundscape, ISO 12913. Part 1 of the standard defines 'soundscape'

as an "acoustic environment as perceived or experienced and/or understood by a person or people, in context" (ISO, 2014). Thus, there is a general agreement that soundscape concerns human perception of the acoustic environment. This is comparable to the European Landscape Convention that defines 'landscape' in similar terms (Council of Europe, 2000). Currently the ISO working group is preparing Part 2 on data collection and reporting requirements in soundscape studies, which include developing soundscape indicators (i.e., acoustic terms used to predict human responses to the acoustic environment).

In order to help European policymakers and authorities to understand and fulfill their responsibilities with regards to the protection of so called 'quiet areas,' the European Environment Agency (EEA) published a good practice guide in 2014 (EEA, 2014). It recommends four complementary methods for identifying quiet areas. The soundscape approach is one of them. EEA also calls for further in-depth research in this field. For example, EEA identifies a need to develop "indicators and measurements of human appreciation of quiet areas and perceived acoustic quality." Thus, EEA provides its support to soundscape research and underlines the need of soundscape indicators.

There have been a few attempts to develop soundscape indicators by identifying relationships between soundscape and established acoustic parameters, such as A-weighted equivalent continuous sound pressure level, and psychoacoustic parameters, such as: Loudness, Roughness, Sharpness, and related percent exceedance levels (Brambilla et al., 2013; Rychtáriková and Vermeir, 2013). The latter are thought to better describes particular auditory sensations which might not be expressed by simple energetic metrics (Genuit and Fiebig, 2006). Detailed information about these three psychoacoustic parameters (including definitions and applications) are found in Fastl and Zwicker (2007). Nevertheless, this approach is not necessarily successful, because many established psychoacoustic parameters are primarily developed for the purpose of single sounds or sound sources and used within a "product sound quality" framework for industrial applications (e.g., automotive sector, domestic appliances industry, etc.). They were not developed for the purpose of soundscape, nor for measuring the acoustic environment holistically. Alternatively, some researchers (Herranz Pascual et al., 2010) have tried to incorporate the human experience of a place in a soundscape index. Yet others believe that "human responses should not be equated to acoustic measures" (Andringa et al., 2013). In fact, the soundscape methodology is far more holistic than mere noise control engineering, shifting from a quantitative to a qualitative approach to the assessment and management of the (urban) acoustic environments. Several studies have pointed out the need for more standardization with regards to these issues (Brown et al., 2011; Aletta et al., 2014). Kang et al. (2016) proposed an overview of the state-of-art in soundscape research, and the challenges this approach is facing.

There is still no consensus about what acoustic properties might be meaningful for describing the perceived properties of the acoustic environments and how the former relate to the latter. Hence, the purpose of the present study was to explore

the acoustic properties of acoustic environments holistically. The main research questions were: (1) whether dimensions describing perceived similarity between acoustic environments, in terms of their acoustic properties, could be identified; and (2) whether those dimensions could be satisfactorily explained by established acoustic metrics. Three successive experiments were conducted. The first experiment mapped the underlying dimensions of perceived similarity among 50 acoustic environments based on their acoustic properties. The second experiment was carried out in order to validate the results from Experiment 1 by a listening experiment. The third experiment replicated Experiment 2 with spliced signals to investigate whether the meaning of the sounds was an important factor. **Figure 1** summarizes the overall methodology of this paper, the details of which will be further discussed in the corresponding sections.

# EXPERIMENT 1: SORTING OF SPECTROGRAMS
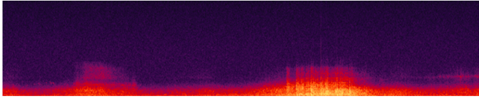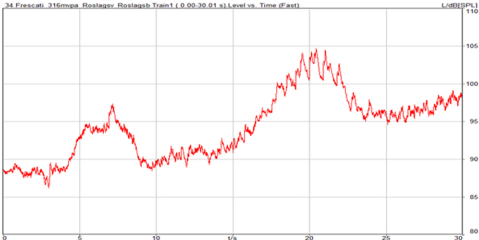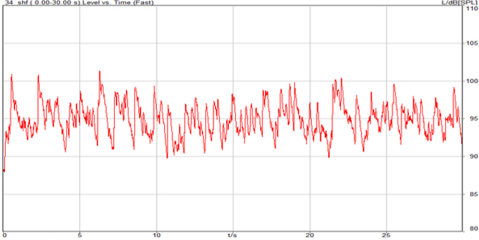
## Method
### Participants
Thirty undergraduates and post-graduates at the University of Sheffield, 18 to 33 years old, participated in the experiment (15 women, 15 men; $M_{age} = 24.2$ years, $SD = 4.8$). The ethnic distribution of the sample was 20 'White or Caucasian' and 10 'Asian or Pacific Islander.' Participants were selected from a group of 100 persons who completed an online survey circulated via the established email list for student volunteers at University of Sheffield. The questions in the online survey were designed to achieve a diverse group of participants in terms of gender, age and ethnic origin. All participants had normal color vision as tested by the "Ishihara test for color deficiency" (Ishihara, 1957). Because the goal was to test only whether or not the participant had a normal color appreciation, a reduced version of the test was used. It included 6 plates, selected according to Ishihara's instructions (Ishihara, 1957). The 30 participants who completed the experiment were rewarded for volunteering with a GBP 10 gift card.

### Stimulus Material
Fifty recordings (30s) from Axelsson et al. (2010) were used for this experiment. They were selected from a library of binaural recordings of outdoor acoustic environments (London and Stockholm) with the aim to achieve a large variation in overall sound-pressure levels and urban/peri-urban locations. **Table 1** presents the A-weighted equivalent continuous sound pressure levels ($L_{Aeq,30s}$) and the main sound sources of the 50 experimental sounds. In order to create visual representations of the acoustic data, the fifty audio files (.wav) were imported in Adobe Audition 3.0. For each binaural recording, the spectrogram (time vs. frequency) was plotted for the right channel. The spectrograms were set to have the time on the $X$-axis (0–30 s, 1 s steps) and the frequency on the $Y$-axis, with a linear scale (0–25 kHz, 1 kHz steps). Regarding the spectral controls for the color scale of the sound-pressure-level dimension, the software default settings were used (132 dB range, 512 frequency

**FIGURE 1 |** Overall experimental methodology of the study, with the different sorting tasks performed by the different groups of participants.

bands resolution, gamma index 2) and the three sampling colors were: yellow (RGB 254, 250, 84 – width 67%), orange (RGB 249, 47, 0 – width 76%) and purple (RGB 45, 7, 69 – width 80%). The 50 spectrograms were printed in color on glossy photo paper (18.5 × 4.5 cm, 150 dpi resolution). **Figure 2** presents three examples (Panels A–C) of the 50 spectrograms used in the experiment.

### Design and Procedure

The experiment took place in an office room at the School of Architecture, University of Sheffield. The design of the experiment consisted of a two-stage data collection procedure: sorting and interview. The participants took part individually. First, the color vision test was performed for each participant. Successful participants were admitted to the following stage. One participant was omitted due to partial color-blindness.

Seated at an office desk, every participant was provided with the 50 color prints of the spectrograms as a stack of photographs mixed in a unique irregular order for each participant. Importantly, they were not informed about what the photographs depicted or what spectrograms represent (i.e., acoustic properties of the recorded acoustic environments). Thus, the participants were expected to treat the photographs as any

abstract images, and were instructed to sort the prints into mutually exclusive groups according to the similarity of the images, and in as many groups as they wanted (2 being the minimum and 25 the maximum). In addition, they were asked to pay attention to whether or not they developed any specific sorting criteria. This information was required in the subsequent interview. Participants were allowed to revise their sorting throughout the experimental session, including the interview.

After completing the sorting task, the participants were interviewed, with the purpose to learn whether or not they had developed any soring criteria, and then which they were. This information was used to interpret the sorting results. During the interview the experimenter took notes (cf. Axelsson, 2007). The 30 experimental sessions lasted between 8 and 45 min each ($M_{time} = 19.5$ min, $SD = 8.9$). There were no time restrictions.

### Results

The participants created between 3 and 17 groups of spectrograms ($M = 8.0$ groups, $SD = 3.7$). The sorting data was used to create a proximity matrix based on how often all possible pairs of the 50 spectrograms appeared in the same group, summed over all 30 participants (cf. Axelsson, 2007).

**TABLE 1** | Description of the 50 experimental sounds with regards to A-weighted equivalent continuous sound pressure levels (dB) and the main sound sources.

| Sound | $L_{Aeq,30s}$ | Main foreground sound sources | Main background sound sources |
|---|---|---|---|
| 1 | 69.03 | Road traffic | Airplane, birdsong |
| 2 | 54.47 | | Birdsong, children, train passing by |
| 3 | 47.64 | | Voices, birdsong, road traffic |
| 4 | 45.18 | Birdsong | Road traffic |
| 5 | 52.61 | Fan | Voices |
| 6 | 67.15 | Motorcycle passing by | Birdsong, wind, footsteps |
| 7 | 58.04 | Fan | Road traffic, birdsong, dripping water |
| 8 | 76.33 | Road traffic, airplane | Car alarm, car horn |
| 9 | 69.80 | Voices, footsteps | Car horns |
| 10 | 63.05 | Pouring water | Road traffic, airplane |
| 11 | 81.20 | Road traffic | |
| 12 | 50.93 | Birdsong | Road traffic, hammering |
| 13 | 60.13 | Airplane | Birdsong, construction works |
| 14 | 65.36 | Road traffic | Hammering, birdsong |
| 15 | 52.12 | Birdsong, footsteps | Voices, children playing, road traffic |
| 16 | 71.74 | Voices | Road traffic |
| 17 | 51.99 | Footsteps, seagulls, wind, rustling leaves | Car passing by |
| 18 | 77.09 | Road traffic | |
| 19 | 77.37 | Pneumatic drill | |
| 20 | 80.31 | Airplane | |
| 21 | 76.62 | Children playing | |
| 22 | 68.31 | Waterfall | Birdsong |
| 23 | 69.99 | Airplane | Birdsong |
| 24 | 74.08 | Children playing | Road traffic, angle grinder |
| 25 | 60.18 | Train passing by | Road traffic, birdsong, footsteps |
| 26 | 72.74 | Fountain | Voices, road traffic |
| 27 | 51.62 | Wind, rustling leaves | |
| 28 | 72.68 | Airplane, road traffic, street sweeper | |
| 29 | 71.03 | Children playing | Road traffic |
| 30 | 74.44 | Angle grinder, road traffic | |
| 31 | 63.14 | Children playing | Road traffic, birdsong, bell |
| 32 | 73.69 | Road traffic | Car alarm, car horn |
| 33 | 67.09 | Road traffic, airplane | Birdsong |
| 34 | 78.91 | Road traffic, train passing by | Birdsong |
| 35 | 60.93 | Rain | Road traffic, voices |
| 36 | 44.16 | | Fan |
| 37 | 72.71 | Fountain, ambulance | Reversing lorry |
| 38 | 72.88 | Children playing | |
| 39 | 57.74 | Birdsong, voices | Road traffic |
| 40 | 45.73 | Birdsong | Ambulance, Airplane |
| 41 | 76.23 | Train passing by | Road traffic, birdsong |
| 42 | 67.97 | Footsteps, road traffic | |
| 43 | 67.40 | Birdsong | Road traffic |
| 44 | 74.27 | Road traffic, Boing 747 landing | |
| 45 | 63.15 | Fountain | Road traffic, voices |
| 46 | 61.91 | Birdsong | Road traffic, recordist hushing |
| 47 | 56.53 | Footsteps | Road traffic, wind, birdsong |
| 48 | 54.00 | Dog playing in water | Road traffic |
| 49 | 63.65 | Fountain, airplane | Voices, birdsong |
| 50 | 70.29 | Chainsaw | Voices, road traffic |
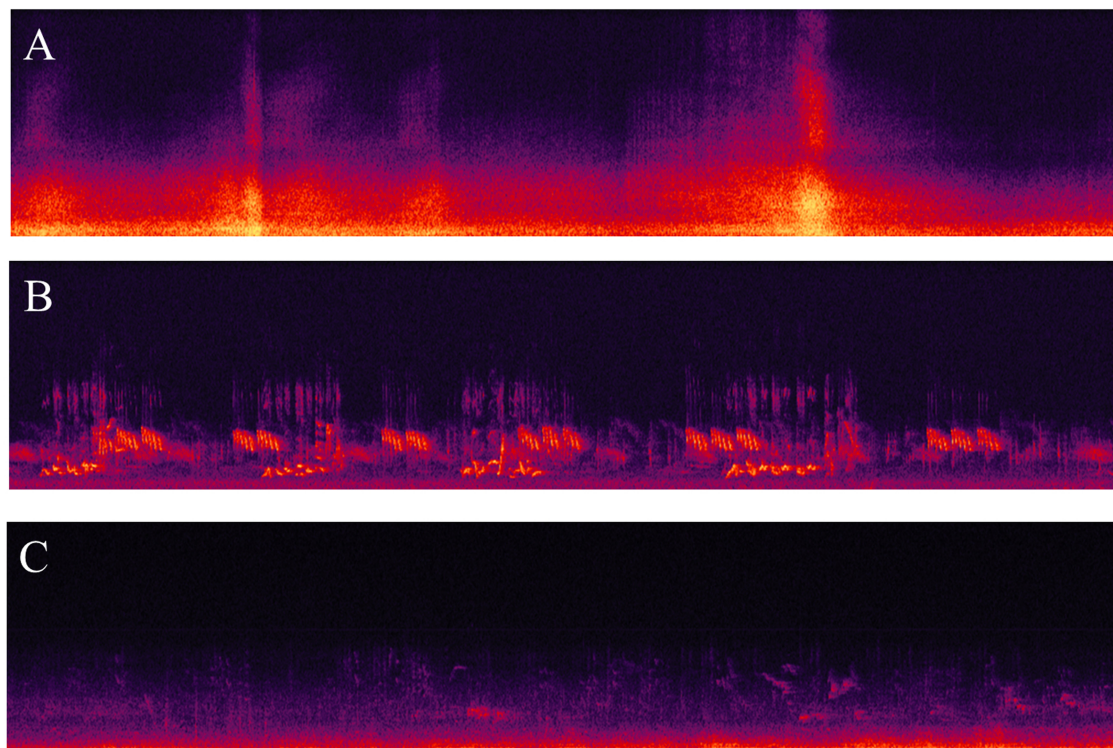
**FIGURE 2 |** Three examples **(A–C)** of spectrograms used in Experiment 1.

The proximity matrix was subjected to MDS (SPSS 21 for Windows). By using the ALSCAL technique (Young and Lewyckyj, 1979), six solutions, with one to six dimensions (stress values: 0.488, 0.257, 0.156, 0.109, 0.088, 0.071), were extracted

(Coxon, 1982). Based on a 'scree' criterion (Cattell, 1966) the three-dimensional solution was selected for further analysis.

**Figure 3** presents the three-dimensional MDS solution. Data points represent the 50 spectrograms, numbered in agreement
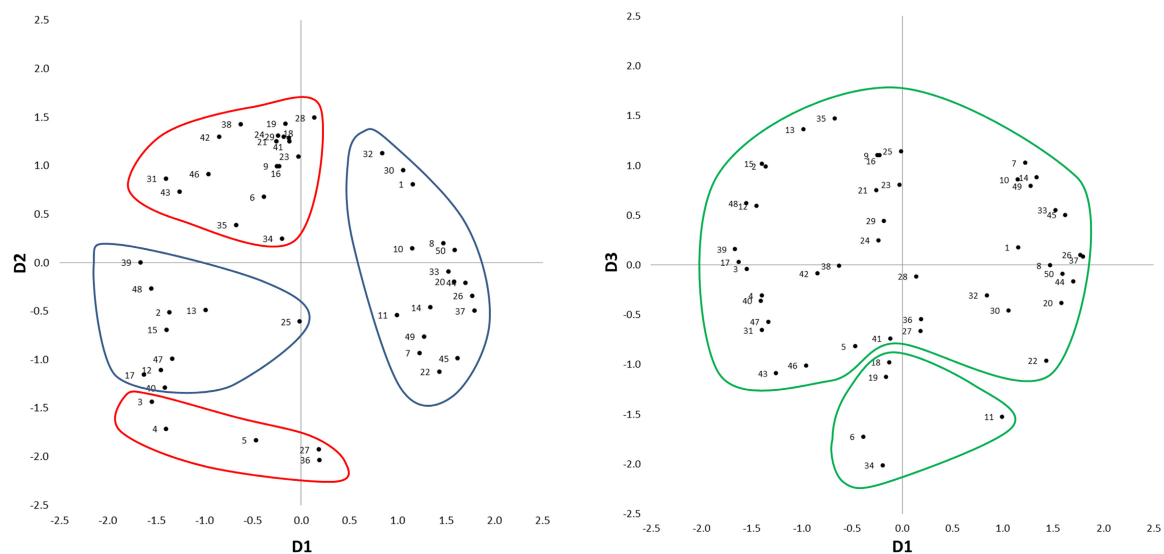


**FIGURE 3 |** Three-dimensional MDS solution for Experiment 1. On the left plot: the blue clusters D1 gather *distinguishable* vs. *indistinguishable* sound sources, while the red clusters on D2 gather *background* vs. *foreground* sounds. On the right plot: the green clusters on D3 gather *intrusive* vs. *smooth* sound sources.

with **Table 1**. In order to aid the interpretation of the three dimensions, the first author created clusters of spectrograms through visual inspection of the spectrograms and by listening to the corresponding audio recordings. In the listening sessions he sought a holistic listening style, aiming to disregard the semantic content, because it was assumed that the information about the 'meaning' of the sources was not available to the participants in sorting the spectrograms.

The first cluster contained spectrograms with positive values in the first dimension (D1). In the interviews they were often described as *"dominated by horizontal stripes," "representing all range of colors"* or *"with colors blurring into each other."* Auditory inspection revealed sounds similar to white noise. Typical dominant sound sources were fountains (e.g., Sounds 26 and 37), road traffic (e.g., Sounds 8 and 33), and aircraft (e.g., Sound 20). Combinations of several noisy sources, often affecting wide frequency ranges, typically provided an acoustic environment where different auditory features were *indistinguishable*.

The second cluster contained spectrograms with negative values in D1. In the interviews they were often described as having *"spikes," "mostly vertical shapes,"* and *"noticeable patterns."* Auditory inspection revealed clearly identifiable sound sources against a generally quiet background: footsteps (e.g., Sounds 17 and 47), birdsong (e.g., Sounds 12 and 15), and a dog playing in the water (Sound 48). Thus, the second cluster represented acoustic environments where the sound sources were *distinguishable*. Consequently, D1 was interpreted as to represent Distinguishable–Indistinguishable sound sources.

The third cluster had positive values in the second dimension (D2) and contained spectrograms that were referred to as *"yellow"* or *"deep red."* Contrariwise, the fourth cluster contained spectrograms with negative values in D2, referred to as *"purplish"* or *"dark."* This suggested that D2 was related to sound-pressure level. Auditory inspection of the corresponding audio files revealed that D2 was associated with distance of the sound sources from the listener. The third cluster represented *foreground* sounds, where sound sources were close (e.g., Sounds 38 and 42); whilst the fourth cluster represented *background* sounds, where sound sources were distant (e.g., Sounds 5 and 27). As a result, D2 was interpreted as to represent Background–Foreground sounds.

For the third dimension (D3), two separate clusters were created. The first of these two clusters contained spectrograms with negative values in D3. These spectrograms were described as *"eventful"* with *"things going on"* and *"aggressive."* The second of the two clusters contained spectrograms with mainly positive values in D3. They were considered as *"even," "smooth,"* and *"generally flat."* In the first case, sounds were characterized by an *intrusive* source, temporarily dominating the acoustic environment (e.g., Sounds 6 and 19). In the second case, sounds were *smooth* and organic, regardless of the temporal or spectral features (e.g., Sounds 7 and 35). The perception was that, regardless of the semantic content of the excerpts and their spectral content, no sound sources were being added to the sound field and this was evolving in time in an even way; D3 was therefore interpreted as to represent Intrusive–Smooth sound sources.

With the intention to provide further material for the interpretation of the three dimensions, the acoustic signals that correspond to the 50 spectrograms were subjected to acoustic analyses. For each acoustic signal (30s) a set of 100 acoustic and psychoacoustic parameters were calculated. This included unweighted, A-weighted and C-weighted equivalent continuous sound pressure levels ($L_{eq}$, $L_{Aeq}$, $L_{Ceq}$), Loudness ($N$), Sharpness ($S$), Roughness ($R$), Fluctuation strength ($Fls$), Tonality ($Ton$), percent exceedance levels for the above mentioned parameters ($P_1$, $P_5$, $P_{10}$, $P_{25}$, $P_{50}$, $P_{75}$, $P_{90}$, $P_{95}$, $P_{99}$), a measurement of the spectral variability ($L_{Ceq}-L_{Aeq}$), and the measurements of the temporal variability ($P_1-P_{99}$, $P_5-P_{95}$, $P_{10}-P_{90}$, $P_{25}-P_{75}$). The rationale for doing this is that there are several studies (Botteldooren et al., 2006; De Coensel and Botteldooren, 2006) in soundscape research suggesting that the way humans construct their auditory perceptual dimensions can be related to three main 'physical features' of the auditory stimuli: the intensity, the spectral content and the temporal structure of sounds. Hence, it seemed reasonable to test a large set of psychoacoustic metrics (which are expected to account for intensity and spectral content) and an equally large combination of differences of their percent exceedance levels (which are expected to account for different degrees of temporal variability).

Data screening revealed curvilinear relationships between the three dimensions and some of the acoustic and psychoacoustic parameters. For this reason the base-10 logarithms were calculated for all of the 100 parameters, except for six of them that included negative values.

Three stepwise multiple linear regression analyses were conducted, using D1, D2, and D3 as dependant variables and the complete set of 194 parameters as independent variables (SPSS 21 for Windows). The strongest predictors for the models of D1 ($F_{4,45} = 42.79$, $p < 0.001$, $R^2 = 0.79$), D2 ($F_{5,44} = 37.07$, $p < 0.001$, $R^2 = 0.81$) and D3 ($F_{3,46} = 9.81$, $p < 0.001$, $R^2 = 0.39$) are reported in **Table 2**.

$L_{A50}$ explained 38.9% of the variance in D1. When controlling for this variable, log measurements of variability in Sharpness [$Log(S_1-S_{99})$] explained an additional 34.9% of the variance.

**TABLE 2 |** The three stepwise linear regression models computed for D1, D2, and D3, with the best predictors, and the corresponding unstandardized coefficients (β), *t* and *p*-values.

| Model | Predictors | β | t | Sig. |
|---|---|---|---|---|
| D1 | $L_{A50}$ | 0.702 | 9.89 | $p < 0.001$ |
| | $Log(S_1-S_{99})$ | −0.937 | −3.85 | $p < 0.001$ |
| | $Log(Fls_{25}-Fls_{75})$ | −0.236 | −2.45 | $p = 0.018$ |
| | $Log(S_{10}-S_{90})$ | 0.531 | 2.28 | $p = 0.028$ |
| D2 | $Log(N_1-N_{99})$ | 0.544 | 6.07 | $p < 0.001$ |
| | $Log(Fls_{95})$ | 0.384 | 4.03 | $p < 0.001$ |
| | $Log(S_1)$ | 0.361 | 3.67 | $p = 0.001$ |
| | $Fls_{10}$ | −0.328 | −3.33 | $p = 0.002$ |
| | $Fls_{99}$ | −0.174 | −2.33 | $p = 0.024$ |
| D3 | $L_{A10}-L_{A90}$ | −1.241 | −4.57 | $p < 0.001$ |
| | $Log(L_{A25}-L_{A75})$ | 0.810 | 3.00 | $p = 0.004$ |
| | $Fls_{99}$ | −0.246 | −2.10 | $p = 0.041$ |

The positive relationship between D1 and $L_{A50}$ shows that there was more acoustic energy associated with the sounds interpreted as indistinguishable, compared to the sounds interpreted as distinguishable. This indicates that, in the former case, several sound sources were present, possibly masking each other. It seems reasonable that several sound sources are louder than one. The negative relationship between D1 and $\text{Log}(S_1 - S_{99})$ shows that as the variability in Sharpness increased, sounds were interpreted as all more distinguishable.

D2 was strongly and positively associated with variability in loudness levels $\text{Log}(N_1 - N_{99})$, which alone explained 66.6% of the variance in D2. This positive relationship indicates that there is a larger variability in Loudness in sounds interpreted as to represent the foreground than in sounds interpreted as to represent the background. This seems plausible, because background sounds at a distance would not vary much in loudness.

D3 was chiefly associated with variability in A-weighted sound-pressure levels: $L_{A10} - L_{A90}$ and $\text{Log}(L_{A25} - L_{A75})$, which explained 21.5 and 11.7% of the variance in D3, respectively. However, the two parameters work in opposite directions, where the former had a negative relationship and the latter a positive relationship with D3. This information is not particularly helpful in moving forward with the interpretation of D3. Thus, the regression analyses resulted in meaningful information for dimensions D1 and D2.

## Discussion

The purpose of Experiment 1 was to map the underlying dimensions of the acoustic properties of acoustic environments considered holistically. Measures of perceived similarity of 50 spectrograms were subjected to MDS analysis. Three dimensions were identified: (D1) Distinguishable–Indistinguishable sounds sources, (D2) Foreground–Background sounds, and (D3) Intrusive–Smooth sound sources. Stepwise multiple linear regression analyses with D1, D2 and D3 as dependent variables and 194 acoustic and psychoacoustic parameters as predictors showed that D1 was positively associated with $L_{A50}$ and negatively associated with $\text{Log}(S_1 - S_{99})$. D2 was positively associated with $\text{Log}(N_1 - N_{99})$. D3 was mainly associated with variability in A-weighted sound-pressure levels, but the percentage of explained variance was low. For this reason it was not worthwhile to give D3 any further attention.

The importance of fore- and background sounds, as well as distinguishable and indistinguishable sounds has been raised previously (Andringa, 2013; Andringa and van den Bosch, 2013). Andringa (2013) argues that these are central dimensions of soundscape and perceived safety. A close or indistinguishable sound source may induce a feeling of threat, whereas a distant or distinguishable sound source may induce a feeling of control.

It is interesting that none of the dimensions (D1–D3) were well-predicted by any single acoustic or psychoacoustic parameter. In all cases a combination of at least two parameters was needed to reach a sizable percentage of variance explained in the dependent variable. This result provide support for the statement in the introduction that acoustic and psychoacoustic

parameters are developed for the purpose of single sounds or sound sources, not for the purpose of soundscape, nor for measuring acoustic environments holistically.

The rationale for the method used in Experiment 1 is that spectrograms represent all acoustic information of an acoustic environment, except the phase angle of the frequencies. Thus, spectrograms were used as a tool for visualizing the acoustic data representing the 50 investigated acoustic environments. By visual inspection of the spectrograms, it was possible to decide to what degree they resembled each other. Spectrograms that look similar should represent acoustic environments that are similar. Consequently, the dimensions that underlie the similarity perceived among the spectrograms should represent holistic acoustic properties. These dimensions can be identified by the aid of Multidimensional Scaling (MDS). Furthermore, the visual sorting task allowed the participants to see and to assess the whole set of stimuli, and to fully compare them with each other.

It is reasonable to ask how many stimuli are necessary to properly map all relevant acoustic dimensions of acoustic environments. The theory behind MDS states that at least nine stimuli are needed to reach a definite MDS solution (Coxon, 1982). SPSS can handle 100 stimuli at most. The stimuli must also be selected to vary with regards to all relevant aspects. For this reason a wide selection is desirable. As specified in the method section, the 50 stimuli used in the present study represent a wide selection of acoustic environments in and around two large cities, which meet the requirements (Axelsson et al., 2010).

With regards to the quality of the present study, it could be argued that it would have been better to calculate the similarity of the spectrograms mathematically, rather than conducting an experiment based on visual perception. However, mathematical calculation of the similarities would have to be based on criteria defined by the experimenter, which could introduce a bias. Using the average response of human participants who unguided develop their own criteria in a sorting task based on what they can see in the spectrograms, and on what makes sense to them, overcomes this potential limitation.

## EXPERIMENT 2: SORTING OF AUDIO RECORDINGS

Considering the outcomes of Experiment 1, it is reasonable to ask to what extent Dimensions 1–3 correspond to how people perceive the acoustic environments. For this reason, a second experiment was conducted in which a new group of participants sorted a subset of the audio recordings.

## Method
### Participants
Ten expert listeners, 22–32 years old (3 women, 7 men; $M_{age} = 26.6$ years, $SD = 3.7$), post-graduates at the Department of Music or the Acoustics Group at the School of Architecture, University of Sheffield, took part in the experiment. Two out of ten persons had also taken part in Experiment 1. Participants attended on a voluntary basis and were not reimbursed.

## Stimulus Material

Based on the MDS solution obtained in Experiment 1, the audio files corresponding to the six most extreme spectrograms (three from the positive, and three from the negative pole) of each of the tree MDS dimensions (D1, D2, and D3 in **Figure 2**) were selected. Thus, there were 18 experimental sounds in total: Sound 17, 39, 48 ($D1^-$); 26, 37, 44 ($D1^+$); 5, 27, 36 ($D2^-$); 19, 28, 38 ($D2^+$); 6, 11, 34 ($D3^-$); 13, 25, 35 ($D3^+$) (see also **Table 1**).

## Equipment

The equipment consisted of a laptop (Asus, Realtek Audio soundcard), and a pair of acoustically open, circumaural headphones (Sennheiser HD 558). The selected audio recordings were played back at the authentic sound-pressure level (Brüel & Kjær Type 4231 sound calibrator).

## Procedure and Design

The experiment took place in the anechoic chamber of the School of Architecture, University of Sheffield. The design consisted of a two-stage data collection procedure: sorting and interview. The participants took part individually.

The experiment was designed to test whether or not the participants would reproduce the six groups that the 18 experimental sounds were selected from. Consequently, the participants were instructed to sort the 18 experimental sounds, presented in the form of icons on a computer screen, into six groups, with the restriction that there had to be exactly three sounds in each group. The sorting had to be based on the similarity of the sounds, so that similar sounds were grouped together. The participants were instructed to engage in holistic listening and assess the similarity of the sounds based on an overall sonic impression, disregarding semantic information. The experimental sounds were presented in a unique random order to every participant. The participants were allowed to play each sound as many times as desired and to revise their sorting throughout the experimental session, including the subsequent interview. Thus, after completing the sorting task, the participants were interviewed about their own sorting criteria. The 10 listening sessions lasted between 20 and 37 min each ($M_{time}$ = 30.2 min, $SD$ = 4.8). There were no time restrictions.

## Results

**Table 3** presents the number of complete, partially complete and incomplete groups that the 10 participants achieved. Two of the participants reproduced the six groups completely. Both were female music students. Two participants reproduced four of the six groups and the remaining two groups partly by 'misallocating' one sound in each. Both were post-graduates in acoustics. One participant reproduced one group completely and three groups partly. The remaining five participants reproduced 1–5 groups partly and none completely.

Eighteen sounds can be organized in 18! (i.e., eighteen factorial) permutations. There is $3!^6 \times 6!$ ways of achieving six complete groups. The probability of achieving six complete groups in the sorting task is $3!^6 \times 6!/18!$, which equals $5.25 \times 10^{-9}$. Thus, it is highly improbably to achieve six complete

**TABLE 3 |** Experiment 2: number of complete, partially complete and incomplete groups that 10 participants achieved.

| Participant | Complete | Partial | Incomplete |
|---|---|---|---|
| 1 | | 1 | 5 |
| 2 | | 2 | 4 |
| 3 | | 3 | 3 |
| 4 | | 4 | 2 |
| 5 | 1 | 3 | 2 |
| 6 | | 5 | 1 |
| 7 | 4 | 2 | |
| 8 | 4 | 2 | |
| 9 | 6 | | |
| 10 | 6 | | |
| Total | 21 | 22 | 17 |

*Complete means that all three experimental sounds that belong in the same group were grouped together. Partial means that 2 out of 3 experimental sounds that belong in the same group were grouped together as expect.*

groups out of pure chance. Still, two participants achieved this result, independently.

To further investigate how likely it is to obtain the results reported above by pure chance, a Monte Carlo experiment was set up. In this experiment, 6 groups of 3 items were sorted at random 10 times, representing 10 participants. For each 'participant,' the six groups were classified as Complete, Partial or Incomplete, counted and recorded. The procedure was repeated 1,000 times. The results show that, on average, 10 participants would together achieve 0.48 complete, 19.75 partially complete, and 39.77 incomplete groups, by chance. For Experiment 2, the result was 21 complete, 22 partially complete, and 17 incomplete groups (**Table 3**). A Chi-Square test shows that the empirical results deviate statistically significantly from chance ($\chi^2_2 = 886.7, p < 0.001$).

## Discussion

It seems that when the 10 expert listeners sorted the 18 experimental sounds in Experiment 2, six of them did something qualitatively different from the 30 participants in Experiment 1 when they sorted the 50 spectrograms. This seems to indicate that perception of acoustic environments chiefly belongs to a different domain compared to the acoustic properties of the same acoustic environments. Thus, dealing with acoustic environments it is necessary to decide if it is the perceived properties that are of interest or the acoustic properties. The two must not be confused. These results are in line with previous findings in soundscape research.

Guastavino (2007) investigated the way in which people categorize environmental sounds in their everyday lives, through a free categorisation task with open-ended verbal descriptions. The presence of human activity emerged as a main clustering criterion, suggesting that environmental sounds are processed and categorized based on their meaning, when such information is available. This seems to be the case in the present Experiment 2, but not in Experiment 1. This is also a potential limitation in the design of Experiment 2. Aucouturier and Defreville (2009) used manipulated ('spliced') acoustic signals, where sound sources

were not identifiable, and found that individuals were still able to judge the similarity of such acoustic signals in a meaningful way. This is probably similar to what the 30 participants did in Experiment 1.

## EXPERIMENT 3: SORTING OF SPLICED SIGNALS

In order to investigate whether the meaning of the sounds affected the results of the sorting task in Experiment 2, a third experiment was conducted. In this listening experiment spliced signals were used in agreement with Aucouturier and Defreville (2009).

## Method
### Participants
Ten expert listeners, 24–33 years old (4 women, 6 men; $M_{age} = 28.4$ years, $SD = 3.6$), post-graduates at the Department of Music or the Acoustics Group at the School of Architecture, University of Sheffield, took part in the experiment. None had taken part in Experiments 1 or 2. Participants attended on a voluntary basis and were not reimbursed.

### Stimulus Material
The same 18 sounds as used in Experiment 2 were used in Experiment 3. However, for Experiment 3 the acoustic signals were spliced in agreement with Aucouturier and Defreville (Coxon, 1982). Every signal was cut into segments of 50 ms, which then were reorganized in a unique random order.

### Equipment, Procedure, and Design
The same equipment as in Experiment 2 was used. The procedure and design was the same as in Experiment 2. The 10 listening sessions lasted between 24 and 38 min each ($M_{time} = 31.6$ min, $SD = 4.6$). There were no time restrictions.

## Results
**Table 4** presents the number of complete, partially complete and incomplete groups that the 10 participants achieved. Two participants achieved five partial and one complete groups. One participant achieved one complete, one partial, and four incomplete groups. The remaining seven participants achieved 3–5 partial groups. A Chi-Square test comparing these results with results expected by chance (see Experiment 2 above), showed that the results deviates statistically significantly from chance ($\chi^2_2 = 40.84$, $p < 0.001$).

Comparing these results with those obtained in Experiment 2 also shows a statistically significant difference between the two results ($\chi^2_2 = 17.88$, $p < 0.01$). Taken together, the results indicate that the 10 participants in Experiment 2 performed better than the participants in Experiment 3. The participants in Experiment 2 achieved 21 complete, 22 partially complete and 17 incomplete groups, compared with the 3 complete, 38 partially complete, and 19 incomplete groups that the participants in Experiment 3 achieved (**Tables 3**, **4**). Thus, the participants in Experiment 3 achieved fewer complete and more partially

**TABLE 4 |** Experiment 3: number of complete, partially complete and incomplete groups that 10 participants achieved.

| Participant | Complete | Partial | Incomplete |
|---|---|---|---|
| 1 | 1 | 1 | 4 |
| 2 | | 3 | 3 |
| 3 | | 3 | 3 |
| 4 | | 4 | 2 |
| 5 | | 4 | 2 |
| 6 | | 4 | 2 |
| 7 | | 4 | 2 |
| 8 | | 5 | 1 |
| 9 | 1 | 5 | |
| 10 | 1 | 5 | |
| Total | 3 | 38 | 19 |

*Complete means that all three experimental sounds that belong in the same group were grouped together. Partial means that 2 out of 3 experimental sounds that belong in the same group were grouped together as expect.*

complete groups than the participants in Experiment 2. In addition, the Chi-Square coefficients show that the participants in Experiment 2 deviated more strongly from chance performance than the participants in Experiment 3.

## Discussion
In Experiment 3, a groups of expert listeners, equivalent to the participants in Experiment 2, achieved a statistically significantly worse result when listening to spliced signals, compared to the results that the participants in Experiment 2 achieved by listening to the authentic acoustic signals. Contrary to expectation and initial assumptions, these results indicate that the spectrograms include information about the meaning of the recorded sounds, not merely meaningless acoustic data.

## GENERAL DISCUSSION AND CONCLUSION

The purpose of the present study was to explore the acoustic properties of acoustic environments holistically. In Experiment 1, spectrograms corresponding to different urban acoustic environments were sorted based on how similar they were. The sorting data was subjected to MDS analysis, and three MDS dimensions were identified: (D1) Distinguishable–Indistinguishable sounds sources, (D2) Foreground–Background sounds, and (D3) Intrusive–Smooth sound sources. None of these dimensions were well-predicted by any single acoustic or psychoacoustic parameter. According to the experimenters' original research plan, Experiment 2 was meant to validate the results of Experiment 1. However, only four of the ten participants achieved the expected result. This raised the question whether or not the spectrograms include information about the meaning of the recorded sounds. Consequently, a new listening experiment was conducted in which ten participants listened to and sorted spliced acoustic signals. Experiment 3 provided a statistically significantly worse result than Experiment 2. These results suggest that there is information about the meaning of

the recorded sounds in the spectrograms, and that the meaning of the sounds may be captured with the aid of holistic features of the acoustic environment. These new, unknown, features remain to be discovered. A possible feature could be the 'noticeability' of events and/or sources. In soundscape research this has often been referred to as 'saliency' of the sounds (Oldoni et al., 2013), which can be defined as the likeliness of a sound event to attract the auditory attention of a listener at unconscious (i.e., biological) level. This can also be applied to the visual domain and would justify how participants were able to attribute 'meaning' to patterns in the spectrograms (e.g., the pneumatic drill in excerpt 19 or the birdsong in excerpt 43). To a large extent, saliency of sources would be lost in spliced signals, which is consistent with the worse performance in Experiment 3 compared to Experiment 2.

Regarding the acoustic properties of the acoustic environments, the main conclusions from this study are related to the results of Experiment 1:

(1) Sound sources interpreted as distinguishable had a lower median sound level ($L_{A50}$) and a higher variability in Sharpness [$Log(S_1 - S_{99})$] than sound sources interpreted as indistinguishable. Thus, "distinguishable" sound sources seem to be related to acoustic environments with relatively low mean sound levels and large changes in the spectral content between low and high frequencies over time.

(2) Sounds interpreted as background had a low variability and sounds interpreted as a high variability in Loudness [$Log(N_1 - N_{99})$]. Thus, acoustic environments with larger loudness variability over time seem to be related to sounds that emerge from the background noise, perceptually.

Taken together, the results of this study show that at present there are no acoustic indicators available that can be used to assess acoustic environments holistically. More specifically, in the linear regression models, none of the considered acoustic metrics alone explained a large amount of variance in the dimensions underlying the perceived similarity of acoustic properties of the investigated acoustic environments. This gap has also been acknowledged by previous research, where it was pointed out that more predictive models for perceptual features are desirable in soundscape research (Aletta et al., 2016). Further in-depth research is needed in this field, which has to include mathematical modeling of the acoustic properties of acoustic environments considered holistically.

A potential limitation in this study is related to the Fast Fourier Transform (FFT) that underlie the spectrograms. The question is how different a spectrogram would be if different settings for the time, frequency and/or amplitude resolution were used, and how this would affect the results of the study. Would spectrograms that were similar in this study—using the default settings—be more or less similar if a different resolution was used? Further studies are needed to validate the present approach to the acoustic properties of the acoustic environment considered holistically.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue of the University of Sheffield with written informed consent from all participants. All participants gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Ethical Committee of the School of Architecture of the University of Sheffield.

## AUTHOR CONTRIBUTIONS

All authors conceived the study and designed the experiments. FA carried out the experiments. FA and ÖA analyzed the results. All authors wrote and critically reviewed the paper.

## SPECIAL NOTE

Portions of this work were presented in "Toward acoustic indicators for soundscape design," Proceedings of Forum Acusticum, Krakow, Poland, 7–12 September 2014.

## FUNDING

## REFERENCES

Aletta, F., Axelsson, Ö., and Kang, J. (2014). "Towards acoustic indicators for soundscape design," in *Proceedings of the Forum Acusticum 2014 Conference*, Krakow.

Aletta, F., Kang, J., and Axelsson, Ö. (2016). Soundscape descriptors and a conceptual model for developing predictive soundscape models. *Landscape Urban Plan* 149, 65–74. doi: 10.1016/j.landurbplan.2016.02.001

Andringa, T. C. (2013). "Soundscape and its Relation to Core Affect, Appraisal, and Motivation," in *Proceedings of AIA-DAGA Conference*, Merano, 1511–1513.

Andringa, T. C., and van den Bosch, K. A. (2013). "Core affect and soundscape assessment. Fore- and background soundscape design for quality of life," in *Proceedings of Internoise*, Innsbruck, 2273–2282.

Andringa, T. C., Weber, M., Payne, S. R., Krijnders, J. D., Dixon, M. N., Linden, R. V. D., et al. (2013). Positioning soundscape research and management. *J. Acoust. Soc. Am.* 134, 2739–2747. doi: 10.1121/1.4819248

Aucouturier, J., and Defreville, B. (2009). Judging the similarity of soundscapes does not require categorization: evidence from spliced stimuli. *J. Acoust. Soc. Am.* 125, 2155–2161. doi: 10.1121/1.3083232

Axelsson, Ö. (2007). Towards a psychology of photography: dimensions underlying aesthetic appeal of photographs. *Percept. Mot. Skills* 105, 411–434. doi: 10.2466/pms.105.2.411-434

Axelsson, Ö., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *J. Acoust. Soc. Am.* 128, 2836–2846. doi: 10.1121/1.3493436

Botteldooren, D., De Coensel, B., and De Muer, T. (2006). The temporal structure of urban soundscapes. *J. Sound Vib.* 292, 105–123. doi: 10.1016/j.jsv.2005.07.026

Brambilla, G., Gallo, V., Asdrubali, F., and D'Alessandro, F. (2013). The perceived quality of soundscape in three urban parks. *J. Acoust. Soc. Am* 134, 832–839. doi: 10.1121/1.4807811

Brown, A. L., Kang, J., and Gjestland, T. T. (2011). Towards standardization in soundscape preference assessment. *Appl. Acoust.* 72, 387–392. doi: 10.1016/j.apacoust.2011.01.001

Cain, R., Jennings, P., and Poxon, J. (2013). The development and application of the emotional dimensions of a soundscape. *Appl. Acoust.* 74, 232–239. doi: 10.1016/j.apacoust.2011.11.006

Cain, R., Jennings, P., Poxon, J., and Scott, A. (2009). "Emotional dimensions of a soundscape," in *Proceedings of Internoise*, Ottawa, ON, 4660–4667.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivar. Behav. Res.* 1, 245–276. doi: 10.1207/s15327906mbr0102_10

Council of Europe (2000). *European Landscape Convention. European Treaty Series, No. 176*. Florence: Council of Europe.

Coxon, A. (1982). *The User's Guide to Multidimensional Scaling*. London: Heinemann Educational, 271.

Davies, W. J. (2013). Editorial to the special issue: applied soundscapes. *Appl. Acoust.* 74, 233. doi: 10.1016/j.apacoust.2012.07.004

De Coensel, B., and Botteldooren, D. (2006). The quiet rural soundscape and how to characterize it. *Acta Acust. Acust.* 92, 887–897.

EEA (2014). *Good Practice Guide on Quiet Areas*. Luxembourg City: Publications Office of the European Union.

Fastl, H., and Zwicker, E. (2007). *Psychoacoustics – Facts and Models*, 3rd Edn. Berlin: Springer Verlag.

Genuit, K., and Fiebig, A. (2006). Psychoacoustics and its benefits for the soundscape approach. *Acta Acta Acust. Acust.* 92, 1–7.

Guastavino, C. (2007). Categorization of environmental sounds. *Can. J. Exp. Psychol.* 61, 54–63. doi: 10.1037/cjep2007006

Herranz Pascual, K., Aspuru Solaga, I., and Garcia Perez, I. (2010). "Proposed conceptual model of environmental experience as framework to study the soundscape," in *Proceedings of Internoise*, Lisbon, 2904–2912.

Ishihara, S. (1957). *Test for Colour Deficiency – 24 Plates Edition*. Tokyo: Kanehara Shuppan, 24.

ISO (2014). *ISO 12913-1:2014 Acoustics — Soundscape — Part 1: Definition and Conceptual Framework*. Geneva: International Organization for Standardization.

Kang, J., Aletta, F., Gjestland, T. T., Brown, A. L., Botteldooren, D., Schulte-Fortkamp, B., et al. (2016). Ten questions on the soundscapes of the built environment. *Build. Environ.* 108, 284–294. doi: 10.1016/j.buildenv.2016.08.011

Oldoni, D., De Coensel, B., Boes, M., Rademaker, M., De Baets, B., Van Renterghem, T., et al. (2013). A computational model of auditory attention for use in soundscape research. *J. Acoust. Soc. Am.* 134, 852–861. doi: 10.1121/1.4807798

Rychtáriková, M., and Vermeir, G. (2013). Soundscape categorization on the basis of objective acoustical parameters. *Appl. Acoust.* 74, 240–247. doi: 10.1016/j.apacoust.2011.01.004

Schulte-Fortkamp, B., and Dubois, D. (2006). Preface to the special issue: recent advances in soundscape research. *Acta Acust. Acust.* 92, I–VIII.

Schulte-Fortkamp, B., and Kang, J. (2013). Introduction to the special issue on soundscapes. *J. Acoust. Soc. Am.* 134, 765–766. doi: 10.1121/1.4810760

Truax, B. (ed.). (1978). *Handbook for Acoustic Ecology, No. 5, Music of the Environment Series, World Soundscape Project*. Vancouver, BC: ARC Publications, 160.

Young, F., and Lewyckyj, R. (1979). *ALSCAL-4 User's Guide*. Chapel Hill, NC: University of North Carolina, 128.

# Motorized Recreation Sounds Influence Nature Scene Evaluations: The Role of Attitude Moderators

*Jacob Benfield[1]\*, B. D. Taff[2], David Weinzimmer[3] and Peter Newman[2]*

[1] *Penn State Abington, Abington Township, PA, United States,* [2] *Recreation, Park, and Tourism Management, Pennsylvania State University, University Park, PA, United States,* [3] *Human Dimensions of Natural Resources, Colorado State University, Fort Collins, CO, United States*

Soundscape assessment takes many forms, including letting the consequences of the soundscape be an indicator of soundscape quality or value. As a result, much social science research has been conducted to better quantify problem soundscapes and the subsequent effects on humans exposed to them. Visual evaluations of natural environments are one area where research has consistently shown detrimental effects of noisy or anthropogenic soundscapes (e.g., those containing noise from motorized recreation), but the potential moderating role of individual attitudes toward elements within the soundscape has not been sufficiently explored. This study demonstrates that both pro-motorized recreation and pro-motorized recreation management attitudes can alter the effect of motorized recreation noise on scenic evaluations in opposing directions. Pro-recreation attitudes lessen the effect of the soundscape, while pro-management attitudes heighten the negative effect of anthropogenic sounds on scenic evaluation. The implications for other areas of soundscape research, especially with regard to soundscape quality assessment through experienced outcomes, are discussed, including possible strategies for prioritizing known or relevant moderating variables.

Keywords: noise, affect, resource management, national park, overflight, motorcycle

## INTRODUCTION

Soundscapes represent a dynamic, complex system of auditory stimuli that can encompass both objective and subjective properties (Bell et al., 2001). For instance, an outdoor concert held at the Acropolis in Athens, Greece creates a soundscape rich with physical stimuli. These include elements such as pitch or intensity, both able to be measured via instrumentation of varied types, while the content of the music itself, combined with the historic location, can embody more subjective properties, such as joy or sorrow, to those in attendance. The measurement of that subjective experience and the assessment of how the soundscape drives the effect utilize completely different methods and instrumentation. As such, the scientific study of soundscapes can be equally as dynamic and complex as the study of visual landscapes.

For example, some research on soundscape focuses heavily on urban soundscapes or transportation effects on residential soundscapes with varying degrees of objective physical measurement or more subjective qualitative interviewing (e.g., Botteldooren et al., 2006; Payne, 2008; Jeon et al., 2010; Craig et al., 2014). Others have emphasized natural or rural soundscape

assessment with varying levels of instrumentation, acoustical monitoring, or survey sampling (e.g., De Coensel and Botteldooren, 2006; Stack et al., 2011; Reed et al., 2012; Jiang et al., 2017). Each of these studies, with varied approaches and built or largely natural environments, contributes to our understanding of soundscapes in meaningful ways but all utilized different assessment methods.

The current study represents one approach—utilizing laboratory simulations of experimentally manipulated soundscapes in a visual landscape evaluation task—aimed at the assessment of subjective qualities and outcomes of soundscapes within a specific type of environment (i.e., protected natural areas). In other words, one method for assessing soundscapes relies less on the physical qualities of the soundscape itself but instead on the outcomes that occur during or immediately following exposure to that soundscape. In this study, the outcome of interest is changes in perception of the visual landscape that occur under differing soundscapes. Perhaps most importantly, this study aims also to highlight an aspect of subjective survey-based soundscape assessment that often goes unstudied or unreported: the moderating role of individual attitudes toward the stimuli.

## Natural Soundscapes as a Resource and Management Priority

Existing soundscape research in natural areas supports the idea that opportunities to experience the sounds of nature are important factors in determining the quality of recreational experiences to visitors of these areas. Foundational research by Driver et al. (1987) demonstrated that in choosing their experiences many outdoor recreationists are motivated to find respite from excessive noise and urban environments. Subsequently, McDonald et al. (1995) found that the vast majority (over 90%) of survey respondents in national parks and protected areas listed the enjoyment of natural quiet and the sounds of nature as important reasons for their visit. In recent years, it has become increasingly apparent to land managers that natural soundscapes are as deserving of protection and careful stewardship as other natural and cultural resources (Newman et al., 2010).

In the United States, federal land management policies over the last few decades have responded to the concern that natural areas are threatened by growing levels of noise from human activities and development. In 1987, the U.S. National Parks Overflights Act was a pioneering piece of legislation that sought for the first time to systematically protect natural soundscapes in National Park Service (NPS) lands. According to Gramann (1999), this law provided the impetus for new directions of investigation into the effects of noise in parks and other protected areas.

The NPS addressed soundscape management directly in 2000 by Director's Order #47 ("Soundscape Preservation and Noise Management"), with the goal "to articulate National Park Service operational policies that will require, to the fullest extent practicable, the protection, maintenance, or restoration of the natural soundscape resource in a condition unimpaired by inappropriate or excessive noise sources" (National Park Service [NPS], 2000, p. 1). This order sought to establish official direction for the preservation of intrinsic park soundscapes by means of better planning, monitoring, and assessment. Also in 2000, the National Parks Air Tour Management Act mandated that the NPS and Federal Aviation Administration work together to identify and mitigate adverse effects of commercial air tours on the soundscapes of parks (National Park Service [NPS], 2000). In 2006, the NPS re-affirmed its commitment to restoring and protecting natural soundscapes by addressing soundscape management in several sections of its official Management Policies, including the statement that "The Service will preserve, to the greatest extent possible, the natural soundscapes of parks" (National Park Service [NPS], 2006, section 4.9). Similar mandates designed to better protect or manage natural soundscapes or to mitigate excessive noise have occurred in many other countries (e.g., Directive 2002/49/EC of the European Parliament, 2017). Thus, the need for clear management standards and mechanisms for assessing soundscape quality and impact in a range of contexts becomes not only a research issue, but a societal one as well.

## Laboratory Research on Soundscapes Assessment for Natural Environments

An important consequence of these legislative and administrative mandates is the pursuit of biological, physical, and social science research on the existing soundscape as well as effects of soundscape on inhabitants and users of natural areas. In other words, these mandates have helped to generate a need for soundscape assessment across several domains of study. One area that has been of particular interest to social scientists and recreation researchers has been the role of natural and anthropogenic soundscapes on visitor evaluations of natural scenes and areas. For example, motorized recreation in the form of snowmobiles, propeller plane overflights, all-terrain vehicle (ATV) excursions, and organized motorcycle rallies are common activities in natural areas but all generate large amounts of high intensity, disruptive anthropogenic noise into the soundscape. Visitors to those locations who are not participating in those activities may have a lessened experience because of that additional vehicle noise. While field-based assessments have also been conducted (e.g., Pilcher et al., 2009; Marin et al., 2011), laboratory simulations have been equally useful and prevalent in assessing the perceived quality and impact of different soundscapes.

For example, early research by Mace et al. (1999) adapted the traditional laboratory-based landscape assessment paradigm to test for impacts from motorized recreation noise on aesthetic ratings of landscape quality in simulated national park settings. This laboratory study looked at potential effects of helicopter tour noise on evaluations of scenic overviews in Grand Canyon National Park. In addition to collecting aesthetic ratings, self-reported changes in affective state in response to the auditory stimuli were also collected. Results of this study showed that negative experiential and aesthetic effects were associated with soundscapes that included helicopter noise, relative to purely

natural soundscapes. This was one of the first studies to directly demonstrate that auditory environments could influence visual ratings.

Follow-up research by Mace et al. (2003) investigated the importance of soundscape source attribution when assessing the presence of motorized helicopter noise in natural settings. To test this, the researchers presented the same auditory stressor (i.e., helicopter overflights) within the soundscape but attributed its purpose to either scenic overflights for tourists, backcountry maintenance activities by park management, or life-saving search and rescue operations. The results indicated that a soundscape dominated by helicopter noise from either tourist overflights or official park management activities was similarly detrimental to the experiences of potential park visitors when compared to natural-only soundscapes. It was also shown that type of visual scenery, such as mountainous or forested, also partially determined the amount of influence different soundscapes had on ratings.

Subsequent laboratory work by Benfield et al. (2010) expanded the findings of Mace et al. (1999, 2003) in several ways. First, the range of soundscape elements investigated was expanded from helicopter noise to include human voices, airplane overflights, and motorized ground vehicles. Second, the range of soundscape settings evaluated was expanded to include three additional national parks (Yellowstone, Everglades, and Olympic National Parks) to demonstrate the robustness of the previously shown location effect. Finally, the number of affective outcomes assessed was expanded to include fatigue, hostility, and other specific states beyond positive or negative affect. Consistent with the prior studies, the anthropogenic soundscapes were each responsible for detriments to both affective state and visual assessment of the landscapes shown. Individual positive affect, attentiveness, and serenity was lowered by the presence of anthropogenic soundscapes, and ratings of hostility increased. Visual assessments of scenic tranquility, beauty and solitude similarly decreased in the presence of anthropogenic noise while ratings of annoyance were higher compared to those in the natural sound condition.

Most recently, Weinzimmer et al. (2014) further refined the assessment of specific soundscape events by comparing directly three common sources of motorized noise in national parks – motorcycles, propeller planes, and snowmobiles. Using a carefully controlled laboratory simulation, Weinzimmer et al. (2014) directly compared different motorized vehicle soundscapes using a within-subjects design. Those direct comparisons replicated previous studies by showing that motorized recreation noise had significant, detrimental effects on both aesthetic and affective dimensions. Those comparisons also demonstrated interesting differences between the three different sources of anthropogenic noise showing that nuanced assessments of the qualities of the sounds themselves were needed.

## Potential Moderators of Subjective Soundscape Assessments

In addition to assessing the main effect of anthropogenic sound on visual evaluations or self-reported mood, research in this domain has occasionally examined different moderators of the effect soundscape has on those outcomes. Some of those efforts have been more successful than others.

As stated previously, Mace et al. (2003) manipulated sound exposure but also examined the role of sound source attribution on subsequent scenic evaluations. By describing the helicopter noise as arising from either legitimate park operations (maintenance, search, and rescue) or from tourist entertainment (scenic overflight), the authors hoped to show that higher perceived legitimacy of the noise source would lessen the detrimental effect previously observed. While subtle differences were shown between the different noise attributions (e.g., scenic beauty was lower for legitimate conditions compared to tourist activity), the overwhelming consensus was that the presence of helicopter noise was detrimental irrespective of attribution given for the sound.

Benfield et al. (2010) conducted *ad hoc* analyses to test the interaction between the visual appeal of the scene being assessed and the soundscape of the scene. They showed that more beautiful scenes, as assessed by participants in the absence of sound, were more affected by the presence of anthropogenic sounds than scenes rated as less beautiful. This effect was shown to generalize across sound types (voice, aircraft, and ground traffic) and the four different parks tested. Essentially, this moderation effect when combined with the overall findings suggested that soundscapes could impact visual quality assessment but that the inherent visual quality of the scene also determined the magnitude of impact any given sound could have on the ratings.

Despite these advances in understanding potential moderators that may influence subjective perceptions of soundscape and subsequent impacts that soundscape has on other scene ratings, much less research has examined how individual attitudes – positive or negative – toward specific elements in the soundscape change those outcomes. This is in spite of a wealth of research on attitudes affecting other aspects of natural resource assessment or management.

Manfredo et al. (2004) suggested that attitudes are some of the most frequently examined and central measures within the assessment of human dimensions of natural resources. For example, over the past 30 years attitudes have been found to predict and influence support of recreational management strategies (Bright, 1997), preferences toward national forest use and management strategies (Clement and Cheng, 2011), perceptions of crowding (Shelby and Heberlein, 1986; Manning, 2007), evaluations of wildlife management strategies (Manfredo, 2008), use of transportation in parks (White et al., 2011; Taff et al., 2013), and perceptions toward resource impacts (Monz, 2009), to name a few. However, only a few studies have explored attitudes toward noise sources and soundscape assessment specifically.

Within the urban setting of Hong Kong, Lam et al. (2009) found that negative attitudes toward railway noise increased annoyance of associated soundscapes but did not significantly affect annoyance toward road-based traffic noise. In the Netherlands, Pedersen et al. (2009) found that negative attitudes toward the visual impact of wind turbines significantly increased annoyance from turbine-associated noise. Tarrant et al.

(1995) conducted surveys with visitors to Wyoming wilderness areas to explore attitudes toward seeing and hearing aircraft, as well as other dimensions of wilderness experience. The authors found that respondents' estimates of noise levels were strongly related to their attitudes toward aircraft overflights, suggesting that wilderness visitors may respond differently to aircraft based in part on their attitudes.

Lai et al. (2009) segmented visitors to a national seashore based on their attitudes toward natural resource management in order to develop marketing strategies. One of the attitudinal items they evaluated related to the elimination of human-caused noises from the seashore, which factored into a dimension the authors termed 'preventing encroachment.' This study discovered three different visitor segments based on respondent attitudes, which included 'conservation-oriented,' 'development-oriented,' and 'status quo' visitors. Results indicated that conservation-oriented respondents were most supportive of 'preventing encroachment,' while development-oriented respondents were least supportive of this action.

Finally, Taff et al. (2014) indirectly manipulated park visitor attitudes toward an existing noise source through the use of messaging. Specifically, prior research had shown that aircraft overflights in a national park in the western United States from a nearby military installation were both frequently noticed and consistently rated as detrimental to the visitor experience. As a follow-up to that finding, these researchers asked park visitors to rate the acceptability of several sound clips taken from inside the park, with some containing a higher prevalence and intensity of military aircraft noise. Half of the surveyed visitors were given no information about the clips while the other half were given information about the overflights' purpose, including the overflights being "in an effort to help keep the United States of America safe." Participants in the "keeping America safe condition" were less likely to rate the overflights as problematic or below minimal levels of acceptability than those who heard the soundscapes without context.

Attitudes are among the most important measures when determining management approaches in parks and protected areas (Manfredo et al., 2004; Vaske, 2008) and should thus be included in the assessment of soundscapes in those areas. However, the potential moderating role of attitudes toward recreation or recreation management within the context of park soundscape experiences deserves additional attention. Specifically, research has not assessed how attitudes toward motorized recreation or the management of motorized recreational noise influences the evaluation of park soundscape experiences.

## The Current Study

Research consistently shows that soundscapes dominated by anthropogenic stimuli has a detrimental effect on visual evaluations of natural landscapes (Mace et al., 1999). That same research has shown that situational aspects, such as noise source attribution (Mace et al., 2003) or the beauty of the scene being evaluated (Benfield et al., 2010), can moderate the effect that sound can have on subsequent evaluations. However, attitudes of the person experiencing the noise and making the evaluations

have not been adequately explored as potential moderators. Considering other research has shown that individual attitudes can impact individual perception of both anthropogenic noise (Taff et al., 2014) and a host of other management policies (Manfredo et al., 2004), a better understanding of the effect of attitudes in relation to recreation noise and soundscape assessment is warranted.

To test the potential moderating role of recreation or management attitudes, an experimental laboratory simulation similar to those cited previously was carried out. Individual attitudes in favor of motorized recreation (i.e., "pro-recreation" attitudes) or in favor of the regulation of motorized recreational noise (i.e., "pro-management" attitudes) were assessed prior to the simulation and the following hypotheses were made:

> H1: Individuals with a higher pro-recreation attitude will be less affected by a recreation noise soundscape when making scene evaluations or reporting affective state following exposure to recreation noise.
>
> H2: Individuals with a higher pro-management attitude will be more affected by a recreation noise soundscape when making scene evaluations or reporting affective state following exposure to recreation noise.

## MATERIALS AND METHODS

### Participants

Seventy-seven undergraduate and graduate students (43 females and 34 males) participated in a laboratory-based study for course research credit. Participants were of mostly of typical college age ($M = 22.38$ years; $SD = 6.89$; range = 16–50) and reported regular visits to national parks within the previous 12 months ($M = 2.94$, $SD = 1.40$).

### Design

This study utilized a 2 (soundscape) × 3 (park setting) repeated measures design. Participants received both soundscape conditions (natural only, natural and motorized recreation noise) in a randomized order. Within each of the soundscape conditions, participants viewed images of three different national park settings (Yellowstone, Glacier, and Denali) in random order. For analysis purposes, scenic evaluations are aggregated across parks and comparisons are made between the aggregate scene score for the two soundscape conditions.

### Materials and Measures
#### Scenic Evaluations
Scenic evaluations were based on an existing landscape assessment paradigm adapted for use in soundscape research (e.g., Mace et al., 1999; Benfield et al., 2010). Evaluations were performed along eight dimensions of aesthetic quality: naturalness, freedom, preference, annoyance, solitude, scenic beauty, tranquility, and acceptability. Prior researchers had chosen these dimensions to incorporate both physical qualities of the scene (e.g., beauty, naturalness) as well as affordances within in (e.g., solitude, freedom) and were retained in this study

to allow for direct comparisons with the most relevant prior literature. Ratings were obtained on a 10-point scale ranging from "1 = very low" to "10 = very high" and a composite score for the eight dimensions was used for all analyses (the annoyance dimension was reverse-coded prior to analysis). Participants were instructed to "evaluate the scene you viewed on the following characteristics" which emphasizes visual perceptions but encompasses the entirety of the scene, including the auditory stimuli.

### Recreation Management Attitudes

Recreation management attitudes were measured using a set of six items intended to measure project relevant attitudes along two opposing dimensions – acceptability of motorized recreation in spite of noise (e.g., "I would be willing to take a motorcycle through a park even if I knew the noise bothered other visitors."; three items; α = 0.84) and acceptability of banning motorized recreational vehicles (e.g., "Snowmobiles should not be allowed in national parks due to the noise they create."; three items; α = 0.86). These items were created based on interview responses of different user groups (motorized recreation users and pro-motorized management users) of one of the parks tested. These two scales, while conceptually representing two related but opposing viewpoints, were judged to be separate from one another. The two scales were negatively correlated with one another but at only moderate levels ($r = -0.57$). Factor analysis of those six items revealed two separate factors, representing 60.81% (pro-motorized recreation) and 16.82% (pro-motorized management) of the total variance. Further, factor loadings for each scale were strong for included items (>0.8) and weak for items in the other scale (<0.29). Within each dimension, items were rated along a seven-point "1 = strongly disagree" to "7 = strongly agree" continuum with an average response across items being calculated.

### Affective Ratings

Affective ratings were collected by self-report using the 20-item Positive and Negative Affect Scale (PANAS; Watson et al., 1988). Prior research (Mace et al., 1999; Benfield et al., 2010) had utilized the PANAS or its extended version and had shown that motorized sounds can also impact affective state of those making the ratings. Again, these measures were retained in this study to allow for direct comparisons with the most relevant prior literature that had not explored the role of moderators. Participants completed the PANAS at three different time points during the experimental procedure: at baseline, following the first natural sound condition, and following the first motorized sound condition. The PANAS consists of a series of words that represent different feelings, and participants use a five-point scale (ranging from 1 = "very slightly or not at all" to 5 = "extremely") to report how much each word describes how they are feeling at that moment. Half of the items are combined to give a positive affect score (α = 0.86–0.91; such as "enthusiastic," "determined," and "interested"), while the other half are combined to provide a negative affect score (α = 0.79–0.89; examples include "upset," "nervous," and "irritable").

### Soundscape Conditions

Soundscape conditions consisted of three conditions with only natural sounds (i.e., birds, wind, and water), and three conditions with motorized sounds added to the natural soundtracks. The motorized sounds consisted of recordings of a propeller plane, snowmobile, and a pair of motorcycles. All sound clips were obtained from the actual parks they were designed to represent in the simulations. Each participant experienced all six conditions (three with natural sounds only; three with overlaid motorized sounds) in one of six pseudo-randomized orders. Sound clips were 45 s in duration, with 7-s fade-in and fade-out effects. All clips were normalized such that the three natural clips had equivalent sound energy levels and the three motorized clips had equivalent sound energy levels. The normalized sound clips were then calibrated so that participants would hear (via headphones) the natural clips at approximately 45 dB(A) and the motorized clips at approximately 60 dB(A); these sound levels were chosen to be representative of those regularly experienced in these locations by visitors.

### Visual Scene Stimuli

Visual scene stimuli were chosen from available landscape photographs of three popular national parks: Yellowstone National Park, Glacier National Park, and Denali National Park and Preserve. These parks were chosen because of on-going and publicly debated issues of motorized recreation management in these areas. Photographic scenes were selected as representative scenic views within the parks. Developments like roads and buildings were not visible in the scenes. Four photographs from each park were included in the experiment, as well as two practice scenes from Grand Canyon National Park. Summer scenes were chosen for Glacier and Denali, while winter scenes were selected for Yellowstone. Seasons corresponded to the specific source of motorized noise from each park (e.g., winter scenes to match snowmobile noise in Yellowstone).

## Procedure

Participants were recruited from courses at a large state funded, public university located in close proximity (45 miles) to a large U.S. National Park. Responses were collected on iPad second generation computers (Apple Inc., Cupertino, CA, United States) programmed with iSURVEY software (Contact Software Limited, Wellington, New Zealand). Prior to participation in the research, all participants provided written consent to participate based on an IRB approved study description.

The landscape assessment task consisted of eight blocks of scenic ratings: two practice blocks, three natural sound blocks, and three motorized recreation sound blocks. Within each block, participants rated four visual scenes from the same park while being exposed to a single soundscape condition. These scenes were shown in random order for 45 s each and participants began making evaluations after 20 s. Thus, the total block time was 3 min of exposure to a single soundscape across four scenes; each block produced four sets of scenic evaluations along the eight aesthetic dimensions.

Affective states from the PANAS were acquired before the two practice blocks, following the first natural sounds block, and

following the first motorized sounds block. This spacing of the PANAS measurement (i.e., every three blocks) corresponds to approximately 9min between each measurement with the most recent 3 min including the corresponding soundscape of interest (baseline, natural, or motorized). The order of the natural or motorized sets of blocks was also randomized. See **Figure 1** for a summary of the procedure.

## RESULTS

The purpose of this study was to examine the potential moderating role that individual attitudes could have on the previously observed relationship between soundscape type and scenic evaluations or affective state. To test this, the average score on attitude items was used as a covariate in repeated measures analysis of covariance (R-ANCOVAs) comparing scene ratings or affective state between natural and motorized soundscape exposures (see **Tables 1**, **2**).

### Results for Scenic Evaluation

Consistent with prior research (e.g., Mace et al., 1999; Benfield et al., 2010), a main effect for sound condition on scenic evaluations was shown for both attitude moderators in the ANCOVAs. Specifically, scene ratings were higher for the natural sound condition ($M = 9.19$, $SD = 0.60$) when compared to scene ratings from motorized soundscapes ($M = 5.94$, $SD = 1.31$) for both the pro-motorized recreation attitude model [$F(1,73) = 201.83$, $p < 0.01$, $\eta_p^2 = 0.734$] and the pro-management attitude model [$F(1,73) = 30.75$, $p < 0.01$, $\eta_p^2 = 0.296$].

Consistent with study hypotheses, a significant interaction between sound type and the attitude moderator was also shown for both the pro-recreation attitude score [$F(1,73) = 10.04$, $p = 0.002$, $\eta_p^2 = 0.121$] and the pro-management attitude score [$F(1,73) = 13.12$, $p = 0.001$, $\eta_p^2 = 0.152$]. As predicted, higher pro-recreation attitudes lessened the negative effect of motorized sound on composite scene ratings (**Figure 2A**). The opposite

**TABLE 1** | Results of repeated measure ANCOVA comparing natural and anthropogenic noise with pro-motorized recreation attitude covariate.

| Outcome | | F | $\eta_p^2$ |
|---|---|---|---|
| **Pro-recreation attitude** | | | |
| Scenic evaluations | Sound type | 201.83** | 0.734 |
| | Attitude | 3.96* | 0.051 |
| | Sound type × attitude | 10.04** | 0.121 |
| Change in positive affect | Affect change | 7.65** | 0.097 |
| | Sound type | 0.53 | 0.007 |
| | Attitude | 3.49† | 0.047 |
| | Affect × sound type | 23.74** | 0.251 |
| | Affect × attitude | 1.05 | 0.015 |
| | Sound type × attitude | 0.15 | 0.002 |
| | Affect × sound type × attitude | 7.60** | 0.097 |
| Change in negative affect | Affect change | 4.76* | 0.063 |
| | Sound type | 3.97* | 0.053 |
| | Attitude | 4.11* | 0.055 |
| | Affect × sound type | 11.37** | 0.138 |
| | Affect × attitude | 6.49* | 0.084 |
| | Sound type × attitude | 1.17 | 0.016 |
| | Affect × sound type × attitude | 2.38 | 0.032 |

†$p < 0.08$, *$p < 0.05$, **$p < 0.01$.

pattern was found for pro-management attitudes: higher pro-management attitudes increased the negative effect of motorized sound on composite ratings (**Figure 2B**). **Table 1** displays the summary results for both repeated measure ANCOVAs.

### Results for Affective Ratings

Positive and negative affect were assessed both before and after each of the sound exposures, and repeated measures ANCOVA was used to test the potential moderating role of attitude on affective state. For each analysis, pre- and post-exposure PANAS ratings were analyzed for each sound condition with the attitude variable added as a covariate. Thus,



| | Sound Condition 1 (Randomly Assigned) | | | Sound Condition 2 (Randomly Assigned) | | |
|---|---|---|---|---|---|---|
| | Park 1 (Randomly Assigned) | Park 2 (Randomly Assigned) | Park 3 (Randomly Assigned) | Park 1 (Randomly Assigned) | Park 2 (Randomly Assigned) | Park 3 (Randomly Assigned) |

Block 1 (Practice) — Block 2 (Practice) — Block 3 — Block 4 — Block 5 — Block 6 — Block 7 — Block 8

Affective Measurements (PANAS)

**FIGURE 1** | Procedural diagram for landscape evaluation task and affective measures.

**TABLE 2 |** Results of repeated measure ANCOVA comparing natural and anthropogenic noise with pro-management of motorized recreation attitude covariate.
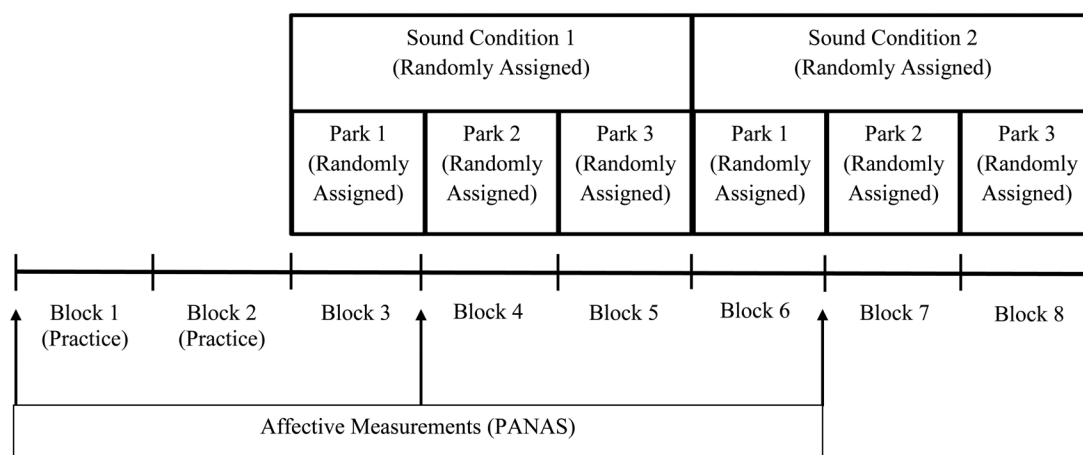
| Outcome | | $F$ | $\eta_p^2$ |
|---|---|---|---|
| **Pro-management attitude** | | | |
| Scenic evaluations | Sound type | 30.75** | 0.296 |
| | Attitude | 3.85* | 0.050 |
| | Sound type × attitude | 13.12** | 0.152 |
| Change in positive affect | Affect change | 0.44 | 0.006 |
| | Sound type | 1.74 | 0.024 |
| | Attitude | 0.26 | 0.004 |
| | Affect × sound type | 0.01 | 0.000 |
| | Affect × attitude | 4.53* | 0.060 |
| | Sound type × attitude | 2.60 | 0.035 |
| | Affect × sound Type × attitude | 4.22* | 0.056 |
| Change in negative affect | Affect change | 3.51† | 0.047 |
| | Sound type | 0.23 | 0.003 |
| | Attitude | 9.31 | 0.116 |
| | Affect × sound type | 0.01 | 0.000 |
| | Affect × attitude | 3.52† | 0.047 |
| | Sound type × attitude | 1.79 | 0.025 |
| | Affect × sound Type × attitude | 2.61 | 0.035 |

$^†p < 0.08$, $^*p < 0.05$, $^{**}p < 0.01$.



**FIGURE 2 |** Moderating effect of **(A)** pro-motorized recreation attitudes or **(B)** pro-management of motorized recreation attitudes on composite scene evaluations in the presence of natural or anthropogenic sounds. Bars represent the minimum and maximum values for scene rating observed for participants with that corresponding attitude score.

a significant three-way interaction between the change in affect rating, the sound condition, and the attitude measure indicated that attitude was moderating the effect on affective state for some sound conditions but not others. Main effects and two-way interactions were of less relevance to the current study aims and not hypothesized, although the interaction between change in affect and sound condition would be consistent with previous research (e.g., Mace et al., 1999; Benfield et al., 2010).

## Pro-motorized Recreation Attitudes

Pro-motorized recreation attitudes were hypothesized to lessen the previously shown deleterious effect of recreation noise on both positive and negative affect (e.g., *less* decrease in positive affect and *less* increase in negative affect). **Table 1** displays the full model results for pro-motorized recreation attitudes as a moderating variable.

For positive affect, full model analyses of the pro-motorized recreation attitude data showed main effects for change in positive affect after exposure [$F(1,71) = 7.65$, $p = 0.007$, $\eta_p^2 = 0.097$], but no main effect for sound condition [$F(1,71) = 0.53$, $p = 0.469$]. However, a two-way interaction between affect and sound condition [$F(1,71) = 23.74$, $p < 0.001$, $\eta_p^2 = 0.251$] indicated that affect scores changed following exposure but differentially depending on sound condition and consistent with prior research. The interaction between change in positive affect and attitude score was also not significant [$F(1,71) = 1.04$, $p = 0.310$]. As hypothesized, the full three-way interaction between change in affect, sound condition, and pro-motorized recreation attitude score was significant indicating that attitude was moderating the previously observed effect that
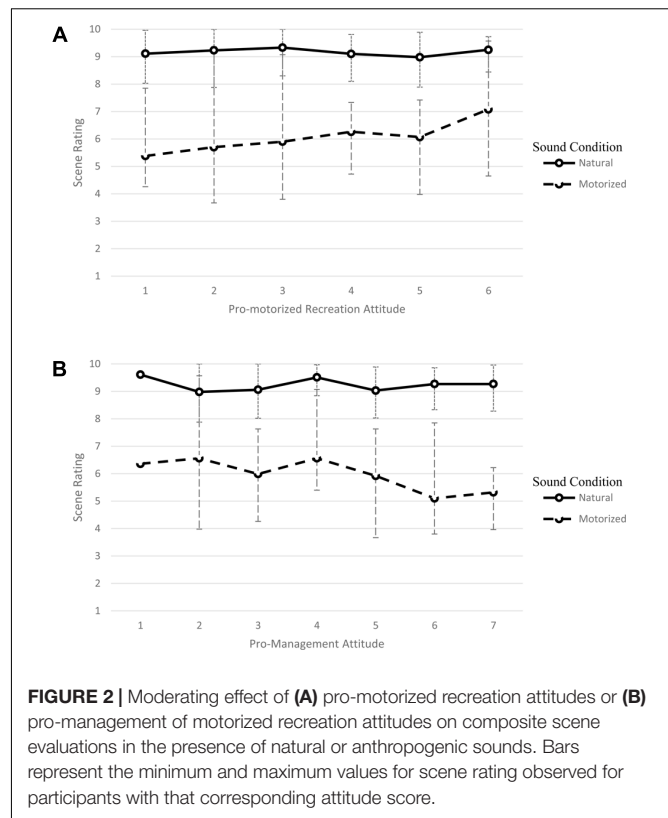
soundscape can have on positive affect, $F(1,71) = 7.60$, $p = 0.007$, $\eta_p^2 = 0.097$.

Follow-up analyses to unpack that three-way interaction were conducted (**Table 3**). Consistent with the hypothesis, pro-recreation attitudes were shown to significantly interact with positive affect ratings for the recreation noise condition ($F = 7.73$, $p = 0.009$, $\eta^2 = 0.181$) but not the natural sound condition ($F = 1.40$, $p = 0.244$). Those with higher pro-motorized recreation attitudes showed less decrease in positive affect following exposure to motorized recreation noise in the scene (**Figure 3A**).

Results for negative affect were similarly consistent with previous research and hypotheses (**Figure 3B**). There was a main effect for change in negative affect following exposure [$F(1,71) = 4.76$, $p = 0.032$] and for sound condition [$F(1,71) = 3.97$, $p = 0.050$]. Additional two-way interactions between change in affect and sound condition [$F(1,71) = 11.37$, $p = 0.001$] and change in affect and attitude [$F(1,71) = 6.49$, $p = 0.013$] were also shown to be significant. In those interactions, negative affect increased after exposure to the motorized sound but not natural sound.

The hypothesized three-way interaction between negative affect, sound condition, and attitude did not reach significance [$F(1,71) = 2.38$, $p = 0.128$]. However, separate R-ANCOVAs, similar to those performed to better understand the significant three-way interaction for positive affect, did show the same pattern of significant two-way interactions between pro-recreation attitude and negative affect ratings following

**TABLE 3 |** Results of repeated measures ANCOVAs for natural and motorized sound exposure conditions separately showing interactions between affect ratings and attitude scores.

| | | $F$ | $\eta_p^2$ |
|---|---|---|---|
| **Pro-recreation attitude: positive affect** | | | |
| Natural | Affect | 2.38 | 0.062 |
| | Attitude | 0.93 | 0.025 |
| | Affect × attitude | 1.40 | 0.037 |
| Motorized | Affect | 27.26** | 0.438 |
| | Attitude | 3.07† | 0.081 |
| | Affect × attitude | 7.73** | 0.181 |
| **Pro-recreation attitude: negative affect** | | | |
| Natural | Affect | 0.99 | 0.027 |
| | Attitude | 0.51 | 0.014 |
| | Affect × attitude | 0.61 | 0.017 |
| Motorized | Affect | 11.73** | 0.251 |
| | Attitude | 4.45* | 0.113 |
| | Affect × attitude | 7.36** | 0.174 |
| **Pro-management attitude: positive affect** | | | |
| Natural | Affect | 0.17 | 0.005 |
| | Attitude | 2.24 | 0.059 |
| | Affect × attitude | 0.00 | 0.000 |
| Motorized | Affect | 0.28 | 0.008 |
| | Attitude | 0.61 | 0.017 |
| | Affect × attitude | 8.55** | 0.196 |
| **Pro-management attitude: negative affect** | | | |
| Natural | Affect | 2.15 | 0.056 |
| | Attitude | 1.82 | 0.048 |
| | Affect × attitude | 0.05 | 0.001 |
| Motorized | Affect | 1.53 | 0.042 |
| | Attitude | 7.97** | 0.185 |
| | Affect × attitude | 4.63* | 0.117 |

†$p < 0.08$, *$p < 0.05$, **$p < 0.01$.

recreation noise ($F = 7.36$, $p = 0.010$, $\eta^2 = 0.174$) but not natural sounds ($F = 0.61$, $p = 0.438$). Thus, the primary analyses for negative affect failed to show the same hypothesized effect on the three-way interaction, but simpler analyses of the two-way interactions did mirror those shown for positive affect (**Table 3**).

## Pro-management

Pro-management attitudes were hypothesized to interact with both PANAS positive and negative affect scores but in a manner contrary to pro-recreation attitudes; it was hypothesized that greater pro-management attitudes would relate to larger affective change in the presence of motorized recreation noise but not in the presence of natural sounds (**Table 2**).

For positive affect, full model analyses of the pro-management attitude data showed no main effects for change in positive affect after exposure [$F(1,71) = 0.44$, $p = 0.508$], sound condition [$F(1,71) = 1.74$, $p = 0.192$], or the attitude moderator [$F(1,71) = 0.26$, $p = 0.612$]. However, a two-way interaction between affect and the attitude moderator [$F(1,71) = 4.53$, $p < 0.037$, $\eta_p^2 = 0.060$] indicated that affect scores changed following exposure but differentially depending on the attitude moderator. The interaction between change in positive affect

**FIGURE 3 |** Moderating effect of pro-motorized recreation attitudes on positive affect **(A)** and negative affect **(B)** or pro-management of motorized recreation attitudes on positive affect **(C)** and negative affect **(D)** in the presence of natural or anthropogenic sounds. Bars represent the minimum and maximum values for scene rating observed for participants with that corresponding attitude score.
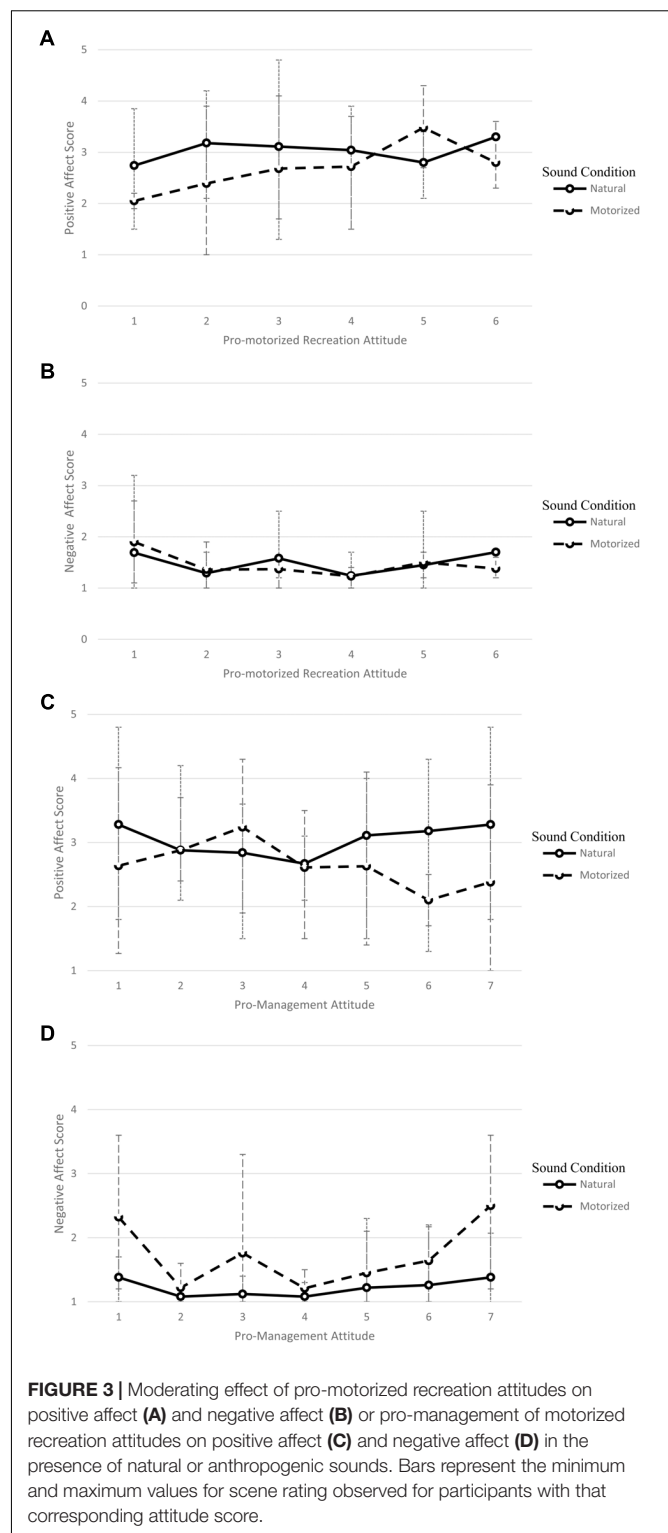
and sound condition was also not significant [$F(1,71) = 0.01$, $p = 0.943$]. However, as hypothesized, the full three-way interaction between change in affect, sound condition, and pro-management of motorized recreation attitude score was significant indicating that attitude was moderating the observed

effect that soundscape can have on positive affect, $F(1,71) = 4.22$, $p = 0.044$, $\eta_p^2 = 0.056$.

Follow-up results were consistent with the hypothesis for recreational noise exposure and pro-management attitudes (**Table 3**). When exposed to motorized recreation noise, a significant interaction between change in positive affect and the attitude moderator was present, $F(1,71) = 8.55$, $p < 0.001$, $\eta_p^2 = 0.196$. The observed decrease in positive affect following exposure to motorized recreation noise was larger for those with high pro-management attitudes (**Figure 3C**). The same interaction and moderating relationship was not present when participants were exposed to natural sounds, $F(1,71) = 0.00$, $p = 0.958$. In short, the moderating effect of pro-management attitudes on change in positive affect following exposure to motorized recreation noise was observed, as hypothesized.

For negative affect, the full model results were not supportive of hypotheses (**Figure 3D**). The main effect for the attitude moderator was the only main effect or interaction shown in the analysis, $F(1,71) = 9.31$, $p = 0.003$, $\eta_p^2 = 0.116$. Similar to the negative affect results for the pro-motorized recreation analyses, simple analyses with two R-ANCOVAs, one for each sound condition, suggested differential effects of the attitude moderator on change in negative affect in the hypothesized manner (i.e., attitude moderation for recreation noise exposure but not for natural sound exposure).

In summary, the hypothesized moderating relationship between changes in affect following sound exposure and the type of sound presented was present for both the pro-recreation and the pro-management attitude moderators, but only for positive affect. In the case of negative affect, the hypothesized interaction failed to reach statistical significance in the full models but demonstrated some support when analyzing each sound condition in isolation (**Table 3**).

## DISCUSSION

Previous research has consistently demonstrated a deleterious effect of anthropogenic noise on scenic evaluations and affect (e.g., Mace et al., 1999; Weinzimmer et al., 2014), but the examination of moderating variables within this type of soundscape assessment has been limited to situational characteristics such as the cause (Mace et al., 2003) or location (Benfield et al., 2010) of the noise. The current study demonstrated that the individual-level characteristics of attitudes toward motorized recreation noise and soundscape management can also affect the severity of anthropogenic noise-related outcomes in simulated natural recreation environments.

Specifically, the data presented show that pro-motorized recreation attitudes reduced the negative impact of motorized recreation noise on scenic evaluations and ratings of positive affective state. While the presence of the noise still reduced ratings of landscape quality and positive affect, the effect was much smaller for those high in pro-recreation attitudes compared to those with lower levels of the same attitude. The reverse was true of pro-soundscape management attitudes. The presence of motorized recreation noise was more problematic to those

holding high pro-management attitudes than for those with a lesser extent of these attitudes. Pro-management attitudes predicted lower scene evaluations in the presence of recreation noise and larger changes in positive affect.

The moderating role of pro-management and pro-motorized recreation attitudes on subsequent perception and evaluation of outdoor recreation and leisure environments under varied sound conditions has not been demonstrated in prior research, thus representing an important addition to our current understanding of how the objectively measurable soundscape and the subjectively experiencing user interact in natural environments. The connection between this set of findings and other research (e.g., Tarrant et al., 1995; Clement and Cheng, 2011) ties the emerging area of natural soundscapes to a larger literature on attitudes in recreation enjoyment and management which has a number of implications for outcomes-based soundscape assessment for both researchers and managers alike.

## Implications

Recreation area managers need to make informed policy and management decisions that impact a wide range of user groups and the current project can aid those management efforts. In this management context, the assessment of soundscape quality often relies of user outcomes relative to competing economic or public goals. For example, the U.S. National Park Service must balance visitor use, which in some locations includes natural soundscape altering, noise-producing motorized recreational activities, while at the same time preserving the natural and cultural resources within these protected areas for visitors to enjoy separately from motorized recreation. Thus, in order to provide quality soundscape experiences while protecting park resources and outside economic interests, it is imperative that managers understand not only the overall soundscape experience but also who visitors are, their motivations, expectations, and how they perceive other aspects of the park experience. Much of this can be determined by assessing visitor attitudes toward recreational settings and management actions (Manfredo et al., 2004), and this study has provided greater understanding of these factors.

By understanding attitudes and the role they play in outcomes-based soundscape assessment, managers can use informational messaging to strengthen attitudes that align with management objectives (e.g., protection of natural sounds) or alter attitudes that misalign with management goals. The extensive body of persuasion literature suggests that effective messaging design requires consideration of many variables (e.g., personal relevance, message source, and timing) that are thought to enhance and motivate understanding in order to alter attitudinal state (Petty and Cacioppo, 1986; Fishbein and Manfredo, 1992; Eagly and Chaiken, 1993; Perloff, 2003; Absher and Bright, 2004). Messages that have the most effect on attitudes contain substantial argument quality, which is thought to stimulate elaboration (Petty and Cacioppo, 1986; Petty and Wegener, 1998; Wood, 2000). In other words, researchers and managers, by relying on a substantial body of literature on persuasion and attitude change, would be potentially able to alter the subjective assessment of

a soundscape rather than altering the physical properties of the soundscape itself.

Interpretive strategists cannot reach or alter the attitudes of all visitors due to situational and/or personal variables, but developing messages that are strong, impactful, and relevant increases attitudinal change or strength (Ham, 2007). Attitudes that align with a message containing impactful arguments are thought to be strengthened, while misaligning attitudes may be altered, if a message enhances consideration and thought about a given topic (Lavine and Snyder, 1996; Wood, 2000; Ziegler et al., 2007; Petty and Wegener, 2008). With consideration of messaging strategies, these results suggest that specific messages emphasizing the impact of motorized recreation noise (e.g., disturbing wildlife or other visitors) could influence those individuals with pro-recreation attitudes to be more cognizant regarding the protection of natural sounds. Alternatively, those individuals with pro-management attitudes could experience strengthened attitudes by receiving this type of targeted message. Future laboratory and field studies related to messaging, perspective taking, or attitude change in the context of soundscape assessment would be justified given the current set of findings.

## Limitations

Participants in this study were not explicitly informed about the sources of noise that were presented within the soundscape, and the sources were not visible during the simulations. It is possible that some participants were not aware that they were hearing motorcycles, snowmobiles, and propeller planes, specifically. Furthermore, the noise sources were not clearly attributed to recreation activities, although other work suggests that attribution may not make a significant impact on assessment (e.g., Mace et al., 2003). Participants were not told that they would be hearing sounds from scenic air tours as opposed to commercial or park administrative flights (e.g., general maintenance or search and rescue operations), which could be perceived similarly but evaluated differently. While these factors can be considered limitations, it is quite possible that explicit attribution of the noise sources to recreational motorized activities would increase the magnitude of the observed effects reported above. For example, changes in landscape assessments and affective response could be underestimated for participants who hold strong attitudes about park management or motorized recreation, but who were not aware that they were hearing sounds generated by those activities.

However, other work has shown that laboratory-based soundscape assessments, particularly in the context of identification and representation, can lead to greater variability in reporting when compared to field-based soundscape assessments (e.g., Guastavino et al., 2005). Similarly, the evaluations taking place are derived from stimuli that is both visual and aural. As such, the landscape context, being natural and remote without the presence of built structures, informs expectations for those soundscapes and subsequent scenic assessments in their presence. Previous research has shown that such visual elements can impact noise and sound assessment, particularly on nature-relevant constructs such as those assessed in this study (e.g., Pheasant et al., 2008; Pheasant and Watts, 2015).

Another potential limitation of the study design relates to the lack of a direct link between recreation attitudes and participation in recreation activities. Participants were not asked to report if they had actually engaged in the motorized activities simulated in this study (or even if they intended to participate in the activities). Rather, they evaluated hypothetical scenarios of motorized vehicle use in national park settings. As discussed further in the next section, it would be informative to test actual members of motorized user groups, who are likely to have strong attitudes about motorized recreation in national parks to see if that indeed alters soundscape assessment. Similarly, the current sample consisted primarily of university students in natural resources classes. Based on their training, these students would be expected to exhibit a bias in favor of park management and resource protection. Natural resources students may also be more knowledgeable than typical park visitors about soundscape-related controversies, assessment strategies, and management objectives in protected areas. Finally, the age of participants may inform both their attitudes and overall response to motorized recreation noise. Younger people, such as the majority this sample, may have their hearing less impacted by external noise sources, view motorized recreation as more appealing, or have less experience in these types of environmental contexts.

Thus, it is reasonable to assume that a more representative sample of potential park visitors would demonstrate a wider range of attitudes and increase the external validity of the findings; however, the combination of laboratory and field-based methodologies offers the strongest approach for investigating the multi-dimensional impacts of motorized noise on visitors (Mace et al., 2013). The present study attempted to isolate through well-controlled experimental manipulation the individual psychological factors that moderate this outcomes-based assessment. The robustness of the observed effect should now be tested in other settings and with other samples of visitors, recreation managers, and non-visiting adults.

## Future Directions

In addition to research designed to address methodological limitations discussed above, the current study provides several avenues for additional research on soundscape assessment, generally. As mentioned previously, the connection between attitude and subsequent soundscape appreciation allows for the wealth of literature on attitude change and persuasion to be utilized as a mechanism for combating problematic noise and/or increasing enjoyment of unique or more pristine soundscapes. Such interventions would run counter to more physical properties-based soundscape assessments because they allow for altering outcomes without changing the actual stimuli.

Showing that individual attitudes can moderate the effect of soundscape on environmental assessment suggests that other individual features need to be more fully incorporated into soundscape assessment research and more fully considered when making management policy. While some work has been done with personality traits (e.g., Benfield et al., 2013), the same cannot be said of other individual visitor variables such as motivation. Research has already demonstrated that visitor motivations for quiet can alter the perceived acceptability of

anthropogenic sounds (Marin et al., 2011), so it seems highly probable that such a motivation for quiet would also affect ratings of scenes in the presence of sounds or changes in affective state caused by the presence of those sounds. Research focused on motivations should be conducted to confirm that connection between acceptability and subsequent changes in scene ratings or affect.

Additionally, little research has effectively demonstrated that anthropogenic noise, in the specific context of natural environments, alters physiological processes related to arousal or stress. Such effects have been demonstrated in wildlife (e.g., Barber et al., 2010), but the connection to park visitors experiencing sounds has not been shown. It may be possible that such physiological effects within outcomes-based soundscape assessment would be moderated by attitude given that some attitudes, such as being in favor of motorized recreation, related to more positive affective responses in the current study. Similarly, emerging research has demonstrated a restorative effect of natural soundscapes (Alvarsson et al., 2010; Benfield et al., 2014; Abbott et al., 2016), but has not examined whether individual attitude, or another variable such as motivation, could moderate that restorative effect.

In summary, the current study showed that outcomes-based soundscape assessment would benefit from additional reliance and focus on moderating variables. In this case, pro-motorized recreation and or management attitudes moderate a well-established set of findings within soundscape assessment research. Such an effect had not been previously shown and, more importantly, has several implications for both management policy and future research as it pertains to soundscape assessment. Based on the current findings, it is reasonable to predict that attitudes may moderate other soundscape-relevant effects, and that other characteristics, such as motivations, may be worth examining in the future and controlling for when making assessment of soundscape quality based on user perceptions, reported experiences, or outcomes.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the American Psychological Association with written informed consent from all subjects in accordance with the Declaration of Helsinki and the Belmont Committee Report. The protocol was approved by the Internal Review Board at the university conducting the research.

## AUTHOR CONTRIBUTIONS

JB conceptualized the RQ, conducted the analysis, and was the primary author of all content in the manuscript. BT contributed significantly to the literature review and analysis and was also the second highest contributor of manuscript content. DW conceptualized the original research project and collected the data. PN supervised DW's data collection and assisted with editing of later drafts of the manuscript.

## FUNDING

## REFERENCES

Abbott, L. C., Taff, B. D., Newman, P., Benfield, J. A., and Mowen, A. J. (2016). The influence of natural sounds on attention restoration. *J. Parks Recreation Adm.* 34, 5–15. doi: 10.18666/JPRA-2016-V34-I3-6893

Absher, J. D., and Bright, A. D. (2004). "Communication research in outdoor recreation and natural resources management," in *Society and Natural Resources: A Summary of Knowledge*, eds M. J. Manfredo, J. J. Vaske, B. L. Bruyere, D. R. Field, and P. J. Brown (Jefferson City, MO: Modern Litho), 117–126.

Alvarsson, J. J., Wiens, S., and Nilsson, M. E. (2010). Stress recovery during exposure to nature sound and environmental noise. *Int. J. Environ. Res. Public Health* 7, 1036–1046. doi: 10.3390/ijerph7031036

Barber, J. R., Crooks, K. R., and Fristrup, K. M. (2010). The costs of chronic noise exposure for terrestrial organisms. *Trends Ecol. Evol.* 25, 180–189. doi: 10.1016/j.tree.2009.08.002

Bell, P. A., Greene, T. C., Fisher, J. D., and Baum, A. (2001). *Environmental Psychology*, 5th Edn. London: Lawrence Erlbaum.

Benfield, J. A., Bell, P. A., Troup, L. J., and Soderstrom, N. C. (2010). Aesthetic and affective effects of vocal and traffic nose on natural landscape assessment. *J. Environ. Psychol.* 30, 103–111. doi: 10.1016/j.jenvp.2009.10.002

Benfield, J. A., Nurse, G. A., Jakubowski, R., Gibson, A. W., Taff, B. D., Newman, P., et al. (2013). Testing noise in the field: a brief measure of individual noise sensitivity. *Environ. Behav.* 46, 353–372. doi: 10.1177/0013916512454430

Benfield, J. A., Taff, B. D., Newman, P., and Smyth, J. (2014). Natural sound facilitates mood recovery. *Ecopsychology* 6, 183–188.

Botteldooren, D., De Coensel, B., and De Muer, T. (2006). The temporal structure of urban soundscapes. *J. Sound Vib.* 292, 105–123. doi: 10.1016/j.jsv.2005.07.026

Bright, A. D. (1997). Attitude strength and support of recreation management strategies. *J. Leis. Res.* 29, 363–379. doi: 10.1080/00222216.1997.11949804

Clement, J. M., and Cheng, A. S. (2011). Using analyses of public value orientations, attitudes and preferences to inform national forest planning in Colorado and Wyoming. *Appl. Geogr.* 31, 393–400. doi: 10.1016/j.apgeog.2010.10.001

Craig, A., Knox, D., and Moore, D. (2014). The perceived annoyance of urban soundscapes. *J. Acoust. Soc. Am.* 136, 2305–2305. doi: 10.1121/1.4900339

De Coensel, B., and Botteldooren, D. (2006). The quiet rural soundscape and how to characterize it. *Acta Acust. United Acust.* 92, 887–897.

Directive 2002/49/EC of the European Parliament (2017). *Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 Relating to the Assessment and Management of Environmental Noise.* Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0049:en:NOT [accessed November 22, 2017].

Driver, B. L., Nash, R., and Haas, G. (1987). "Wilderness benefits: a state-of-knowledge view," in *Proceedings of the National Wilderness Research Conference: Issues, State of Knowledge, Future Directions*, Ogden, UT, 294–319.

Eagly, A. H., and Chaiken, S. (1993). *The Psychology of Attitudes.* Fort Worth, TX: Harcourt Brace Jovanovich.

Fishbein, M., and Manfredo, M. J. (1992). "A theory of behavior change," in *Influencing Human Behavior: Theory and Application in Recreation, Tourism, and Natural Resources Management*, ed. M. J. Manfredo (Champaign, IL: Sagamore Publishing Inc.).

Gramann, J. (1999). The effect of mechanical noise and natural sound on visitor experiences in units of the National Park System. *Soc. Sci. Res. Rev.* 1, 1–16.

Guastavino, C., Katz, B. F., Polack, J. D., Levitin, D. J., and Dubois, D. (2005). Ecological validity of soundscape reproduction. *Acta Acust. United Acust.* 91, 333–341.

Ham, S. H. (2007). From interpretation to protection: Is there a theoretical basis? *J. Assoc. Herit. Interpret.* 12, 20–23.

Jeon, J. Y., Lee, P. J., You, J., and Kang, J. (2010). Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds. *J. Acoust. Soc. Am.* 127, 1357–1366. doi: 10.1121/1.3298437

Jiang, J., Zhang, J., Zhang, H., and Yan, B. (2017). Natural soundscapes and tourist loyalty to nature-based tourism destinations: the mediating effect of tourist satisfaction. *J. Travel Tourism Mark.* 35, 162–177.

Lai, P.-H., Sorice, M. G., Nepal, S. K., and Cheng, C.-K. (2009). Integrating social marketing into sustainable resource management at Padre Island National Seashore: an attitude-based segmentation approach. *Environ. Manage.* 43, 985–998. doi: 10.1007/s00267-009-9293-9

Lam, K.-C., Chan, P.-K., Chan, T.-C., Au, W.-H., and Hui, W.-C. (2009). Annoyance response to mixed transportation noise in Hong Kong. *Appl. Acoust.* 70, 1–10. doi: 10.1016/j.apacoust.2008.02.005

Lavine, H., and Snyder, M. (1996). Cognitive processing and the functional matching effect in persuasion: the mediating role of subjective perceptions of message quality. *J. Exp. Soc. Psychol.* 32, 580–604. doi: 10.1006/jesp.1996.0026

Mace, B. L., Bell, P. A., and Loomis, R. J. (1999). Aesthetic, affective, and cognitive effects of noise on natural landscape assessment. *Soc. Nat. Resour.* 12, 225–242. doi: 10.1080/089419299279713

Mace, B. L., Bell, P. A., Loomis, R. J., and Haas, G. E. (2003). Source attribution of helicopter noise in pristine national park landscapes. *J. Park Recreation Adm.* 21, 97–119.

Mace, B. L., Corser, G. C., Zitting, L., and Denison, J. (2013). Effects of overflights on the national park experience. *J. Environ. Psychol.* 35, 30–39. doi: 10.1016/j.jenvp.2013.04.001

Manfredo, M. J. (2008). *Who Cares about Wildlife: Social Science Concepts for Exploring Human-wildlife Relationships and Conservation Issues.* New York, NY: Springer Press. doi: 10.1007/978-0-387-77040-6

Manfredo, M. J., Teel, T. L., and Bright, A. D. (2004). "Applications of the concepts of values and attitudes in human dimensions of natural resources research," in *Society and Natural Resources: A Summary of Knowledge*, eds M. J. Manfredo, J. J. Vaske, B. L. Bruyere, D. R. Field, and P. Brown (Jefferson City, MO: Modern Litho), 271–282.

Manning, R. E. (2007). *Studies in Outdoor Recreation: Search and Research for Satisfaction*, 3rd Edn. Corvallis, OR: Oregon State University Press.

Marin, L. D., Newman, P., Manning, R., Vaske, J. J., and Stack, D. (2011). Motivation and acceptability norms of human-caused sound in Muir Woods National Monument. *Leis. Sci.* 33, 147–161. doi: 10.1080/01490400.2011.550224

McDonald, C. D., Baumgartner, R. M., and Iachan, R. (1995). *National Park Service Aircraft Management Studies.* USDI Report No. 94-2. Denver, CO: National Park Service.

Monz, C. A. (2009). Climbers' attitudes toward recreation resource impacts in the Adirondack Park's Giant Mountain Wilderness. *Int. J. Wilderness* 15, 26–33.

National Park Service [NPS] (2000). *Director's Order #47: Soundscape Preservation and Noise Management.* Available at: http://www.nps.gov/policy/DOrders/DOrder47.html [accessed November 24, 2011].

National Park Service [NPS] (2006). *National Park Service Management Policies.* Washington, DC: U.S. Government Printing Office.

Newman, P., Manning, R., and Trevino, K. (2010). From landscapes to soundscapes: introduction to the special issue. *Park Sci.* 26, 2–5.

Payne, S. R. (2008). Are perceived soundscapes within urban parks restorative? *J. Acoust. Soc. Am.* 123, 5519–5524. doi: 10.1121/1.2935525

Pedersen, E., Berg, F., Bakker, R., and Bouma, J. (2009). Response to noise from modern wind farms in the Netherlands. *J. Acoust. Soc. Am.* 126, 634–643. doi: 10.1121/1.3160293

Perloff, R. M. (2003). *The Dynamics of Persuasion: Communication and Attitudes in the 21st Century*, 2nd Edn. Mahwah, NJ: Lawrence Erlbaum Associates.

Petty, R. E., and Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change.* New York, NY: Springer-Verlag. doi: 10.1007/978-1-4612-4964-1

Petty, R. E., and Wegener, D. (2008). "Matching versus mismatching attitude functions: implications for scrutiny of persuasive messages," in *Attitudes Their Structure, Function and Consequences: Key Readings in Social Psychology*, eds R. H. Fazio and R. E. Petty (New York, NY: Psychology Press).

Petty, R. E., and Wegener, D. T. (1998). Matching versus mismatching attitude functions: implications for scrutiny of persuasive messages. *Pers. Soc. Psychol. Bull.* 24, 227–240. doi: 10.1177/0146167298243001

Pheasant, R. J., Horoshenkov, K., Watts, G., and Barret, B. T. (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

Pheasant, R. J., and Watts, G. R. (2015). Towards predicting wildness in the United Kingdom. *Landsc. Urban Plan.* 133, 87–97. doi: 10.1016/j.landurbplan.2014.09.009

Pilcher, E. J., Newman, P., and Manning, R. E. (2009). Understanding and managing experimental aspects of soundscapes at Muir Woods National Monument. *Environ. Manage.* 43, 425–435. doi: 10.1007/s00267-008-9224-1

Reed, S. E., Boggs, J. L., and Mann, J. P. (2012). A GIS tool for modeling anthropogenic noise propagation in natural ecosystems. *Environ. Modell. Softw.* 37, 1–5. doi: 10.1016/j.envsoft.2012.04.012

Shelby, B., and Heberlein, T. A. (1986). *Carrying Capacity in Recreational Settings.* Corvallis, OR: Oregon State University Press.

Stack, D. W., Peter, N., Manning, R. E., and Fristrup, K. M. (2011). Reducing visitor noise levels at Muir Woods National Monument using experimental management. *J. Acoust. Soc. Am.* 129, 1375–1380. doi: 10.1121/1.3531803

Taff, B. D., Newman, P., Pettebone, D., White, D. D., Lawson, S. R., Monz, C., et al. (2013). Dimensions of alternative transportation experience in Yosemite and Rocky Mountain National Parks. *J. Transp. Geogr.* 30, 37–46. doi: 10.1016/j.jtrangeo.2013.02.010

Taff, D., Newman, P., Lawson, S. R., Bright, A., Marin, L., Gibson, A., et al. (2014). The role of messaging on acceptability of military aircraft sounds in Sequoia National Park. *Appl. Acoust.* 84, 122–128. doi: 10.1016/j.apacoust.2013.09.012

Tarrant, M. A., Haas, G. E., and Manfredo, M. J. (1995). Factors affecting visitor evaluations of aircraft overflights of wilderness areas. *Soc. Nat. Resour.* 8, 351–360. doi: 10.1080/08941929509380927

Vaske, J. J. (2008). *Survey Research and Analysis: Applications in Parks, Recreation and Human Dimensions.* State College, PA: Venture Publishing.

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Weinzimmer, D., Newman, P., Taff, D., Benfield, J., Lynch, E., and Bell, P. (2014). Human responses to simulated motorized noise in national parks. *Leis. Sci.* 36, 251–267. doi: 10.1080/01490400.2014.888022

White, D. D., Aquino, J. F., Budruk, M., and Golub, A. (2011). Visitors' experiences of traditional and alternative transportation in Yosemite National Park. *J. Park Recreation Adm.* 29, 38–57.

Wood, W. (2000). Attitude change: persuasion and social influence. *Annu. Rev. Psychol.* 51, 539–570. doi: 10.1146/annurev.psych.51.1.539

Ziegler, R., Dobre, B., and Diehl, M. (2007). Does matching versus mismatching message content to attitude functions lead to biased processing? The role of message ambiguity. *Basic Appl. Soc. Psychol.* 29, 269–278. doi: 10.1080/01973530701503366

Check for updates

# Influence of Music on the Behaviors of Crowd in Urban Open Public Spaces

Qi Meng[1,2], Tingting Zhao[1] and Jian Kang[1,2]*

[1] Heilongjiang Cold Region Architectural Science Key Laboratory, School of Architecture, Harbin Institute of Technology, Harbin, China, [2] UCL Institute for Environmental Design and Engineering, University College London, London, United Kingdom

Sound environment plays an important role in urban open spaces, yet studies on the effects of perception of the sound environment on crowd behaviors have been limited. The aim of this study, therefore, is to explore how music, which is considered an important soundscape element, affects crowd behaviors in urban open spaces. On-site observations were performed at a 100 m × 70 m urban leisure square in Harbin, China. Typical music was used to study the effects of perception of the sound environment on crowd behaviors; then, these behaviors were classified into movement (passing by and walking around) and non-movement behaviors (sitting). The results show that the path of passing by in an urban leisure square with music was more centralized than without music. Without music, 8.3% of people passing by walked near the edge of the square, whereas with music, this percentage was zero. In terms of the speed of passing by behavior, no significant difference was observed with the presence or absence of background music. Regarding the effect of music on walking around behavior in the square, the mean area and perimeter when background music was played were smaller than without background music. The mean speed of those exhibiting walking around behavior with background music in the square was 0.296 m/s slower than when no background music was played. For those exhibiting sitting behavior, when background music was not present, crowd density showed no variation based on the distance from the sound source. When music was present, it was observed that as the distance from the sound source increased, crowd density of those sitting behavior decreased accordingly.

Keywords: music, crowd behavior, movement, soundscape, urban open space

## INTRODUCTION

The phrase "urban open space" can describe many types of open areas (Marcus and Francis, 1998). One definition holds that, as the counterpart of development, urban open space is a natural and cultural resource, synonymous with neither unused land nor park and recreation areas. Another definition is that open space is land and/or water area with its surface open to the sky which has been consciously acquired or publicly regulated to serve conservation and urban shaping functions in addition to providing recreational opportunities (Myers, 1975; Thompson, 2002). In modern cities, the benefits that urban open space provides to citizens can be separated into three basic categories: recreation, ecology, and esthetic value (Brander and Koetse, 2011). Sound quality is considered to be a key part of the ecological/sustainable development of urban open spaces (Zhang et al., 2006). However, the sound environment of urban open

spaces is often not satisfactory because of a lack of consideration for human behavior during the planning and managing of the spaces (Meng and Kang, 2016). Therefore, research on the effect of perception of the sound environment on crowd behavior will be of importance to landscape research in this field. According to the International Standards Organization, a "soundscape" is defined as an acoustic environment as perceived/experienced in context (ISO 12913-1, 2014). Behavior comes into play in soundscape assessment in that the activities and behaviors of surrounding people form a key facet of context.

Individual behaviors generally refer to the attitude or performance of a person in certain situations; their actions can be largely random, subject to the effect of the environment (Jia, 2012). In contrast, crowd behaviors refer to the attitude or performance of a crowd in an environment; they can be composed of certain regularities, subject to the effect of the environment (Yuan and Tan, 2011; Xie et al., 2013). Thus, instead of individual behaviors, crowd behaviors are usually examined in studies on urban open spaces (Marušić, 2011; Lepore et al., 2016; Meng and Kang, 2016). Lewin et al. (1936) presented the following formula, which indicates the interaction between an individual and his/her environment: B = f (P, E), in which B represents behavior; P represents persons, including individuals and groups; and E represents the environment in which those persons live. Based on this formula, both users' social characteristics and local environment must be considered in human behavior studies. Previous studies pointed out that recreational behavior can be affected by users' cultural background, age, and different local areas (Floyd, 1998; Payne et al., 2002; Guéguen et al., 2008).

Many different aspects of crowd behavior can be examined to draw conclusions. These can include characteristics of behaviors, such as movement and action (Wang, 2014); characteristics of movement, for example, characterized as movement or non-movement, with the former including passing by and walking around and the latter including sitting (Chen, 2009); and characteristics of actions, such as sitting, standing, watching, and loitering (Lepore et al., 2016). The number of participants is also an important factor, that is, whether the behavior involves one person, two people, or multiple people (Jia, 2012); the intrinsic properties of such behaviors can also be examined, such as whether they are necessary, spontaneous, or social (Gehl, 1987). The frequency and location of the behavior are also important, such as whether it is neighborhood or urban behavior (Chen, 2009). Additionally, factors such as crowd behavior in the sound environment, participation behavior, tendency behavior, avoiding behavior, and other behaviors which are not affected by the environment are all significant angles that reveal crucial information regarding crowd behaviors (Jia, 2012).

The sound environment can affect human perception, and human perception can influence crowd behavior in both indoor and outdoor spaces. For example, previous studies have demonstrated that environmental music affects the pace of shopping and amount of time spent in shopping malls (Milliman, 1982). Other studies have also shown that eating and talking behavior can be affected by background music in dining spaces (Fiegel et al., 2014; Meng et al., 2017a). In urban open spaces,

studies have found that people who pass by will stop to stand and watch music-related activities, whereas the amount of exercising behavior will be changed a little by music-related activities (Meng and Kang, 2016). Another study indicated that the presence of music can prolong the duration of stay in a tunnel when compared with silence, and classical music caused the longest duration of stay (Aletta et al., 2016). It has also been found that in the case of sound stimulation in the audio–visual environment of the countryside, study participants' gazing range was demonstrated to be significantly more dispersed than when no sound stimulation was present (Ren and Kang, 2015). Previous studies have mainly focused on the effect of the sound environment on one's action (Zakariya et al., 2014; Aletta et al., 2016; Lepore et al., 2016). However, studies on the effects of certain typical sound sources on crowd behaviors classified as movement or non-movement have been limited.

Musical sound is a common sound source in urban open spaces (North et al., 2004; Styns et al., 2007). Studies have shown that when people listen to music, their emotions fluctuate, and the effect is to change their behavior (Orr et al., 1998). Studies have shown that different languages, tempos, tones, and sound levels of music can cause different effects on emotions, mental activities, and physical reactions. Overall, languages and tempos are the two most important factors (Sakharov et al., 2005; Carpentier and Potter, 2007). Other studies have found that fast music is associated with more activation than slow music (Gomez and Danuser, 2004; Natarajan et al., 2004). For example, a study researching participants with headphones found that fast music increases walking speed, while slow music causes slower walking speeds (Franěk et al., 2014).

The cited studies indicate that the sound environment can affect crowd behaviors; building on this finding, the present research focuses on the effects of music, an important soundscape element, on specific crowd behaviors, classified as movement, including passing by and walking around, or non-movement, including sitting. Previous studies indicate that path and speed are significant characteristics that describe movement behavior, while crowd density is important in describing non-movement behavior (Ye et al., 2012; Lavia et al., 2016). Therefore, the aims of this study are to find out: (1) whether music can change the path or speed of passing by or walking around behavior; we hypothesize that the speed of passing by or walking around behavior will increase with music, the path of passing by behavior will shift closer to the music, and the area or perimeter of walking around behavior will decrease with the music, since some previous studies have pointed out that music-related activities can increase the speed of passing by or walking around behavior in some urban open spaces and (2) whether music can decrease or increase sitting behaviors in urban open spaces; we hypothesize that the sitting behavior will increase with decreasing distance of music, since eating and talking behaviors can be affected by music. An urban leisure square was chosen as the case site, and music was chosen. In addition, three typical behaviors were selected for further analysis at the case site, and on-site observations were used for data collection. To achieve the aim of the study, several different approaches were explored. First, this study examined the effect of music on the path and speed of

passing by behavior. Second, it determined the effect of music on the path and speed of walking around behavior. Third, it observed the effect of music on the location of sitting behavior.

## MATERIALS AND METHODS

### Survey Site

Previous studies have indicated that an urban street is a kind of linear space, where people have little choice but for their paths to be confined to the pavement by buildings or motor vehicles (Hwang et al., 2011). In contrast, an urban square is an "areal" type of space, where people are free to choose their direction and path of travel (Marcus and Francis, 1998; Zakariya et al., 2014). In this study, a typical urban leisure square named "LANDSCAPE" square, located in Harbin, in Northeast China, was selected as the case site. Maps of the square and survey site can be found in **Figure 1**.

This site was chosen for the following three reasons. First, it is located at the crossing of Changjiang Road and Hongxiang Road; roads are present on three sides of the square, and vegetation is found on the fourth side. The square's surrounding environment (determining its scale and format) resembles that of squares encountered in Europe and Japan (Ashihara, 1985; Whitlock, 2004), and thus has typicality. Next, the urban leisure square is nearly 100 m long and 70 m wide and covers an area of about 7000 m$^2$, which is typical for modern cities (Dai, 2014). Finally, the square has various sports facilities and lush vegetation; thus, a large number of local residents visit the square and it is an important urban open space in this area for people to relax and interact. Overall, the square provides convenient conditions and the opportunity to gather a large number of samples to study crowd movement behaviors as well as non-movement behaviors.

Studies indicate that environmental changes, such as changes in temperature and humidity, influence subjective acoustic perception (Thwaites et al., 2005; dela Fuente de Val et al., 2006). To avoid the effects of these environmental factors, measurements were performed on workdays in May and September 2017. The mean monthly temperatures (18–26°C) and the relative humidities of these months are approximately the same. Other studies indicate that illuminance, which may also affect the perception or behaviors of users, changes with the time of day (Liu et al., 2013; Meng et al., 2017b). Thus, in the present study, measurements were performed between 9:00–11:00 a.m. and 14:00–16:00 p.m. daily on workdays to avoid the effects of time of day on the light environment (Meng et al., 2017a).

### Sound Source

In previous music studies, musical sound can be classified by sound level, tempo, genre, context, familiarity, and so on (Husain et al., 2002; Sakharov et al., 2005; Kang, 2017). Based on the method used in the experiment by Lavia et al. (2016), the music excerpts were designed to be "inclusive," "non-aversive," and sound good in a highly reverberant environment. In this study, a typically familiar pop song with lyrics, named "Free to Fly," was selected as the stimulus for intervention in the acoustic environment of the square; the tempo of this song is 120 bpm.

A loudspeaker was used as the sound source; its location is shown in **Figure 1C** (S means sound source). The loudspeaker location was chosen for the following three reasons. First, the music played by the loudspeaker can be clearly heard at any point in the square. Second, the distance between the loudspeaker and the walls and other major reflective surfaces was ensured to be at least 20 m (Zahorik, 2002). Finally, to avoid any influence caused by the visual presence of the loudspeaker, it was placed near the water feature fence to avoid identification. During the experiment, the musical excerpt and silence were reproduced cyclically (Husain et al., 2002; Carpentier and Potter, 2007). The sound level was 88–90 dBA, exceeding background sound level.

### Measurement of Sound Environment

Previous studies indicate that acoustic perception of urban open spaces can be affected by sound pressure level (Yu and Kang, 2010; Xie et al., 2012). Since the measurement time for the current study was 9:00–11:00 a.m. and 14:00–16:00 p.m., its crowd density was less than 0.05/m$^2$; thus, the influence of the number of people on the square on the acoustic environment could be ignored (Meng and Kang, 2015). Therefore, acoustic environmental measurement was carried out point by point, not simultaneously.

To measure the sound environment, the area was divided into 6 m × 6 m units (Li and Meng, 2015). The equivalent continuous A-weighted sound pressure level ($L_{Aeq}$) was immediately recorded using an 801 sound-level meter after each observation was completed. During the measurement, the sound-level meter was adjusted to the slow speed (Kang and Zhang, 2010). Additionally, the distance between the measurement location and walls and other major reflective surfaces was ensured to be at least 1 m, and the distance between the measurement location and the ground was 1.2–1.5 m (Barron and Foulkes, 1994; Zahorik, 2002). One measurement was performed every 10 s. The data for each location were recorded for 5 min. A mean value was calculated to obtain the corresponding $L_{Aeq}$ (Zhang et al., 2016).

The sound field in the urban open space can be seen in **Figure 2** (S means sound source). When there was no music sound source in the square, the background sound pressure level was 56.7 dB. When music was present, the sound pressure level at 2 m away from the music sound source was 88.3, or 31.6 dB higher than that without music. With the effect of musical sound source considered, the equivalent A-weighted sound pressure level reduced constantly with increasing distance in the square. From 2 to 12 m away from the sound source, the sound pressure level reduced at 18.3 dB; from 12 to 24 m away from the sound source, the sound pressure level reduced at 6.4 dB; and from 24 to 36 m away from the sound source, the sound pressure level reduced at 3.4 dB. The attenuation degree of the equivalent A-weighted sound level was mainly determined by the degree of enclosure for the space.

### Observation of Crowd Behaviors

In previous soundscape research, the investigation methods were classified as questionnaires and observations (Meng and Kang, 2016; Meng et al., 2017a). The questionnaires
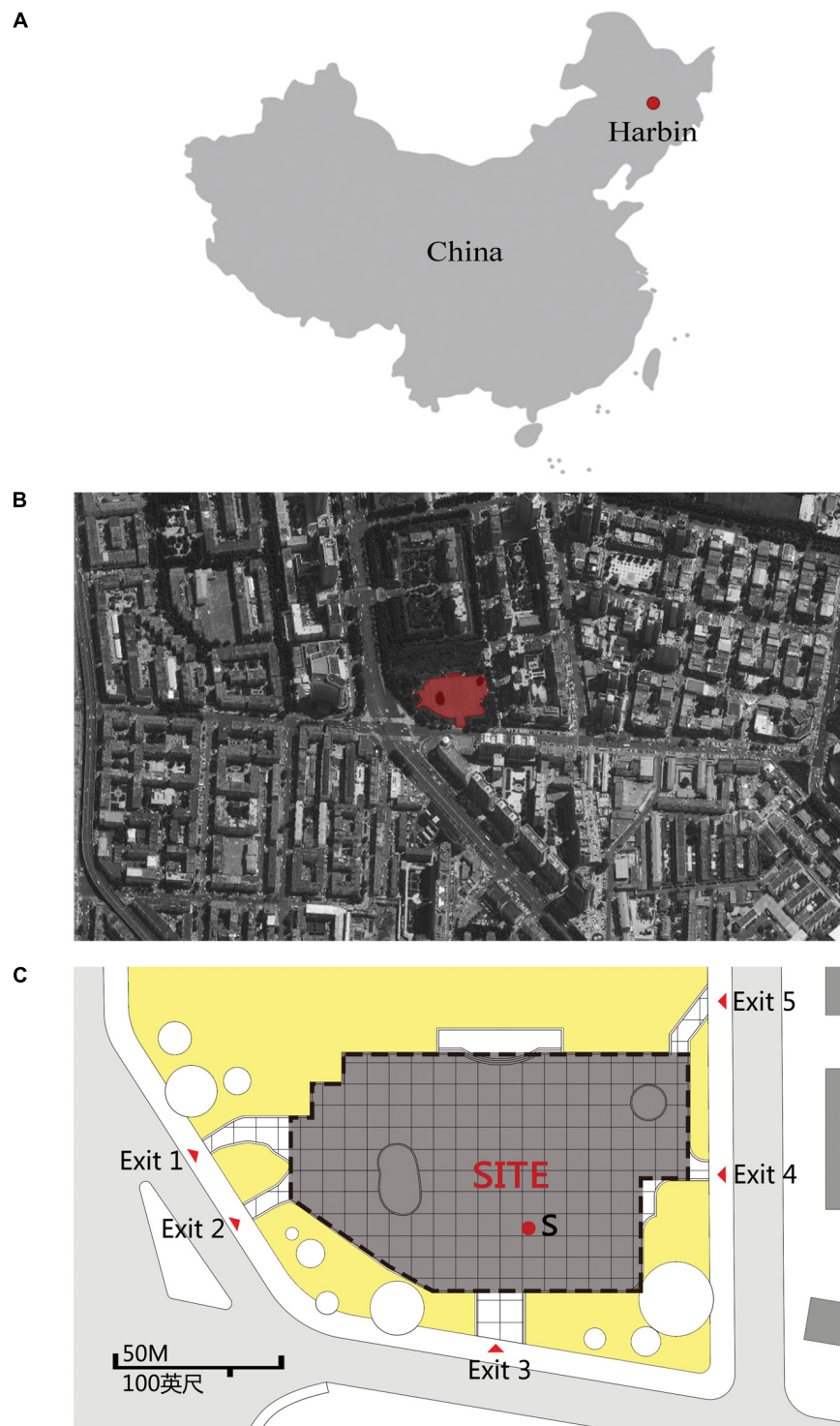
**FIGURE 1 |** The map of LANDSCAPE square and the case site: **(A)** location of LANDSCAPE square, **(B)** map of LANDSCAPE square, and **(C)** the survey site.

mainly focused on subjective evaluation indexes such as sound comfort, subjective loudness, and sound preference. This study involves measurements of crowd movement and non-movement behaviors, including path, speed, location of the stop points,

etc., which is difficult to assess by questionnaire interview; thus, observation was the method used. To avoid any biases in the observation process, this study used an unmanned aerial vehicle (UAV; Oakes and North, 2008); the UAV was flown at a height

**FIGURE 2 |** The distribution of sound pressure level in the case site (S means sound source): **(A)** with music and **(B)** without music.
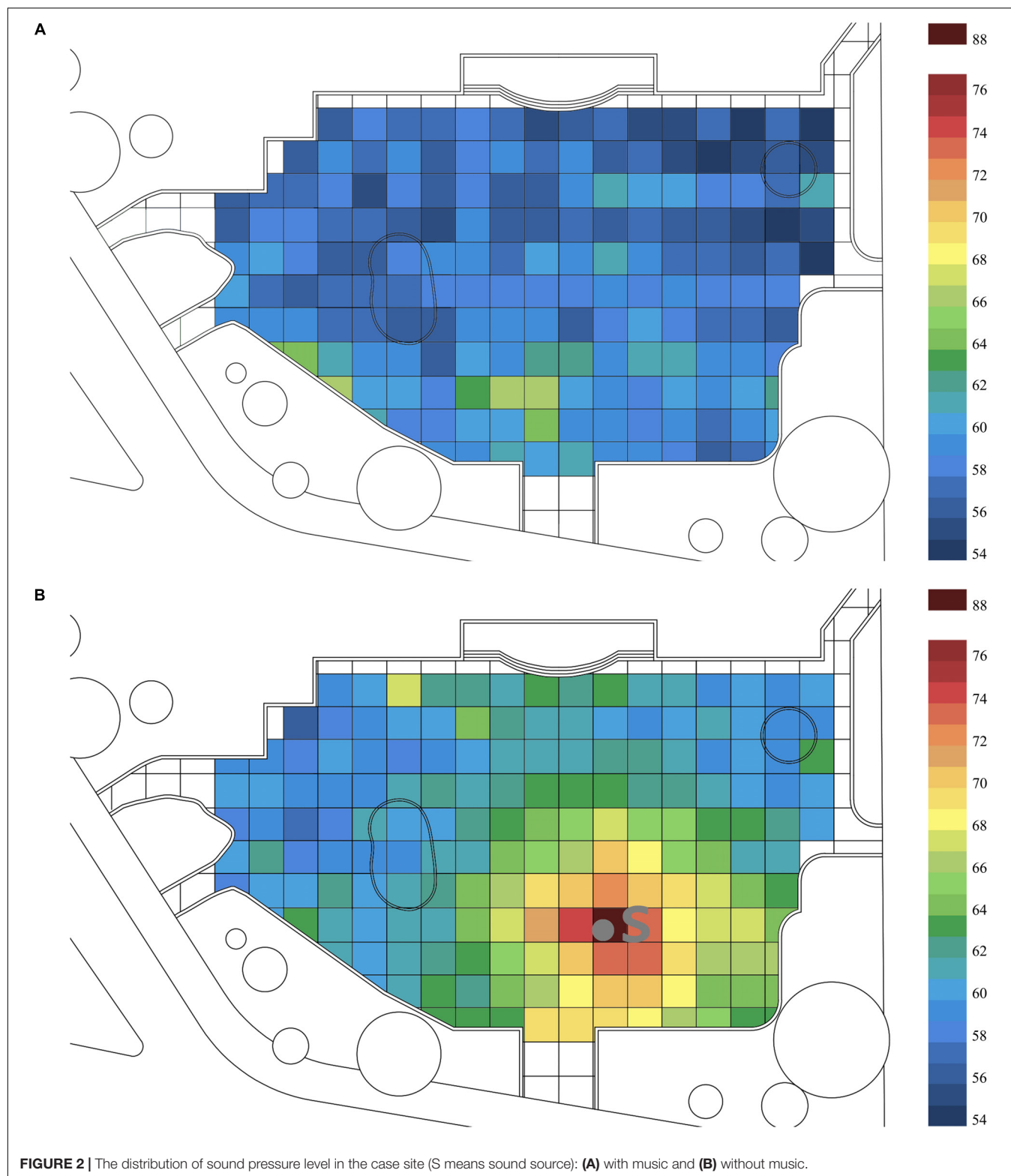
of 100 m, since at that height, people on the square could not hear the noise of the UAV (Sinibaldi and Marino, 2013). The observations were made under completely natural conditions, and since the subjects generally did not know that they were being observed, their behaviors were genuine; thus, the results were more reliable.

Each video shot by the UAV lasted 20–25 min. Videos of 10–15 groups for every situation were shot to ensure stochastic behavior

in the measurement (Meng and Kang, 2013). Meanwhile, one photograph was taken every 10 s. The behaviors of the subjects were then classified and analyzed statistically, based on the review of videos and photos in the laboratory.

## Statistical Analysis of Crowd Behaviors

In this study, different samples were used for different behaviors, to capture the different times of collection of the behaviors. For instance, the period needed to collect two samples of passing by behaviors was the same as that for three samples of walking around behavior and of five samples for sitting behavior. In all, 51 samples were collected for passing by behavior: 26 samples (12 males and 14 females) without music and 25 (13 males and 12 females) with music; 84 samples were collected for walking around behavior, of which 43 samples (20 males and 23 females) were without music and 41 samples (20 males and 21 females) with music; and 123 samples were collected for sitting behavior, 63 without music and 60 with music. In preliminary study, it was found that proportions of males and females engaging in given behaviors at the case site are generally equal (Meng et al., 2017a). In order to use the $T$-test to compare the samples, therefore, 24 samples (12 males and 12 females) without music and 24 samples (12 males and 12 females) with music were randomly selected for passing by behavior, 40 samples (20 males and 20 females) with music and 40 samples (20 males and 20 females) without music for walking around behavior, and 60 samples without music and 60 with music for sitting behavior.

In the present study, the power analysis was used to test sample sizes (Carpentier and Potter, 2007). The results showed that the power of samples for passing by behavior is 0.60, $p = 0.04$ with effect size 0.6; for walking around behavior, power is 0.77, $p = 0.03$, with effect size 0.6; and for sitting behavior, power is 0.87, $p = 0.01$ with effect size 0.6. This indicates that all samples were sufficient.

### The Path and Speed of Passing by Behaviors

To study the effect of music on the path and speed of passing by behaviors in urban open spaces, 48 samples, including 24 without music and 24 with, were selected for observation from the videos shot by the UAV. The locations of the entrances and exits are shown in **Figure 3**. In previous studies, the path was represented by a set of dots, and each dot was considered a relatively independent process (Ye et al., 2012). The entire process of passing by behavior was thus viewed as a collection of data flows between the many dots. As revealed in **Figure 3A** using passing by behavior as an example, the path points were labeled with round dots representing the subject's position as observed every 10 s in the photography taken by the UAV. Thus, as **Figure 3B** shows, the path was in turn conceived by connecting all of the points.

The calculation process of the mean speed was as follows (Marušić, 2011; Ye et al., 2012):

$$Vn = \Delta Ln / \Delta Tn$$

where $\Delta Ln$ is the distance of dot Cn and dot Cn+1, $\Delta Tn$ is the time of dot Cn and dot Cn+1, $Vn$ is the mean speed of distance
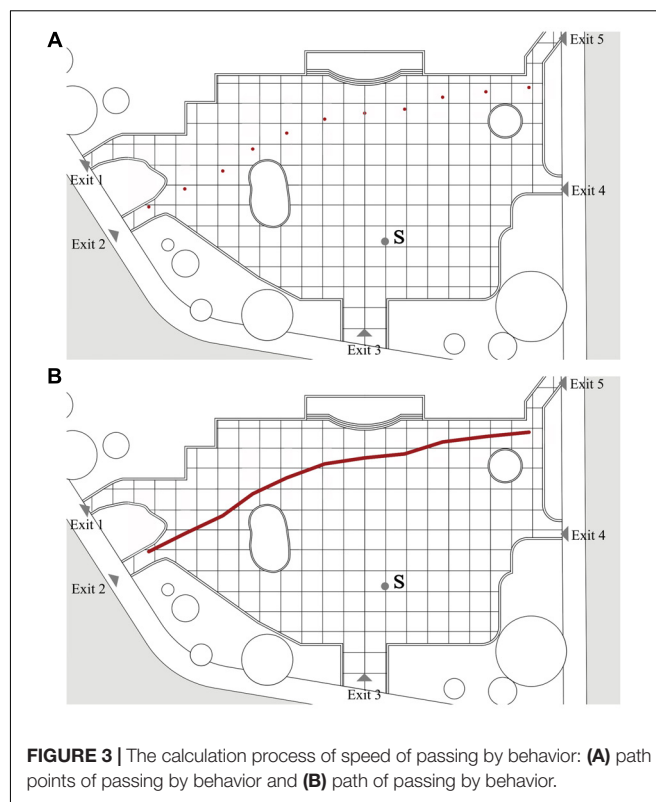


**FIGURE 3 |** The calculation process of speed of passing by behavior: **(A)** path points of passing by behavior and **(B)** path of passing by behavior.

of dot Cn and dot Cn+1, and mean speed: $\Delta V = (V1 + V2 + V3. \ldots. + Vn)/n$.

The mean speed was a superimposition of data from 24 samples (with or without music). The unit was m/s.

### The Path and Speed of Walking Around Behaviors

To study the effect of music on the path and speed of walking around behaviors in urban open spaces, 80 samples, including 40 samples without music and 40 samples with music, were selected for observation from the videos shot by the UAV. Based on the observations, the paths of walking around behavior were classified into four types. The data from each walking behavior were calculated as the mean of five occurrences, since previous studies indicate that the error of the mean of more than or equal to five times could be ignored. The calculation process for mean speed of walking behavior was the same as with passing by behavior. The results were the superimposition of data from 10 samples for each kind of path. The units used included m$^2$ (area), m (perimeter), and m/s (speed).

### The Crowd Density of Sitting Behaviors

To study the effect of music on the crowd density of a non-movement behavior in urban open spaces, crowd location was measured using the same photography method. Using sitting behavior as an example, one photograph shot by the UAV was selected every 2 min (Westover, 1989; Meng and Kang, 2015). In the laboratory, the locations of the crowd in the picture were labeled with round dots, and a 6 m × 6 m grid was used. The value obtained was divided by the measurement area to

determine a mean value of crowd density as the average number of persons per square meter. The unit used for measurement was persons/m². A total of 60 samples with music and 60 samples without music were used. The unit used for measurement was persons/m² (Meng et al., 2017b).

## RESULTS

## Effects of Music on Movement Behavior: Passing by Behavior

### Path

This section addresses the effects of music on the path of passing by behavior, which is shown in **Figure 4**, both with and without background music; the squares with different colors in **Figures 4A,B** indicate the numbers of users passing by, from 0 to 24 persons.

#### Without background music

As **Figure 4A** shows, 79.2% of people with passing by behavior selected a relatively short walking path. One possible reason for this is that when moving with a clear goal, passers-by often tended to choose the shortest path. This is usually a straight line approximately toward the goal, unless there is an obstacle (Gehl, 1987; Chen, 2009). It was found that 20.8% of people engaging in passing by behavior selected a relatively longer walking path, and even that 8.3% of them walked near the edge of the square. The areas covered by passing by behaviors were approximately 1436 m².

#### With background music

As **Figure 4B** shows, 0% of people with passing by behavior walked near the edge of the square, and 87.5% of them selected the relatively shorter path to walk; 12.5% of people with passing by behavior passed by the square while walking close to the sound source. The areas covered by passing by behaviors were approximately 876 m².

The area of passing by behavior, both with and without background music, is marked in **Figure 4C**, for the number of persons per grid without music, and **Figure 4D**, with music. Comparing these two group of numbers, it can be seen that the path boundaries, with and without background music, were generally significantly different, with independent-samples $T$-test $t = 0.848$, $p = 0.018$, and effect size = 0.412. The number of observations of people with passing by behaviors with a relatively short path close to the music sound source in the square was 8.3% higher with background music than without. Passing by behaviors near the edge of the square with background music were zero when compared to the square without background music. Passing by behaviors closer to the music sound with background music were 12.5% higher when compared to the square without background music. The areas covered by passing by behaviors can also be reduced with music. This means that the presence of music caused people to be more centralized and walk closer to the sound source when passing by.
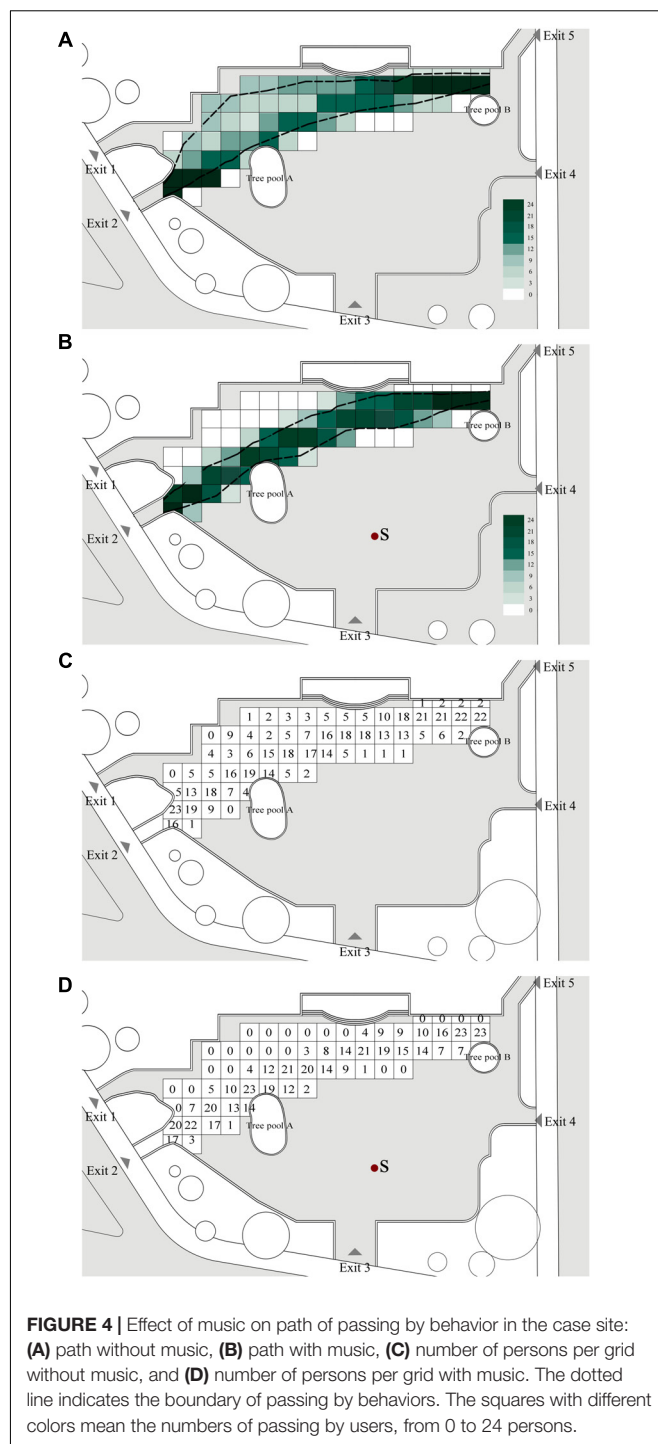


**FIGURE 4 |** Effect of music on path of passing by behavior in the case site: **(A)** path without music, **(B)** path with music, **(C)** number of persons per grid without music, and **(D)** number of persons per grid with music. The dotted line indicates the boundary of passing by behaviors. The squares with different colors mean the numbers of passing by users, from 0 to 24 persons.

### Speed

In terms of the speed of the passing by behavior, the square was considered both with and without background music.

#### Without background music

The mean speed of the walking around behavior in the square was 1.30 m/s. The minimum speed was 1.09 m/s, and the maximum speed was 1.57 m/s.

### With background music

The mean speed of the walking around behavior in the square was 1.30 m/s. The minimum speed was 1.06 m/s, and the maximum speed was 1.59 m/s.

The mean speed of the passing by behavior in the square with music was generally not significantly different from that of without music, with independent-samples $T$-test $t = -0.208$, $p = 0.836$, and effect size = 0.032.

Furthermore, exploring gender effects indicated that the path and speed of the passing by behavior for males and females, with and without background music, were generally not significantly different, with $T$-test $t = 0.132$, $p = 0.732$, and effect size = 0.051.

## Effects of Music on Movement Behavior: Walking Around Behavior

### Path

This section addresses the effects of music on the path of walking around behavior. Previous studies indicate that the paths from movement behavior are not random, but rather they are regular and directional. In this case, the users' paths in the square were influenced by environmental factors. Thus, based on the observations, the paths of walking around behavior were classified into four categories according to the location of boundaries and the water feature fence in the square. As **Figure 5** shows, path "a" represents walking around the fountain; path "b" implies walking around the fountain and tree pool A; path "c" represents walking around the boundary of the square except tree pool B; and path "d" implies walking around the boundary of the square including tree pool B. There were significant differences between the four paths, and therefore they were separated for a comparative analysis in which the square with and without background music was considered.

### Without background music

As **Figure 6A** shows, the mean areas of the paths from walking around behavior in the square were 1535 (a), 3024 (b), 4668 (c), and 5259 m$^2$ (d). As **Figure 6B** shows, the mean perimeters of the paths from walking around behavior in the square was 137 (a), 181 (b), 252 (c), and 274 m (d).
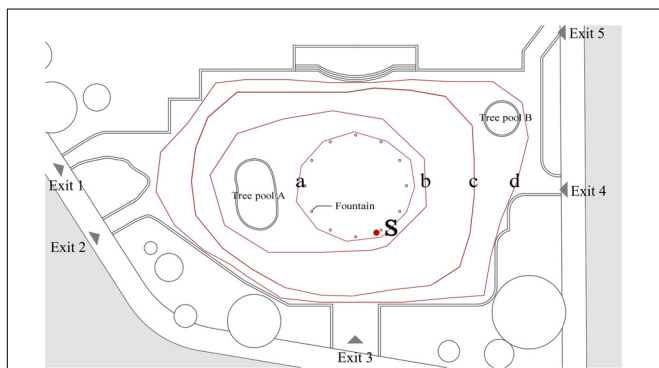


**FIGURE 5 |** Four kinds of paths of walking around behavior.
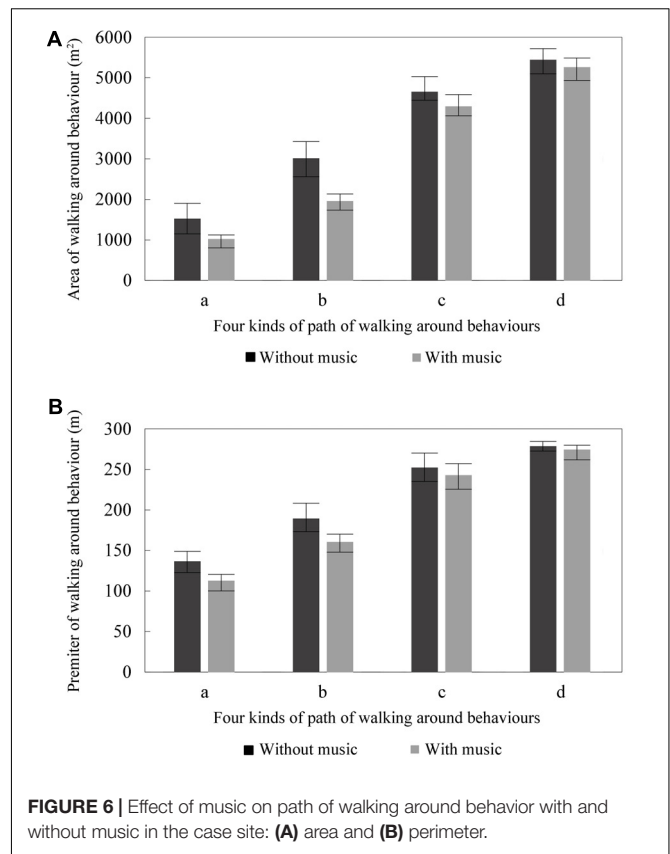


**FIGURE 6 |** Effect of music on path of walking around behavior with and without music in the case site: **(A)** area and **(B)** perimeter.

### With background music

As **Figure 6A** shows, the mean areas of the paths from walking around behavior in the square were 1038 (a), 1973 (b), 4311 (c), and 5338 m$^2$ (d). As **Figure 6B** shows, the mean perimeters of the paths from walking around behavior in the square were 113 (a), 160 (b), 243 (c), and 278 m (d).

The results indicated that the mean areas of walking around behavior in the square with background music were 32.36 (a), 30.74 (b), 7.66 (c), and 4.30% (d) smaller than that of the square without background music. Comparing the cases with and without background music indicated that the mean perimeters of walking around behavior in the square with background music were 17.34 (a), 15.14 (b), 3.68 (c), and 1.54% (d) smaller than that of the square without background music.

The ANOVA test was used to analysis the significance among music, category, and characteristics of walking around behavior, as shown in **Table 1**.

Compared with walking around behavior, with music and without, there were significant differences in areas at categories a and b, with ANOVA $p = 0.000$ and effect size = 0.592 (a) and 0.529 (b); and there were no significant differences in areas at categories c and d, with ANOVA $p = 0.165$ (c) and 0.489 (d) and effect size = 0.104 (c) and 0.027 (d). Similarly, there were significant differences in perimeters at categories a and b, with ANOVA $p = 0.000$ (a) and 0.006 (b), and effect size = 0.621 (a) and 0.355 (b), and no significant differences

in perimeters at categories c and d, with ANOVA $p = 0.296$ (c) and 0.249 (d), and effect size = 0.060 (c) and 0.073 (d). A possible reason for these results in categories a and b is that the crowd may have tended to move toward sound stimuli and then walk around at a shorter distance away from the music; this is then similar to the results found for passing by behavior. Compared with categories a and b, a possible reason for the results in categories c and d is that the crowd at categories c and d was relatively far away from the music sound source and therefore the effect of music was not significant in these situations.

## Speed

The mean speed of the four paths was analyzed first. The maximum difference of mean speeds among the four paths was 0.26 m/s without background music and 0.19 m/s with background music. It can be seen, from **Table 1**, that the speed of the paths in walking around behavior with background music was significantly slower than without background music in the four categories, with ANOVA $p = 0.010$ (a), 0.000 (b), 0.002 (c), and 0.002 (d), and effect size = 0.317 (a), 0.539 (b), 0.425 (c), and 0.410 (d). There were no significant differences between the four categories, with ANOVA $p = 0.590$ (without music) and 0.965 (with music), and effect size = 0.051 (without music) and 0.007 (with music). Therefore, the paths were merged to analyze the speed of walking around behavior. **Figure 7** shows the speed of walking around behavior in squares with and without background music.

### Without background music

The mean speed of the paths for walking around behavior in the square was 1.43 m/s. The minimum speed was 1.15 m/s, and the maximum speed was 1.74 m/s.
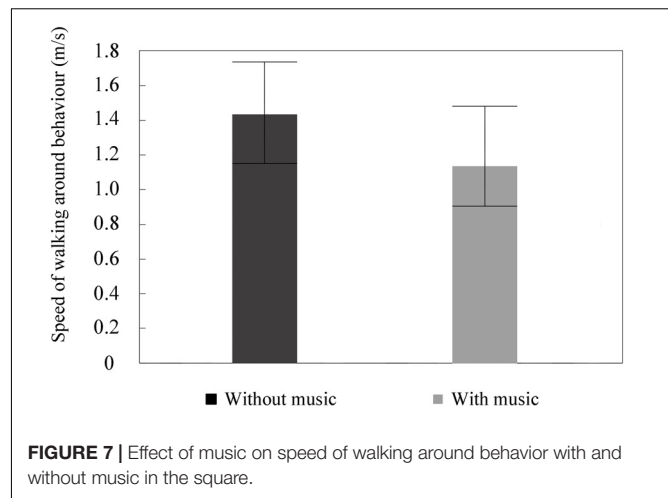


**FIGURE 7 |** Effect of music on speed of walking around behavior with and without music in the square.

### With background music

The mean speed of the paths for walking around behavior in the square was 1.14 m/s. The minimum speed was 0.93 m/s, and the maximum speed was 1.49 m/s.

Furthermore, exploring gender effects indicated that the path and speed of walking around behavior for males and females, with and without background music, was generally not significantly different with $T$-test $t = 0.211$, $p = 0.932$, and effect size = 0.005.

## Effects of Music on Non-movement Behavior: Sitting Behavior

This section addresses the effects of music of sitting behavior on crowd density. According to the statistical analyses, the number of those exhibiting sitting behavior ranged from 0 to 30. The relationship between crowd density and distance away from

**TABLE 1 |** ANOVA test for music, category, path and speed of walking around behavior.

| Factor | Variables | Mean | SD | F | Sig. | Effect size |
|--------|-----------|------|-----|---|------|-------------|
| Path | A0 | 3621.9443 | 1524.38340 | 177.963 | 0.000 | 0.937 |
|  | A1 | 3165.3397 | 1794.34592 | 248.789 | 0.000 | 0.954 |
|  | P0 | 211.5420 | 57.27472 | 224.300 | 0.000 | 0.949 |
|  | P1 | 199.1993 | 67.67651 | 296.166 | 0.000 | 0.961 |
|  | S0 | 1.4300 | 0.16733 | 0.647 | 0.590 | 0.051 |
|  | S1 | 1.1403 | 0.17839 | 0.090 | 0.965 | 0.007 |
| Music | Aa | 1286.7890 | 350.52975 | 20.195 | 0.000 | 0.529 |
|  | Ab | 2498.7895 | 700.22086 | 26.140 | 0.000 | 0.592 |
|  | Ac | 4490.0245 | 567.88599 | 2.100 | 0.165 | 0.104 |
|  | Ad | 5298.9650 | 245.73979 | 0.499 | 0.489 | 0.027 |
|  | Pa | 125.2505 | 15.48015 | 29.500 | 0.000 | 0.621 |
|  | Pb | 171.1615 | 17.66377 | 9.909 | 0.006 | 0.355 |
|  | Pc | 248.2220 | 19.45561 | 1.158 | 0.296 | 0.060 |
|  | Pd | 276.8485 | 8.06845 | 1.419 | 0.249 | 0.073 |
|  | Sa | 1.2835 | 0.21607 | 8.339 | 0.010 | 0.317 |
|  | Sb | 1.3075 | 0.25367 | 21.037 | 0.000 | 0.539 |
|  | Sc | 1.2620 | 0.20776 | 13.300 | 0.002 | 0.425 |
|  | Sd | 1.2875 | 0.23637 | 12.503 | 0.002 | 0.410 |

*A, area; P, perimeter, S, speed; a, b, c, and d, four categories; 0, without music; 1, with music.*

FIGURE 8 | Relationship between crowd density of those with sitting behavior and distances away from music sound source in the square: **(A)** with music and **(B)** without music, where the solid line means 0–10 persons, the dotted line means 11–20 persons, and the chain line means 21–30 persons.
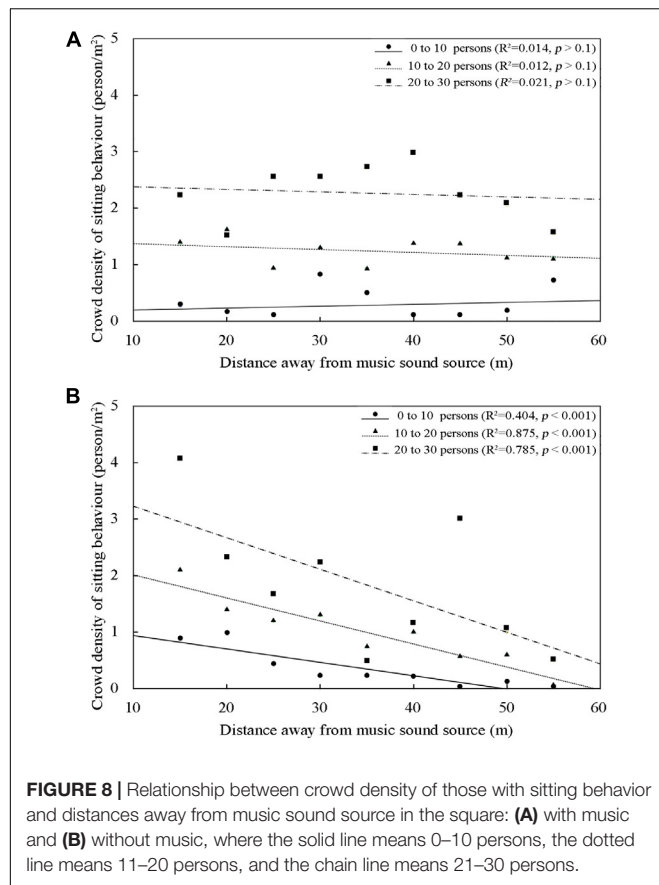
the music sound source in the square is shown in **Figure 8**, where the solid line means 0–10 persons, the dotted line means 11–20 persons, and the chain line means 21–30 persons, along with the linear regression and the coefficient of determination $R^2$. Results for observations with and without background music are discussed.

### Without Background Music

As **Figure 8A** shows, there were no significant differences in sitting behavior by distance away from the music sound source, with linear regression $R^2$ of 0.014 (0–10 persons), 0.012 (11–20 persons), and 0.021 (21–30 persons) and $p > 0.1$. The results indicated that sitting behavior remained randomly distributed over the case site with the increase of crowd density, and was generally not changed with different distance of sound sources. When the number of persons engaged in sitting behavior ranged from 0 to 10, 11 to 20, and 21 to 30, the crowd densities were respectively about 0.46, 1.27, and 1.96 persons/m$^2$ within 15–20 m of the music sound source, and generally the same at 25–30 and 35–40 m.

### With Background Music

It can be seen that sitting behavior increased with decreasing distance away from the music sound source, with linear regression $R^2$ of 0.404 (0–10 persons), 0.875 (11–20 persons), and 0.785 (21–30 persons) and $p < 0.001$. It can be seen from

**Figure 8B** that the crowd of persons engaged in sitting behavior decreased with increasing distance of music sound source. When the number of those exhibiting sitting behavior ranged from 0 to 10, the crowd densities were 0.89 persons/m$^2$ within 15–20 m of the music sound source, 0.56 persons/m$^2$ within 25–30 m, and 0.41 persons/m$^2$ within 35–40 m. When the number of those exhibiting sitting behavior ranged from 10 to 20, the crowd densities were about 1.82 persons/m$^2$ within 15–20 m of the music sound source. When the number of those exhibiting sitting behavior ranged from 20 to 30, the crowd densities were about 2.95 persons/m$^2$ within 15–20 m of the music sound source. One possible reason for these results is that the frequency of the music heard is reduced as the distance from music sound source is increased. It is interesting to note that when the numbers exhibiting increase in sitting behavior, the inclination of the three corresponding linear trend curves fell faster. For example, when the number of those with sitting behavior ranged from 0 to 10, crowd density in the square with background music was reduced by 0.12 persons/m$^2$ for every 5 m away from the source; when that number ranged from 20 to 30, crowd density was reduced by 0.28 person/m$^2$ every 5 m.

The comparison reveals that the crowd density of those exhibiting sitting behavior in the square with background music was higher than that without background music, when the distance to sound source was relatively shorter, while the crowd density of sitting behavior in the square with background music was lower than that without background music, when the distance was relatively long. For example, when the number of those with sitting behaviors ranged from 20 to 30, at 15–20 m away from the music sound source, crowd density was 0.99 persons/m$^2$ higher with background music than without, while at 35–40 m away from the music sound source, crowd density of those with sitting behaviors with background music was 0.99 persons/m$^2$ lower than without.

## DISCUSSION

The purpose of the present study was to explore the effect of music on movement behaviors, such as passing by behavior and walking around behavior, and non-movement behaviors, such as sitting behavior, in urban open spaces.

Regarding the effect of music on passing by behavior, as discussed in Section "Effects of Music on Movement Behavior: Passing by Behavior," the speed of passing by behavior was generally not significantly affected by music, while the path of passing by behavior shifted closer to the music sound source. This is in contrast to Lavia et al. (2016), who found that when music was deployed, people's walking speed through the other open spaces was slower. One possible reason for this is that the aims of passing by behavior are different in the two studies. In Lavia et al.'s (2016) study, the users are just intending to walk in the street. Some previous studies have pointed that when walkers do not have a clear purpose, their speed can be changed by landscape or environmental factors (Chen, 2009; Jia, 2012; Xie et al., 2013). In contrast, in the present study, the users are passing to go to work or school, and thus have a clear purpose, making it

reasonable that the speed of their behavior was not affected by the background music. An investigation among students also pointed out that visual differences do not change the speed of going to school (Fiegel et al., 2014). As for the effect of sound on path of passing by behavior, some previous studies have indicated that some animals and people will change their path to be far away from traffic noise (Lambert et al., 1984; Lengagne, 2008), whereas in the present study, it can be seen that the path of the crowd can be changed to be nearer the music. These results reinforce that behaviors can be effectively changed using the urban soundscape (Husain et al., 2002; Kang and Zhang, 2010).

Regarding the effect of music on walking around behavior, as discussed in Section " Effects of Music on Movement Behavior: Walking Around Behavior," the area, perimeter, and speed of the walking around behavior decrease with music. This result was the same as another study in which crowd behavior tended to move toward music (Jia, 2012). It is also interesting to note that a third study found that the children have different play behaviors with increasing distance from music (Holmes and Willoughby, 2005). In addition, the difference by the presence or absence of background music decreased as the area and perimeter increased. This was different from the effect of music on speed of passing by behavior, as the mean speed during walking around behavior with background music in the square was 0.29 m/s slower than without background music. These results show once more that behaviors without aims can be changed by environmental and landscape factors (Chen, 2009; Jia, 2012; Xie et al., 2013). Therefore, it can be concluded that the music drew people closer to the sound source and slowed their speed of walking. These results were the same as those found by Lavia et al. (2016) for other urban open spaces. A possible reason for this is that when hearing background music, users feel more comfortable; thus, their speed of walking around slows. Another reason may be that the presence of background music improves the feeling of safety in the environment; thus, background music contributes to building an urban slow space, which is more suitable for residents' health (Ye et al., 2012).

Third, regarding the effect of music on the crowd density of those exhibiting sitting behavior, as discussed in Section "Effects of Music on Non-movement Behavior: Sitting Behavior," when there was no music, there was no significant difference in density no matter how close to the sound source they were located. However, as the distance from the sound source increased, crowd density of those with sitting behavior decreased accordingly. Some previous studies have pointed that when there is no music, there is no significant difference in the crowd density of those with sitting behavior in indoor spaces such as railway stations and underground shopping streets (Debrezion et al., 2009; Meng et al., 2013). In urban open spaces, Meng and Kang (2016) also found that human sound-related activities generally have little effect on the sitting behaviors of pedestrians. On the effect of music, this result proves once more the finding that music-related activities increased the number of persons who passed by who stood and watched (Meng and Kang, 2016). As with music, users can also be attracted to a location by some nature sounds, such as sounds of bird or water (Liu et al., 2013). Some previous studies have also pointed that the acoustic perception of music is usually more salient than that of nature sounds (Aletta et al., 2016); this may lead to the different changes in sitting behaviors.

There are a number of possible implications for the applied value of the present study. Certain soundscapes, such as some music, may lead pedestrians to different paths in urban open spaces; it will be useful in landscape design to further investigate ways to lead walkers to suitable paths in gardens, for instance. Moreover, in leisure spaces such as parks, music can be used to decrease the speed of users and help them enjoy the landscape carefully. Furthermore, in rest areas in squares, a public-address system can be used to broadcast music to increase non-movement behavior, which can effectively increase interactions of citizens.

As demonstrated in the literature review, there are many classifications of musical sounds; however, only typical musical sounds were used in this study. In future studies, different tempos, genres, contexts, and levels of familiarity of musical sounds could also be investigated for comparison. Also, the location of the sound source was fixed in the present study; it can be seen from other studies that different locations of sound sources may lead to varying acoustic perceptions (Kang and Zhang, 2010). Therefore, in future studies by the present authors, different locations of music sound sources will be designed to find out their different effects on behaviors. Regarding movement and non-movement behaviors, only their speed and path were investigated in this work, whereas some previous studies have also pointed that characteristics of these behaviors such as duration and location also have effects that will be important for landscape design and urban planning (Lepore et al., 2016); therefore, in future studies, the present authors will further explore and explain these factors also.

## AUTHOR CONTRIBUTIONS

All authors carried out the study, designed the experiments, and wrote and critically reviewed the paper. QM and JK carried out the experiments. TZ analyzed the results.

## FUNDING

## REFERENCES

Aletta, F., Lepore, F., Kostara-Konstantinou, E., Kang, J., and Astolfi, A. (2016). An experimental study on the influence of soundscapes on people's behaviour in an open public space. *Appl. Sci.* 6:276. doi: 10.3390/app6100276

Ashihara, Y. (1985). *External Space Design*. Beijing: China Architecture & Building Press.

Barron, M., and Foulkes, T. J. (1994). Auditorium acoustics and architectural design. *J. Acoust. Soc. Am.* 96:612. doi: 10.1121/1.410457

Brander, L. M., and Koetse, M. J. (2011). The value of urban open space: meta-analyses of contingent valuation and hedonic pricing results. *J. Environ. Manage.* 92, 2763–2773. doi: 10.1016/j.jenvman.2011.06.019

Carpentier, F. R. D., and Potter, R. F. (2007). Effects of music on physiological arousal: explorations into tempo and genre. *Media Psychol.* 10, 339–363. doi: 10.1080/15213260701533045

Chen, S. (2009). *Study on Behaviour Vitality of in City Square*. Master's thesis, Central South University, Changsha.

Dai, Y. (2014). *Study on User Behaviour in an Urban Square Space*. Master's thesis, Central South University, Changsha.

Debrezion, G., Pels, E., and Rietveld, P. (2009). Modelling the joint access mode and railway station choice. *Transp. Res. Part E* 45, 270–283. doi: 10.1016/j.tre.2008.07.001

dela Fuente de Val, G., Atauri, J. A., and Lucio, J. V. (2006). Relationship between landscape visual attributes and spatial pattern indices: a test study in Mediterranean-climate landscapes. *Landsc. Urban Plan.* 77, 393–407. doi: 10.1016/j.landurbplan.2005.05.003

Fiegel, A., Meullenet, J. F., Harrington, R. J., Humble, R., and Seo, H. S. (2014). Background music genre can modulate flavor pleasantness and overall impression of food stimuli. *Appetite* 76, 144–152. doi: 10.1016/j.appet.2014.01.079

Floyd, M. F. (1998). Getting beyond marginality and ethnicity: the challenge for race and ethnic studies in Leisure Research. *J. Leis. Res.* 30, 3–22. doi: 10.1080/00222216.1998.11949816

Franěk, M., van Noorden, L., and Režný, L. (2014). Tempo and walking speed with music in the urban context. *Front. Psychol.* 5:1361. doi: 10.3389/fpsyg.2014.01361

Gehl, J. (1987). *Life between Buildings: Using Public Space*. New York, NY: Van Nostrand Reinhold Company, Inc.

Gomez, P., and Danuser, B. (2004). Affective and physiological responses to environmental noises and music. *Int. J. Psychophysiol.* 53, 91–103. doi: 10.1016/j.ijpsycho.2004.02.002

Guéguen, N., Jacob, C., Guellec, H. L., Morineau, T., and Lourel, M. (2008). Sound level of environmental music and drinking behaviour: a field experiment with beer drinkers. *Alcohol. Clin. Exp. Res.* 32, 1795–1798. doi: 10.1111/j.1530-0277.2008.00764.x

Holmes, E., and Willoughby, T. (2005). Play behaviour of children with autism spectrum disorders. *J. Intellect. Dev. Disabil.* 30, 156–164. doi: 10.1080/13668250500204034

Husain, G., Thompson, W. F., and Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Percept.* 20, 151–171. doi: 10.1525/mp.2002.20.2.151

Hwang, R. L., Lin, T. P., and Matzarakis, A. (2011). Seasonal effects of urban street shading on long-term outdoor thermal comfort. *Build. Environ.* 46, 863–870. doi: 10.1016/j.buildenv.2010.10.017

ISO 12913-1 (2014). *Acoustics- Soundscape-Part1: Definition and Conceptual Framework*. Geneva: International Organisation for Standardization.

Jia, S. Q. (2012). *Crowd Behaviour in Soundscape*. Master's thesis, Harbin Institute of Technology, Harbin.

Kang, J. (2017) From dBA to soundscape indices: managing our sound environment. *Front. Eng. Manag.* 4, 184–192. doi: 10.15302/J-FEM-2017026

Kang, J., and Zhang, M. (2010). Semantic differential analysis of the soundscape in urban open public spaces. *Build. Environ.* 45, 150–157. doi: 10.1016/j.buildenv.2009.05.014

Lambert, J., Simonnet, F., and Vallet, M. (1984). Patterns of behaviour in dwellings exposed to road traffic noise. *J. Sound Vib.* 92, 159–172. doi: 10.1016/0022-460X(84)90553-4

Lavia, L., Witchel, H. J., Kang, J., and Aletta, F. (2016). "A preliminary soundscape management model for added sound in public spaces to discourage anti-social and support pro-social effects on public behaviour," in *Proceedings of DAGA Conference*, (Aachen), 1–4. doi: 10.13140/RG.2.1.1392.4242

Lengagne, T. (2008). Traffic noise affects communication behaviour in a breeding anuran. *Hyla arborea. Biol. Conserv.* 141, 2023–2031. doi: 10.1016/j.biocon.2008.05.017

Lepore, F., Kostara-Konstantinou, E., Aletta, F., Astolfi, A., and Kang, J. (2016). "A preliminary investigation about the influence of soundscapes on people's

behaviour in an open public space," in *Proceedings of Internoise*, (Hong Kong), 5219–5224.

Lewin, K., Heider, F., and Heider, G. M. (1936). *Principles of Topological Psychology*. New York, NY: McGraw-Hill. doi: 10.1037/10019-000

Li, J. N., and Meng, Q. (2015). Study on the soundscape in commercial pedestrian streets. *Tech. Acoust.* 34, 326–329.

Liu, J., Kang, J., Luo, T., Behm, H., and Coppack, T. (2013). Spatiotemporal variability of soundscapes in a multiple functional urban area. *Landsc. Urban Plan.* 115, 1–9. doi: 10.1016/j.landurbplan.2013.03.008

Marcus, C. C., and Francis, C. (1998). *People Places: Design Guidelines for Urban Open Space*. New York, NY: John Wiley & Sons, Inc.

Marušić, B. G. (2011). Analysis of patterns of spatial occupancy in urban open space using behaviour maps and GIS. *Urban Des. Int.* 16, 36–50. doi: 10.1057/udi.2010.20

Meng, Q., and Kang, J. (2013). Influence of social and behavioural characteristics of users on their evaluation of subjective loudness and acoustic comfort in shopping malls. *PLoS One* 8:e54497. doi: 10.1371/journal.pone.0054497

Meng, Q., and Kang, J. (2015). The influence of crowd density on the sound environment of commercial pedestrian streets. *Sci. Total Environ.* 511, 249–258. doi: 10.1016/j.scitotenv.2014.12.060

Meng, Q., and Kang, J. (2016). Effect of sound-related activities on human behaviours and acoustic comfort in urban open spaces. *Sci. Total Environ.* 573, 481–493. doi: 10.1016/j.scitotenv.2016.08.130

Meng, Q., Kang, J., and Jin, H. (2013). Field study on the influence of spatial and environmental characteristics on the evaluation of subjective loudness and acoustic comfort in underground shopping streets. *Appl. Acoust.* 74, 1001–1009. doi: 10.1016/j.apacoust.2013.02.003

Meng, Q., Sun, Y., and Kang, J. (2017a). Effect of temporary open-air markets on the sound environment and acoustic perception based on the crowd density characteristics. *Sci. Total Environ.* 60, 1488–1495. doi: 10.1016/j.scitotenv.2017.06.017

Meng, Q., Zhang, S., and Kang, J. (2017b). Effects of typical dining styles on conversation behaviours and acoustic perception in restaurants in China. *Build. Environ.* 121, 148–157. doi: 10.1016/j.buildenv.2017.05.025

Milliman, R. E. (1982). Using background music to affect the behavior of supermarket shoppers. *J. Mark.* 46, 86–91. doi: 10.2307/1251706

Myers, M. (1975). Decision making in allocating metropolitan open space: state of the art. *Trans. Kansas Acad. Sci.* 78, 149–153. doi: 10.2307/3627339

Natarajan, K., Rajendra, A. U., Alias, F., Tiboleng, T., and Puthusserypady, S. K. (2004). Nonlinear analysis of EEG signals at different mental states. *Biomed. Eng. Online* 3:7. doi: 10.1186/1475-925X-3-7

North, A. C., Hargreaves, D. J., and Hargreaves, J. J. (2004). Uses of music in everyday life. *Music Percept.* 22, 41–77. doi: 10.1525/mp.2004.22.1.41

Oakes, S., and North, A. C. (2008). Reviewing congruity effects in the service environment musicscape. *Int. J. Serv. Ind. Manage.* 19, 63–82. doi: 10.1108/09564230810855716

Orr, T. J., Myles, B. S., and Carlson, J. K. (1998). The impact of rhythmic entrainment on a person with autism. *Focus Autism Other Dev. Disabl.* 13, 163–166. doi: 10.1177/108835769801300304

Payne, L. L., Mowen, A. J., and Orsega-Smith, E. (2002). An examination of park preferences and behaviors among urban residents: the role of residential location, race, and age. *Leis. Sci.* 24, 181–198. doi: 10.1080/01490400252900149

Ren, X., and Kang, J. (2015). Interactions between landscape elements and tranquility evaluation based on eye tracking experiments. *J. Acoust. Soc. Am.* 5, 3019–3022. doi: 10.1121/1.4934955

Sakharov, D. S., Davydov, V. I., and Pavlygina, R. A. (2005). Intercentral relations of the human EEG during listening to music. *Hum. Physiol.* 31, 27–32. doi: 10.1007/s10747-005-0065-5

Sinibaldi, G., and Marino, L. (2013). Experimental analysis on the noise of propellers for small UAV. *Appl. Acoust.* 74, 79–88. doi: 10.1016/j.apacoust.2012.06.011

Styns, F., Van, N. L., Moelants, D., and Leman, M. (2007). Walking on music. *Hum. Mov. Sci.* 26, 769–785. doi: 10.1016/j.humov.2007.07.007

Thompson, C. W. (2002). Urban open space in the 21st century. *Landsc. Urban Plan.* 60, 59–72. doi: 10.1016/S0169-2046(02)00059-2

Thwaites, K., Helleur, E., and Simkins, I. M. (2005). Restorative urban open space: exploring the spatial configuration of human emotional fulfilment in urban open space. *Landsc. Res.* 30, 525–547. doi: 10.1080/01426390500273346

Wang, C. (2014). *Simulation Research on Occupant Energy-Related Behaviours in Building*. Doctoral dissertation, Tsinghua University, Beijing.

Westover, T. N. (1989). Perceived crowding in recreational settings an environment-behavior model. *Environ. Behav.* 21, 258–276. doi: 10.1177/0013916589213002

Whitlock, S. (2004). The market and the city: square, street, and architecture in early modern Europe. *J. Mod. Hist.* 79, 649–651. doi: 10.1086/523218

Xie, H., Kang, J., and Mills, G. H. (2012). Sound power levels of typical medical equipment in intensive care units. *Acta Acust. United Ac.* 98, 659–666. doi: 10.3813/AAA.918545

Xie, H., Kang, J., and Mills, G. H. (2013). Behaviour observation of major noise sources in critical care wards. *J. Crit. Care* 28, 1109.e5–1109.e18. doi: 10.1016/j.jcrc.2013.06.006

Ye, P., Wang, H., and Gao, F. (2012). A preliminary study on the research method of urban public space environment behaviour based on GPS: a Case Study of Shengli Square in Hefei. *Archit. J.* 2012, 28–33.

Yu, L., and Kang, J. (2010). Factors influencing the sound preference in urban open spaces. *Appl. Acoust.* 71, 622–633. doi: 10.1016/j.apacoust.2010.02.005

Yuan, W., and Tan, K. H. (2011). A model for simulation of crowd behaviour in the evacuation from a smoke-filled compartment. *Phys. A* 390, 4210–4218. doi: 10.1016/j.physa.2011.07.044

Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *J. Acoust. Soc. Am.* 111, 1832–1846. doi: 10.1121/1.1458027

Zakariya, K., Harun, N. Z., and Mansor, M. (2014). Spatial characteristics of urban square and sociability: a review of the city square. *Melb. Proc. Soc. Behav. Sci.* 2014, 678–688. doi: 10.1016/j.sbspro.2014.10.099

Zhang, D., Zhang, M., Liu, D., and Kang, J. (2016). Soundscape evaluation in Han Chinese Buddhist temples. *Appl. Acoust.* 111, 188–197. doi: 10.1016/j.apacoust.2016.04.020

Zhang, Y., Yang, Z., and Yu, X. (2006). Measurement and evaluation of interactions in complex urban ecosystem. *Ecol. Model.* 196, 77–89. doi: 10.1016/j.ecolmodel.2006.02.001

# Personal Audiovisual Aptitude Influences the Interaction Between Landscape and Soundscape Appraisal

Kang Sun[1]*, Gemma M. Echevarria Sanchez[1], Bert De Coensel[1], Timothy Van Renterghem[1], Durk Talsma[2] and Dick Botteldooren[1]

[1] Department of Information Technology, Ghent University, Ghent, Belgium, [2] Department of Experimental Psychology, Ghent University, Ghent, Belgium

It has been established that there is an interaction between audition and vision in the appraisal of our living environment, and that this appraisal is influenced by personal factors. Here, we test the hypothesis that audiovisual aptitude influences appraisal of our sonic and visual environment. To measure audiovisual aptitude, an auditory deviant detection experiment was conducted in an ecologically valid and complex context. This experiment allows us to distinguish between accurate and less accurate listeners. Additionally, it allows to distinguish between participants that are easily visually distracted and those who are not. To do so, two previously conducted laboratory experiments were re-analyzed. The first experiment focuses on self-reported noise annoyance in a living room context, whereas the second experiment focuses on the perceived pleasantness of using outdoor public spaces. In the first experiment, the influence of visibility of vegetation on self-reported noise annoyance was modified by audiovisual aptitude. In the second one, it was found that the overall appraisal of walking across a bridge is influenced by audiovisual aptitude, in particular when a visually intrusive noise barrier is used to reduce highway traffic noise levels. We conclude that audiovisual aptitude may affect the appraisal of the living environment.

Keywords: audiovisual interactions, landscape, soundscape, environmental perception, personal factor

## INTRODUCTION

The phrase 'soundscape' used in this study is as defined by International Organization for Standardization (ISO): an "acoustic environment as perceived or experienced and/or understood by a person or people, in context" (ISO, 2014). The subjective appraisal of our living environment is influenced by landscape and soundscape alike. It is well known that these influences are not independent. This interaction partly originates at a low level of auditory and visual perception. In soundscape theory, the importance of visual context on soundscape appraisal has been stressed (Weinzimmer et al., 2014; Botteldooren et al., 2015). Using virtual reality, it was likewise shown that the sonic environment affects overall pleasantness of the public space even when the participants in the experiment focused on visual designs and were kept unaware of the sound (Echevarria Sanchez et al., 2017). In the home environment, it has been shown that vegetation as seen through a window affects the self-reported noise annoyance at home (Li et al., 2010; Van Renterghem and Botteldooren, 2016; Leung et al., 2017). The visibility of a sound source may also affect the

awareness of sound. On the one hand, it has been shown that people get more annoyed when the sound source is visible (Zhang et al., 2003), while other studies found that sound is actually less annoying when the source is visible (Maffei et al., 2013). It remains currently unknown what drives these differences. In this paper, we forward the hypothesis that a personal factor or multiple personal factors influence the interaction between landscape and soundscape appraisal. Personal traits and beliefs are known to influence the perception and appraisal of the sonic environment both at home [e.g., noise sensitivity (Miedema and Vos, 2003; Heinonen-Guzejev, 2009)] and in public spaces [e.g., meaning given to tranquility (Filipan et al., 2017) and recreation (Pilcher et al., 2009; Miller et al., 2014)]. So it is not unlikely that this additional personal factor would indeed exist.

Previous studies have already shown that considerable individual differences exist in the way humans process audiovisual information, ranging from differences in connectivity between auditory and visual pathways (e.g., van den Brink et al., 2013), to selective preferences in processing auditory or visual material (Giard and Peronnet, 1999). More generally, when engaged in a visual task, participants tend to ignore auditory stimuli, as demonstrated by the well-known Colavita effect (Colavita, 1974). One striking result from many studies on the Colavita effect is that when participants are presented with either auditory or audiovisual stimuli, and have to respond to a change in the auditory stimulus, they usually do so accurately on the auditory-only trials, but fail to detect this change when an audio–visual stimulus is presented to them. A main question is why participants miss such an auditory change.

One possible answer comes from Simons and Chabris, who explored how an unexpected object could go unnoticed during a monitoring task, in a phenomenon they described as inattentional blindness (Simons and Chabris, 1999). Recent research also demonstrates that a single discrete visual distractor can improve the detectability of an unexpected object in an inattentional blindness task (Pammer et al., 2014). Visual distractor processing tends to be more pronounced when the perceptual load of a task is low compared to when it is high [perpetual load theory (Lavie, 1995)]. Sandhu and Dyson studied the effect of auditory load on visual distractors and vice versa. They found that in both attend auditory and attend visual conditions, the distractor processing was evident, especially when the distractors were visual (Sandhu and Dyson, 2016). Perpetual load theory has been supported from assessing the impact of perceptual load on the flanker task (Eriksen and Eriksen, 1974), as well as behavioral paradigms, such as negative priming (Lavie and Fox, 2000), implicit learning (Jiang and Chun, 2001) and inattentional blindness (Cartwright-Finch and Lavie, 2007).

A possible explanation for inattentional blindness based on perpetual load theory is that conscious perception of task-irrelevant stimuli critically depends upon the level of task-relevant perceptual load rather than intentions or expectations (Cartwright-Finch and Lavie, 2007). Aging could increase the susceptibility to inattentional blindness (Graham and Burke, 2011). Likewise, individual differences in cognitive ability related to working memory and executive functions affect inattentional blindness (Fougnie and Marois, 2007). Several studies have

shown that this phenomenon could be associated with general fluid intelligence (O'Shea and Fieo, 2015) and executive attentional control (Kahneman, 1973). Moreover, an explanation in terms of attention and working memory capacity can explain individual differences in perceiving audiovisual stimuli.

As a counterpart to inattentional blindness, Macdonald and Lavie reported that people could also miss sounds in high-visual-load condition; a phenomenon which they described as "inattentional deafness" (Macdonald and Lavie, 2011). It stands in parallel with inattentional blindness, following the same procedure of reducing perceptual processing of task-irrelevant information in high-load tasks. Therefore, one could expect various forms of "inattentional deafness" resembling the known forms of "inattentional blindness" (Mack and Rock, 1998), ranging from failing to recognize meaningful distractor objects (Lavie et al., 2009) to failing to notice the presence of stimuli (Neisser and Becklen, 1975).

Earlier research has also shown the benefit of vision in speech-reception (Musacchia et al., 2007). By contrast, it has also been shown that in situations of uncertainty, observers tend to follow the more reliable auditory cue (Apthorp et al., 2013). Very mild forms or hearing damage might lead to reduced speech intelligibility (Bharadwaj et al., 2014; Füllgrabe et al., 2015) and thus a stronger reliance on visual cues. But, it was also observed that some persons are simply more auditory dominated while others are more visual dominated (Giard and Peronnet, 1999).

The above discussion indicates that there might be individual differences in the way people perceive audiovisual stimuli that would be more pronounced in a rather complicated audiovisual environment, possibly due to individual differences in distractibility. Individual levels of distractibility can vary from slight facilitation from a noisy background to severe disruption (Ellermeier and Zimmer, 1997). It has been suggested that individual differences in working memory capacity underlie individual differences in susceptibility to auditory distraction in most tasks and contexts (Sörqvist and Rönnberg, 2014). The findings on working memory capacity reflect individual differences in the ability to control attention and avoid distraction (Conway et al., 2001). It has been shown that high-working memory capacity individuals are less susceptible to the effects of auditory distractors (Beaman, 2004; Sörqvist, 2010). A recent study showed that attention restoration is achieved through increased exposure to natural sounds, while conversely, human-caused sounds reduce attention restoration (Abbott et al., 2016).

Throughout this article, the personal factor which was discussed above and that is expected to influence how persons perceive and appraise a combined auditive and visual stimulus will be labeled *audiovisual aptitude*. The term *aptitude* was chosen to highlight our hypothesis that this personal factor reflects a natural ability to process audiovisual scenes. This ability includes focusing on either (the visual or auditory) part of the scene and its composition in both simple and complex scenes. Its detailed meaning will further be explored in the discussion section.

This paper uses an audiovisual deviant detection experiment, with real-life scenes containing multiple visual and audio elements, to categorize persons according to their auditory acuity and their distractibility by incongruent visual stimuli. Two

previously conducted experiments (labeled experiments 2 and 3 in the following sections) have been reanalyzed by including audiovisual aptitude as a personal factor. Audiovisual aptitude is expected to modify the effect of the view from the window on reported noise annoyance in Experiment 2. In Experiment 3, it modifies the effect of sonic and visual stimuli on pleasantness of walking across a bridge.

The audiovisual deviant detection experiment was designed to focus on the skills and sensitivities that matter for environmental sound perception. Previous research has shown that sounds that can be recognized relate to the overall appraisal of soundscapes in public places such as parks (Pilcher et al., 2009; Axelsson et al., 2010; Miller et al., 2018). Likewise, it was shown that noticing sounds from outside influences annoyance at home (De Coensel et al., 2009). In general, perception is a comprehensive process, in which a single factor sometimes cannot explain the final result (Botteldooren et al., 2006; Brown, 2012). Thus, the first part was designed to test the participant's ability to analyze complex auditory scenes and identify individual sounds in it. An ecologically valid setting assures that participants can also rely on personal experience and context-related expectation, factors that will also influence the appraisal of the environment in everyday life. A deviant detection task is chosen where the deviant is a complex auditory scene in which one sound is missing. To explore the influence of visual information on sound perception that is explained above, the second part of the test adds the visual context that matches the auditory scene. Congruent visual information on the deviant (missing sound) would be beneficial in general for the deviant detection task. Yet, as people are in general expected to be more visually guided (Colavita effect), participants could then simply detect the visual deviant, which would not be very instructive for identifying their audiovisual aptitude. Hence, the information on the deviant was made incongruent between the visual and the auditory information, making distraction and perceptual load dominant mechanisms.

## METHODOLOGY

### Overview

This study uses three experiments conducted by the same participants to identify the personal differences in audiovisual aptitude (Experiment 1) and to explore how these differences influence perception of the environment (Experiments 2 and 3).

The first experiment explores audiovisual aptitude. It consists of a blind audio test (Part 1) and audiovisual test (Part 2) sharing the same audio track. During both tests, participants were requested to detect the deviant auditory stimulus amongst three fragments. This experiment contained four scenarios, in which either the audio or visuals altered. This ecologically valid alternative to simple psychological stimuli is intended to investigate whether a person's visual attention mechanism dominates auditory attention.

Meanwhile, the same participants joined the other two experiments, one focusing on road traffic annoyance at home and the other on the perceived quality of the public space. These have

been analyzed in view of the audiovisual aptitude. This setting allows to explore whether the personal audiovisual aptitude identified in Experiment 1 can be used to explain differences in response in the other two experiments.

With the criteria of good (peripheral) hearing and completing the whole experiment, this study collected 68 participants (28 Female, $M_{age}$ = 27.9, $SD$ = 5.05, range: 20–46 years, 48 obtained a master degree or higher). In later analysis, participants were classified based on gender, age (divided into two groups by median value 27, group 1: 20–27 years, 37 participants, $M_{age}$ = 24.2, $SD$ = 1.8; group 2: 31 participants, 28–46 years, $M_{age}$ = 32.5, $SD$ = 3.9) and education. All the principles outlined in the Helsinki Declaration of 1975, as revised in 2000 (World Medical Association, 2001), have been followed in all the experiments involving human subjects. All participants signed an informed consent form before the start of the experiments.

## Experiment 1: Audiovisual Aptitude
### Layout of the Paired Test

As shown in **Table 1**, the audio test (Part 1) only contains the audio content, while the video test (Part 2) contains both sound and vision. In each part, participants were asked a single question after experiencing the three items: 'Which of the three items sounds most differently from the other two?'. In Part 1, item 2 was the correct answer, whereas in Part 2 item 5 was the correct answer. During the analysis stage, in Part 1, choosing item 2 will be marked as correct, and consequently, choosing item 1 or 3 will be considered as mistake 1 (M1). In Part 2, item 5 is correct, and 4 and 6 mistakes (M2).

### Scenarios Content

This study uses four different scenarios. Content details of the videos are listed in **Table 2**. **Figure 1** shows screenshots of the four scenarios.

In **Figure 1**, the object (VAO) that is absent in one of the videos in each scenario is indicated with a circle, while its path and moving direction are shown with the solid lines and arrows. Scenario (a) shows a view of a tarmac through a terminal window, with several aircrafts and a few shuttle busses far in the scene. The background sound consists of terminal announcements and people talking. Scenario (b) is a crowded student restaurant, with people eating, talking and laughing (forming the background sound). The attention attracting object in scenario (b) is a tapping finger, with its small movement within the range of the solid line circle as shown in **Figure 1b**. Scenario (c) shows an aircraft runway in front of a terminal window with many shuttle busses and vans moving around. Differently from scenario (a), the background of this scenario is an outdoor site with various mechanical sounds. The attention attracting object, a departing aircraft, occurs in the background of the scene. Scenario (d) shows a small city in a city outskirt, containing chickens on the left side of the screen, as well as a few cars passing by behind the park. The background sound here consists in chicken sounds, park sounds and city background sound. All four scenarios were recorded with a stable camera.

For each scenario, item 6 is the stimulus where the attracting object was removed from the visual. In scenario (a), (c), and (d),
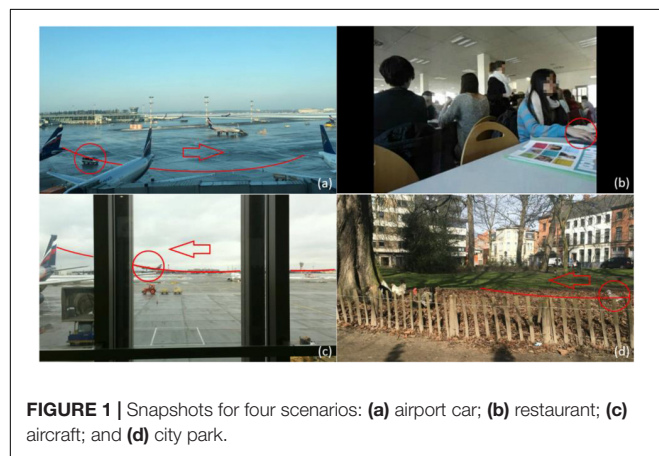
**TABLE 1 |** Overview of audio-visual scenarios studied in Experiment 1.

| | Item Number | File format | Content | | Mistake type |
|---|---|---|---|---|---|
| | | | Auditory | Vision | |
| Part 1 | 1 | Audio | Background sound + AAO | Black screen | M1 |
| | 2 | Audio | Background sound | Black screen | |
| | 3 | Audio | Background sound + AAO | Black screen | M1 |
| Part 2 | 4 | Video | Background sound + AAO* | Background view + VAO* | M2 |
| | 5 | Video | Background sound | Background view + VAO | |
| | 6 | Video | Background sound + AAO | Background view | M2 |

*Congruent visual attention attracting object (VAO) and matching auditory attention attracting object (AAO).

**TABLE 2 |** Visual and auditory context for each of the scenarios used in the audiovisual aptitude experiment together with congruent visual attention attracting object (VAO) and matching auditory attention attracting object (AAO).

| Number | a | b | c | d |
|---|---|---|---|---|
| Scenario | Airport car | Restaurant | Aircraft | City park |
| Main visual context (background view) | Terminal window view to parking apron | Student restaurant at sitting position | Terminal window view to airport runway | A bunch of chicken in the park |
| Main auditory context (background sound) | Broadcasting, people talking, aircraft engine | People talking, eating, forks and plates | Airport outside sound, wind, shuttle bus passing | Chicken crowing and walking on fallen leaves |
| VAO | Shuttle bus passing | Tapping finger | Departing aircraft | Walking pigeon |
| AAO | Shuttle bus sound | Finger tapping sound | Aircraft departing sound | Pigeon cooing, walking on leaves |
| Total duration | 0:27 | 0:35 | 1:00 | 0:55 |
| AO duration | 0:12 | 0:12 | 0:24 | 0:11 |
| (percentage) | (44.4%) | (34.3%) | (40%) | (20%) |



**FIGURE 1 |** Snapshots for four scenarios: **(a)** airport car; **(b)** restaurant; **(c)** aircraft; and **(d)** city park.

the (visually) attracting objects were removed. In scenario (b), the tapping finger was replaced by a stable hand lying on the table.

## Procedure

This experiment was conducted scenario by scenario. In part 1 of the test, participants were asked to listen to items 1, 2, and 3 presented with audio only (black screen). In part 2, participants were asked to watch items 4, 5, and 6 from the same scenario. Once they finished a particular scenario, they could move on to the next one until all four scenarios were experienced.

The four scenarios were presented in random order and also the order of presenting the items was randomized. Each item could be played only once, and there was no backtrack and alteration once a single scenario was completed. All participant finished this experiment with the same headphones in the same quiet room (with a background noise of about 30 dBA).

In addition, personal information like age, gender and education level, as well as noise sensitivity [via Weinstein's questionnaire (Weinstein, 1978)] were recorded ($M_{sensitivity}$ = 79.40, $SD$ = 10.95, participants were split into two groups with midpoint 73.5 afterwards). The hearing status of all participants was assessed via pure tone audiometry (PTA) carried out in a quiet but not sound-proof room using a regularly calibrated AC5Clinical Computer Audiometer.

## Experiment 2: Annoyance in Living Room

In a mock-up living room (**Figure 2**), participants were asked to engage in some light activities for 10 min while hearing highway traffic sounds. After 10 min, the standard ICBEN noise annoyance question was asked using an 11-point answering scale, referring to the past 10 min. This experiment was conducted with four sound pressure levels [45 dB(A), 50 dB(A), 55 dB(A), and 60 dB(A), measured in the center of the living room] corresponding to four different acoustical window insulation cases. The following 3 days, the same experimental procedure was repeated. However, while participants were led to believe that they simply evaluated again four window types, what actually changed was the video playing in the background to simulate a

FIGURE 2 | The mock-up living room with hidden loudspeakers indicated next to the mock-up window.



FIGURE 3 | (a) Equipment used for calibration. (b) Equipment used for virtual reality experiment.
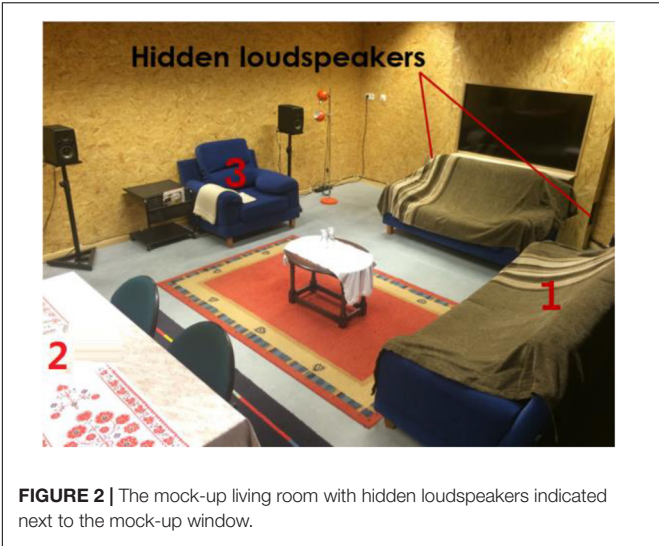
window view (**Table 3**). With this experimental design, we aimed to go beyond simple loudness evaluation (as can be expected by playing a short sound fragment only). In addition, we hid the true purpose, especially regarding our interest in the visuals displayed as a window view. More details on this experiment can be found in (Sun et al., 2018).

## Experiment 3: Perception of Public Space

The third experiment is complementary to the second one in two ways. Firstly, it considers the public space, more specifically the perceived environmental quality of a bridge crossing a ring road giving access to a park. Secondly, four visual designs were evaluated, hiding the fact that our interest is now in the effect of the noise coming from the highway below the bridge on audiovisual quality assessment. To achieve this, on each day of the experiment the participants evaluated a walk across the bridge in a virtual environment displayed to them using oculus rift
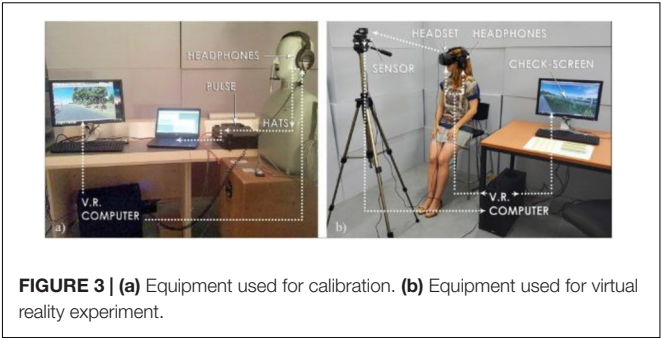
(**Figure 3**). A sequence of four rather different visual designs were displayed to them each day (**Figure 4**), yet the sound coming from the highway under the bridge stayed the same. Participants were asked to rate the pleasantness of the total experience without specifically referring to sound. On the subsequent days, they evaluated visually identical environments yet the sound changed without informing the participants. More details on this experiment can be found in (Echevarria Sanchez et al., 2017).

In this experiment, participants were virtually moving across the bridge following a pre-defined path, but they could freely move their head. An important and interesting aspect that could be analyzed with this setup is the head movement, which is a proxy for their looking behavior, reflecting where people's (visual) attention is directed to (Gibson and Pick, 1963). Recording the looking behavior allows assessing the frequency and total duration of gazing at the highway during the walk. This counting is based on the head movement of the participants and the screen middle point is used as a proxy for the visual focus point. This recording in only performed with the four matching situations (visual designs with the corresponding sonic environments).

## Statistical Analysis

To test whether the personal factors have an impact on the results of part 1 and 2 in Experiment 1, a repeated analysis of

TABLE 3 | Snapshots from the videos played in the mock-up window.

| | Green elements | No green elements |
|---|---|---|
| Sound source visible |  |  |
| Sound source invisible |  |  |

FIGURE 4 | Snapshot of the virtual reality display of the four bridge designs; the barrier seen on the right progressively increases in height when going from V1 to V4, reducing the highway noise level.
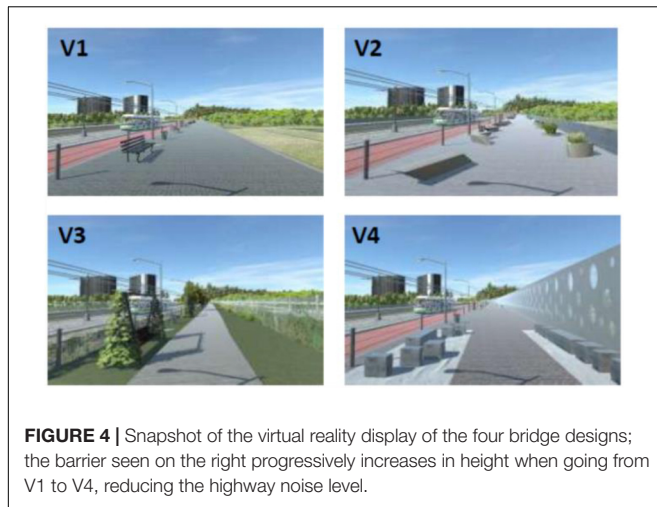
variance (anova) test was conducted. To observe the relation between a sound factor (the duration of the attention attracting object) and the overall result of part 1 and disparity between overall results in part 1 and 2, a linear regression was performed. Furthermore, in Experiments 2 and 3, first, a generalized linear model is built to find the fittest classification of participants through Experiment 1 – that is the classification that results in the best model quality. Then, a mixed-effect generalized linear model targeting at noise annoyance (Experiment 2) and pleasantness (Experiment 3) is conducted, using 'participant' as a random factor to generalize the results, accounting for various factors including the fittest personal factor via Experiment 1. The Akaike Information Criterion (AIC) is used to rate the model quality (models with smaller AIC values fit better). At last, an anova test is conducted to check the impact of personal factors on the gazing time in Experiment 3. The statistics analysis in this study was conducted in SPSS statistics (version 25).

## RESULTS AND ANALYSIS

## Audiovisual Aptitude
### Overview
**Figure 5** shows the percentage of the participants that made a mistake in different parts of the audiovisual aptitude experiment. In part 1 (M1), scenario 'park' is where people made most mistakes while scenario 'airport car' led to the smallest number of mistakes. Despite the scenario differences, task performance in general decreases by adding a visual setting containing incongruent information on the deviant. Comparing the differences between M1 and M2, visual information makes the task performance significantly worse in some scenarios ('airport car' and 'aircraft'), while in other scenarios, it has less effect. Further analysis will focus on personal factors that can be deduced.

### Effect of Personal Factor
Aiming at M1, an anova test with factor scenario and various personal factors was made. The result shows that the factor

education ($F_{1,264}$ = 2.31; $p$ > 0.05), gender ($F_{1,264}$ = 1.25; $p$ > 0.05), noise sensitivity ($F_{1,264}$ = 0.052; $p$ > 0.05) and age ($F_{1,264}$ = 0.11; $p$ > 0.05) are not significant. Interestingly, the interaction between the factors scenario and age is significant ($F_{3,264}$ = 2.97; $p$ < 0.05), as shown in **Figure 6**.

On the other hand, the same procedure applied to M2 reveals that the factors education ($F_{1,264}$ = 1.11; $p$ > 0.05), gender ($F_{1,264}$ = 0.46; $p$ > 0.05) and noise sensitivity ($F_{1,264}$ = 0.054; $p$ > 0.05) are not significant, while age ($F_{1,264}$ = 9.98; $p$ < 0.01) is a significant factor, as shown in **Figure 7**.

As can be seen in part 1, factor age itself has no statistical significance on M1. Still there is a very strong interaction between age and scenario. Younger participants made more errors in scenario 'park' (**Figure 6**). In part 2 of the experiment, age is a statistically significant factor, namely older participants made more mistakes than younger ones in all scenarios (**Figure 7**).

Furthermore, **Figure 8** shows the difference between results in part 1 and part 2, which suggests the effect of visual distraction on each age group in the four scenarios. A rather smaller variation among all four scenarios occurs in older participants.



FIGURE 5 | Proportion of the participants making mistakes in different scenarios of the aptitude experiment.



FIGURE 6 | Interaction between scenario and age on M1 mistakes (Age Group 2 is older than Age Group 1). ×: population marginal means significantly different.
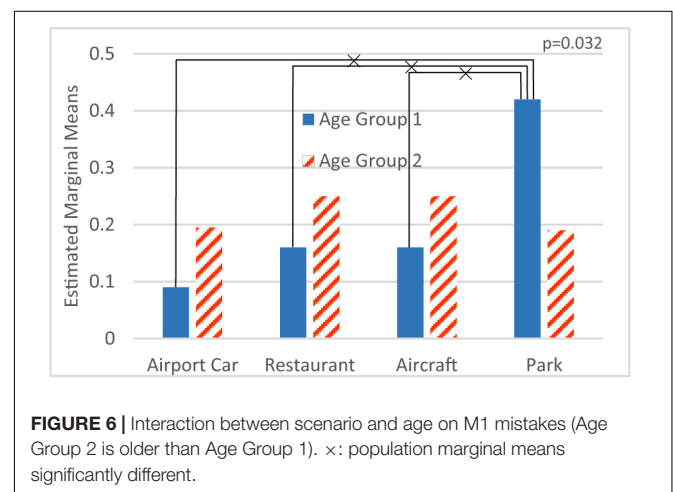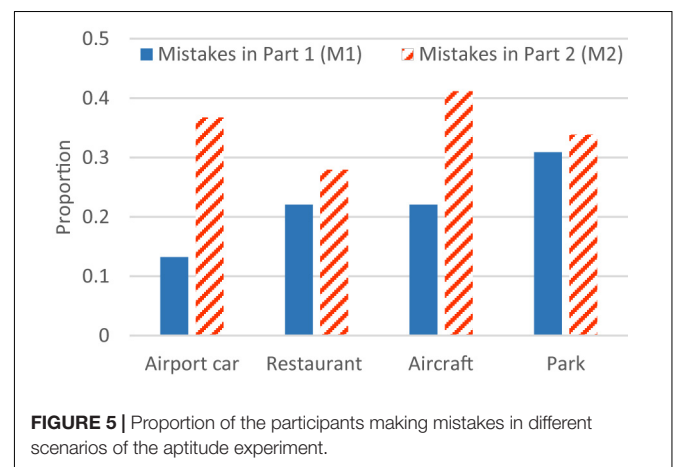
## Effect of Sound Features

The observation task in part 1 could be described as a pure sound deviant detection. The variation of results between each scenario (M1, **Figure 5**) should be ascribed to the sound itself. One feature that differs between scenarios is the total duration (%) of the attracting object (AO) stimuli, as shown in **Table 2**. A one-way anova test involving duration (%) as a factor on the results of M1 (on each participant) shows it has statistical

significance ($F_{3,264} = 2.54$; $p < 0.05$). In **Figure 9**, the correlation between AO duration (%) and M1 also supports the hypothesis that longer AO duration (%) decreases the difficulty of the sonic deviant detection task; the chance of making errors increases with decreasing duration.

In **Figure 5**, the difference between M1 and M2 suggests that the mistakes caused by the incongruent visual information also span a wide range: scenario 'airport car' has the biggest [$\Delta(M2 - M1) = 0.24$] and scenario 'park' has the smallest ($\Delta = 0.03$) effect. This trend (**Figure 10**) also applies to the other two scenarios – scenario 'aircraft' (duration of AO = 40%; $\Delta = 0.19$) and scenario 'restaurant' (duration of AO = 34.3%; $\Delta = 0.06$). Despite the correlation between the duration (%) of AO and M1 (**Figure 9**). **Figure 11** further shows the correlation between M1 and $\Delta$.

## Clustering by Audiovisual Aptitude

Combining the results of part 1 and part 2 in two dimensions (**Figure 12**) gives a clear view of the distribution of the participants. Participants were categorized into four groups. Group 1 (29.4%) are participants who made no mistakes in Part 1 but made at least one mistake after introducing the visual information (Part 2). Participants in group 2 (44.1%) made at least one mistake in both tests. On the contrary, group 3
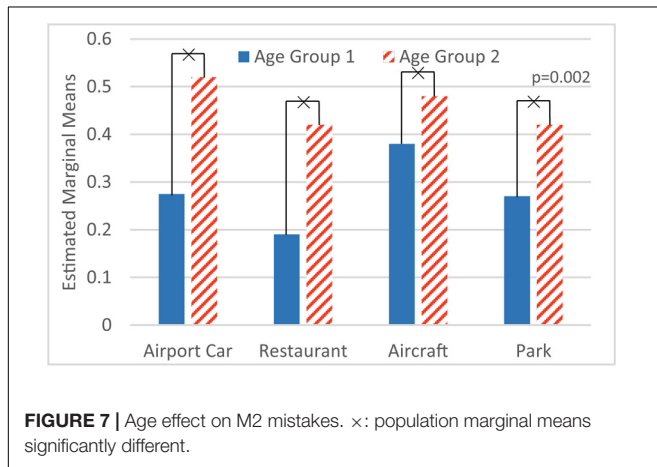


**FIGURE 7 |** Age effect on M2 mistakes. ×: population marginal means significantly different.
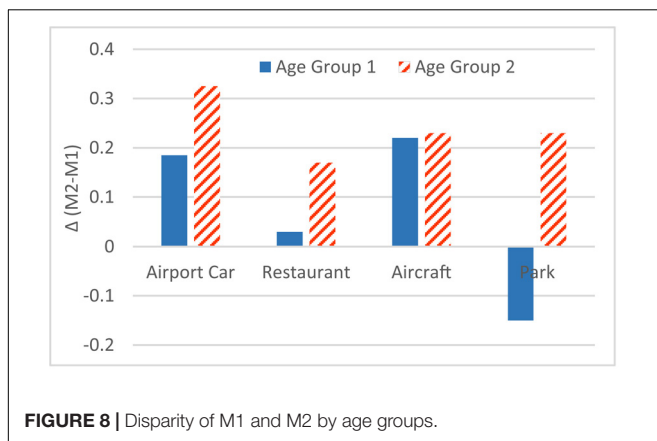


**FIGURE 8 |** Disparity of M1 and M2 by age groups.



**FIGURE 9 |** Correlation between duration (%) of AO stimuli and M1.



**FIGURE 10 |** The correlation between AO duration (%) and $\Delta$ (M2-M1) (disparity of M1 and M2).
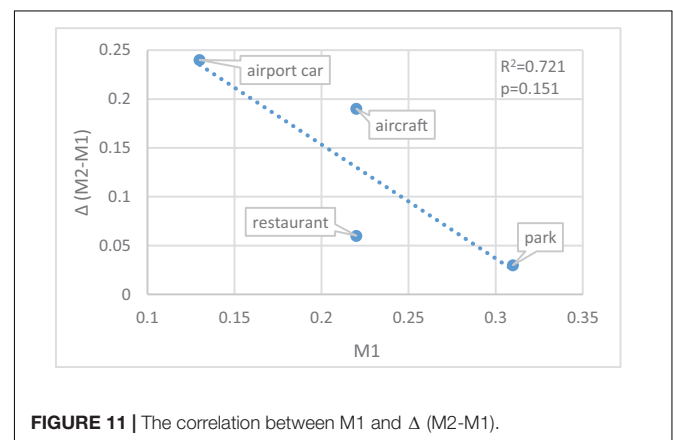


**FIGURE 11 |** The correlation between M1 and $\Delta$ (M2-M1).

FIGURE 12 | Participants grouping in the audiovisual aptitude experiment.

TABLE 4 | Comparison between models in living room experiment.

| Model | Aptitude clustering | | | | df | Information criterion (Akaike corrected) |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | | |
| 1 | A | B | B | B | 4 | 3961.255 |
| 2 | B | A | B | B | 4 | 3964.488 |
| 3 | B | B | A | B | 4 | 3961.430 |
| 4 | B | B | B | A | 4 | 3989.188 |
| 5 | A | A | B | B | 4 | 3990.073 |
| 6 | A | B | A | B | 4 | 3989.473 |
| 7 | A | B | B | A | 4 | 3988.186 |
| 8 | A | A | B | C | 5 | 3960.111 |
| 9 | A | B | A | C | 5 | 3987.032 |
| 10 | A | B | C | A | 5 | 4014.913 |
| 11 | A | B | B | C | 5 | 3991.336 |
| 12 | A | B | C | B | 5 | 3960.627 |
| 13 | A | B | C | C | 5 | 3991.185 |
| 14 | A | B | C | D | 6 | 3957.773 |
| 14+ | | | | | | 3934.948 |

TABLE 5 | Details of model 14+ in living room experiment.

| Fixed effects | Target: annoyance at home | | | |
|---|---|---|---|---|
| Source | F | df1 | df2 | Sig. |
| Intercept | 58.739 | 13 | 1.073 | 0.000 |
| Noise sensitivity | 6.663 | 1 | 1.073 | 0.010 |
| SPL | 242.440 | 3 | 1.073 | 0.000 |
| Noise sensitivity*Sound source | 6.003 | 2 | 1.073 | 0.003 |
| Audiovisual aptitude*Green | 2.451 | 7 | 1.073 | 0.017 |

* 'Participant' is used as random factor.

(14.7%) are participants who made no mistake in any of the tests. Participants in group 4 (11.8%) made at least one mistake in Part 1, but flawlessly performed after introducing the visual information (Part 2).

These four groups generally represent different reactions toward the audiovisual stimuli, which would affect the perception as in the task performance. In the following analysis of the second and third experiment, this classification of participants will be referred to as audiovisual aptitude.

## Effect of Audiovisual Aptitude on Annoyance at Home

Previous analysis of this experiment showed the dominating effect of the sound level on noise annoyance and a smaller influence of the window view (Sun et al., 2018). To test the effect of audiovisual aptitude, a generalized linear model was built targeting annoyance and involving only sound pressure levels and various ways of categorizing the four groups that were identified before. **Table 4** shows the comparison of models with different groupings, aiming at searching for the best model (with lowest information criterion). Model 14 is better than other models, even though it increases the degrees of freedom. More factors and interactions are included to model 14 using a stepwise adding/removing methodology. Statistical significance of model deviance reduction when including an additional variable has been checked by likelihood ratio testing (based on the Chi-square distribution). **Table 5** shows details of the best model (model 14+) with all statistically significant factors.

Even though audiovisual aptitude is not significant as a single effect due to the presence of more important factors (namely SPL and noise sensitivity), there is a strong interaction between audiovisual aptitude and visibility of green elements (see the window scenes of the living room, section "Experiment 2: Annoyance in Living Room"). Details of this interaction are shown in **Figure 13**. Persons from all aptitude groups are slightly



FIGURE 13 | The interaction between audiovisual aptitude and green elements visibility on annoyance. ×: population marginal means significantly different.

less annoyed when green elements are visible from the windows except in group 1. On the contrary, these persons that score very well on the purely auditory deviant detection task (Part 1, Experiment 1), but fail when an incongruent visual element is added (Part 2, Experiment 1), are less annoyed when a window scene without green elements is present.

## Effect of Audiovisual Aptitude on Perceived Quality of the Public Space

### Models for Perceived Quality

Analysis of the third experiment showed the strong effect of the visual bridge design and a more moderate effect of highway sound on the pleasantness rating (Echevarria Sanchez et al., 2017). In this it should be noted that sound was only changed in between days to deliberately hide changes. The same procedure as in the previous experiment is applied, using a generalized linear model now targeting pleasantness and involving only sound environment, bridge design, and audiovisual aptitude. As in the previous experiment, statistical significance of model deviance reduction has been checked by likelihood ratio testing. Model 14+ adding more interactions to model 14 using subsequent adding and removing of factors, further improved the model quality. Details are shown in **Tables 6**, **7**.

A strong interaction occurs between audiovisual aptitude and both bridge design and sound environment. In **Figure 14**, only people from aptitude group 2 have an increasing pleasantness rating with lower contribution of highway sound. Group 1 and 3 have a special preference for the sound environment with the 2nd and 3rd strongest contribution of highway sound, 68.6 dB(A)

TABLE 6 | Comparison between models in public space experiment.

| Model | Aptitude clustering | | | | df | Information criterion (Akaike corrected) |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | | |
| 1 | A | B | B | B | 7 | 4161.258 |
| 2 | B | A | B | B | 7 | 4134.640 |
| 3 | B | B | A | B | 7 | 4160.538 |
| 4 | B | B | B | A | 7 | 4160.429 |
| 5 | A | A | B | B | 7 | 4161.331 |
| 6 | A | B | A | B | 7 | 4161.570 |
| 7 | A | B | B | A | 7 | 4161.065 |
| 8 | A | A | B | C | 8 | 4160.176 |
| 9 | A | B | A | C | 8 | 4164.030 |
| 10 | A | B | C | A | 8 | 4160.841 |
| 11 | A | B | B | C | 8 | 4213.013 |
| 12 | A | B | C | B | 8 | 4160.962 |
| 13 | A | B | C | C | 8 | 4161.575 |
| 14 | A | B | C | D | 9 | 4133.550 |
| 14+ | | | | | | 4123.957 |

TABLE 7 | Details of model 14+ in public space experiment.

| Fixed effects | Target: pleasantness in public space | | | |
|---|---|---|---|---|
| Source | F | df1 | df2 | Sig. |
| Intercept | 12.582 | 27 | 1.060 | 0.000 |
| Bridge design | 63.038 | 3 | 1.060 | 0.000 |
| Sound environment | 2.670 | 3 | 1.060 | 0.046 |
| Audiovisual aptitude*Bridge design | 2.516 | 9 | 1.060 | 0.007 |
| Audiovisual aptitude*Sound environment | 2.502 | 9 | 1.060 | 0.008 |

*'Participant' is used as random factor.
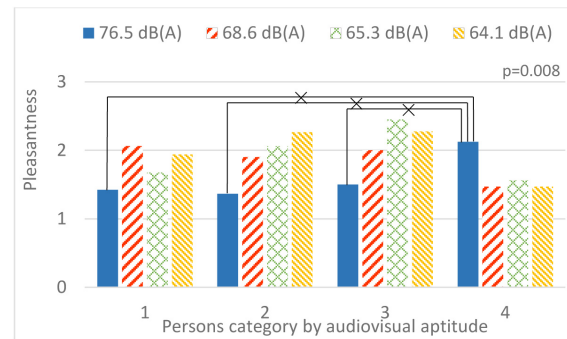


FIGURE 14 | The interaction between audiovisual aptitude and sound environment (highway SPL is used as a label) on pleasantness. ×: population marginal means significantly different.
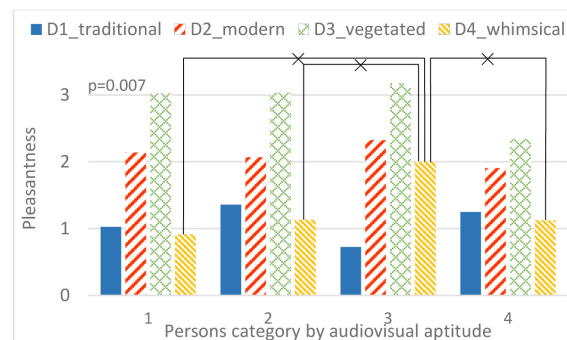


FIGURE 15 | The interaction between audiovisual aptitude and bridge design on pleasantness. ×: population marginal means significantly different.

and 65.3 dB(A), respectively. Oddly, people from group 4 prefer the sound environment with the strongest highway sound more than any others. In **Figure 15**, people in all aptitude groups show a common high appraisal of bridge design 3 (including vegetation, **Figure 4**, V3), followed by design 2. Designs 1 and 4 lead to relatively low pleasantness ratings, with design 4 being even slightly worse than design 1 for most people. However, the only exception is group 3 (those who performed without errors in the aptitude experiment, in both parts 1 and 2): design 4 is much higher rated than design 1. In addition, **Figure 16** shows the effect of audiovisual aptitude on pleasantness of the matching audiovisual combinations, namely the bridge design with the corresponding sonic environment. Persons from group 1, 2, and 3 share the similar trend, except for people from group 3 slightly preferring bridge 4 rather than bridge 2. However, for persons in group 4, bridge 4 is clearly the worst and the other three bridges do not differ from each other very much.

### Looking Behavior Study: The Gazing Time

A one-way anova test with factor bridge design and gazing time (total time, **Table 8**) shows this is a statistical significant factor ($F_{3,224} = 8.84$; $p < 0.01$). It reveals that at bridges 1 and 2 (**Figure 4**, V1 and V2), people tend to look more often and
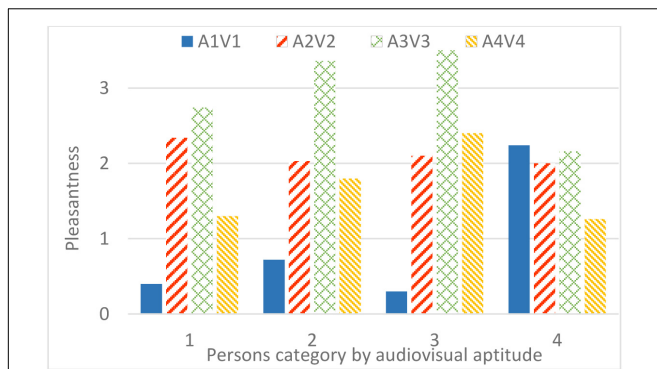
FIGURE 16 | Effect of audiovisual aptitude on pleasantness of matching audiovisual designs.
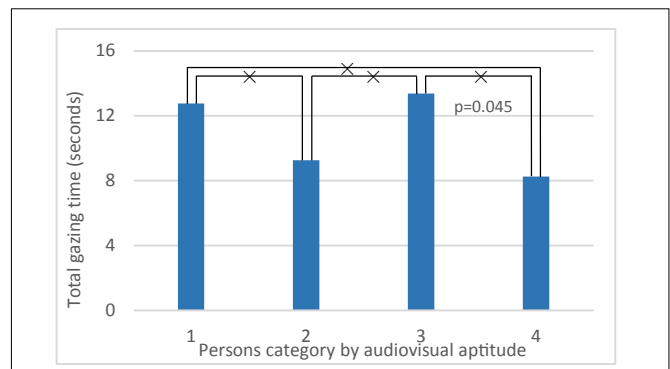


FIGURE 17 | Effects of audiovisual aptitude on total gazing time. ×: population marginal means significantly different.



FIGURE 18 | Effects of bridge designs on total gazing time. ×: population marginal means significantly different.

longer at the highway. These two bridges both contain rather low edge barriers, visually exposing the sound source directly. Also, in all four bridge designs, the average gazing time is longer than the median gazing time, which shows that participants who actually look at the highway traffic do this for a longer time.

An anova test targeting at total gazing time involving the factor bridge design and personal factors shows that education ($F_{1,220} = 3.03$; $p > 0.05$), gender ($F_{1,220} = 2.50$; $p > 0.05$), age ($F_{1,220} = 3.77$; $p > 0.05$), and noise sensitivity ($F_{1,220} = 0.04$; $p > 0.05$) have no statistical significance, while audiovisual aptitude ($F_{3,212} = 2.73$; $p < 0.05$) is significant. However, there is no strong interaction between the factors bridge design and audiovisual aptitude ($F_{9,212} = 0.72$; $p > 0.05$). Moreover, looking back at the overall pleasantness, no clear correlation between total gazing time and pleasantness is found ($F_{113,228} = 0.64$; $p > 0.05$).

Note that in this section, the four bridges not only differ from each other by visual design, but also the sound level from the highway is decreasing from bridge 1 (highest) to bridge 4 (lowest). **Figure 17** shows that persons in aptitude groups 1 and 3, who made no errors in Part 1 of audiovisual aptitude experiment (Experiment 1), look at traffic longer than the other two groups. **Figure 18** shows that bridge 1 and 2, which have a rather low barrier and thus higher highway noise levels, result in more gazing time than in case of the other two bridges.

## DISCUSSION

The goal of current study was to provide evidence for the existence of a personal factor that could influence the perception of landscape and soundscape and their interaction. For this purpose, an experiment (Experiment 1) was designed to explore the individual difference in capability for unraveling audiovisual stimuli and its distractibility from auditory acuity. This personal factor was labeled *audiovisual aptitude*. Two other experiments (Experiments 2 and 3) were re-analyzed involving this personal factor. We found that in Experiment 2, this individual difference

TABLE 8 | Total gazing time for each bridge design.

| Bridge designs | Gazing time | | | | | |
|---|---|---|---|---|---|---|
| | Total time (seconds) | | Number of times | | Average time (seconds) | |
| | Average | Median | Average | Median | Average | Median |
| 1 | 14.58 | 11.9 | 2.84 | 3 | 4.85 | 4 |
| 2 | 14.48 | 11.6 | 2.88 | 3 | 4.50 | 4.06 |
| 3 | 7.81 | 4.6 | 1.72 | 1 | 2.97 | 3.05 |
| 4 | 7.19 | 5.7 | 1.53 | 1 | 3.83 | 2.95 |

modified the impact of window views on self-report noise annoyance in a living room context. In Experiment 3, this individual difference altered the impact of highway sound pressure level and visual bridge design on the pleasantness rating of a public space. It also affected the looking behavior during the perception of the public space.

Our audiovisual aptitude test categorizes people according to their ability to perform the purely auditory test at one hand and the audiovisual test at the other. It is a rather strict way of grouping participants in four groups. For instance, aptitude group 3 does not allow a single mistake. Each of the groups identified in **Figure 12** can be characterized in more detail and the underlying reasons for people to belong to this group may be explored. This also makes the definition of the factor audiovisual aptitude more precise.

For persons in aptitude group 1, incongruent visual information interferences the performance on the auditory task for the average person. They perform very well on the blind auditory test but start making mistakes once incongruent visual information is presented to them simultaneously. Macdonald and Lavie highlighted the level of perceptual load in a visual task as a critical determinant of inattentional deafness, an equivalent of inattentional blindness (Macdonald and Lavie, 2011). Persons in this group were successful in the sound deviant task with a low visual perceptual load (black screen, Part 1), but failed when the visual perceptual load increased (Part 2) which could be explained by being more vulnerable to inattentional deafness. Collignon et al. (2008) suggested the possibility of visual dominance in emotional processing under incongruent auditory and visual stimuli. However, this visual dominance in affect perception does not occur in a rigid manner, namely the visual dominance will disappear if the reliability of the visual stimuli is diminished (Collignon et al., 2008). The reliability of visual and auditory information influences the cross-modal asymmetry effects in temporal perception (Wada et al., 2003).

Group 2 contains most of the participants in this study. Although they often detect deviant auditory stimuli correctly with or without visual information, they make at least one error in both tasks with a slight tendency of making more errors when visual incongruent information is present (**Figure 12**). The complexity of the test arises either from the cocktail party effect (Conway et al., 2001) or the visual distraction effect on perception (Simons and Chabris, 1999). Both phenomena have been identified before. Hearing damage, even at a level where people would not report hearing problems or tonal audiometry does not show significant threshold shifts, could still cause reduced auditory scene analysis capacity (Füllgrabe et al., 2015). Auditory neuropathy has recently been identified as one possible cause (Bharadwaj et al., 2014). Although the age of the participants in this study does not warrant expecting a high incidence of hearing damage, some participants could clearly have more difficulties in performing the test. Also at the cognitive level we can expect some groups to perform worse (Edwards, 2016).

Persons in group 3 succeed in detecting the deviant sound in each of the four situations regardless of the presence of

incongruent visual information. They could be labeled hearing specialists and are probably auditory dominated. Noise sensitivity was found before to be moderately stable and associated with current psychiatric disorder and a disposition to negative affectivity (Stansfeld, 1992), which is at least partly inherited (Heinonen-Guzejev, 2009). The present study included the Weinstein noise sensitivity survey. Persons in this group do not answer consistently different on this noise sensitivity questionnaire, which seems to indicate that another characteristic is measured by the proposed test. Other authors also noted that despite the fact that noise sensitivity has been established and widely applied in noise-related studies, it reveals only one personality trait. Miedema and Vos (2003) questioned the validity of ascribing noise sensitivity to a general negative affectivity among people. Recent research also showed that the personality had an independent effect on noise sensitivity (Shepherd et al., 2015).

Finally, group 4 contains people that seem to be helped by the incongruent visual information while detecting deviant sound environments. They are the smallest group in this study. For purely visual tasks, it was demonstrated that a single discrete visual distraction can improve the detectability of an unexpected object (Pammer et al., 2014). Yet, it is equally likely that the visual information gives them a clue on what sounds they need to listen for in the auditory deviant detection task. Some people may have acquired the skill to compensate for their inability to form auditory objects in an auditory scene analysis task via top down mechanisms grounded in visual information.

The usefulness of the personality factor identified by the proposed audiovisual test for understanding the perception of the soundscape, and specifically the interaction between the visual and the sonic environment in it, is illustrated with two experiments.

Experiment 2 focused on road traffic noise annoyance in a living room environment. Comparing predictive models showed that keeping the four groups identified above (as separate groups) explained the observations best. **Figure 13** further shows that participants belonging to aptitude groups 2, 3, and 4 reported less noise annoyance when green elements were visible from the window, which is consistent with many studies (Maffei et al., 2013; Van Renterghem and Botteldooren, 2016). However, persons belonging to group 1 behaved significantly differently. They reported more annoyance at the same noise exposure when green elements were shown in the window pane (**Table 3**). To explain these observations, it should first be noted that the green views in this case did not provide an appealing and readable green area following the reasoning in (Kaplan and Kaplan, 1989). Instead, it only served as a visual barrier between the window and a highway. For this reason, the positive effect found in other studies may be less pronounced or even reversed. The deviating influence of a green window view on the annoyance response in group 1 may be explained in several ways. Persons in this group were identified as visual dominant and the mediocre quality of the green may have a stronger negative effect on them. Such a green view is also incongruent with the sonic environment. Persons in aptitude group 1, which are easily distracted by incongruent visual information,

may value congruence more and experience the expectation gap more strongly. This expectation gap could confuse them and push them to reporting more annoyance by the traffic noise.

The evaluation of the pleasantness of crossing a bridge over the highway using virtual reality (Experiment 3) also revealed significant differences between the audiovisual aptitude groups. **Figure 16** shows that the most obvious group with deviant pleasantness evaluation is group 4. These participants value the audiovisual design 1 (without barrier) much more than other participants and at the same time they seem to find less pleasure in the green design (A3V3). To investigate further the reasons for this deviant rating, a closer investigation of **Figures 14**, **15** reveals that it is not the visibility of the source that makes the original situation (A1V1) more pleasurable but to some extent the higher highway noise level. However, the magnitude of the effect is much more pronounced in the physically matching situation. Thus, congruency of the audiovisual information seems to play a role. In the perceived restorativeness soundscape scale (PRSS) study, Payne pointed out that specific types of sounds and their associated meanings were more important in influencing the perceived restorativeness of the soundscape than its overall sound pressure level (Payne, 2013). Considering the relatively lower pleasantness rating of the green design (A3V3) in group 4 compared to the other groups, the effect in this case seems better explained by the lower pleasure rating of the visual design (D3) as seen in **Figure 15**. Combining all of these observations leads to the hypothesis that persons belonging to group 4 value congruency of audiovisual information and moreover prefer to see the highway that produces the sound they hear. This matches what could be expected by the description of possible traits within this group 4 given above: these people need visual information to understand the auditory scene. Not having this information leads to a lower pleasantness rating.

Also group 3 shows deviant pleasantness ratings, in particular they value the design including a high noise barrier (A4D4) more than others (**Figure 16**). Looking at **Figures 14**, **15** it becomes clear that this is caused by a significantly higher pleasantness rating of visual design 4 even if averaged over combinations with different highway sound levels. Earlier, this group was identified as hearing specialists, persons that are very skillful in identifying deviant sounds and that do not get misled by incongruent visual information. At first sight, this may contradict the observation that the bridge design 4 is rated more pleasantly even if combined with different highway noise levels. However, the hypothesis is forwarded that seeing the high noise barrier already induces the feeling that highway noise will be mitigated, a fact that is highly appreciated by this group.

In addition, **Figure 14** shows that most participants (aptitude groups 1, 2, and 3) are following a trend of higher pleasantness rating with decreasing highway sound pressure level, despite the small difference between them. Even though the experiment was conducted on different days and the level difference can be as low as 1.2 dB(A), such a trend was still obtained. The presence of sounds that can create a frame of reference such as footsteps and a tram pass by could explain this (Echevarria Sanchez et al., 2017).

The virtual reality method used in Experiment 3 also allows to monitor the head movement of the participants in the study. Participants belonging to groups 1 and 3 turned their head significantly longer toward the cars on the highway. Participants in these groups make no errors on the auditory deviant detection task but may fail in the presence of incongruent visual information. Head movement is helpful in auditory scene analysis (Kondo et al., 2014), yet persons belonging to groups 1 and 3 are not expected to need this information as they are performing very well on the purely auditory test. A more plausible explanation for the observed difference between groups might be that it reflects a stronger focus on environmental sound.

Hence Experiments 2 and 3 show that the personal factor obtained from the aptitude experiment modifies perception of the audiovisual environment, both in a home setting and in the public space. This consistent and stable personal factor could be a potential modifier in studies on the interaction between visual and auditory information in perception experiments and could affect the way the urban environment is designed.

The core strength of the categorization should be ascribed to the aptitude experiment itself, so this experiment is analyzed in more detail. The test has been designed to assess the aptitude of participants in the auditory scene analysis step in auditory perception and to measure resistance against incongruent visual information. Indirectly it integrates an assessment of peripheral hearing status and attention focusing and gating capabilities of the person. For this reason, the test was based on ecologically valid and complex auditory and visual scenes rather than on more abstract test that are commonly used in psychology. This choice was made to maximize the probability of finding significant associations to the noise annoyance and public space perception. An appropriate test should be sensitive, reproducible, and easy to understand.

To guarantee sensitivity for all persons, the test consisted of four different contexts and deviants that could be more or less easily detected: then scenario 'airport car' would be the easiest one while scenario 'park' the hardest. This range in difficulty is mainly achieved by the duration (%) of AO stimuli as shown in Section "Effect of Sound Features." **Figure 10** indicates that in scenario 'airport car,' the monitoring task is relatively easy (perceptual load of the task is low), the visual distraction is sufficiently working. While vice versa, in scenario 'park,' the monitoring task is rather hard (perceptual load of the task is high), the visual distractor processing tends to be less pronounced. This comparison agrees with perceptual load theory (Lavie, 1995). **Figure 11** confirms that the more difficult the purely auditory task, the lower the influence of the visual distractor.

Furthermore, the sensitivity of the test for age of the participant reflects the sensitivity of the test. Earlier research suggested that older adults were more affected by irrelevant speech in a monitoring task (Bell et al., 2008). The age deficits occurred in many conditions and increased with the similarity of distractor and target (Scialfa et al., 1998). Cohen and Gordon-Salant (2017) also stated that older adults may be more susceptible to irrelevant auditory and visual competition in a real-world environment. Some research has shown that older and younger persons obtained similar performance with purely

auditory stimuli, but older adults have poor performance with audiovisual modality (Sommers et al., 2005). These findings are congruent with the presented study, as stated in Section "Effect of Personal Factor." However, in part 1 of the audiovisual aptitude experiment, younger participants made less mistakes in all scenarios except for scenario 'park' (**Figure 6**). In **Figure 8**, the smaller variation in older participants suggests that the visual distraction tends to have a more equalized effect on them. However, for younger participants, there's a bigger difference between scenarios, which might indicate that the visual distraction process highly depends on the context for younger people. Early research showed the effect of sound familiarity on recognition (Cycowicz and Friedman, 1998), which could suggest a large part of younger participants in this experiment were unfamiliar with a natural sonic environment.

The latter observation could lead to poor reproducibility of the test in another group of persons with different familiarity with the audiovisual scenes that are presented. This could be a plie for choosing a more abstract audiovisual test. The reported experiments were intended to show the existence of a difference in audiovisual aptitude between persons that could affect perception of the sonic and visual environment. It nevertheless has some limitations. An auditory deviant detection test with a limited number of scenarios will not reveal the full truth of above-mentioned hypothesis. The scenarios may not have been optimally chosen to balance familiarity with the environment amongst all participants. In addition to the age influence, other demographic factors may lead to a change in behavior in specific scenarios. For such an experiment, the number of participants matches widespread practice. However, using larger test populations may uncover other and more subtle influences and relationships. Also the verification – Experiments 2 and 3 – has certain shortcomings. In Section "Looking Behavior Study: The Gazing Time," for instance, the head movement was used as a proxy for eye movement since no eye tracer, compatible with the VR headset, was available at the time of the experiment.

## CONCLUSION

Our study provides evidence for the existence of a personal factor that influences the effect of the view from a living room window on perceived noise annoyance by highway traffic noise and the effect of both the visual design and the highway noise level on perceived pleasantness of crossing a bridge over a highway. This personal factor, which we labeled audiovisual aptitude, may explain differences in perception of the (audiovisual) environment observed in other studies. It was shown that this personal factor differs from noise sensitivity, a

known personality trait. It could become as important as noise sensitivity in understanding differences in perception of the living environment when both landscape and soundscape matter.

In this work, a deviant detection experiment was used to categorize persons according to their audiovisual aptitude. It was shown that categorization in four groups resulted in more performant models for predicting the above-mentioned influences than using less groups. Each group could be linked to personal factors identified previously in literature. Nevertheless, it can be expected that such an extensive test resulting in four groups might not be necessary. Based on the insights gained in this work, an audiovisual aptitude questionnaire may be constructed.

Future research may also focus on finding the neurological basis for the difference in audiovisual aptitude between persons. Recent research shows that high noise sensitivity is associated with altered sound feature encoding and attenuated discrimination of sound noisiness in the auditory cortex (Kliuchko et al., 2016). Audiovisual aptitude is expected to be related to attention moderated auditory scene analysis.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Good Clinical Practice (ICH/GCP), Commission for Medical Ethics [registration number BE670201628136 (31-03-2016)] with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Commission for Medical Ethics.

## AUTHOR CONTRIBUTIONS

KS and GES carried out the experiments under the supervision of BDC, TVR, and DB. KS performed the analytic calculations. KS took the lead to wrote the manuscript. All authors provided critical feedback and helped to shape the research, analysis and manuscript.

## FUNDING

## REFERENCES

Abbott, L. C., Taff, B. D., Newman, P., Benfield, J. A., and Mowen, A. J. (2016). The influence of natural sounds on attention restoration. *J. Park Recreat. Admi.* 34, 5–15. doi: 10.18666/JPRA-2016-V34-I3-6893

Apthorp, D., Alais, D., and Boenke, L. T. (2013). Flash illusions induced by visual, auditory, and audiovisual stimuli. *J. Vis.* 13:3. doi: 10.1167/13.5.3

Axelsson, Ö., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *J. Acoust. Soc. Am.* 128, 2836–2846. doi: 10.1121/1.3493436

Beaman, C. P. (2004). The irrelevant sound phenomenon revisited: what role for working memory capacity? *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 1106–1118. doi: 10.1037/0278-7393.30.5.1106

Bell, R., Buchner, A., and Mund, I. (2008). Age-related differences in irrelevant-speech effects. *Psychol. Aging* 23, 377–391. doi: 10.1037/0882-7974.23.2.377

Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., and Shinn-Cunningham, B. G. (2014). Cochlear neuropathy and the coding of supra-threshold sound. *Front. Syst. Neurosci.* 8:26. doi: 10.3389/fnsys.2014.00026

Botteldooren, D., Andringa, T., Aspuru, I., Brown, A. L., Dubois, D., Guastavino, C., et al. (2015). "From sonic environment to soundscape," in *Soundscape and the Built Environment,* Vol. 36, eds K. Jian, S. F. Brigitte, and R. Boca (Boca Raton, FL: CRC Press), 17–42. doi: 10.1201/b19145-3

Botteldooren, D., De Coensel, B., and De Muer, T. (2006). The temporal structure of urban soundscapes. *J. Sound Vib.* 292, 105–123. doi: 10.1016/j.jsv.2005.07.026

Brown, A. L. (2012). A review of progress in soundscapes and an approach to soundscape planning. *Int. J. Acoust. Vib.* 17, 73–81. doi: 10.20855/ijav.2012.17.2302

Cartwright-Finch, U., and Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition* 102, 321–340. doi: 10.1016/j.cognition.2006.01.002

Cohen, J. I., and Gordon-Salant, S. (2017). The effect of visual distraction on auditory-visual speech perception by younger and older listeners. *J. Acoust. Soc. Am.* 141, EL470–EL476. doi: 10.1121/1.4983399

Colavita, F. B. (1974). Human sensory dominance. *Atten. Percept. Psychophys.* 16, 409–412. doi: 10.3758/BF03203962

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., et al. (2008). Audio-visual integration of emotion expression. *Brain Res.* 1242, 126–135. doi: 10.1016/j.brainres.2008.04.023

Conway, A. R., Cowan, N., and Bunting, M. F. (2001). The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychon. Bull. Rev.* 8, 331–335. doi: 10.3758/BF03196169

Cycowicz, Y. M., and Friedman, D. (1998). Effect of sound familiarity on the event-related potentials elicited by novel environmental sounds. *Brain Cogn.* 36, 30–51. doi: 10.1006/brcg.1997.0955

De Coensel, B., Botteldooren, D., De Muer, T., Berglund, B., Nilsson, M. E., and Lercher, P. (2009). A model for the perception of environmental sound based on notice-events. *J. Acoust. Soc. Am.* 126, 656–665. doi: 10.1121/1.3158601

Echevarria Sanchez, G. M., Van Renterghem, T., Sun, K., De Coensel, B., and Botteldooren, D. (2017). Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space. *Landsc. Urban Plan.* 167, 98–107. doi: 10.1016/j.landurbplan.2017.05.018

Edwards, B. (2016). A model of auditory-cognitive processing and relevance to clinical applicability. *Ear Hear.* 37, 85S–91S. doi: 10.1097/AUD.0000000000000308

Ellermeier, W., and Zimmer, K. (1997). Individual differences in susceptibility to the "irrelevant speech effect". *J. Acoust. Soc. Am.* 102, 2191–2199. doi: 10.1121/1.419596

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Atten. Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Filipan, K., Boes, M., De Coensel, B., Lavandier, C., Delaitre, P., Domitrovi⊠, H., et al. (2017). The personal viewpoint on the meaning of tranquility affects the appraisal of the urban park soundscape. *Appl. Sci.* 7:91. doi: 10.3390/app7010091

Fougnie, D., and Marois, R. (2007). Executive working memory load induces inattentional blindness. *Psychon. Bull. Rev.* 14, 142–147. doi: 10.3758/BF03194041

Füllgrabe, C., Moore, B. C., and Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6:347. doi: 10.3389/fnagi.2014.00347

Giard, M. H., and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490. doi: 10.1162/089892999563544

Gibson, J. J., and Pick, A. D. (1963). Perception of another person's looking behavior. *Am. J. Psychol.* 76, 386–394. doi: 10.2307/1419779

Graham, E. R., and Burke, D. M. (2011). Aging increases inattentional blindness to the gorilla in our midst. *Psychol. Aging* 26, 162–166. doi: 10.1037/a0020647

Heinonen-Guzejev, M. (2009). *Noise Sensitivity Medical, Psychological and Genetic Aspects.* Doctoral dissertation, University of Helsinki, Helsinki.

ISO (2014). *ISO 12913-1:2014 Acoustics — Soundscape — Part 1: Definition and Conceptual Framework.* Geneva: International Organization for Standardization.

Jiang, Y., and Chun, M. M. (2001). Selective attention modulates implicit learning. *Q. J. Exp. Psychol. A* 54, 1105–1124. doi: 10.1080/713756001

Kahneman, D. (1973). *Attention and Effort*, Vol. 1063. Englewood Cliffs, NJ: Prentice-Hall.

Kaplan, R., and Kaplan, S. (1989). *The Experience of Nature: A Psychological Perspective.* New York, NY: Cambridge University Press.

Kliuchko, M., Heinonen-Guzejev, M., Vuust, P., Tervaniemi, M., and Brattico, E. (2016). A window into the brain mechanisms associated with noise sensitivity. *Sci. Rep.* 6:39236. doi: 10.1038/srep39236

Kondo, H. M., Toshima, I., Pressnitzer, D., and Kashino, M. (2014). Probing the time course of head-motion cues integration during auditory scene analysis. *Front. Neurosci.* 8:170. doi: 10.3389/fnins.2014.00170

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 451–468. doi: 10.1037/0096-1523.21.3.451

Lavie, N., and Fox, E. (2000). The role of perceptual load in negative priming. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1038–1052. doi: 10.1037/0096-1523.26.3.1038

Lavie, N., Lin, Z., Zokaei, N., and Thoma, V. (2009). The role of perceptual load in object recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1346–1358. doi: 10.1037/a0016454

Leung, T. M., Xu, J. M., Chau, C. K., and Tang, S. K. (2017). The effects of neighborhood views containing multiple environmental features on road traffic noise perception at dwellings. *J. Acoust. Soc. Am.* 141, 2399–2407. doi: 10.1121/1.4979336

Li, H. N., Chau, C. K., and Tang, S. K. (2010). Can surrounding greenery reduce noise annoyance at home? *Sci. Total Environ.* 408, 4376–4384. doi: 10.1016/j.scitotenv.2010.06.025

Macdonald, J. S., and Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Atten. Percept. Psychophys.* 73, 1780–1789. doi: 10.3758/s13414-011-0144-4

Mack, A., and Rock, I. (1998). *Inattentional Blindness.* Cambridge: MIT Press.

Maffei, L., Masullo, M., Aletta, F., and Di Gabriele, M. (2013). The influence of visual characteristics of barriers on railway noise perception. *Sci. Total Environ.* 445, 41–47. doi: 10.1016/j.scitotenv.2012.12.025

Miedema, H. M., and Vos, H. (2003). Noise sensitivity and reactions to noise and other environmental conditions. *J. Acoust. Soc. Am.* 113, 1492–1504. doi: 10.1121/1.1547437

Miller, Z. D., Hallo, J. C., Sharp, J. L., Powell, R. B., and Lanham, J. D. (2014). Birding by ear: a study of recreational specialization and soundscape preference. *Hum. Dimens. Wildl.* 19, 498–511. doi: 10.1080/10871209.2014.921845

Miller, Z. D., Taff, B. D., and Newman, P. (2018). Visitor experience of wilderness soundscapes in Denali national Park and Preserve. *Int. J. Wilderness* 2.

Musacchia, G., Sams, M., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15894–15898. doi: 10.1073/pnas.0701498104

Neisser, U., and Becklen, R. (1975). Selective looking: attending to visually specified events. *Cogn. Psychol.* 7, 480–494. doi: 10.1016/0010-0285(75)90019-5

O'Shea, D. M., and Fieo, R. A. (2015). Individual differences in fluid intelligence predicts inattentional blindness in a sample of older adults: a preliminary study. *Psychol. Res.* 79, 570–578. doi: 10.1007/s00426-014-0594-0

Pammer, K., Korrel, H., and Bell, J. (2014). Visual distraction increases the detection of an unexpected object in inattentional blindness. *Vis. Cogn.* 22, 1173–1183. doi: 10.1080/13506285.2014.987859

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

Pilcher, E. J., Newman, P., and Manning, R. E. (2009). Understanding and managing experiential aspects of soundscapes at Muir Woods National Monument. *Environ. Manag.* 43, 425–435. doi: 10.1007/s00267-008-9224-1

Sandhu, R., and Dyson, B. J. (2016). Cross-modal perceptual load: the impact of modality and individual differences. *Exp. Brain Res.* 234, 1279–1291. doi: 10.1007/s00221-015-4517-0

Scialfa, C. T., Esau, S. P., and Joffe, K. M. (1998). Age, target-distractor similarity, and visual search. *Exp. Aging Res.* 24, 337–358. doi: 10.1080/036107398244184

Shepherd, D., Heinonen-Guzejev, M., Hautus, M. J., and Heikkilä, K. (2015). Elucidating the relationship between noise sensitivity and personality. *Noise Health* 17, 165–171. doi: 10.4103/1463-1741.155850

Simons, D. J., and Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception* 28, 1059–1074. doi: 10.1068/p281059

Sommers, M. S., Tye-Murray, N., and Spehar, B. (2005). Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. *Ear Hear.* 26, 263–275. doi: 10.1097/00003446-200506000-00003

Sörqvist, P. (2010). Effects of aircraft noise and speech on prose memory: what role for working memory capacity? *J. Environ. Psychol.* 30, 112–118. doi: 10.1016/j.jenvp.2009.11.004

Sörqvist, P., and Rönnberg, J. (2014). Individual differences in distractibility: an update and a model. *Psych J.* 3, 42–57. doi: 10.1002/pchj.47

Stansfeld, S. A. (1992). Noise, noise sensitivity and psychiatric disorder: epidemiological and psychophysiological studies. *Psychol. Med. Monogr. Suppl.* 22, 1–44. doi: 10.1017/S0264180100001119

Sun, K., De Coensel, B., Echevarria Sanchez, G. M., Van Renterghem, T., and Botteldooren, D. (2018). Effect of interaction between attention focusing capability and visual factors on road traffic noise annoyance. *Appl. Acoust.* 134, 16–24. doi: 10.1016/j.apacoust.2018.01.001

van den Brink, R. L., Cohen, M. X., van der Burg, E., Talsma, D., Vissers, M. E., and Slagter, H. A. (2013). Subcortical, modality-specific pathways contribute to multisensory processing in humans. *Cereb. Cortex* 24, 2169–2177. doi: 10.1093/cercor/bht069

Van Renterghem, T., and Botteldooren, D. (2016). View on outdoor vegetation reduces noise annoyance for dwellers near busy roads.

*Landsc. Urban Plan.* 148, 203–215. doi: 10.1016/j.landurbplan.2015.12.018

Wada, Y., Kitagawa, N., and Noguchi, K. (2003). Audio–visual integration in temporal perception. *Int. J. Psychophysiol.* 50, 117–124. doi: 10.1016/S0167-8760(03)00128-4

Weinstein, N. D. (1978). Individual differences in reactions to noise: a longitudinal study in a college dormitory. *J. Appl. Psychol.* 63, 458–466. doi: 10.1037/0021-9010.63.4.458

Weinzimmer, D., Newman, P., Taff, D., Benfield, J., Lynch, E., and Bell, P. (2014). Human responses to simulated motorized noise in national parks. *Leis. Sci.* 36, 251–267. doi: 10.1080/01490400.2014.888022

World Medical Association (2001). World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. *Bull. World Health Organ.* 79, 373–374.

Zhang, B., Shi, L., and Di, G. (2003). The influence of the visibility of the source on the subjective annoyance due to its noise. *Appl. Acoust.* 64, 1205–1215. doi: 10.1016/S0003-682X(03)00074-4

Check for
updates

# The Evolution of Soundscape Appraisal Through Enactive Cognition

Kirsten A.-M. van den Bosch[1]*, David Welch[2] and Tjeerd C. Andringa[1,3]

[1] SoundAppraisal Ltd., Groningen, Netherlands, [2] School of Population Health, University of Auckland, Auckland, New Zealand, [3] University College Groningen, University of Groningen, Groningen, Netherlands

We propose a framework based on evolutionary principles and the theory of enactive cognition ("being by doing"), that addresses the foundation of key results and central questions of soundscape research. We hypothesize that the two main descriptors (measures of how people perceive the acoustic environment) of soundscape appraisal ('pleasantness' and 'eventfulness'), reflect evolutionarily old motivational and affective systems that promote survival through preferences for certain environments and avoidance of others. Survival is aimed at ending or avoiding existential threats and protecting viability in a deficient environment. On the other hand, flourishing occurs whenever survival is not an immediate concern and aims to improve the agent's viability and by co-creating ever better conditions for existence. As such, survival is experienced as unpleasant, and deals with immediate problems to be ended or avoided, while flourishing is enjoyable, and therefore to be aimed for and maintained. Therefore, the simplest, safety-relevant meaning attributable to soundscapes (audible safety) should be key to understanding soundscape appraisal. To strengthen this, we show that the auditory nervous system is intimately connected to the parts of our brains associated with arousal and emotions. Furthermore, our theory demonstrates that 'complexity' and 'affordance content' of the perceived environment are important underlying soundscape indicators (measures used to predict the value of a soundscape descriptor). Consideration of these indicators allows the same soundscape to be viewed from a second perspective; one driven more by meaning attribution characteristics than merely emotional appraisal. The synthesis of both perspectives of the same person–environment interaction thus consolidates the affective, informational, and even the activity related perspectives on soundscape appraisal. Furthermore, we hypothesize that our current habitats are not well matched to our, evolutionarily old, auditory warning systems, and that we consequently have difficulty establishing audible safety. This leads to more negative and aroused moods and emotions, with stress-related symptoms as a result.

Keywords: soundscapes, enactive cognition, evolutionary psychology, soundscape descriptors, soundscape indicators, audible safety, tranquility

## INTRODUCTION

In this paper, we will use the conceptual framework of enactive cognition to address the foundation of key results and central questions of soundscape research. We will propose a theory based on evolutionary psychology, which underlies the identification of pleasantness and eventfulness as important soundscape descriptors.

Traditionally, research on noise (defined as unwanted sound) has focused on the relation between adverse effects and acoustic parameters such as level in decibels, dB(A). Cardiovascular diseases are one of the most studied adverse effects of noise exposure and include: hypertension, high blood pressure, ischaemic heart disease, and myocardial infarction (Ising and Kruppa, 2004; World Health Organization [WHO], 2011). These effects tend to be predicated (albeit implicitly) on the noise-stress hypothesis, under which noise is a non-specific stressor that activates the autonomic nervous system and endocrine system. This stress response elicits changes in stress hormones such as cortisol and (nor)epinephrine, affecting the individual's metabolism, and increasing the risk for cardiovascular diseases. These effects seem to occur above noise levels around 65 dB(A) (Babisch, 2002; Ising and Kruppa, 2004). While these are valuable observations, they lack a suitable framework to explain the origins, effects, and workings of the noise-stress hypothesis. This theoretical basis is important, since it is becoming clear that auditory appraisal is greater than the sum of its decibels. In fact, the very definition of 'noise' as unwanted sound entails appraisal on the dimension of desirability that has no obvious relation to decibels.

The soundscape approach contributes to a growing body of research indicating that, for noise, it is not just objectively measurable signal properties, but the *meaning* attributed to it that has the most prominent effect on health (Ising and Kruppa, 2004). This coheres with phenomenological approaches to the relationship between individuals (or groups) and their environment (Von Uexküll, 1992/1934; Graumann, 2002) that focus on how meaning is constructed. From this perspective it is not surprising that merely one third of noise disturbance can be accounted for by acoustics alone (Guski, 2001). Research shows that sounds may be unpleasant due to the meaning attributed to them rather than their measurable energetic properties. Qualitatively unpleasant sounds (such as metal scraping on slate) can seem worse than electric shocks or neutral sounds presented at much higher levels (Neumann et al., 2008) and emotionally laden sounds elicit greater physiological responses (e.g., startle reflex, skin conductance) than neutral sounds of similar level (Bradley and Lang, 2000). Similarly, the mere reduction of noise levels does not necessarily lead to more positive appraisals of that environment (Adams et al., 2006; Dubois et al., 2006); on the contrary, it can even lead to (more) anxiety (Stockfelt, 1991).

By targeting the meaning of sound, soundscape research goes beyond the traditional focus on noise (Schulte-Fortkamp, 2002; Botteldooren et al., 2006; Cain et al., 2013) including both positive and negative effects on the perceiver. These effects could be attributed to very basic aspects of our perception. Auditory appraisal can even be seen as a basic requirement of life for humans as we have evolved, meaning it must be based on the environmental conditions for which our nervous systems evolved. The domain of enactive cognition (Varela et al., 1993; Thompson, 2007; Froese and Ziemke, 2009; Di Paolo et al., 2010) provides a conceptual framework to address questions related to the basic properties and role of soundscape, such as why pleasantness and eventfulness are crucial soundscape descriptors.

## COGNITIVE FOUNDATIONS OF THE SOUNDSCAPE CONCEPT

The enactive approach of cognition sets out with the observation that life on Earth consists of individuals that remain alive because they *do* things to avoid premature death. This can be summarized as "being by doing" (Froese and Ziemke, 2009) and an entity that does this, within the domain of enactive cognition, is referred to as 'an agent.' This holds for all life: in single or multicellular living agents (organisms like humans and plants) this same basic function requirement of "being by doing" needs to be fulfilled. According to Barandiaran et al. (2009, p. 367) agency is *"an autonomous organization that adaptively regulates its coupling with its environment and contributes to sustaining itself as a consequence."* This formal definition is a succinct formulation of a number of properties that living agents exhibit to remain alive and functioning. It entails the following:

1. An agent must not only be able to respond to the here and now, but be able to deal with its future demands as well. Agency is essentially anticipatory (Louie, 2010; Vernon, 2010).
2. Agents are continually adapting to the environment to ensure that they can sustain themselves. In practice, this entails that they satisfy their needs. Yet they are free to self-select their need satisfying behavior. This is called 'needful freedom' (Froese and Ziemke, 2009).
3. Unlike a rock or a hurricane, agents display (through their behavior) a measure of control over how they respond to and interact with the environment.

The last property goes to the core of what it entails to be alive: agents act differently in different situations and the decision to act resides within the agent itself. While an inanimate object is only subject to external forces, an agent is a source of self-controlled modifications of its relation to the environment. In other words, it is agentic (**Figure 1**) (Barandiaran et al., 2009).

Agents sense the environment via specialized sensory systems which alter the internal state in response to the relevant observations of the environment (Egbert et al., 2010). Depending on a combination of what is sensed and the agent's needs, behavior is selected. For example, this may entail the uptake of nutrients or a movement up or down some perceived gradient (Egbert et al., 2010). Evolution dictates that agents tend to optimize the functions of sensing and behavior so that outcomes are beneficial for survival. The combined process of sensing, behavior selection, and behavior enaction contributing to the agent's continued existence and flourishing, is known as 'cognition' (Di Paolo and Thompson, 2014). From the

**FIGURE 1 |** Agency. Agency is an organism's ability to self-regulate its coupling to its habitat/environment through sensing and self-initiated activities (behavior). Adapted from Barandiaran et al. (2009).
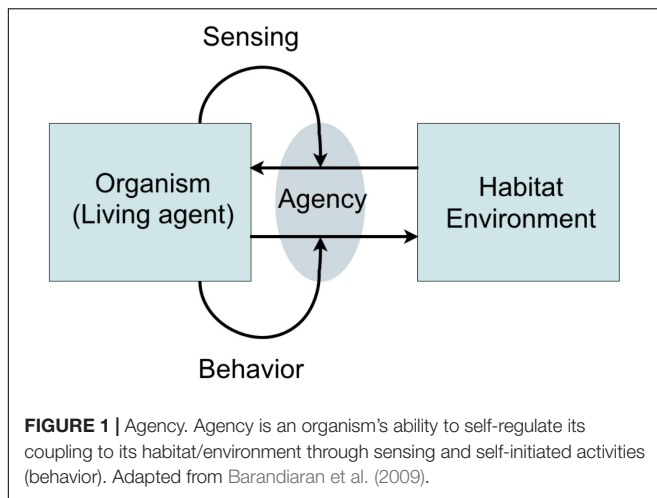
perspective of cognition, the environment may be described as the combination of potential benefit or harm to the agent, and the investments the agent must make to respond. This constitutes the 'affordances' of the environment. The term affordance was first coined by Gibson (1979, p. 127), where he defined it as follows: *"The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. [. . .] It implies the complementarity of the animal and the environment."*

Soundscape can be seen as the human sonic analog of this: for humans, the soundscape represents what the acoustic environment offers, provides, or furnishes the individual for good or ill. Definitions of soundscape refer accordingly to an *"acoustic environment as perceived or experienced and/or understood by a person or people, in context"* (International Organization for Standardization [ISO], 2014), or as Truax (1999, p. 126) described it, *"an environment of sound (sonic environment) with emphasis on the way it is perceived and understood by the individual, or by a society. It thus depends on the relationship between the individual and any such environment."* An environment's influence on agents depends upon the cognition it causes and the resulting meaning attribution in terms of affordances and the investments to realize them. It will be clear that the physical signals in the environment are a necessary precondition for meaning attribution. However, species specific innate processing capabilities, individual histories, social relations, and cultural knowledge usually dominate meaning attribution (Schafer, 1977). This implies that soundscape descriptors (measures of how people perceive the acoustic environment; Aletta et al., 2016) should reflect meaning attribution, as opposed to merely describing the physical properties of the sound itself (Cain et al., 2013). Such descriptors are addressed in the next section.

## SOUNDSCAPE DESCRIPTORS: PLEASANTNESS VERSUS EVENTFULNESS

In parallel with the arguments based on enactive cognition, Bradley and Lang (2000) found that the principal variance in

emotional meaning people give to sounds, can be explained by two (appetitive and defensive) motivational systems that underlie affective judgment; valence indicates which system is active, and arousal indicates the intensity of activation of these systems. Semantic descriptors of soundscapes appear to reflect a similar two-dimensional model for the underlying perceptual factors (Cain et al., 2013). Axelsson et al. (2010) named these 'Pleasantness' and 'Eventfulness.' Davies and Murphy (2012, p. 4) suggest that *"the weight of evidence in the literature is now sufficient for the first two dimensions of calmness/pleasantness and activity/eventfulness to be regarded as a 'standard model' for the perceptual dimensions of soundscapes."* which is supported by a recent review on soundscape descriptors by Aletta et al. (2016).

It is important here to note the subtle yet substantial difference between descriptors of the affective quality of the environment (pleasantness and eventfulness) and descriptors of emotional responses to the environment (valence and arousal). The soundscape depends upon a combination of environmental influences on our senses (especially hearing), the process of ascribing meaning to the sensation of those influences (which may be termed perception), and the cognitive-emotional responses to the perception. By the definition we have used, the perception is the soundscape. Therefore, the soundscape depends on acoustical environmental cues and gives rise to psychological responses such as affective states, feelings, and cognitions. Furthermore, these psychological responses can in turn influence the perception of that environment (Schafer, 1977). The notion of this reciprocal relationship is supported by *in vivo* research on the way humans appraise their (current) environment and in what way that influences how they feel, plan, and act (Kuppens et al., 2012). From the perspective of an agent, the soundscape is the internal representation of the (mostly) acoustic environment, and the psychological responses to it are not necessarily clearly distinguishable from the soundscape itself. Thus, it can be difficult to separate these elements when considering the response of a person to an acoustic environment. To illustrate this distinction (and the similarities), **Figure 2** shows the two different elements or categories of descriptors.

Since the main descriptors of affective quality of the environment (Pleasantness and Eventfulness) closely resemble the concept of 'core affect,' this concept is used here to depict descriptors of emotional responses to the environment. Furthermore, both are often visualized as a circumplex model allowing for a side by side comparison. Core affect defines basic affective feelings that are always present and is an integral blend of the dimensions Pleasantness (valence) and Activation (arousal) (Russell, 2003). Core affect is a relation to the world as a whole and not a relation with something specific in that world. Like moods, it does not have (or need) the intentionality (directedness) of emotions and it is, unlike emotions, continually present to self-report. Following Kuppens et al. (2012), core affect reflects one part of the bidirectional relationship, the appraisal of the environment the other part.

Until now we have shown that pleasantness and eventfulness emerge as key soundscape descriptors from scientific literature. However, we argue that our theoretical basis allows to derive the same result from first principles. According to Andringa et al.

**FIGURE 2 |** Core affect and appraisal. **(A)** Depicts core affect (Russell, 2003), while **(B)** depicts appraisal of the environment (Axelsson et al., 2010). Adapted from Andringa and Lanser (2013). The safety related remarks in the circle are addressed in Section "Audible Safety."
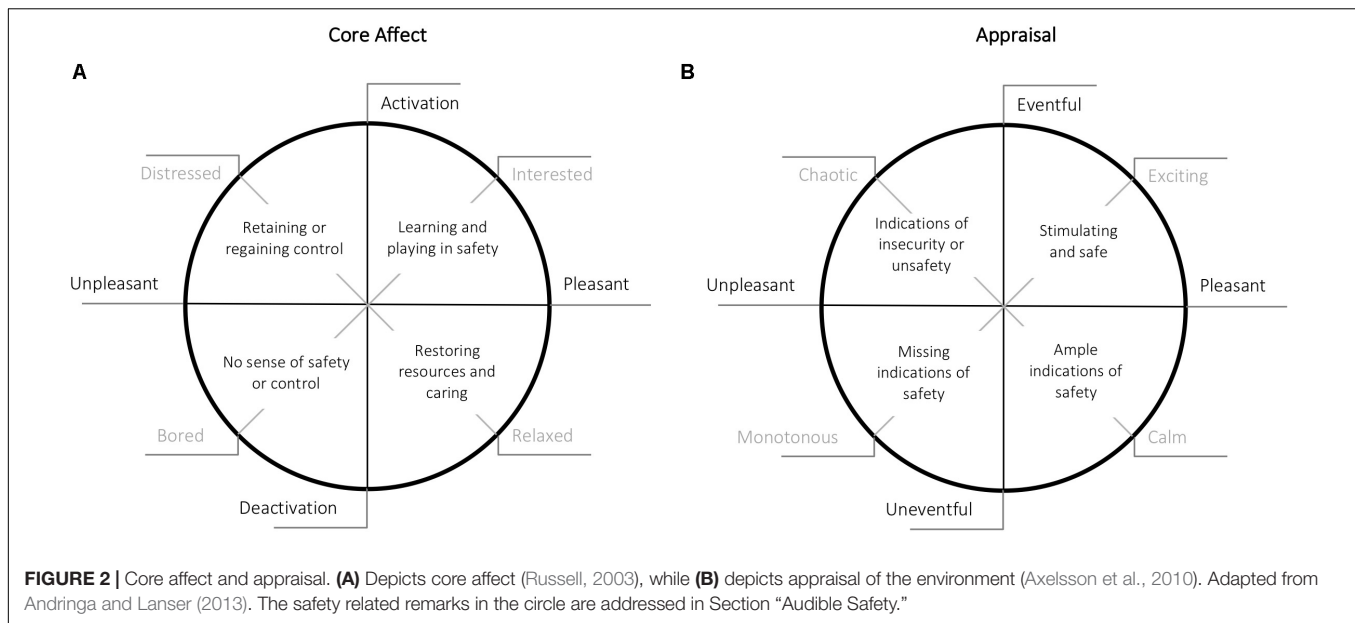
(2015), agents exist in a superposition of two modes of being: (1) Survival (coping mode) and (2) Flourishing (co-creation mode). Survival is aimed at ending or avoiding threats to existence and protecting viability in a deficient environment. It is essentially problem-oriented, reactive, and self-centered. Flourishing occurs when survival is not an immediate concern and aims to improve the agent's viability and to create ever better conditions for living (Fredrickson, 2001; Andringa et al., 2015). This corresponds to pervasive optimization, proactivity, and is environment-oriented, which has been connected to positive emotions and in particular to the broaden and build hypothesis (e.g., Fredrickson, 2001; Fredrickson and Branigan, 2005) using observations that positive emotions do not have a clear focus and broaden the scope of attention (Andringa et al., 2015).

We argue that the reactive survival (coping) mode is thus prevalent in low viability situations while the proactive flourishing (co-creation) mode is prevalent in high viability environments. As such, survival mode is experienced as unpleasant, and deals with immediate problems to be ended or avoided, while flourishing is enjoyable, and therefore to be aimed for and perpetuated (Andringa et al., 2015). These modes may be considered in terms of the two main descriptors of soundscapes: pleasantness and eventfulness. The absence of threats to survival and flourishing are perceived as pleasant states, whereas threats to survival or a lack of opportunities are unpleasant. Eventfulness is a dimension orthogonal to pleasantness and reflects the investment required to respond adequately to threats or opportunities. High investment environments lead to a high arousal level, while low investment environments allow low arousal.

To promote survival, our surroundings constantly influence our perception, cognition and emotions, even when we are not aware of it (Bitner, 1992). Therefore, as noted before, perception and the affective responses it elicits should not be considered separately: they are essentially intertwined (Kuppens et al., 2012).

Perception impels our basic emotions (Izard, 2007) and our emotions serve to establish our position in our environment; they attract us toward places, situations, and people, where we can flourish, and they repel us from situations where survival is threatened or where it is difficult to flourish (Levenson, 1999). This push and pull, attraction and rejection, evaluation in terms of positive and negative, beautiful and ugly, good, and bad, is a central part of our lives and a cross-cultural phenomenon (Osgood, 1975). Wundt (1897) referred to this as affect, and he argued that these subjective experiences, or impressions of the world, in terms of good or bad (valence) are the most pervasive aspect of human perception. Similarly, Russell (2003) has described core affect as the heart of all affective experiences. The full range of highly positive and deeply negative emotional meanings that people attribute to sounds (Bradley and Lang, 1999) arises from an interaction between the listener, the listener's attitude toward the sound source, the sound source itself, and other context (Tajadura-Jiménez, 2008). These insights describe a deep and essential mutual influencing of the state of the individual and the appraised environment, which implements the notion of agency as defined in **Figure 1**. In fact, the variety of relations between individual and environment as described in the previous paragraph and the key results of soundscape research are all perspectives on agency.

## AUDIBLE SAFETY

The abovementioned findings suggest that environments are processed based on characteristics beneficial for survival and below we outline why this assumption holds for auditory perception. Hearing is a universal sense (Horowitz, 2012) since no animal species has evolved without an acute sense of hearing (unlike vision), and it has an evolutionary history of several hundreds of millions of years (much older than vision; Hester,

2005). Considering that, the auditory system's most important function and original *raison d'être* (with respect to other senses) would then be to estimate danger and safety (Juslin and Västfjäll, 2008; Andringa and Lanser, 2013; Andringa and van den Bosch, 2013). Sound is perceived omnidirectionally, independently of lighting, physical obstructions, or wakefulness, and has strong attention-capturing power (Fritz et al., 2007). Furthermore, humans have an attentional bias for sounds outside the visual field (Tajadura-Jiménez et al., 2010b), with such sounds eliciting more arousal and larger physiological responses (Tajadura-Jiménez et al., 2010a), and humans have faster reaction times to auditory than to visual stimuli (Jasìkowski et al., 1990). These findings, together with our proposed evolutionary perspective, imply that audible safety might be the central element in the appraisal of our acoustic environments.

In line with this, the auditory nervous system is intimately connected to the parts of our brains associated with arousal and emotions (**Figure 3**). The reticular formation is a distributed network of nuclei in the brainstem and has control over arousal and many aspects of brain activity (Brown et al., 2012). Inputs from the most peripheral nuclei in the auditory pathway, the cochlear nucleus and superior olivary complex, innervate the reticular system's caudal pontine nucleus (Koch and Schnitzler, 1997). This operates in parallel and interactively with the classical auditory pathway to influence our experience of sound, and is also involved in other sensory systems, initiation and control of motor activity, autonomic arousal, sleep and wakefulness, and emotions (Siegel and Sapru, 2011). The system provides one mechanism for the emotional impact of sound, and it may influence the physiological and thus the emotional response to stimuli that have salience for survival and are treated as important.

The midbrain mediates freezing and flight in the face of alarming sounds as well as containing the limbic system, where emotional responses are mediated (Spreng, 2000; Ising and Kruppa, 2004; Kraus and Canlon, 2012). The auditory system projects information via the inferior colliculus and the medial geniculate nucleus of the thalamus, to the auditory cortex. The inferior colliculus, with involvement from the auditory cortex, directs flight from sudden, loud sounds via the superior colliculus and the periaqueductal gray (Xiong et al., 2015). The medial geniculate nucleus also projects to the amygdala, where emotional valence is attributed to sound (Kraus and Canlon, 2012). Furthermore, the amygdala itself has projections back into the auditory system (the inferior colliculus), implying that there may be modulation of auditory signals depending upon the emotional/meaningful/safety-related content in them (Marsh et al., 2002).

These observations allow us to propose that the brain constantly responds not only to the acoustic aspects of sounds but also to deeply programmed affectual, arousing, and attention-grabbing aspects of sounds. These two aspects of the response to sound occur in parallel and with feedback. From the perspective of the (human) agent, the two aspects of the percept (perception of the acoustics and meaning attribution) are inextricable. This allows to design for desired forms of audible comfort by separating the attention-grabbing foreground from a background

that continually provides us with a sense of place. If this is a sense of a *safe* place – because the midbrain is able to estimate ample indicators of safety – the listener is allowed full freedom and control to self-regulate mind-states according to needs (Andringa and Lanser, 2013; Andringa and van den Bosch, 2013).

In **Figure 2**, the relation between indicators of (audible) safety, affective appraisal of soundscapes, and core affect is illustrated. Here, it can be seen that pleasantly appraised environments co-occur with a pleasant inner affective state, proactive behavior, and (at least) ample indicators or safety. In the absence of indications of safety or presence of indications of unsafety, an environment is perceived as unpleasant, on which the agent will reply with reactive behavior, to avoid or end an unpleasant inner affective state (core affect). More specifically: a calm environment affords ample indications of safety that allow us to restore our resources and to care for self and environment; a lively environment is a stimulating and safe place that allows us to learn and play; a boring environment misses indications of safety, which does not afford a sense of safety or control; and a chaotic environment contains clear indications of insecurity or unsafety and forces to retain or regain control.

To summarize, we hypothesize that our appraisal of soundscapes is based on old survival-driven strategies, and we propose that the first (subconscious) decision made in the processing of auditory information is an assignment of safety by subcortical processes. Only in the case of a predominance of positive indicators of safety, will listeners have full freedom and control over mind-states. If not, part of the cognitive resources will be involved in general vigilance or directed attention to potential threats (Andringa and Lanser, 2013).

## SOUNDSCAPE INDICATORS: COMPLEXITY VERSUS AFFORDANCES

By accepting pleasantness and eventfulness as the main affective descriptors of soundscape appraisal, "*the hunt*" for the underlying indicators has begun; these are defined as "*measures used to predict the value of a soundscape descriptor*" (Aletta et al., 2016). Our evolutionary perspective and the concept of audible safety provide clues about them.

We propose that the second set of soundscape descriptors, calmness and excitement, as proposed by Axelsson et al. (2010) (or calmness and vibrancy as found by Cain et al., 2013), actually reflect the indicators 'Complexity' and 'Affordance Content,' respectively (Andringa and van den Bosch, 2013). This interpretation allows for an explanation that draws on our proposed evolutionary theory, while maintaining the essential two-dimensional space. Here, affordances are the threats and opportunities in an environment (Gibson, 1979) and indicate the extent to which the environment offers options for self-selected behavior and self-regulation (Andringa et al., 2015). The complexity of an environment refers to the number of competing auditory streams (Bregman, 1994), and thus how difficult it is to process the available affordance content (Axelsson, 2011) and to choose situationally appropriate behavior. The larger the search space and the smaller the set of beneficial options, the

FIGURE 3 | Schematic diagram of the classical auditory pathway (green arrows) and its associations with structures related to psychophysiological responses to soundscapes (red arrows).
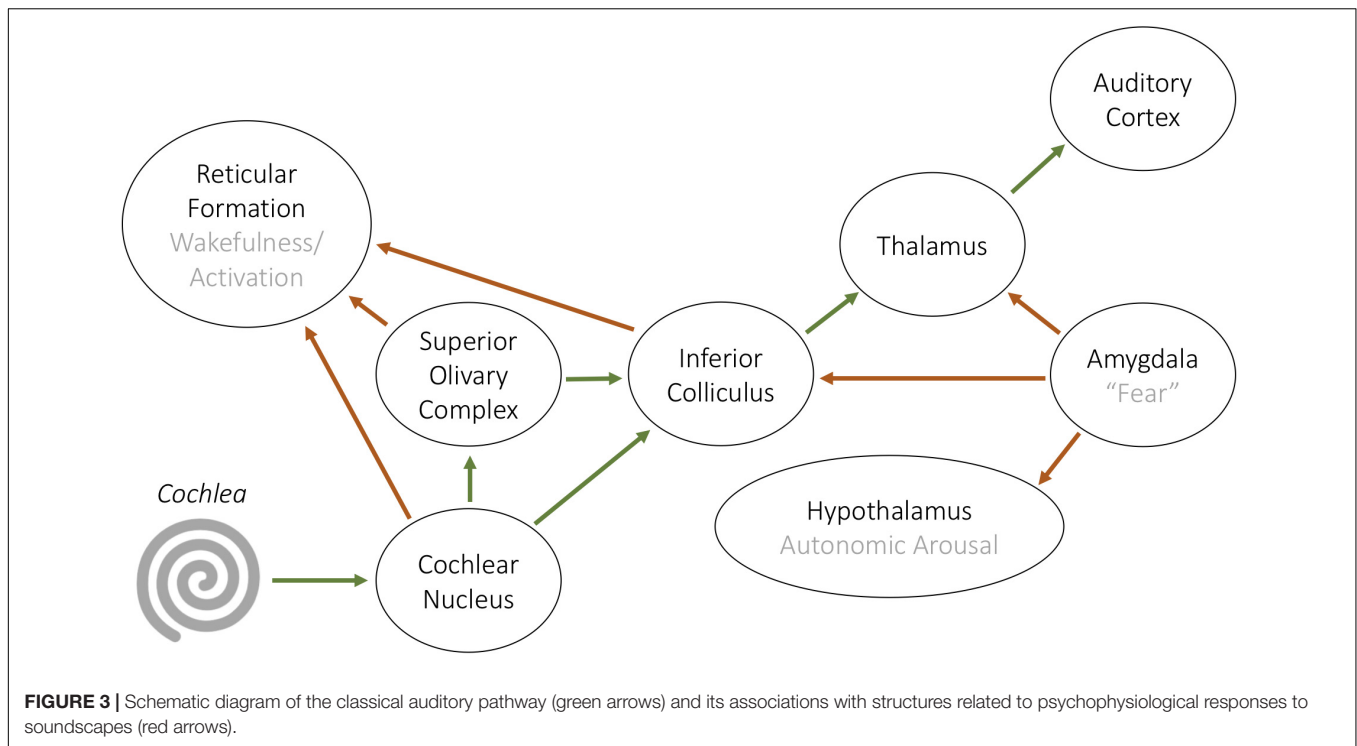
more complex and demanding the decision-making process that we refer to as 'meaning attribution.' The observation that the appraisal of the environment in part depends on the degree of perceived control (Russell, 2003) is illustrative of the influence complexity has on perception.

The new dimensional structure of indicators can be seen to have parallels in the prospect refuge theory (Appleton, 1975). Natural environments may be (visually) analyzed based on structural aspects such as, depth, threats, and opportunities (e.g., navigability, concealment), which elicit affective responses mediating adaptive behavior, and as such promote survival (Appleton, 1975; Ulrich, 1983; Greene and Oliva, 2009). Although the prospect refuge hypothesis was originally formulated for landscapes, soundscapes help us just as much in characterizing different environments (Pheasant et al., 2010) and determining survival relevant meaning. Schafer's (1977) definitions of high-fi and low-fi soundscapes was already suggestive of this function. A hi-fi soundscape has little overlap of the foreground sounds, and the sounds from the wider surroundings. This allows for a distant sonic horizon and a high signal-to-noise, or foreground-to-background, ratio. Alternatively, low-fi soundscapes are associated with an industrial, mechanized world and have sonic horizons that are much closer (Schafer, 1977). As such, a high-fi (often natural) soundscape is favorable for survival purposes since it makes the signal easier to process (Andringa, 2002), which reduces the processing complexity of its analysis.

To illustrate the above-mentioned findings, **Figure 4** integrates the main descriptors of soundscapes with the proposed underlying indicators and the relation to meaning attribution in terms of enactive cognition's central notion of



FIGURE 4 | Two-dimensional model of soundscapes with the main descriptors pleasantness and eventfulness, the underlying indicators complexity and affordances, and the relation to meaning attribution. The indicated degrees are used as a guideline for further explanation in **Table 1**.
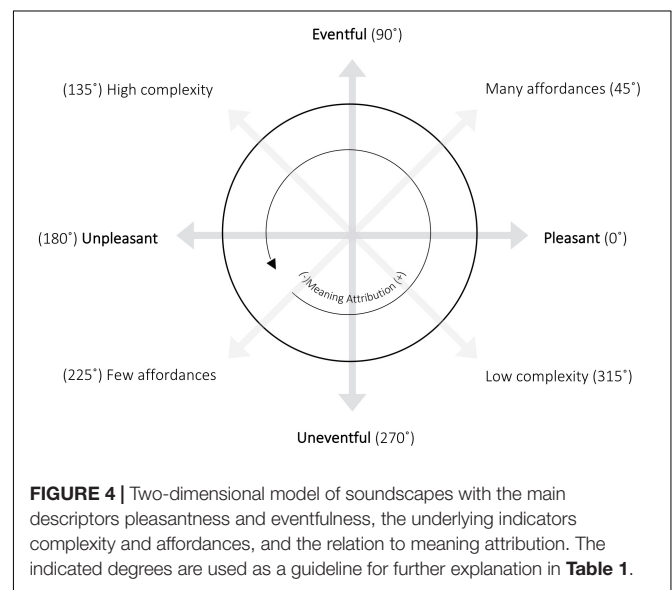
"being by doing." The horizontal axis represents the soundscape descriptor pleasantness (a measure of 'being'). The vertical axis represents the soundscape descriptor eventfulness (associated with 'doing'). The diagonal axes represent the indicators affordance content (need satisfaction benefits) and complexity (of action selection). Meaning attribution is a function of the indicators affordance content and complexity, as described by Axelsson (2011), and as such could actually be viewed as a (compound) soundscape indicator in itself, influencing the perceived quality of soundscapes.

**Table 1** elaborates on eight possible positions in this two-dimensional model and interprets them in terms of the meaning attribution. It uses the degrees as depicted in **Figure 2** and starts from 225°: what Eckblad (1981) referred to as the 'empty sector.' We do that because this sector is interpreted slightly differently when approached from the 180° direction, where it corresponds to an inability to attribute meaning, than from the 270° direction, where it corresponds to the absence of useful affordances (hysteresis). Both interpretations lead to the lowest rate of useful meaning to be estimated from the signal, in terms of satisfying agentic needs.

Note that the agent should always remain responsive to possible developments in the environment. This entails that it cannot spend more resources than it can muster before the situation changes. Perception is always under time pressure and hence processing resources are finite. Highly complex environments may change before meaning can be attributed reliably, which puts the agent under time-pressure to decide on the basis of insufficient information. If this is the case, the agent is unable to reliably determine audible safety and/or other relevant affordances. Alternatively, in an environment devoid of affordances, the perceiver is equally unable to determine audible safety and other meaning, however, much it searches for these. Hence, from 225° we go anti-clockwise via environments that become progressively more complex to environments after 90° that are so complex that they cannot be processed in full, and finish back at 225° in environments of which only superficial real-time meaning can be attributed.

## A PRACTICAL IMPLICATION: THE NEED FOR TRANQUILITY

Our evolutionary perspective on soundscapes allows the formulation of some practical implications which should be considered in the design of soundscapes. For example, we

**TABLE 1 |** Soundscape indicator-based descriptions of the two-dimensional model of soundscape appraisal as depicted in **Figure 4**.

| Angle | Description in terms of affordances (threats and opportunities), complexity (to analyze the sonic environment and select behavior), indicators of audible safety, viability, and investment | Meaning attribution interpretation |
|---|---|---|
| 225° | An absence of useful perceived affordances leads to a minimal search space for situationally appropriate behavior. Minimal agency and problematic viability. Medium complexity environment due to unsuccessful estimation of audible safety, for example because indicators of safety might be masked by other sounds. | Minimal meaning attribution in combination with unsuccessful safety estimation |
| 270° | Few, neither safe nor dangerous perceived affordances, deep relaxation associated with the absence of an urge to invest in interaction. Low complexity environment. Audible safety indicators either somewhat present or indicators of unsafety absent yet in principle easily audible because they are not masked by other sounds (silence). | Very little meaning to be attributed |
| 315° | Normal level perceived affordances with abundant indicators of safety and an (audible) absence of threats, which makes it very easy to select behavior. Minimal investment required in environmental interaction. Minimally complex environment. Allowing for full freedom and control to self-regulate mind-states according to needs and desires. | Meaning attribution very easy |
| 0° | Many perceived opportunities, ample indications of safety, maximally beneficial, highest viability, easy to select behaviors, yet requiring investment in environmental interaction. Low complexity environment. | Meaning attribution easy |
| 45° | Maximum level of perceived opportunities, ample indications of safety. This leads to a large search space in which a beneficial choice is neither crucial nor harmful. However, exploration of the rich affordances space requires a fairly high investment. Medium complexity environment. | Rich meaning attribution |
| 90° | High perceived affordances, but now not necessarily with safe outcomes and/or weaker indications of safety. Still large space for behavior selection, but with a smaller set of beneficial outcomes, which makes it difficult to select beneficial behaviors and avoid harmful outcomes. Requires maximal investment in environmental interaction. High complexity environment. | Meaning attribution challenging |
| 135° | Focus on potentially or actually *unsafe* perceived affordances (threats - indications of danger). Ignoring parts of the sonic environment, without sufficient indicators that these parts are safety irrelevant. Behavior selection is crucial to select the few options that are not harmful (and perhaps find the even fewer that are beneficial). Environmental interaction focused on these options. Maximally high complexity environment requiring a high investment. | Information overload, focus on subset, meaning attribution incomplete |
| 180° | Low level of perceived affordances. Threats and indications of danger are dominant and prevent an adequate analysis of potentially relevant content. Only few behavioral options are not dangerous. Behavioral choices become limited to the few that are not beneficial or even harmful. High complexity environment in which analysis efforts do not pay off, leading to a sense of agentic inadequacy. | Information overload and processing inability, attributed meaning attribution crucial but not satisfactory |
| 225° | Minimal level of perceived affordances all with minimal options, the behavioral selection search space may not include solutions so that the individual feels trapped and is subject to environmental influences. This is a medium complexity environment because the event rate to attend is not high, yet unable to address unfulfilled needs. This leads to a minimal sense of agency in an environment in which investment opportunities are low. | Meaning attribution unable to satisfy needs despite efforts |

propose that environments which are dominated by mechanical sounds, will effectively mask natural sounds that are preferred by our auditory sensory system to estimate audible safety. This is supported by findings indicating that mechanical sounds decrease perceived tranquility, and natural sounds enhance it (Pheasant et al., 2010). Similarly, findings by Darner (1966) demonstrated that mechanical sounds elicited unpleasant and alert feelings (as opposed to the sound of birds), and more recently Buxton et al. (2012) found that electronic sounds are more arousing than other sounds of similar loudness.

Our urbanized societies have become more mechanical, less harmonious, less predictable and controllable, leading to more negative appraisals of the (urban) soundscapes we live in (Davies et al., 2009). This results in a universal need for quietness (Pheasant et al., 2010; Booi and van den Berg, 2012), which can be explained by the Attention Restoration Theory of Kaplan (1995). The Attention Restoration Theory states that prolonged periods of (subconscious) directed attention lead to attentional fatigue, which needs to be recovered in restorative environments. This gains support from findings that restorative environments offering relief from sustained directed attention (associated with high complexity processing) are known to reduce stress and increase well-being (Hartig et al., 1997). For restoration, we need an alternate mode of attention, one that benefits recovery: fascination. It is proposed that natural environments are ideally suited for fascination because they are tranquil, leave a harmonic impression, and are rich, yet do not demand directed attention (Kaplan, 1995; Booi and van den Berg, 2012; Payne, 2013). We suggest this is due to the high redundancy of easy to process indications of audible safety in most natural environments. Therefore, our environments should offer more diversity through better access to green and natural spaces, especially in busy cities, so that people have access to tranquil (and audibly safe) soundscapes where they can recover from our cacophonous habitats (Booi and van den Berg, 2012).

## FUTURE DIRECTIONS

Pleasantness and eventfulness, and their indicators affordances and complexity, are predicted by the evolutionary cognitive theory we have described above. However, does this two-dimensional model truly and fully describe the soundscape, or might there be other important dimensions that could be predicted by the framework? It should account for all descriptors that would contribute to evolution, which includes dimensions of perceived affective quality such as pleasantness and eventfulness, but also descriptors from other categories. One candidate is 'appropriateness,' which has been mentioned in research Aletta et al. (2016). Soundscape appraisal is highly variable across intended activities (Nielbo et al., 2013; Steele et al., 2015), and expectations and appropriateness seem to play a significant role in the evaluation of soundscapes Aletta et al. (2016). Any activity or state draws on information schemes (Eckblad, 1981) and the encountered situation is matched against the existing cognitive schemes of information. A match between the scheme and the real-world situation leads to pleasant affective

responses to stimuli based on affirmation and security, whereas a mismatch (inappropriateness) leads to negative affective responses, confusion and insecurity. Appropriateness thus makes for very personal assessments of environments which are only suitable for specific situations or places (Brown et al., 2011). In the context of a soundscape, appropriateness would reflect the extent to which aspects of the acoustic environment matched the scheme in the mind of the listener. Sound elements which did not match (for example a car motor in a wilderness) would be perceived as inappropriate. In terms of evolutionary theory, the capacity to detect inappropriate elements would indeed be crucial for survival, and thus such a soundscape dimension may be expected to exist.

There are many other possible factors which appear to play a role in our appraisal of soundscapes. For example, a sense of pace or the passage of time, feelings of spirituality associated with the sonic environment, and an awareness of spaciousness, have all been identified using an essentially atheoretical approach to observing the soundscape (Welch et al., unpublished). Other research and theoretical work relating to the appreciation of loud music represents an understanding of an (artificial) soundscape, and concepts such as feelings of power or personal strength, and an experience of being transported to other worlds or imaginary realities have been reported (Blesser, 2007; Welch and Fremaux, 2017a,b). These qualities of the soundscape do not seem to be captured by the pleasantness/eventfulness dimensions and nor are they yet incorporated into the theoretical stance we have proposed here. Widening our understanding of the soundscape may be possible on both a practical and a theoretical level. On a practical level, we may gradually increase the dimensionality, or else learn how to apply different dimensionalities according to the physical/perceptual environment to allow these qualities to be incorporated.

On a theoretical level, we may be able to apply the evolutionary/cognitive approach we have proposed here to some of these other qualities. Alternatively, a compound theory which also draws upon other positions than the evolutionary may be necessary. Application of enactive cognition theory to explain the (apparently) more fundamental aspects of the soundscape (e.g., time) seems feasible. For example, any agent must operate with time constraints and we have therefore evolved to be able to do this. Our awareness of time passing would reflect an ability which evolved to allow us to make judgments about probabilities of survival and flourishing in future: this then may represent another theoretical dimension of our emotional appraisal, and may therefore provide a theoretical basis for future explorations of the soundscape. More careful thinking will be necessary to consider these possibilities, but the development for a strong theoretical basis to help drive and interpret soundscape research is crucial.

## CONCLUSION

Based on an evolutionary theory in which agents are motivated to seek pleasant and avoid unpleasant environments with the intention to flourish, we have bolstered the theoretical

underpinning to a two-dimensional model of soundscape appraisal. We have shown that, according to our theory, (1) the main soundscape descriptors pleasantness and eventfulness arise by necessity and (2) that affordance content and complexity of behavior selection are underlying indicators of these soundscape descriptors. Our theoretical basis comprises the defining properties of life and cognition (as formulated in the domain of enactive cognition), which lead to the formulation of constraints and opportunities afforded by living in a sonic world that underpin the science of soundscapes. Since our auditory sensory system can be regarded as an important warning system, and people appraise their soundscapes based on the level of safety they attribute to them, we propose that the simplest, safety-relevant meaning attributable to soundscapes is of central importance in understanding human perception.

Our approach allows the same soundscape to be formulated from a second perspective; one driven more by meaning attribution characteristics than merely emotional appraisal. The synthesis of the proposed indicators and the most common descriptors of soundscapes provides both perspectives of the same person–environment interaction, which consolidates the affective, informational, and the activity related perspectives on soundscape appraisal. Furthermore, we hypothesize that our current habitats are not well matched to our, evolutionarily old, auditory warning systems, and that we consequently have difficulty establishing audible safety. This leads to more negative and aroused moods and emotions, with stress-related symptoms as a result. A return to more natural sounding environments, or the design of non-natural environments with less threatening and less impoverished qualities, is the best guarantee for providing environments that are optimized for human inhabitants.

## AUTHOR CONTRIBUTIONS

All authors contributed equally to this manuscript and developed the theory interactively. KvdB initiated the collaboration, wrote the first draft, and managed the writing process. DW provided most of the neurological information. TA added the perspective from enactive cognition.

## ACKNOWLEDGMENTS

## REFERENCES

Adams, M., Cox, T., Moore, G., Croxford, B., Refaee, M., and Sharples, S. (2006). Sustainable soundscapes: noise policy and the urban experience. *Urban Stud.* 43, 2385–2398. doi: 10.1080/00420980600972504

Aletta, F., Kang, J., and Axelsson, Ö. (2016). Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landsc. Urban Plann.* 149, 65–74. doi: 10.1016/j.landurbplan.2016. 02.001

Andringa, T. C. (2002). *Continuity Preserving Signal Processing*. Ph.D. thesis, University of Groningen, Groningen.

Andringa, T. C., and Lanser, J. J. L. (2013). How pleasant sounds promote and annoying sounds impede health: a cognitive approach. *Int. J. Environ. Res. Public Health* 10, 1439–1461. doi: 10.3390/ijerph100 41439

Andringa, T. C., and van den Bosch, K. A. (2013). Core affect and soundscape assessment: fore- and background soundscape design for quality of life. *Paper Presented at the 42nd International Congress and Exposition on Noise Control Engineering, Innsbruck*, Innsbruck.

Andringa, T. C., Van Den Bosch, K. A., and Wijermans, N. (2015). Cognition from life: the two modes of cognition that underlie moral behavior. *Front. Psychol.* 6:362. doi: 10.3389/fpsyg.2015.00362

Appleton, J. (1975). *The Experience of Landscape*. London: Wiley.

Axelsson, Ö. (2011). *Aesthetic Appreciation Explicated*. Ph.D. thesis, Stockholm University, Stockholm.

Axelsson, Ö., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *J. Acoust. Soc. Am.* 128, 2836–2846. doi: 10.1121/1.3493436

Babisch, W. (2002). The noise/stress concept, risk assessment and research needs. *Noise Health* 4, 1–11.

Barandiaran, X., Di Paolo, E. A., and Rohde, M. (2009). Defining agency: individuality, normativity, asymmetry, and spatio-temporality in action. *Adapt. Behav.* 17, 367–386. doi: 10.1177/1059712309343819

Bitner, M. J. (1992). Servicescapes: the impact of physical surroundings on customers and employees. *J. Mark.* 56, 57–71. doi: 10.2307/1252042

Blesser, B. (2007). *The Seductive (Yet Destructive) Appeal of Loud Music*. Available at: http://www.blesser.net

Booi, H., and van den Berg, F. (2012). Quiet areas and the need for quietness in Amsterdam. *Int. J. Environ. Res. Public Health* 9, 1030–1050. doi: 10.3390/ ijerph9041030

Botteldooren, D., De Coensel, B., and De Muer, T. (2006). The temporal structure of urban soundscapes. *J. Sound Vib.* 292, 105–123. doi: 10.1016/j.jsv.2005.07.026

Bradley, M., and Lang, P. J. (1999). *The International Affective Digitized Sounds (IADS)[: Stimuli, Instruction Manual, and Affective Ratings*. Technical Report B-2. Gainesville, FL: University of Florida.

Bradley, M. M., and Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology* 37, 204–215. doi: 10.1111/1469-8986.3720204

Bregman, A. S. (1994). *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

Brown, A. L., Kang, J., and Gjestland, T. (2011). Towards standardization in soundscape preference assessment. *Appl. Acoust.* 72, 387–392. doi: 10.1016/j. apacoust.2011.01.001

Brown, R. E., Basheer, R., McKenna, J. T., Strecker, R. E., and McCarley, R. W. (2012). Control of sleep and wakefulness. *Physiol. Rev.* 92, 1087–1187. doi: 10.1152/physrev.00032.2011

Buxton, O. M., Ellenbogen, J. M., Wang, W., Carballeira, A., O'Connor, S., Cooper, D., et al. (2012). Sleep disruption due to hospital noises: a prospective evaluation. *Ann. Intern. Med.* 157, 170–179. doi: 10.7326/0003-4819-157-3-201208070-00472

Cain, R., Jennings, P., and Poxon, J. (2013). The development and application of the emotional dimensions of a soundscape. *Appl. Acoust.* 74, 232–239. doi: 10.1016/j.apacoust.2011.11.006

Darner, C. L. (1966). Sound pulses and the heart. *J. Acoust. Soc. Am.* 39, 414–416.

Davies, W. J., Adams, M. D., Bruce, N. S., Marselle, M., Cain, R., Jennings, P., et al. (2009). The positive soundscape project: a synthesis of results from many disciplines. *Paper Presented at the 38th International Congress and Exposition on Noise Control Engineering*, Ottawa.

Davies, W. J., and Murphy, J. (2012). Reproducibility of soundscape dimensions. *Paper Presented at the 41st International Congress and Exposition on Noise Control Engineering*, New York, NY.

Di Paolo, E. A., and Thompson, E. (2014). "The enactive approach," in *The Routledge Handbook of Embodied Cognition*, ed. L. Shapiro (London: Routledge Press), 1–14.

Di Paolo, E. A., Rohde, M., De Jaegher, H., University of Sussex. (2010). "Horizons for the enactive mind," in *Enaction Towards a New Paradigm for Cognitive Science*, eds J. Stewart, O. Gapenne, and E. A. Di Paolo (Cambridge, MA: MIT Press), 34–87.

Dubois, D., Guastavino, C., and Raimbault, M. (2006). A cognitive approach to urban soundscapes: using verbal data to access everyday life auditory categories. *Acta Acust. U. Acust.* 92, 865–874.

Eckblad, G. (1981). *Scheme Theory: A Conceptual Framework for Cognitive-Motivational Processes*. London: Academic Press.

Egbert, M. D., Barandiaran, X., and Di Paolo, E. A. (2010). A minimal model of metabolism-based chemotaxis. *PLoS Comput. Biol.* 6:17. doi: 10.1371/journal.pone.0063617

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. *Am. Psychol.* 56, 218–226. doi: 10.1037/0003-066X.56.3.218

Fredrickson, B. L., and Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cogn. Emot.* 19, 313–332. doi: 10.1080/02699930441000238

Fritz, J. B., Elhilali, M., David, S. V., and Shamma, S. A. (2007). Auditory attention - focusing the searchlight on sound. *Curr. Opin. Neurobiol.* 17, 437–455. doi: 10.1016/j.conb.2007.07.011

Froese, T., and Ziemke, T. (2009). Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artif. Intell.* 173(3–4), 466–500. doi: 10.1016/j.artint.2008.12.001

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin Harcourt (HMH).

Graumann, C. F. (2002). "The phenomenological approach to people-environment studies," In *Handbook of Environmental Psychology*, R. B. Bechtel and A. Churchman (Hoboken, NJ: John Wiley & Sons, Inc).

Greene, M., and Oliva, A. (2009). Recognition of natural scenes from global properties: seeing the forest without representing the trees. *Cogn. Psychol.* 58, 137–176. doi: 10.1016/j.cogpsych.2008.06.001

Guski, R. (2001). "Moderatoren der Lärmwirkung," in *Handbuch der Umweltmedizin*, eds H. E. Wichmann, H.-W. Schlipköter, and G. Fülgraff (Landsberg: Kent H. Landsberg Company).

Hartig, T., Kaiser, F. G., and Bowler, P. A. (1997). Further development of a measure of perceived environmental restorativeness. *Working Paper No. 5. Institute for Housing Research, Uppsala Universitet*.

Hester, E. (2005). The evolution of the auditory system: a tutorial. *Contemp. Issues Commun. Sci. Disord.* 32, 5–10.

Horowitz, S. (2012). *The Universal Sense*. London: Bloomsbury Publishing.

International Organization for Standardization [ISO] (2014), *Acoustics — Soundscape — Part 1: Definition and Conceptual Framework*. Geneva: International Organization for Standardization.

Ising, H., and Kruppa, B. (2004). Health effects caused by noise: evidence in the literature from the past 25 years. *Noise Health* 6, 5–13.

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspect. Psychol. Sci.* 2, 260–280. doi: 10.1111/j.1745-6916.2007.00044.x

Jasìkowski, P., Jaroszyk, F., and Hojan-Jezierska, D. (1990). Temporal-order judgments and reaction time for stimuli of different modalities. *Psychol. Res.* 52, 35–38. doi: 10.1007/BF00867209

Juslin, P. N., and Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci.* 31, 559–575. doi: 10.1017/S0140525X08005293

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

Koch, M., and Schnitzler, H. U. (1997). The acoustic startle response in rats - circuits mediating evocation, inhibition and potentiation. *Behav. Brain Res.* 89(1-2), 35–49. doi: 10.1016/s0166-4328(97)02296-1

Kraus, K. S., and Canlon, B. (2012). Neuronal connectivity and interactions between the auditory and limbic systems. Effects of noise and tinnitus. *Hear. Res.* 288(1-2), 34–46. doi: 10.1016/j.heares.2012.02.009

Kuppens, P., Champagne, D., and Tuerlinckx, F. (2012). The dynamic interplay between appraisal and core affect in daily life. *Front. Psychol.* 3:380. doi: 10.3389/fpsyg.2012.00380

Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cogn. Emot.* 13, 481–504. doi: 10.1080/026999399379159

Louie, A. H. (2010). Robert Rosen's anticipatory systems. *Foresight* 12, 18–29. doi: 10.1108/14636681011049848

Marsh, R. A., Fuzessery, Z. M., Grose, C. D., and Wenstrup, J. J. (2002). Projection to the inferior colliculus from the basal nucleus of the amygdala. *J. Neurosci.* 22, 10449–10460. doi: 10.1523/JNEUROSCI.22-23-10449.2002

Neumann, D. L., Waters, A. M., and Westbury, H. R. (2008). The use of an unpleasant sound as the unconditional stimulus in aversive Pavlovian conditioning experiments that involve children and adolescent participants. *Behav. Res. Methods* 40, 622–625. doi: 10.3758/BRM.40.2.622

Nielbo, F. L., Steele, D., and Guastavino, C. (2013). "Investigating soundscape affordances through activity appropriateness," In *Proceedings of the Meetings on Acoustics ICA2013*. Dhaka: ASA.

Osgood, C. E. (1975). *Cross-Cultural Universals of Affective Meaning*. Chicago, IL: University of Illinois Press.

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

Pheasant, R. J., Fisher, M. N., Watts, G. R., Whitaker, D. J., and Horoshenkov, K. V. (2010). The importance of auditory-visual interaction in the construction of 'tranquil space'. *J. Environ. Psychol.* 30, 501–509. doi: 10.1016/j.jenvp.2010.03.006

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172. doi: 10.1037/0033-295X.110.1.145

Schafer, R. M. (1977). *The Tuning of the World: Toward a Theory of Soundscape Design*. New York, NY: Knopf.

Schulte-Fortkamp, B. (2002). The meaning of annoyance in relation to the quality of acoustic environments. *Noise Health* 4, 13–18.

Siegel, A., and Sapru, H. N. (2011). *Essential Neuroscience*. Philadelphia, PA: Wolters Kluwer Health.

Spreng, M. (2000). Central nervous system activation by noise. *Noise Health* 2, 49–58.

Steele, D., Steffens, J., and Guastavino, C. (2015). The role of activity in urban soundscape evaluation. *Paper Presented at the 10th European Congress and Exposition on Noise Control Engineering*, Maastricht, 1507–1512.

Stockfelt, T. (1991). Sound as an existential necessity. *J. Sound Vib.* 151, 367–370. doi: 10.1016/0022-460X(91)90533-P

Tajadura-Jiménez, A. (2008). *Embodied Psychoacoustics: Spatial and Multisensory Determinants of Auditory-Induced Emotion*. Gothenburg: Chalmers University of Technology.

Tajadura-Jiménez, A., Larsson, P., Väljamäe, A., Västfjäll, D., and Kleiner, M. (2010a). When room size matters: acoustic influences on emotional responses to sounds. *Emotion* 10, 416–422. doi: 10.1037/a0018423

Tajadura-Jiménez, A., Väljamäe, A., Kitagawa, N., and Ho, H. N. (2010b). Whole-body vibration influences on sound localization in the median plane. *Proc. Inst. Mech. Eng. D* 224, 1311–1320. doi: 10.1243/09544070JAUTO1600

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge MA: The Belknap Press of Harvard University Press.

Truax B. (ed.) (1999). *Handbook for Acoustic Ecology*, 2nd Edn. Vancouver: Cambridge Street Publishing.

Ulrich, R. S. (1983). "Aesthetic and affective response to natural environment," in *Behavior and the Natural Environment*, eds I. Altman, and J. Wohlwill, (Boston, MA: Springer), 85–125.

Varela, F. J., Thompson, E., and Rosch, E. (1993). *The Embodied Mind : Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Vernon, D. (2010). Enaction as a conceptual framework for developmental cognitive robotics. *Paladyn* 1, 89–98. doi: 10.2478/s13230-010-0016-y

Von Uexküll, J. (1992/1934). A stroll through the worlds of animals and men: a picture book of invisible worlds. *Semiotica* 89, 5–85.

Welch, D., and Fremaux, G. (2017a). Understanding why people enjoy loud sound. *Semin. Hear.* 38, 348–358. doi: 10.1055/s-0037-1606328

Welch, D., and Fremaux, G. (2017b). Why do people like loud sound? A qualitative study. *Int. J. Environ. Res. Public Health* 14:E908. doi: 10.3390/ijerph140 80908

World Health Organization [WHO] (2011). *Burden of Disease from Environmental Noise*. Geneva: WHO.

Wundt, W. (1897). *Grundriss der Psychologie*. Leipzig: W. Engelmann.

Xiong, X. R. R., Liang, F. X., Zingg, B., Ji, X. Y., Ibrahim, L. A., Tao, H. Z. W., and Zhang, L. I. (2015). Auditory cortex controls sound-driven innate defense behaviour through corticofugal projections to inferior colliculus. *Nat. Commun*, 6:7224. doi: 10.1038/ncomms 8224

ORIGINAL RESEARCH

# Sound Categories: Category Formation and Evidence-Based Taxonomies

Oliver Bones*, Trevor J. Cox and William J. Davies

*Acoustics Research Centre, University of Salford, Salford, United Kingdom*

Five evidence-based taxonomies of everyday sounds frequently reported in the soundscape literature have been generated. An online sorting and category-labeling method that elicits rather than prescribes descriptive words was used. A total of $N = 242$ participants took part. The main categories of the soundscape taxonomy were people, nature, and manmade, with each dividing into further categories. Sounds within the nature and manmade categories, and two further individual sound sources, dogs, and engines, were explored further by repeating the procedure using multiple exemplars. By generating multidimensional spaces containing both sounds and the spontaneously generated descriptive words the procedure allows for the interpretation of the psychological dimensions along which sounds are organized. This reveals how category formation is based upon different cues – sound source-event identification, subjective-states, and explicit assessment of the acoustic signal – in different contexts. At higher levels of the taxonomy the majority of words described sound source-events. In contrast, when categorizing dog sounds a greater proportion of the words described subjective-states, and valence and arousal scores of these words correlated with their coordinates along the first two dimensions of the data. This is consistent with valence and arousal judgments being the primary categorization strategy used for dog sounds. In contrast, when categorizing engine sounds a greater proportion of the words explicitly described the acoustic signal. The coordinates of sounds along the first two dimensions were found to correlate with fluctuation strength and sharpness, consistent with explicit assessment of acoustic signal features underlying category formation for engine sounds. By eliciting descriptive words the method makes explicit the subjective meaning of these judgments based upon valence and arousal and acoustic properties, and the results demonstrate distinct strategies being spontaneously used to categorize different types of sounds.

Keywords: soundscape, everyday sounds, taxonomy, categories, category formation, valence, arousal, acoustic correlates

## INTRODUCTION

Categorization is a fundamental process by which meaning is applied to sensory experience (Dubois, 2000) based upon the correlational structure of the attributes of objects in the environment (Rosch, 1978). Knowledge about the environment is parsed and organized according to category structures. This simplifies the environment and gleans information with less cognitive

effort ("cognitive economy"; Rosch, 1978), with inferences assuming that category members have similar attributes.

Which attributes of the many sounds experienced in everyday life are used to form categories? One approach to answering this question is the semantic differential method, whereby participants score concepts and events on a number of attribute rating scales. Typically this is followed by factor analysis in order to extract the principal dimensions which are then interpreted according to the attribute scales with which they most strongly correlate. Classically the results from this type of method are said to demonstrate that the factors 'evaluation,' 'potency,' and 'activity' (EPA) characterize the affective components of meaning (Osgood, 1952, 1969), and that this occurs universally across cultures (e.g., Heise, 2001). A number of studies have used this method for soundscapes, finding dimensions analogous to EPA, such as 'pleasantness' (Björk, 1985; Payne et al., 2007; Axelsson et al., 2010; Hong and Jeon, 2015), 'preference' (Kawai et al., 2004; Yu et al., 2016), 'calmness' (Cain et al., 2013), 'relaxation' (Kang and Zhang, 2010), 'dynamic' (Kang and Zhang, 2010), 'vibrancy' (Cain et al., 2013), 'playfulness' (Yu et al., 2016), and 'eventfulness' (Axelsson et al., 2010; Hong and Jeon, 2015). A number of other, possibly more sound-specific components are also reported, such as 'sense of daily life' (Kawai et al., 2004), 'familiarity' (Axelsson et al., 2010), 'spatiality' (Kang and Zhang, 2010), 'harmony' (Hong and Jeon, 2015), 'communication' (Kang and Zhang, 2010), 'loudness,' and 'richness' (Yu et al., 2016). A related framework is that of 'core affect' (for a review see Russell, 2003); this is a dimensional model of affective states as the linear combination of valence (a pleasure–displeasure continuum) and arousal (an alertness continuum).

Cluster analysis of semantic differential data identifies groups of sounds that are considered similar in terms of the attribute scales used, and factor analysis identifies the underlying dimensions of ratings on those scales. As these attributes are prescribed *a priori* by the experimenter, however, the dimensions may lack ecological validity for understanding categorization. An alternative approach is to generate similarity date by pairwise comparison (e.g., Gygi et al., 2007) or by sorting tasks. This approach avoids prescribing attributes on which to rate sounds, although in the case of pairwise comparisons the time required to perform comparisons can be prohibitively long. Moreover, since the procedure does not generate semantic labeling the meaning of the resulting categories must be interpreted by the researcher. Sorting tasks on the other hand produce similarity data which can be interpreted by linguistic analysis of the category descriptions. Since this approach allows participants to form categories using their own criteria and to provide their own descriptors, this method provides insight into how categories are formed with greater ecological validity than the semantic differential method.

Dubois et al. (2006) used a sorting method to investigate soundscapes consisting of sounds containing human activity. The results produced categories formed principally by similarity of sound sources and places. For categories consisting of sounds that were identified as containing noises, these were categorized by similar sources or actions. Another study by Guastavino (2007) asked participants to sort ambient urban noise. Similar to Dubois et al. (2006), categories were principally differentiated by those

that contained sounds consisting of mostly human activity or those that contained sounds consisting mostly of traffic noise. Subcategories were formed around type of activity. Likewise, Morel et al. (2012) found that categories of road traffic noise were formed based upon vehicle type (sound source) and driving condition (action).

These categorizations and dimensions relate to complex environmental sounds, and are consistent with Guastavino (2006). In this study a linguistic analysis of interview data found that descriptions of sound sources accounted for 76% of the descriptions of the soundscapes. With respect to detached sounds, using a similar sorting and labeling procedure Houix et al. (2012) identified categories of domestic noises based on temporal extent, which resembled those previously proposed by Gaver (1993), based upon the type of material (e.g., solid) and events (e.g., impact) producing the sound. Previous work using semantic differentials has identified dimensions, such as 'identifiability,' 'timbre,' and 'oddity' (Ballas, 1993), EPA (Björk, 1985), and 'harshness,' 'complexity,' 'appeal,' and 'size' (Kidd and Watson, 2003). As noted above, these are not necessarily an ecologically valid representation of the criteria by which categories are formed. Using a hierarchical sorting paradigm, Giordano et al. (2010) found evidence for symbolic (acoustic) properties predicting similarity of environmental sounds from non-living objects, whereas iconic (semantic) meaning predicted similarity of environmental sounds from living things (see also results of neuroimaging studies by, e.g., Lewis et al., 2005). Finally, a recent study by Bergman et al. (2016) found evidence for valence and arousal contributing to the first dimension of data from a pairwise dissimilarity rating task with everyday sounds, suggesting that emotional response may also play a role in categorization.

Our study explored the formation of categories for a set of everyday sounds that are frequently reported in the soundscape literature. Evidence-based taxonomies were developed in order to explore the formation of categories at different levels of hierarchy. In order to test the hypothesis that the use of cues for category formation would differ both between levels of the emergent taxonomy and between different sounds within levels, we performed a statistical analysis of verbal correlates of sound category formation. The different ways that people use cues to form sound categories have important implications for research in everyday sound. The relationship between sound category formation and emergent sound taxonomies sheds light on the perception of everyday sound.

## MATERIALS AND METHODS

### Procedure

Multiple measurements revealed the formation of categories at different levels of the emergent taxonomy. The top level experiment tested 'soundscape,' the middle 'nature' and 'manmade,' and the bottom 'dogs' and 'engines.' The categories formed at the 'top' level of the taxonomy informed the selection of sounds for studies at the 'middle' level, and individual sound sources from the 'middle' level were selected for a study of

'bottom' level sounds. Each study was conducted via a web interface on Sound101[1], a website hosted by one of the authors. Each sound was represented by a tile containing a single word descriptor (e.g., 'Road_1'), all of which were arranged in a random order in a 'sound bank' panel on the left hand side of the screen at the onset of the study. In the case of the dogs and engines studies, tiles were labeled as 'Dog_1,' 'Dog_2' etc. Instructions at the top of the screen directed participants to: click the tiles to hear the sound; group similar sounds together by dragging them from the sound bank into one of five categories; use all five categories; give each category a name describing the sounds in the category. In addition, participants were instructed not to use category names, such as 'miscellaneous,' 'random,' or 'sounds' etc. A pilot study found that five was the mean number of categories used when freely sorting the 60 sounds from the top level study. No time limit was imposed, and the average time taken was approximately 20 min. The procedure was approved by the University of Salford Science & Technology Research, Innovation and Academic Engagement Ethical Approval Panel.

## Stimuli and Participants

All participants completed a short web form prior to the experiment consisting of questions on age, sex, and audio expertise ('Are you an audio engineer, an acoustician, a proficient musician, or similar?') and main language. Participants were screened so as to only include those aged 18 and over and with English as their main language. Demographic data is displayed in **Table 1**: as can be seen, participants for each study were broadly similar, with the exceptions that there were more participants aged 18–29 in the dog study, and fewer participants who self-identified as being audio experts in the engine study. These two features are addressed in the discussion section.

All stimuli (see Supplementary Table 1) were taken from Freesound[2]. Some were sourced directly from Freesound, others were sourced from ESC-50, a database of audio clips collected

[1]http://www.sound101.org

[2]www.freesound.org

from Freesound and curated into categories by Piczak (2015). Where audio clips were sourced directly from Freesound, they were identified by searching filenames and descriptions using keywords corresponding to the sound names. Search results were sorted by number of downloads. Audio clips of synthesized sounds were rejected. Files were selected based upon subjective audio quality and duration: preference was given to clips that were ≤5 s, but where necessary clips were manually edited in duration. All stimuli were normalized to maximum amplitude of 3 dB below full-scale.

### Top Level: Soundscape

$N = 50$ participants completed the initial study. Sixty stimuli (Supplementary Table 1) were selected so as to be representative of sounds described in a number of studies from the soundscape literature (Kawai et al., 2004; Gygi et al., 2007; Brown et al., 2011; Yang and Kang, 2013; Salamon et al., 2014). Brown et al. (2011) in particular place an emphasis on sounds occurring in multiple environmental contexts. Therefore an effort was made to include examples of sounds recorded indoors and outdoors where this was possible. In some cases these were recordings of sounds occurring outside, recorded from indoors, e.g., 'Fireworks_2.' In other cases these were recordings which were audibly recorded in different sized spaces, e.g., 'Laughter_1' sounded like it was recorded in a large room due to the audible reverberation, whereas 'Laughter_2' did not contain audible reverberation. All stimuli had duration of ≤5s.

### Middle Level: Nature and Manmade Sounds

Analysis of top level sounds generated three principal categories, people, nature, and manmade. Of these, nature and manmade were considered the most interesting to explore further, since classification of the vocal and music sounds of the people category have been well studied previously (e.g., Pachet and Cazaly, 2000; Ververidis et al., 2004; Li and Ogihara, 2005; Giordano et al., 2010).

Each of the nature and manmade sound studies consisted of five exemplars of 13 sounds. All stimuli had duration of ≤5s. $N = 45$ participants completed the nature study; $N = 48$ completed the manmade study.

**TABLE 1 |** Demographic data of participants for all studies.

|              |               | Soundscape | Nature | Manmade | Dogs | Engines |
|--------------|---------------|------------|--------|---------|------|---------|
| Age          | 18–29         | 41         | 35     | 35      | 70   | 33      |
|              | 30–39         | 37         | 41     | 38      | 12   | 49      |
|              | 40–49         | 18         | 11     | 13      | 14   | 10      |
|              | 50–59         | 2          | 16     | 10      | 4    | 6       |
|              | 60–69         | 2          | 2      | 2       | 0    | 2       |
|              | 70–79         | 0          | 2      | 2       | 0    | 0       |
| Sex          | Male          | 39         | 36     | 63      | 46   | 43      |
|              | Female        | 61         | 61     | 38      | 54   | 57      |
|              | Rather not say| 0          | 2      | 0       | 0    | 0       |
| Audio expert | Yes           | 16         | 16     | 15      | 18   | 2       |
|              | No            | 84         | 84     | 85      | 82   | 98      |

*All values are percentages rounded up to the nearest whole percent.*

### Bottom Level: Dog and Engine Sounds

To investigate category formation for single sound sources, an individual sound from each of the nature and manmade categories was selected, dog and engine sounds, respectively. In the interests of ecological validity, dog sounds were not restricted to 5 s duration; rather, clips were selected so as to sound natural (mean = 5.8 s, SD = 3.0 s). In some cases this meant selecting a section that sounded like a complete dog bark from a longer clip. $N = 50$ participants completed the dog's study, whilst $N = 49$ completed the engines study.

## Analysis

### Contingency Table

For each experiment the data from each participant were initially collected as a contingency table of 1s and 0s, where rows corresponded to individual sounds and columns corresponded to category names and where a 1 indicated that a sound had been placed in a given category, before being collated into a sounds $\times$ $5N$ categories contingency table of data from all participants. Each contingency table was consolidated by summing data where category names were the same or synonymous. Category names were initially processed by removing white space; removing special characters; removing the words 'sound' and 'sounds'; removing numbers; converting to lower-case; and correcting spelling. Category names were then stemmed (e.g., 'natural' and 'nature' were reduced to 'natur-') before restoring each stem to the most common pre-stemming version of that word (e.g., 'nature'). Categories which had either the same name following this process, or which were identified as synonyms by Microsoft's synonym checker were then summed. This resulted in a contingency table which contained numbers other than 1 and 0 (see Supplementary Table 2 for details of which data were summed this way). Hereafter category names are referred to as 'descriptive words.' Consolidating the contingency tables reduced the number of descriptive words for soundscape sounds from 250 to 94; from 225 to 75 for nature; from 240 to 78 for manmade; from 250 to 59 for dogs; and from 245 to 96 for engines. A Pearson's Chi-squared test found a dependence between rows and columns for all resulting contingency tables, demonstrating a significant relationship between sounds and descriptive words: soundscape, $\chi^2(5487) = 7813.5$, $p < 0.001$; nature, $\chi^2(4736) = 8227.4$, $p < 0.001$; manmade, $\chi^2(4928) = 8989.7$, $p < 0.001$; dogs, $\chi^2(2494) = 3977.3$, $p < 0.001$; engines, $\chi^2(3705) = 3915.0$, $p < 0.001$.

### Correspondence Analysis

Each consolidated contingency table was submitted to a correspondence analysis (CA; see Greenacre, 1984; Lê et al., 2008), a method similar to principal component analysis but suitable for categorical rather than continuous data, in order to identify the principal dimensions of the data. CA was performed using the FactoMineR package (Lê et al., 2008) in R V3.3.3. This step was used to denoise the data prior to clustering (Husson et al., 2010), and to extract the dimensions of the similarity data so that sounds and descriptive words could be plotted in the same space. Dimensions with eigenvalues greater than would be the case were the data random were retained. For example, the top level soundscape contingency table had 60 rows (sounds) and 94 columns (descriptive words). Therefore were the data random the expected eigenvalue for each dimension would be 1.7% in terms of rows [1/(60-1)] and 1.1% in terms of columns [1/(94-1)], so all dimensions with eigenvalues greater than 1.7% were retained. The number of dimensions retained during correspondence analysis of each contingency table and the variance explained is displayed in Supplementary Table 3.

### Cluster Analysis and Category Naming

Agglomerative hierarchical cluster analysis of the dimensions resulting from CA was performed using Ward's criterion (see Husson et al., 2010), using FactoMineR. Taxonomies were derived by 'slicing' the resulting dendrograms at different heights and giving each resulting cluster a category name according to the descriptive words that contributed to that cluster. For all taxonomies apart from the dog taxonomy slices were performed so as to create all possible clusters above the height of the dendrogram at which the ratio of between-cluster inertia to total inertia was 0.1. Between-cluster inertia describes the deviation of the center of gravity of all clusters from the overall center of gravity, and total inertia describes this value summed with within-cluster inertia, i.e., the deviation of individuals from the center of gravity of each cluster. This ratio becomes greater with slices at higher levels of the dendrogram and cluster members become less similar. At slices at lower levels of the dendrogram this value becomes smaller and cluster members become more similar. The value of 0.1 was selected to allow populating the taxonomy with enough labels so as to be meaningful without compromising the quality of the labeling. In the case of the dog taxonomy, a ratio of 0.15 was chosen for the same reason.

The contribution of each descriptive word to each cluster was assessed by comparing global frequency (the total number of times sounds were assigned to a descriptive word) to the internal frequency for a given cluster (the number of times sounds within a cluster were assigned to that descriptive word). Significance of over-representation of each descriptive word within each cluster was assessed using a hypergeometric distribution (see Lê et al., 2008). The hypergeometric distribution describes the number of times an event occurs in a fixed number of trials, where each trial changes the probability for each subsequent trial because there is no replacement. Since the total number of descriptive words and the total number of times a given descriptive word was used is known, the probability $p$ of a given descriptive word being used to describe sounds within a given cluster can be calculated. To illustrate this, consider the descriptive words applied to the soundscape dendrogram sliced into three clusters. Descriptive words that were over-represented in the first cluster are displayed in **Table 2** (see Supplementary Tables 4–8 for descriptive words corresponding to other clusters and other sounds). 'People' is the most over-represented descriptive word in this cluster: the sounds in this cluster were assigned to this descriptive word 257 times, out of a total of 357 times that sounds were assigned to this word. This first category of the soundscape taxonomy was therefore named 'people.'

**TABLE 2 |** Descriptive words that were significantly over-represented in the first cluster of the soundscape categorization data.

| Descriptive word | Internal Freq. | Global Freq. | p | v-test |
|---|---|---|---|---|
| People | 257 | 357 | <0.001 | 22.109 |
| Music | 63 | 121 | <0.001 | 7.438 |
| Vocal | 16 | 16 | <0.001 | 6.608 |
| Entertainment | 18 | 20 | <0.001 | 6.352 |
| Chatter | 10 | 10 | <0.001 | 5.060 |
| Changes | 9 | 10 | <0.001 | 4.316 |
| Harmony | 9 | 10 | <0.001 | 4.316 |
| Social | 9 | 11 | <0.001 | 3.974 |
| Alive | 9 | 11 | <0.001 | 3.974 |
| Enjoying | 8 | 12 | 0.002 | 3.096 |
| Marine | 7 | 10 | 0.003 | 2.993 |
| Species | 9 | 16 | 0.005 | 2.801 |
| Pleasant | 8 | 14 | 0.008 | 2.658 |
| Events | 6 | 9 | 0.009 | 2.606 |
| Relaxing | 5 | 8 | 0.029 | 2.184 |

*This cluster was given the category name 'people'.*

This method of objectively naming taxonomic categories was sufficient in the majority (31 out of 56) cases. However, in other cases it was necessary to subjectively choose a descriptive word that was significantly over-represented but ranked lower to avoid repetition of category names (see Supplementary Tables 3–7). For example, in constructing the manmade taxonomy, a category was created with the name 'home' within a higher-level category also named 'home.' In these instances a name was subjectively chosen from a descriptive word lower down the table that better represented the content of the category. In this example 'daily life' was chosen for the category within 'home' that contained subcategories named 'toilet' and 'food.'

## Category Formation
### Multinomial logit regression of descriptive words
The main aim of this study was to explore differences in how categories were formed between and within each level of the taxonomy. In order to examine this, each of the descriptive words (pre-consolidation) used in each of the studies were independently coded by three people: the first author and two acoustics doctoral students. All three are native speakers of English. Words were coded as describing either the *source-event* (referring to the inferred source of the sound), the *acoustic* signal (explicitly referring to the sound itself), or a *subjective-state* (describing an emotional state caused by the sound or of the sound source). Word types were determined by agreement between at least two of the three coders: this criteria was met for all words (see Supplementary Table 9). Multinomial logit regression models were used to compare the likelihood of each type of descriptive word being used to describe sounds at each level of the taxonomy and for each group of sounds. In each case the dependent variable was the type of descriptive word used (e.g., subjective-state vs. source), and the independent variables were level of the taxonomy (e.g., top vs. middle) or the sound type (e.g., nature vs. manmade). Multinomial logit regression models

produce log-odds coefficients ($B$) that can be expressed as an odds ratio ($e^B$). These describe how many times more likely a type of descriptive word is used relative to another type of descriptive word, at a given level of the taxonomy relative to another level, or for a sound type relative to another sound type.

In order to assess the effect that providing labels for the tiles had on how categories were formed, a supplementary top level study was performed in which tiles were labeled with pseudorandomized numbers. Multinomial logit regression models demonstrated that providing text labels did not significantly change the proportion of word types used (see Supplementary Tables 10, 11).

### Post hoc analysis
To explore strategies for categorization further, the arrangement of sounds and descriptive words within the space created by the dimensions elicited by CA were examined. Based upon the results of the multinomial logit regression models a *post hoc* decision was taken to explore arousal and valence for the descriptive words used for dog sounds. A correlation between the coordinates of words describing subjective-states and measures of valence and arousal for those words was calculated. The arousal and valence values for the words were taken from a scored dataset of 13915 lemmas (Warriner et al., 2013; see Supplementary Table 12).

Similarly, the multinomial logit regression models indicated further analysis of engine sounds should use acoustic features. This was based upon the finding that explicit assessment of the acoustic signal accounted for categorization. The coordinates of engine sounds and two simple acoustic features commonly used by industry to assess product sounds, fluctuation strength and sharpness, were tested for correlation. Fluctuation strength is a measure of amplitude modulation below 20 Hz, whilst sharpness is a measure of high-frequency content. Both measures account for the perceptual distance between frequencies by dividing the signal into critical bands using the Bark scale. Fluctuation strength is measured in units of vacil where 1 vacil is defined as the fluctuation strength produced by a 1000 Hz tone with a sound pressure level of 60 dB, 100% amplitude modulated at 4 Hz. Sharpness is measured in acum, where 1 acum has the equivalent sharpness of a narrow-band noise with a center frequency of 1000 Hz, a bandwidth of 1 critical band, and a sound pressure level of 60 dB. Both fluctuation strength and sharpness were evaluated with dBFA software using the criteria of Zwicker and Fastl (2013). Since the presentation level of the stimuli in

**TABLE 3 |** Percentages of different types of descriptive words used at each level of the taxonomy and for each type of sound.

| | Top | Middle | | Bottom | |
|---|---|---|---|---|---|
| Source | 81.2 | 75.1 | | 42.0 | |
| Acoustic | 14.8 | 17.7 | | 35.6 | |
| Subjective | 4.0 | 7.2 | | 22.4 | |
| | Soundscape | Nature | Manmade | Dogs | Engines |
| Source | 81.2 | 75.6 | 74.6 | 24.0 | 60.0 |
| Acoustic | 14.8 | 13.8 | 21.7 | 34.0 | 37.1 |
| Subjective | 4.0 | 10.7 | 3.8 | 42.0 | 2.9 |

**TABLE 4 |** Results of the multinomial logit regression models.

| | | B | $e^B$ | SE | p |
|---|---|---|---|---|---|
| Middle vs. Top | Subjective vs. Source | 0.65 | 1.9 | 0.37 | 0.08 |
| | Acoustic vs. Source | 0.26 | 1.30 | 0.22 | 0.22 |
| | Acoustic vs. Subjective | −0.39 | 0.68 | 0.41 | 0.35 |
| Bottom vs. Top | Subjective vs. Source | 2.40 | 11.0 | 0.34 | <0.001* |
| | Acoustic vs. Source | 1.54 | 4.7 | 0.21 | <0.001* |
| | Acoustic vs. Subjective | −0.86 | 0.4 | 0.38 | 0.022* |
| Bottom vs. Middle | Subjective vs. Source | 1.74 | 5.7 | 0.22 | <0.001* |
| | Acoustic vs. Source | 1.27 | 3.6 | 0.16 | <0.001* |
| | Acoustic vs. Subjective | −0.47 | 0.6 | 0.24 | 0.049* |
| Nature vs. Manmade | Subjective vs. Source | 1.03 | 2.8 | 0.41 | 0.011* |
| | Acoustic vs. Source | −0.47 | 0.6 | 0.25 | 0.063 |
| | Acoustic vs. Subjective | −1.50 | 0.2 | 0.45 | <0.001* |
| Dogs vs. Engines | Subjective vs. Source | 3.60 | 36.7 | 0.42 | <0.001* |
| | Acoustic vs. Source | 0.83 | 2.3 | 0.22 | <0.001* |
| | Acoustic vs. Subjective | −2.78 | 0.1 | 0.42 | <0.001* |

*In each case the dependent variable was the type of descriptive word used (e.g., subjective-state vs. source), and the independent variables were level of the taxonomy (e.g., top), or the sound type (e.g., nature).*

this study was not controlled due to participants being recruited online, both sharpness and fluctuation strength calculations were referenced to a 1000 Hz sine wave with an amplitude of 1 Pa, which equates to a sound pressure level of 71 dB at full-scale. Note that of interest here is the relative rather than absolute fluctuation strength and sharpness.

Association between the coordinates of descriptive words (dogs) and sounds (engines), respectively, are reported using one-tailed Spearman's Rho ($r_s$) and Pearson's product-moment ($r$) correlations. No attempt was made to identify acoustic correlates of the dimensions of other data, since categorization in these cases was accounted for by other cues.

# RESULTS

## Category Formation
The main purpose of the current study was to explore differences in the way that sound categories are formed between and within different levels of category hierarchy. The types of words used



**FIGURE 1 |** Soundscape sounds **(A)** and descriptive words **(B)** plotted on the first two dimensions of categorization data. Note that the dimensions are the same in both panels. Sounds **(A)** are colored according to which of the main categories they belong to, and descriptive words **(B)** are colored according to type. Labels are displaced from their corresponding data point, indicated by a connecting line, to avoid overlapping.

**FIGURE 2 |** Nature sounds **(A)** and descriptive words **(B)** plotted on the first two dimensions of categorization data. Note that the dimensions are the same in both panels. Sounds **(A)** are colored according to which of the main categories they belong to, and descriptive words **(B)** are colored according to type. Labels are displaced from their corresponding data point, indicated by a connecting line, to avoid overlapping.



**FIGURE 3 |** Manmade sounds **(A)** and descriptive words **(B)** plotted on the first two dimensions of categorization data. Note that the dimensions are the same in both panels. Sounds **(A)** are colored according to which of the main categories they belong to, and descriptive words **(B)** are colored according to type. Labels are displaced from their corresponding data point, indicated by a connecting line, to avoid overlapping.

to describe sounds at each level of the taxonomy and for each type of sound are presented in **Table 3**. A series of multinomial logit regression models were fitted to the descriptive word data (see **Table 4**). The likelihood of using words describing

source-event, signal, and subjective-states did not significantly differ between the middle and top levels of the taxonomy: at these levels the majority of words used described source-events (top, 81%; middle, 75%). However, there were significant
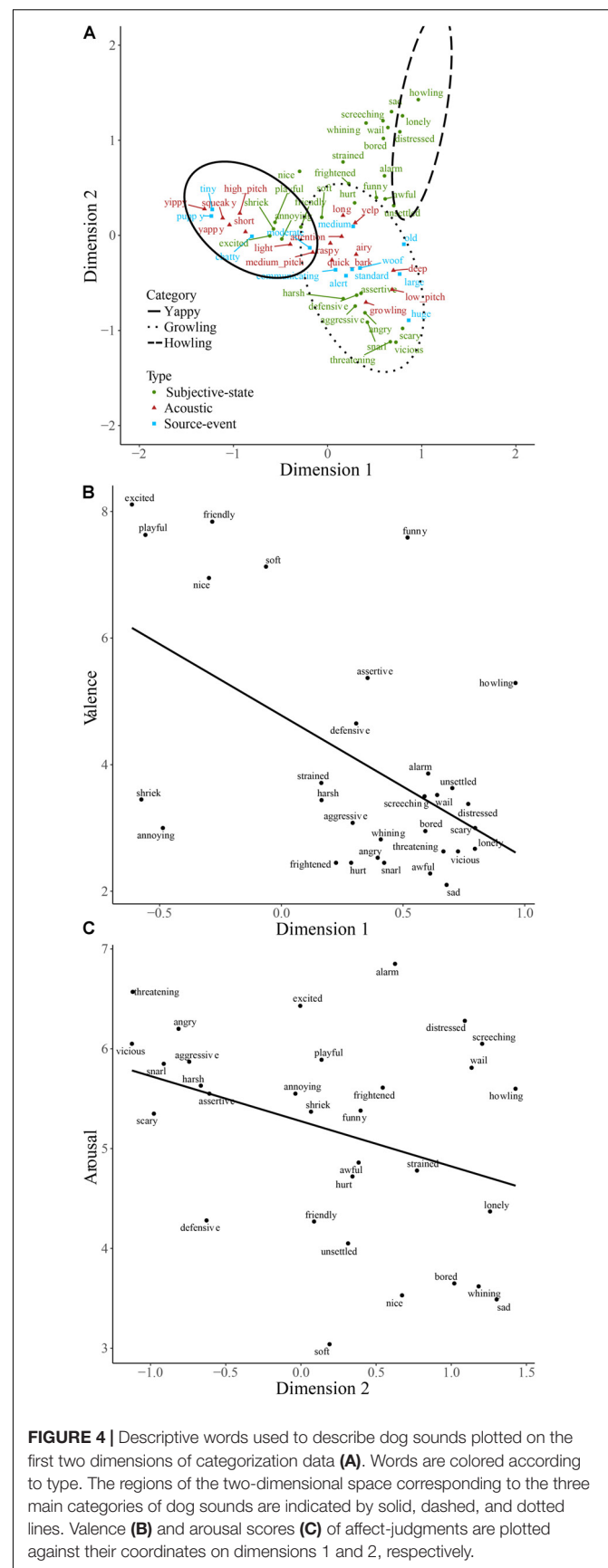
differences between bottom and top, and bottom and middle levels. Expressed as an odds ratio ($e^B$), there were 11 times the odds of using a word that described a subjective-state rather than the source-event at the bottom level compared to the top level, and 4.7 times the odds of using a word describing the acoustic signal rather than the source-event. On the other hand, there were 0.4 times the odds of using a word describing the acoustic signal rather than a subjective-state at the bottom level compared to the top level.
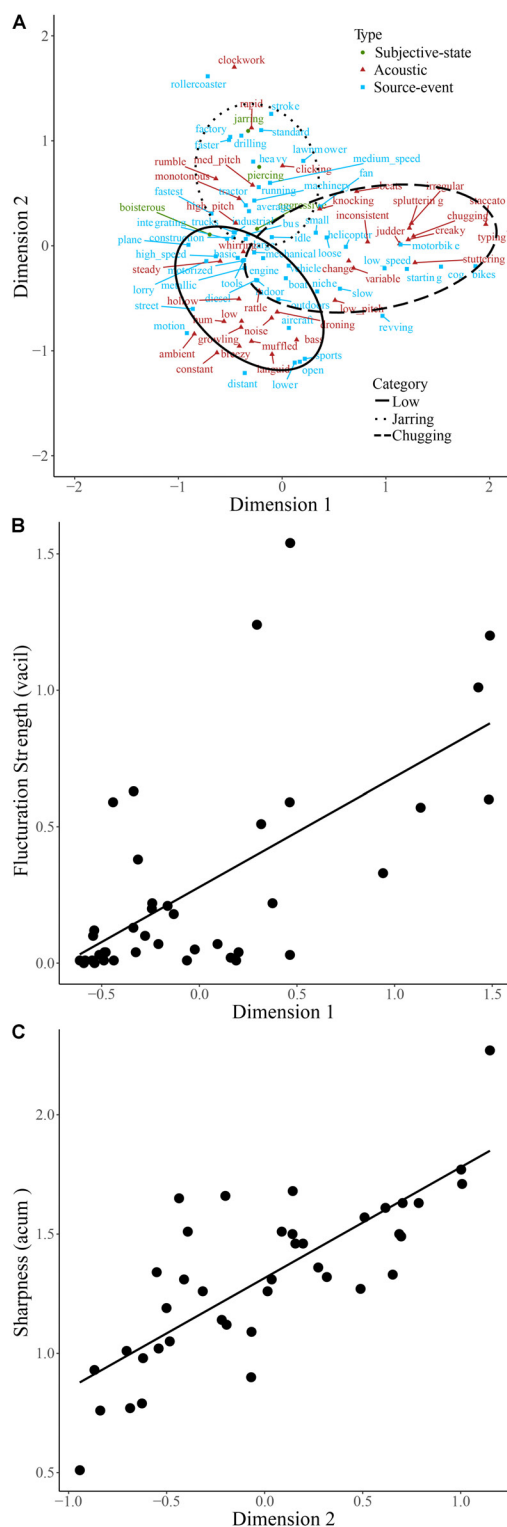
There were also 5.7 times the odds of using a word describing a subjective-state rather than the source-event at the bottom level compared to the middle level, and 3.6 times the odds of using a word describing the acoustic signal rather the source-event. On the other hand there were 0.6 times the odds of using a word describing the acoustic signal rather than a subjective state at the bottom level compared to the middle level.

Within the middle level there were 2.8 times the odds of using a word that described a subjective-state rather than the source-event when describing nature sounds compared to manmade sounds. However, there were only 0.2 times the odds of using a word describing the acoustic signal rather than a subjective-state. Within the bottom level there was 36.7 times the odds of using a word that described a subjective-state rather than the source-event when describing dog sounds compared to engine sounds, and 2.3 the odds of using a word describing the acoustic signal rather than the source-event. However, there were only 0.1 times the odds of using a word describing the acoustic signal rather than a subjective-state when describing dog sounds compared to engine sounds.

## Top and Middle Level Sounds: Soundscape, Nature, and Manmade

Sounds and descriptive words for soundscape, nature, and manmade sounds are plotted on the first two dimensions resulting from correspondence analysis of each of the contingency tables in **Figures 1–3**. Note that the dimensions are the same in both panels of each plot. Note also that here and in other two-dimensional plots the descriptive words are those retained following consolidation of the contingency table, and therefore the ratio of descriptive term types differs from that described above. Some insight into category formation is gained by inspecting sounds at the boundaries of the categories. In the top-level study in **Figure 1**, sounds such as footsteps and cutlery are categorized as manmade, though they are closer to the people category than manmade sounds like helicopter and ventilation. This suggests that at the top level, category formation is based upon identification of the sound source-event. Similarly, as might be expected, the rain sounds in **Figure 2** are, despite being part of the weather category, close in space to the water category. In **Figure 3**, the footstep sounds are closer to the home and transport categories than are other industrial sounds. That footsteps are categorized as industrial within the manmade taxonomy, and as manmade within the soundscape taxonomy, suggests that in this instance the sounds were categorized by their acoustic features (e.g., impacts) rather than by sound source-event *per se*.



**FIGURE 4 |** Descriptive words used to describe dog sounds plotted on the first two dimensions of categorization data **(A)**. Words are colored according to type. The regions of the two-dimensional space corresponding to the three main categories of dog sounds are indicated by solid, dashed, and dotted lines. Valence **(B)** and arousal scores **(C)** of affect-judgments are plotted against their coordinates on dimensions 1 and 2, respectively.

**FIGURE 5 |** Descriptive words used to describe engine sounds plotted on the first two dimensions of categorization data **(A)**. Words are colored according to type. The regions of the two-dimensional space corresponding to the three main categories of engine sounds are indicated by solid, dashed, and dotted lines. Fluctuation strength **(B)** and sharpness **(C)** of engine sounds are plotted against their coordinates on dimensions 1 and 2, respectively.
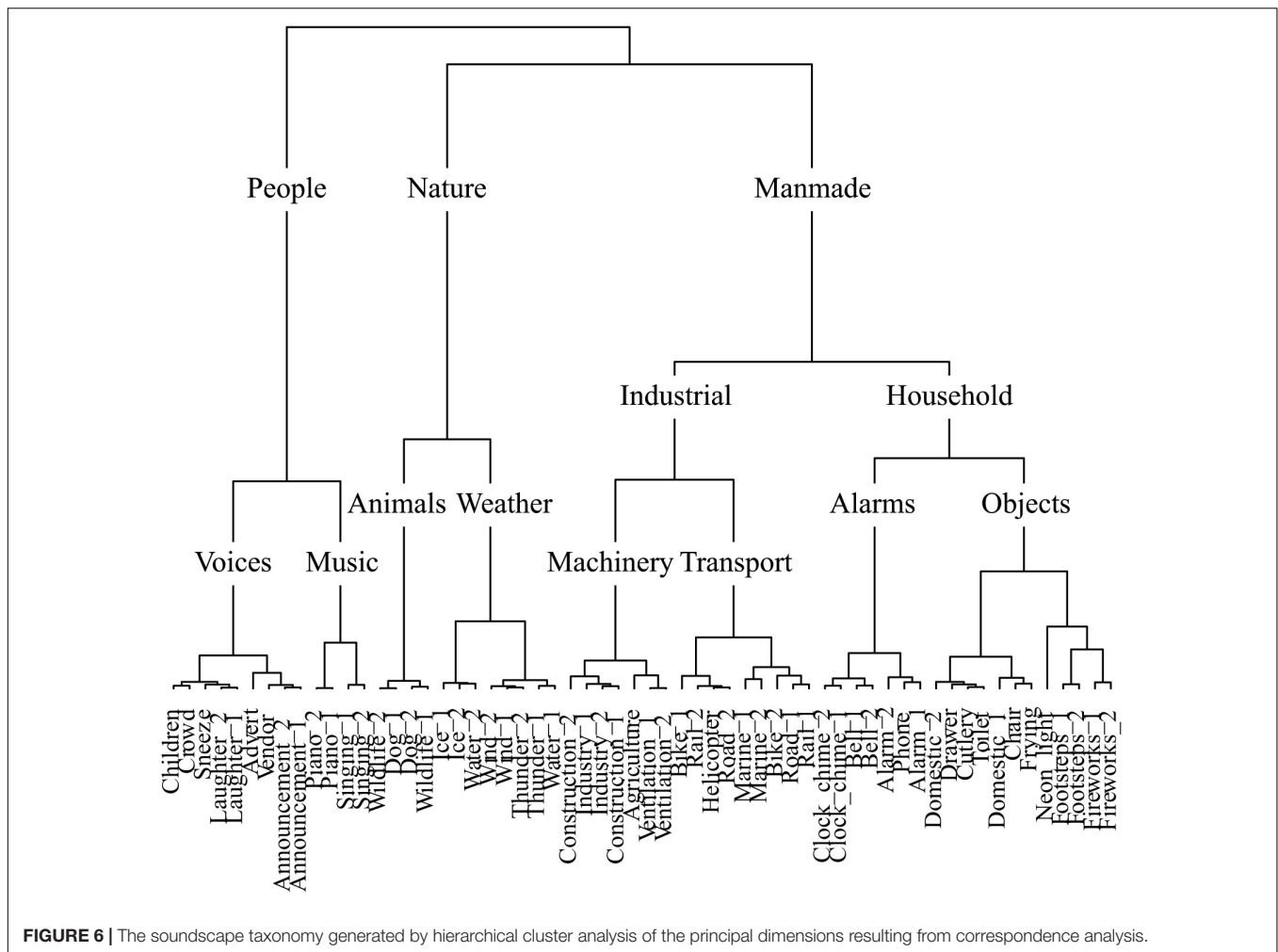
## Bottom Level

### Dogs

The majority of words used to describe dog sounds were those describing subjective-states (**Table 3**), and the odds of using this type of word rather than words describing source-event or acoustic signal was far greater than for engine sounds (**Table 4**). In order to explore this further, dog sounds and descriptive words are plotted on the first two dimensions resulting from correspondence analysis of the contingency table in **Figure 4A**. The first two dimensions accounted for 50.5% of the total variance. The space populated by the howling category contains descriptive words, such as 'sad,' 'lonely,' and 'distressed;' that populated by the yappy category contains descriptive words, such as 'puppy,' 'squeaky,' and 'excited;' and the space populated by the growling category contains descriptive words, such as 'aggressive,' 'snarling,' and 'scary.' More generally, the descriptive words change from being broadly positive to broadly negative along the first dimension, and from describing states of higher to lower arousal along the second dimension. The coordinates of subjective-states on the first dimension were found to correlate with valence scores (**Figure 4B**; $r_s(29) = -0.53$, $p < 0.001$), and their coordinates on the second dimension were found to correlate with arousal scores (**Figure 4C**; $r_s(29) = -0.35$, $p = 0.03$). This is consistent with participants using subjective-states corresponding to valence and arousal to differentiate the dog sounds.

### Engines

Engine sounds and descriptive words are plotted on the first two dimensions of the categorization data, accounting for 33.8% of the total variance, in **Figure 5A**. The chugging category is located on the positive half of dimension 1 and at approximately 0 on dimension 2. The low and jarring categories cover areas from approximately −1 to +0.5 on dimension 1, located below and above 0 on dimension 2, respectively. Since words describing subjective-states made up just 2.5% of descriptive words, category formation of engine sounds differs from dog sounds. Compared to dog sounds the odds of using words explicitly describing the acoustic signal rather than a subjective-state were significantly greater for engine sounds. Visual inspection of **Figure 5A** shows that words relating to temporal regularity (e.g., 'constant,' 'steady,' and 'rumble') are located to the left of the plot and that those relating to temporal irregularity (e.g., 'staccato,' 'stuttering,' and 'chugging') are located to the right. This suggests that the first dimension relates to the fluctuation of the sound. Likewise, dimension 2 of **Figure 5A** may relate to the sharpness of the sound, with terms, such as 'jarring,' 'drilling,' and 'piercing' located toward the top of the plot and terms, such as 'languid,' 'muffled,' and 'hum' toward the bottom. Consistent with these being the basis for category formation of engine sounds, fluctuation strength and sharpness of the engine sounds were found to correlate with the coordinate of each sound on dimension 1 (**Figure 5B**; $r_s(38) = 0.81$, $p < 0.001$) and dimension 2 (**Figure 5C**; $r(38) = 0.83$, $p < 0.001$), respectively.

**FIGURE 6 |** The soundscape taxonomy generated by hierarchical cluster analysis of the principal dimensions resulting from correspondence analysis.

## Taxonomies

**Figure 6** displays the taxonomy derived from cluster analysis of the dimensions of the soundscape contingency table. Sounds are initially partitioned into three categories: people, nature, and manmade. Note that **Figure 6** is limited in depth by the number of sounds used in the top level soundscape experiment (60). Thus the music category, for example, contains only piano and singing sounds. However, the depth of any branch could be expanded by applying the same experimental method to a restricted set of sounds; for example, to 60 different music sounds.

Figure 7 displays the taxonomy derived from cluster analysis of the dimensions of the nature contingency table. The three main categories are animals, water, and nature. **Figure 8** displays the taxonomy derived from cluster analysis of the dimensions of the manmade contingency table. The first division is between outside and home sounds. Outside sounds consist of two categories, transport and industrial. The home category divides into time and daily-life.

Taxonomies derived from cluster analysis of the dimensions of the dogs and engines contingency tables are displayed in **Figures 9, 10**, respectively. Dog sounds are initially partitioned into howling, yappy, and growling. Engine sounds are initially partitioned into chugging and humming. Chugging sounds are further divided into motor-bike and revving sounds; humming sounds are further divided into jarring and low sounds.

## DISCUSSION

### Category Formation

The main aim of the study was to use verbal correlates of sound categorization to explore differences between how categories are formed between and within different levels of category hierarchy. The results demonstrate a significant difference between the types of words used to describe categories of sounds, between the bottom and top levels, and between the bottom and middle levels of the emergent taxonomy. The findings are consistent with source-event identification being the principal cue for category formation at the top and middle levels of the taxonomy. This agrees with previous suggestions that at this level of differentiation, sounds are typically categorized by perceived similarities between sound sources rather than by abstracted acoustic features *per se* (e.g., Gaver, 1993; Marcell et al., 2000; Dubois et al., 2006; Guastavino, 2006; Houix et al., 2012). It
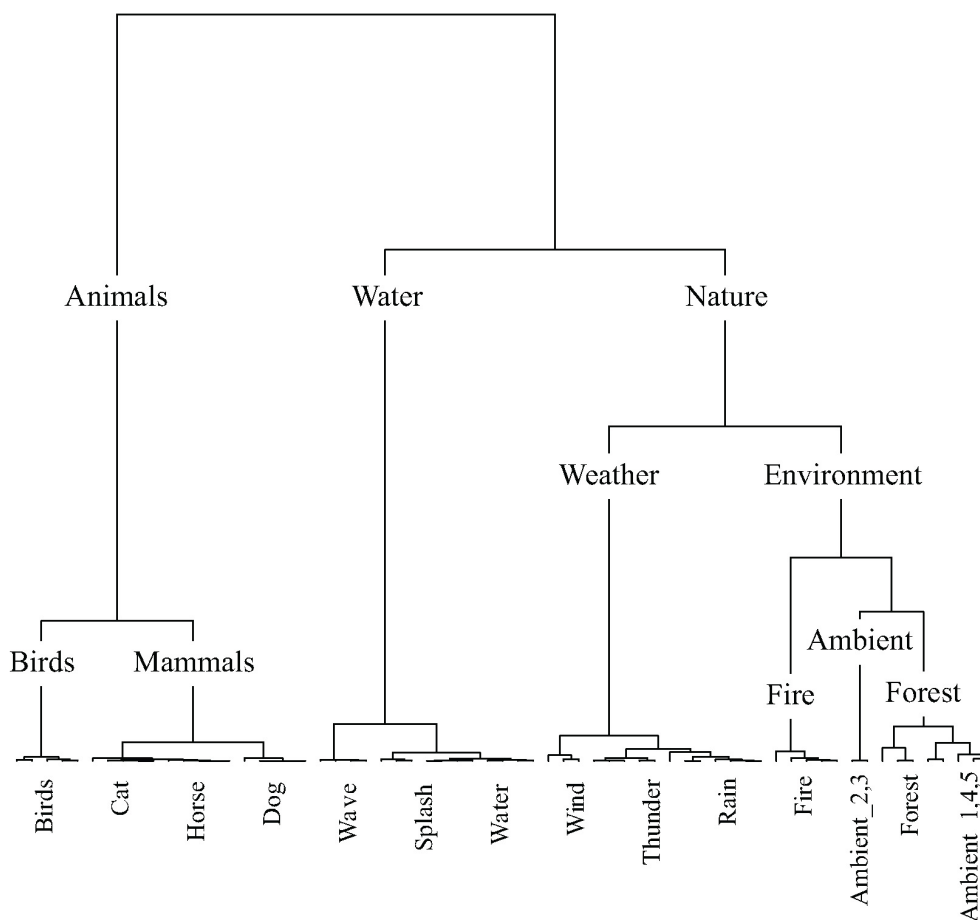
**FIGURE 7 |** The nature taxonomy generated by hierarchical cluster analysis of the principal dimensions resulting from correspondence analysis. Note that five exemplars of each sound were used in this study. All five exemplars of each sound were categorized together, except for the case of ambient where two exemplars were categorized together in a category by the same name and three were categorized as forest sounds.

also concurs with everyday listening being primarily concerned with gathering information about sound sources (Schubert, 1975; Gaver, 1993). However, despite evidence for source-event identification being the principal cue by which categories were formed within the middle level, it was found that nature sounds were more likely than manmade sounds to be described by a subjective-state compared to a source-event, and less likely to be described by explicit reference to the acoustic signal compared to a subjective-state. When categorizing multiple examples of a specific sound source from the nature category (dogs), participants were even more likely to use words describing a subjective-state compared to a source-event, relative to when categorizing multiple examples of a specific sound-source from the manmade category (engines), and even less likely to use words describing the acoustic signal compared to a subjective-state.
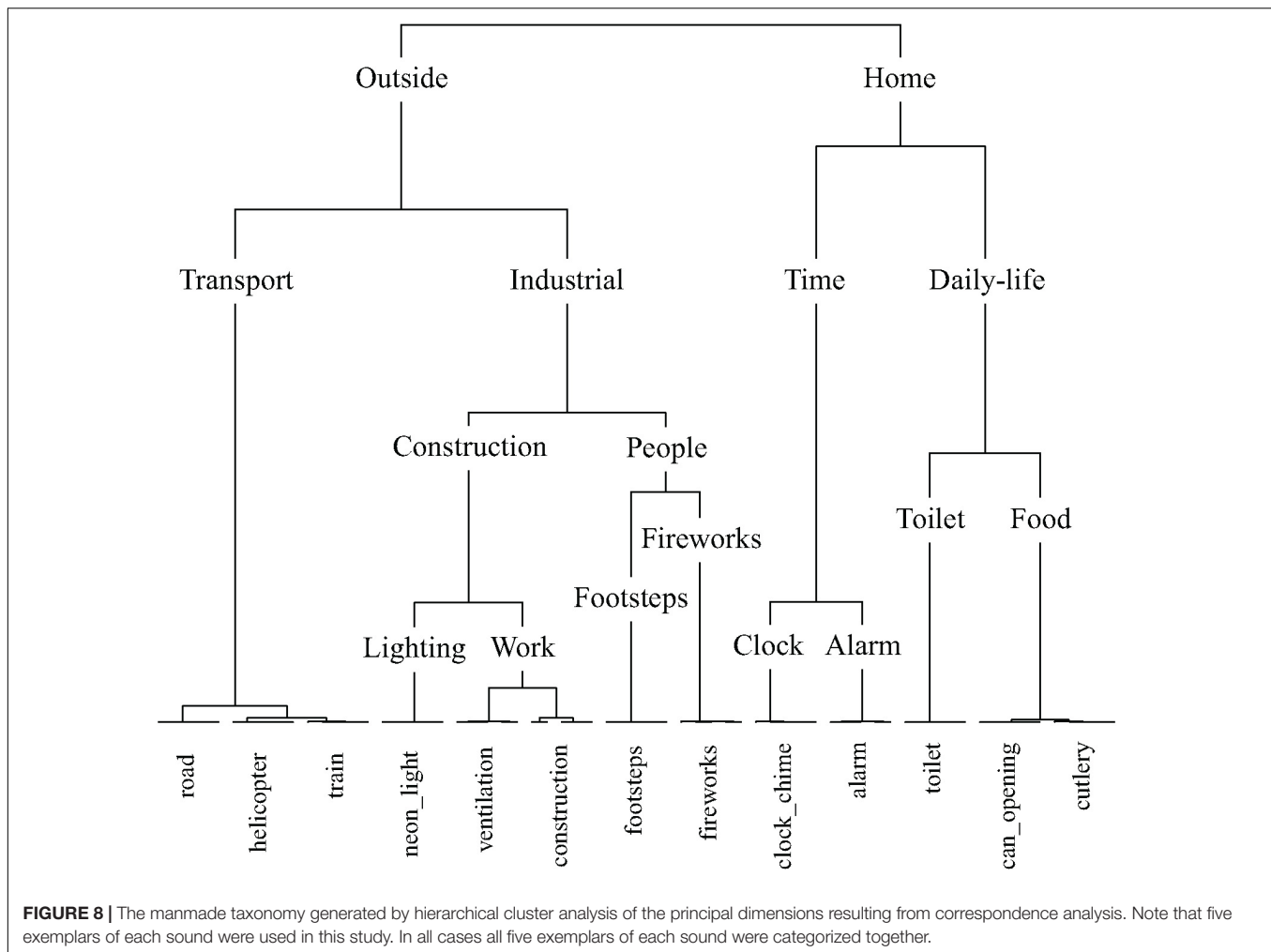
Taken together, these results provide strong evidence that the use of cues for forming categories differs both between and within levels of hierarchy. It is likely that in the case of dog sounds subjective-states represent the greatest potential for differentiating sounds, whereas for engine sounds this strategy

is insufficient or meaningless, and a strategy based upon explicit assessment of the acoustic properties of the sounds is employed.

## Categorization Based Upon Explicit Judgment of the Acoustic Signal

In the case of engine sounds, although the amount of variance explained by two dimensions was low (relative to, e.g., Kawai et al., 2004; Gygi et al., 2007; Axelsson et al., 2010; Hong and Jeon, 2015) they strongly correlated with fluctuation strength and sharpness, respectively, suggesting that these acoustic features were used to differentiate and categorize these sounds. It is notable that despite these acoustic properties being regularly used in product sound evaluation within the automotive industry (e.g., Nor et al., 2008; Wang et al., 2014) to the authors' knowledge this is the first time that a *spontaneous* strategy for differentiating engine sounds using sharpness and fluctuation strength cues has been demonstrated, providing ecological validity to these measures.

One important feature of the approach taken in the present study is that it is possible to interpret the perceptual correlates of fluctuation strength and sharpness of engine sounds using

**FIGURE 8 |** The manmade taxonomy generated by hierarchical cluster analysis of the principal dimensions resulting from correspondence analysis. Note that five exemplars of each sound were used in this study. In all cases all five exemplars of each sound were categorized together.

the spontaneously generated descriptive words. For example, as fluctuation strength increases the engine sounds become more 'chugging' and 'judder'-like etc. This is to say, the data represents a mapping between these acoustic features and their subjective meaning in relation to engine sounds.

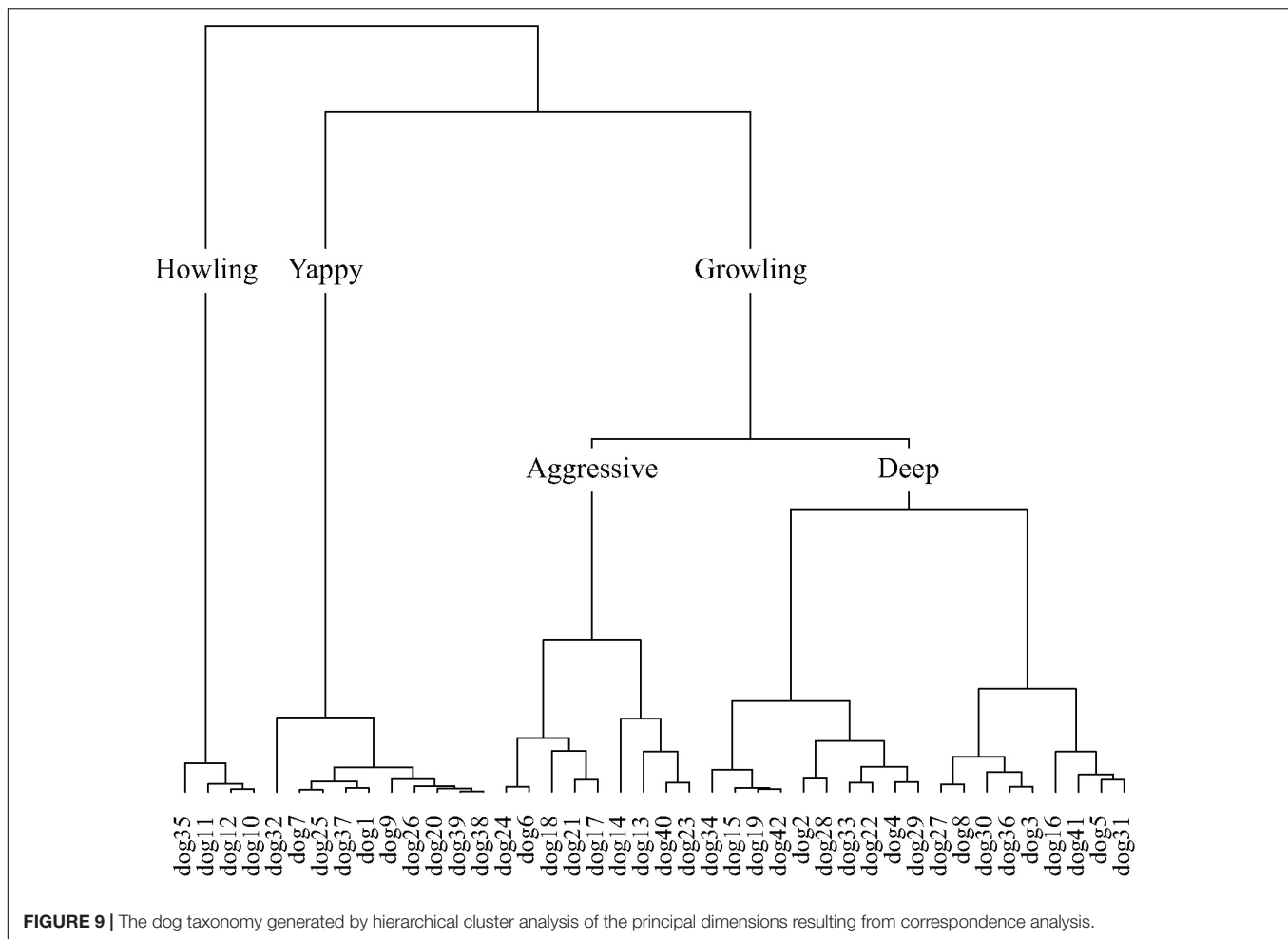### Categorization Based Upon Valence and Arousal

The circumplex model of affect regards valence and arousal as being 'core affect' (Russell, 1980, 2003; Posner et al., 2005) and emotions as being the perceived potential for a stimulus to cause a change in this core affect. Rather than having discrete borders, emotions are understood as being instantiated out of the subjective interpretation of patterns of neurophysiological activity in the mesolimbic system and the reticular formation, responsible for the sensations of valence and arousal, respectively. Previous work has employed the concept of core affect as, for example, an organizing principle for musical sounds (e.g., Gomez-Marin et al., 2016), and as the basis for automatic classification of sounds (Fan et al., 2016). Our finding of an association between the first two dimensions of the dog categorization data and valence and arousal lends support to the circumplex model of affect. It appears to be a meaningful

framework for understanding human categorization of *some* sound types.

Whilst Bergman et al. (2016) found that valence and arousal ratings together mapped onto a dimension of dissimilarity data of everyday sounds explicitly chosen so as to produce an emotional response, we have shown that valence and arousal independently correspond to the first two dimensions of the data from a task where participants were free to categorize by whichever cues they chose. An interesting feature of the method presented here is in the potential for using the spontaneously generated descriptive words that are mapped onto dimensions corresponding to valence and arousal to interpret the perception of affective qualities within the context of dog sounds. For example, it can be said that dog sounds that cause a large valence response are those that are perceived as 'excited,' 'playful,' and 'friendly' (**Figure 4B**), and those that cause a large arousal response are those perceived as 'vicious,' 'snarling,' and 'threatening' (**Figure 4C**).

## Taxonomies

The present study has produced five sound taxonomies using a method where participants were free to use whichever cues they prefer to form categories: a 'top level' soundscape taxonomy,
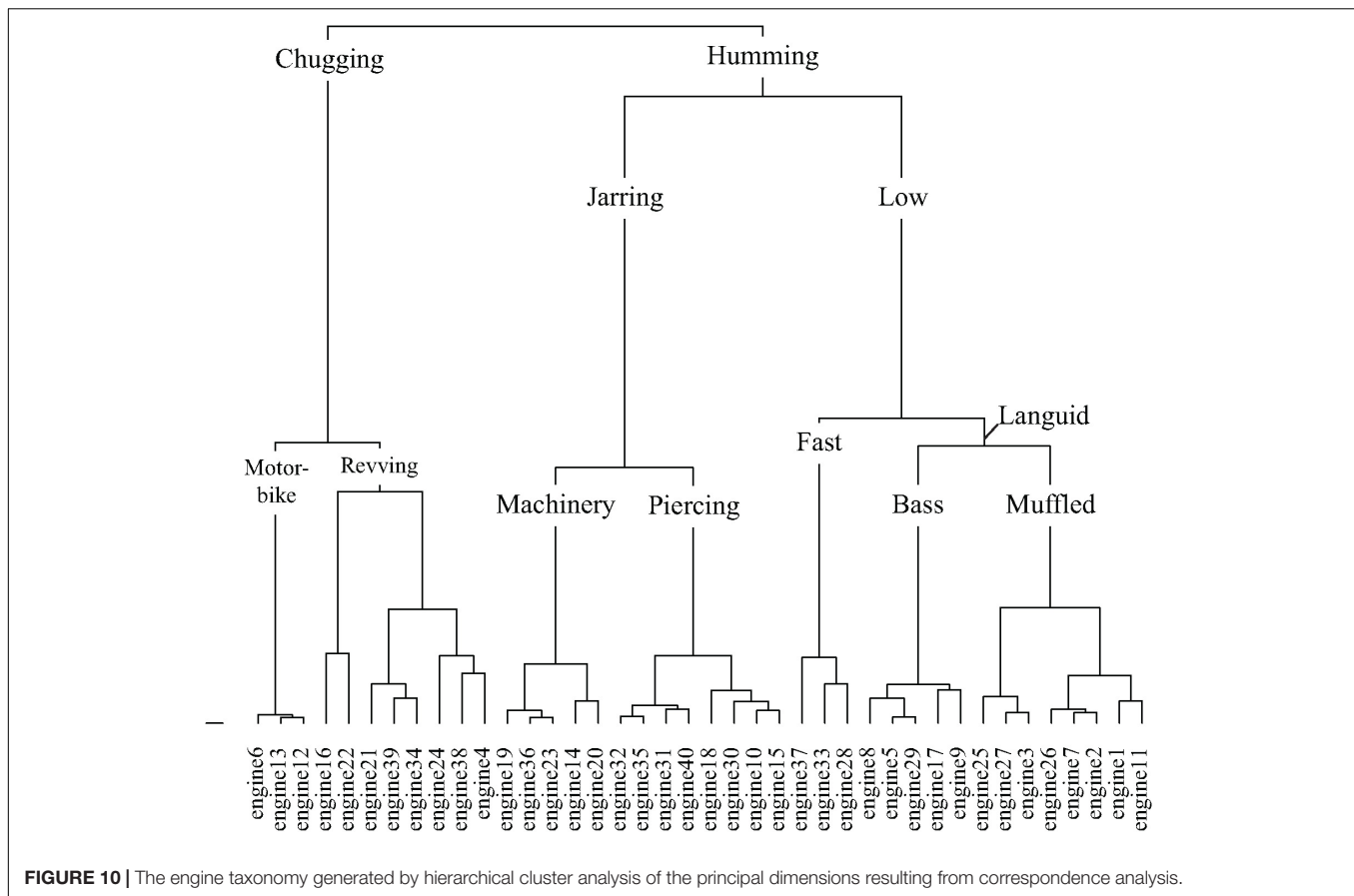
**FIGURE 9 |** The dog taxonomy generated by hierarchical cluster analysis of the principal dimensions resulting from correspondence analysis.

'middle level' nature and man-made sounds taxonomies, and 'bottom level' dog and engine sounds taxonomies. Previous attempts to taxonomize environmental sounds have taken a variety of approaches (e.g., Gaver, 1993; Gygi et al., 2007; Brown et al., 2011; Lemaitre and Heller, 2013; Salamon et al., 2014; Lindborg, 2015). The framework for standardized reporting of events within a soundscape based upon expert opinion produced by Brown et al. (2011) has proven particularly influential in soundscape research, although strictly speaking it is not a taxonomy *per se*. The taxonomies we presented here improve on previous accounts because they are generated experimentally using statistical modeling, being based on the responses of the general public.

The soundscape presented by Brown et al. (2011) is initially divided into indoor and outdoor sounds, with sounds within both further divided into urban, rural, wilderness, and underwater environments. Sounds are then categorized by sound source the same way within each environment. Of the taxonomies presented here, the manmade taxonomy is the only one to have a principal division between environmental contexts, outside and home; for these sounds the environment with which they are most commonly associated was a strong organizing principle. The soundscape taxonomy presented here does not have the same

initial division by environment; rather, sounds are categorized by source-event. It is note-worthy that the categories of sounds prescribed by Schafer (1993), based upon a review of descriptions of sounds in literature, anthropological reports, and historical documents, bear resemblance to a number of categories to have spontaneously emerged here. Schafer's categories: natural, human, society, mechanical, and indicators, are similar to the categories in the soundscape taxonomy: nature, people, manmade, machinery, and alarms, respectively.

It is interesting to note which of the sounds that were used in both the top level soundscape study and the middle level manmade study were categorized differently: the neon-light sound was categorized as 'manmade – household – objects' within the soundscape taxonomy, but as 'outside – industrial – construction' within the manmade taxonomy. Both footstep sounds and fireworks sounds were categorized as 'manmade – household – objects' within the soundscape taxonomy, but as 'outside – industrial – people' within the manmade taxonomy. This is likely to be due in part to an effect of context; within the context of the set of sounds used in the soundscape study the impact sound of footsteps, the snapping and cracking sound of the fireworks, and the popping sound of the neon-light led to these being deemed as belonging together in an objects

**FIGURE 10 |** The engine taxonomy generated by hierarchical cluster analysis of the principal dimensions resulting from correspondence analysis.

category with other sounds with similar acoustic properties such as the sound of rattling cutlery. However, within the context of the sounds used in the manmade study the meaning of the cutlery sounds was more strongly associated with the sound of a can opening, whilst the neon-light sound was more strongly associated with industrial sounds and the footsteps and fireworks sounds were grouped together in a separate people category. Notably, the people category of the manmade taxonomy is somewhat out of place, due to the meaning of the footstep and firework sounds arguably being least similar to the other sounds used in the manmade study, and the least clearly manmade.

## Methodology Considerations and Implications for Soundscape Research

Contrary to the results presented here, a number of soundscape studies have reported principal dimensions relating to subjective-states. This may be due in part to the use of prescriptive semantic differentials ratings in previous studies. Cain et al. (2013) assessed the perception of soundscapes in their entirety and found the principal dimensions 'calmness' and 'vibrancy,' but specifically asked people to rate soundscapes for their 'calmness,' 'comfort,' how fun they were, how confusing they were, and how intrusive they were. Yu et al. (2016) found a principal dimension 'preference,' but used semantic differential scales containing terms, such as 'beautiful,' 'relaxing,' and 'comfortable.' Kang

and Zhang (2010) reported 'relaxation' as the first dimension, using semantic differential scales, such as 'agitating,' 'comfort,' 'pleasant,' and 'quiet.' Similarly, in work more comparable to the present study, Payne et al. (2007) assessed the perception of individual sounds heard within the soundscape and found a dimension 'pleasantness,' but again explicitly used the semantic scales 'pleasantness' and 'stressful.' Unlike the studies mentioned above, Kawai et al. (2004) inferred 'preference' and 'activity' as the two principal dimensions resulting from PCA based upon a semantic differential task using terms generated by participants to describe the sound groupings. However, these were not the original, spontaneously generated names given to groups of sounds, which described the identified sound sources ('sounds of nature,' 'sounds of water'), rather participants were then instructed to further describe sounds within the group with a word that 'best represented the overall representation' and a word with the opposite meaning, in order to construct semantic differential scales. It is likely that this instruction to provide opposing descriptors biased the participants to produce adjectives rather than sound sources.

While these previous studies demonstrate that it is possible to differentiate soundscapes and the quotidian sounds in terms similar to valence and arousal when instructed to do so, our study indicates that this strategy is unlikely to be used spontaneously. Valence and arousal does not reflect the cognitive processes used in sound categorization for four of the taxonomies. This is

consistent with Osgood (1969). He noted that findings generated by his own EPA framework may be a phenomenon that only occurs with forced use of adjectives, e.g., he notes that the concept 'tornado' is regularly rated as highly 'unfair,' despite this making no literal sense.

As noted previously, there was a significant difference in participant age between the dogs and engines studies. It cannot therefore be ruled out that the differences in categorization strategy were due to the larger proportion of 18–29 year olds that took part in the dogs study employing a strategy based upon subjective-states. However, it is suggested that the effect is much more likely to have been caused by the availability and utility of the strategies, reflecting the difference between an animate object with agency and an inanimate machine. In the case of dog sounds, the range and perceived magnitude of affective qualities meant that categorization was easiest based upon this measure. In the case of engine sounds affective qualities were less distinct, whereas the meaning could be better described by the acoustic signals themselves. It is also notable that the strategy used in the engines study was one based upon explicit judgments about the acoustic properties of the sounds despite the participants of this study having the smallest proportion of audio experts.

The taxonomies presented here represent the meanings attributed to the sounds in each study. This differs conceptually to the taxonomy presented by Brown et al. (2011), which was presented as a framework for standardizing soundscape reporting, and so tried to account for as many combinations of source and context as possible. It is noted that differences between the two taxonomies might reflect differences in the sounds selected. Take for example the distinction between 'nature – wildlife' and 'domestic animals' in Brown et al. (2011). This was not found in our study, although maybe a greater sample of both domestic and wild animal sounds would have changed this finding. More generally, the categories presented here are not intended to be taken as absolute. Although the sounds used were chosen to represent sounds frequently reported in the soundscape literature, it must be acknowledged that a different selection of sounds could have resulted in different categories emerging. Context is doubtless an important component of sound perception; for example, one's activity within the context of the soundscape is likely to affect the way in which individual sounds are evaluated (Cain et al., 2008). The procedure described here did not account for such contextual factors; rather the presented taxonomies reflect categories of detached sounds. Similarly, it is likely that in real-world situations perception of the soundscape is shaped by interactions between acoustic and visual cues (e.g., Ge and Hokao, 2005).

## CONCLUSION

Taxonomies of sounds commonly found in soundscape studies, nature sounds, manmade sounds, dog sounds, and engine sounds are presented. Statistical analysis of the frequency with which types of descriptive terms were used demonstrate that whilst participants primarily categorized soundscape, nature, and manmade sounds based upon sound source-event, two further

strategies were used to categorize dog and engine sounds based upon subjective-states and explicit assessment of the acoustic signal, respectively. The dimensions of the dog categorization data corresponded to valence and arousal scores. The dimensions of the engine categorization data corresponded to descriptive terms relating to fluctuation strength and sharpness, and were found to correlate with these two acoustic features. The method used here allows for the interpretation of the subjective meaning of these features within the context of engine sounds: fluctuation strength was perceived as 'chugging' and 'stuttering,' whilst sharpness was perceived as 'jarring' and 'piercing.' Similarly, it can be said that valence is perceived as 'yappy' and 'excited' within the context of dog sounds, and arousal as 'aggressive' and 'growling.' The results of the present study suggest that careful consideration should be given to the appropriateness of the use of prescriptive semantic differential methods in future work.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01277/full#supplementary-material

# REFERENCES

Axelsson, O., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *J. Acoust. Soc. Am.* 128, 2836–2846. doi: 10.1121/1.3493436

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *J. Exp. Psychol.* 19, 250–267. doi: 10.1037/0096-1523.19.2.250

Bergman, P., Västfjäll, D., Tajadura-Jiménez, A., and Asutay, E. (2016). Auditory-Induced emotion mediates perceptual categorization of everyday sounds. *Front. Psychol.* 7:1565. doi: 10.3389/fpsyg.2016.01565

Björk, E. A. (1985). The perceived quality of natural sounds. *Acustica* 57, 185–190.

Brown, A. L., Kang, J., and Gjestland, T. (2011). Towards standardization in soundscape preference assessment. *Appl. Acoust.* 72, 387–392. doi: 10.1016/j.apacoust.2011.01.001

Cain, R., Jennings, P., Adams, M., Bruce, N., Carlyle, A., Cusack, P., et al. (2008). "An activity-centric conceptual framework for assessing and creating positive urban soundscapes," in *Proceedings of the Institute of Acoustics Spring Conference* (Reading: Institute of Acoustics Spring Conference).

Cain, R., Jennings, P., and Poxon, J. (2013). The development and application of the emotional dimensions of a soundscape. *Appl. Acoust.* 74, 232–239. doi: 10.1016/j.apacoust.2011.11.006

Dubois, D. (2000). Categories as acts of meaning: the case of categories in olfaction and audition. *Cogn. Sci. Q.* 1, 35–68.

Dubois, D., Guastavino, C., and Raimbault, M. (2006). A cognitive approach to urban soundscapes: using verbal data to access everyday life auditory categories. *Acta Acust. United Acust.* 92, 865–874.

Fan, J., Thorogood, M., and Pasquier, P. (2016). Automatic soundscape affect recognition using a dimensional approach. *J. Audio Eng. Soc.* 64, 646–653. doi: 10.17743/jaes.2016.0044

Gaver, W. W. (1993). What in the world do we hear?: an ecological approach to auditory event perception. *Ecol. Psychol.* 5, 1–29. doi: 10.1207/s15326969eco0501\_1

Ge, J., and Hokao, K. (2005). Applying the methods of image evaluation and spatial analysis to study the sound environment of urban street areas. *J. Environ. Psychol.* 25, 455–466. doi: 10.1016/j.jenvp.2005.10.003

Giordano, B. L., Mcdonnell, J., and Mcadams, S. (2010). Hearing living symbols and nonliving icons: category specificities in the cognitive processing of environmental sounds. *Brain Cogn.* 73, 7–19. doi: 10.1016/j.bandc.2010.01.005

Gomez-Marin, D., Jorda, S., and Herrera, P. (2016). "Rhythm spaces," in *Proceedings of the 4th International Workshop on Musical Metacreation*, Paris.

Greenacre, M. (1984). *Theory and Application of Correspondence Analysis*. London: Academic Press.

Guastavino, C. (2006). The ideal urban soundscape: investigating the sound quality of French cities. *Acta Acust. United Acust.* 92, 945–951.

Guastavino, C. (2007). Categorization of environmental sounds. *Can. J. Exp. Psychol.* 61, 54–63. doi: 10.1037/cjep2007006

Gygi, B., Kidd, G. R., and Watson, C. S. (2007). Similarity and categorization of environmental sounds. *Percept. Psychophys.* 69, 839–855. doi: 10.3758/BF03193921

Heise, D. R. (2001). Project Magellan: collecting cross-cultural affective meanings via the internet. *Electron. J. Sociol.* 5.

Hong, J. Y., and Jeon, J. Y. (2015). Influence of urban contexts on soundscape perceptions: a structural equation modeling approach. *Landsc. Urban Plan.* 141, 78–87. doi: 10.1016/j.landurbplan.2015.05.004

Houix, O., Lemaitre, G., Misdariis, N., Susini, P., and Urdapilleta, I. (2012). A lexical analysis of environmental sound categories. *J. Exp. Psychol. Appl.* 18, 52–80. doi: 10.1037/a0026240

Husson, F., Josse, J., and Pages, J. (2010). *Principal Component Methods-Hierarchical Clustering-Partitional Clustering: Why Would We Need to Choose for Visualizing Data?* Technical Reports, Applied Mathematics Department. Rennes: Argocampus.

Kang, J., and Zhang, M. (2010). Semantic differential analysis of the soundscape in urban open public spaces. *Build. Environ.* 45, 150–157. doi: 10.1016/j.buildenv.2009.05.014

Kawai, K., Kojima, T., Hirate, K., and Yasuoka, M. (2004). Personal evaluation structure of environmental sounds: experiments of subjective evaluation using

subjects' own terms. *J. Sound Vib.* 277, 523–533. doi: 10.1016/j.jsv.2004.03.013

Kidd, G. R., and Watson, C. S. (2003). The perceptual dimensionality of environmental sounds. *Noise Control Eng. J.* 51:216. doi: 10.3397/1.2839717

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* 25, 1–18. doi: 10.18637/jss.v025.i01

Lemaitre, G., and Heller, L. M. (2013). Evidence for a basic level in a taxonomy of everyday action sounds. *Exp. Brain Res.* 226, 253–264. doi: 10.1007/s00221-013-3430-7

Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., and Deyoe, E. A. (2005). Distinct cortical pathways for processing tool versus animal sounds. *J. Neurosci.* 25, 5148–5158. doi: 10.1523/JNEUROSCI.0419-05.2005

Li, T., and Ogihara, M. (2005). "Music genre classification with taxonomy," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 197, Calgary, AB.

Lindborg, P. (2015). A taxonomy of sound sources in restaurants. *Appl. Acoust.* 110, 297–310. doi: 10.1016/j.apacoust.2016.03.032

Marcell, M. M., Borella, D., Greene, M., Kerr, E., and Rogers, S. (2000). Confrontation naming of environmental sounds. *J. Clin. Exp. Neuropsychol.* 22, 830–864. doi: 10.1076/jcen.22.6.830.949

Morel, J., Marquis-Favre, C., Dubois, D., and Pierrette, M. (2012). Road traffic in urban areas: a perceptual and cognitive typology of pass-by noises. *Acta Acust. United Acust.* 98, 166–178. doi: 10.3813/AAA.918502

Nor, M. J. M., Fouladi, M. H., Nahvi, H., and Ariffin, A. K. (2008). Index for vehicle acoustical comfort inside a passenger car. *Appl. Acoust.* 69, 343–353. doi: 10.1016/j.apacoust.2006.11.001

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychol. Bull.* 49, 197–237. doi: 10.1037/h0055737

Osgood, C. E. (1969). On the whys and wherefores of E, P, and A. *J. Pers. Soc. Psychol.* 12, 194–199. doi: 10.1037/h0027715

Pachet, F., and Cazaly, D. (2000). "A taxonomy of musical genres," in *Proceedings of the Content-Based Multimedia Information Access* (Lafayette, LA: Centre de Hautes Études Internationales D'informatique Documentaire), 1238–1245.

Payne, S. R., Devine-Wright, P., and Irvine, K. N. (2007). "People's perceptions and classifications of sounds heard in urban parks: semantics, affect and restoration," in *Proceedings of the Inter-Noise 2007*, Istanbul.

Piczak, K. J. (2015). "ESC: dataset for environmental sound classification," in *Proceedings of the 23rd ACM international Conference on Multimedia*, New York, NY, 1015–1018.

Posner, J., Russell, J. A., and Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* 17, 715–734. doi: 10.1017/S0954579405050340

Rosch, E. H. (1978). "Principles of categorization," in *Cognition and Categorization*, eds E. H. Rosch and B. B. Lloyd (Hillsdale, NJ: Lawrence Erlbaum Associates).

Russell, A. R. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714

Russell, A. R. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 100, 145–172. doi: 10.1037/0033-295X.110.1.145

Salamon, J., Jacoby, C., and Bello, J. P. (2014). "A dataset and taxonomy for urban sound research," in *Proceedings of the 22nd ACM International Conference on Multimedia*, New York, NY, 1041–1044.

Schafer, R. M. (1993). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Ottawa, ON: Bear & Co.

Schubert, E. D. (1975). "The role of auditory perception in language processing," in *Reading, Perception and Language: Papers from the World Congress on Dyslexia*, eds D. D. Duane and M. B. Rawson (Oxford: Dyslexia).

Ververidis, D., Kotropoulos, C., and Pitas, I. (2004). "Automatic emotional speech classification," in *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, QC, 593–596.

Wang, Y. S., Shen, G. Q., and Xing, Y. F. (2014). A sound quality model for objective synthesis evaluation of vehicle interior noise based on artificial neural network. *Mech. Syst. Signal Process.* 45, 255–266. doi: 10.1016/j.ymssp.2013.11.001

Warriner, A. B., Kuperman, V., and Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* 45, 1191–1207. doi: 10.3758/s13428-012-0314-x

Yang, M., and Kang, J. (2013). Psychoacoustical evaluation of natural and urban sounds in soundscapes. *J. Acoust. Soc. Am.* 134, 840–851. doi: 10.1121/1.4807800

Yu, B., Kang, J., and Ma, H. (2016). Development of indicators for the soundscape in urban shopping streets. *Acta Acust. United Acust.* 102, 462–473. doi: 10.3813/AAA.918965

Zwicker, E., and Fastl, H. (2013). *Psychoacoustics: Facts and Models.* Berlin: Springer Science & Business Media.

# Public Space Users' Soundscape Evaluations in Relation to Their Activities. An Amsterdam-Based Study

Edda Bild[1]*, Karin Pfeffer[2], Matt Coler[3], Ori Rubin[1] and Luca Bertolini[1]

[1] Urban Planning Group, Department of Human Geography, Planning and International Development Studies, Faculty of Social and Behavioural Sciences, University of Amsterdam, Amsterdam, Netherlands, [2] Department of Urban and Regional Planning and Geo-Information Management, Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede, Netherlands, [3] University of Groningen, Leeuwarden, Netherlands

Understanding the relationship between people and their soundscapes in an urban context of innumerable and diverse sensory stimulations is a difficult endeavor. What public space users hear and how they evaluate it in relation to their performed or intended activities can influence users' engagement with their spaces as well as their assessment of suitability of public space for their needs or expectations. While the interaction between the auditory experience and activity is a topic gaining momentum in soundscape research, capturing the complexity of this relationship in context remains a multifaceted challenge. In this paper, we address this challenge by researching the user-soundscape relationships in relation to users' activities. Building on previous soundscape studies, we explore the role and interaction of three potentially influencing factors in users' soundscape evaluations: level of social interaction of users' activities, familiarity and expectations, and we employ affordance theory to research the ways in which users bring their soundscapes into use. To this end, we employ a mixed methods design, combining quantitative, qualitative and spatial analyses to analyze how users of three public spaces in Amsterdam evaluate their soundscapes in relation to their activities. We documented the use of an urban park in Amsterdam through non-intrusive behavioral mapping to collect spatial data on observable categories of activities, and integrated our observations with on site questionnaires on ranked soundscape evaluations and free responses detailing users' evaluations, collected at the same time from park users. One of our key findings is that solitary and socially interactive respondents evaluate their soundscapes differently in relation to their activities, with the latter offering higher suitability and lower disruption ratings than the former; this points to qualitatively different auditory experiences, analyzed further based on users' open-ended justifications for their evaluations. We provide a methodological contribution (adding to existing soundscape evaluation methodologies), an empirical contribution (providing insight on how users explain their soundscape evaluations in relation to their activities) and a policy and design-related contribution, offering additional insight on a transferable methodology and process that practitioners can employ in their work on the built environment to address the multisensory experience of public spaces.

Keywords: soundscape evaluation, activity, public space, familiarity, expectation, affordance

# INTRODUCTION

Research shows that urban sound affects the health and well-being of urbanites in a significant manner, at the same time influencing the use and appreciation of public spaces (Mehta, 2014; Van Kempen et al., 2014). Given this demonstrated importance of sound as part of the urban experience, scientists and practitioners alike have sought to develop strategies to research and influence the relationship between urbanites and their soundscapes, on the one hand to minimize the potential negative effects of sound on urban life, and on the other hand to maximize the opportunities for enjoyment or relaxation that urban sound offers. Whilst extensive attention has been paid to aspects of soundscape evaluation that could potentially feed into effective urban sound policies (Andringa et al., 2013), capturing the complexity of the qualitative urban auditory experience in context (in a real life setting) remains a challenge.

The challenge has both methodological and empirical dimensions, as well as policy and design implications. Strategies focused on how users of various urban spaces evaluate their soundscapes are relatively common in both soundscape research as well as urban policy or practice-related initiatives (see e.g., Axelsson et al., 2010; Booi and van den Berg, 2012; Lercher et al., 2016). However, the conventional methods and tools to study those evaluations are limited in their scope. For example, with regards to public spaces, evaluations are currently mostly collected using questionnaires, largely disregarding other (potentially less invasive) methods that can contribute to a more holistic understanding of the relationship between public space users and their soundscapes, in context. *In situ* methods like field observation (and behavioral mapping) are still rarely used in soundscape research and are currently in the "experimental" stage of implementation with inconsistent results (see Steele et al., 2016; Aletta et al., 2016b; Bild et al., 2018; Lavia et al., 2018, for different approaches). Furthermore, the questionnaires used as tools to gain insight on users' soundscape evaluations mostly employ categorical-based assessments and rarely include open-ended questions (see Yang and Kang, 2005; Raimbault, 2006, the work in the "Positive Soundscapes Project"[1], Nielbo et al., 2013; Bild et al., 2018 for examples), thus representing a limited understanding of users' soundscape evaluations. Finally, these methods minimize or do not adequately account for the role of moderating factors, like activity, in influencing how people evaluate what they hear, despite increasing evidence on activity as a moderating activity for users' soundscapes (e.g., Aspuru et al., 2011; Bild et al., 2015, 2018; Steffens et al., 2015). The challenge has implications for sound-related urban practice and design initiatives, as it affects the adequate and comprehensive collection and implementation of soundscape knowledge in everyday projects.

In this paper we propose to address these shortcomings in a large-scale, multi-sited urban study based in Amsterdam (Netherlands), where we used a mixed methods approach combining fieldwork observations with questionnaires to capture both reported and "enacted" soundscape evaluations (materialized through public space use). In examining users' evaluations of their soundscapes in urban public outdoor spaces, we rely on *users' activities* as a key variable that can influence their evaluations, and, through that, the current and future use of the urban public space (see Nielbo et al., 2013; Steffens et al., 2017; Bild et al., 2018 for comparable approaches). With this in mind, this paper aims to understand the factors that can influence and moderate, both separately and together, how users of three different public spaces evaluate their soundscapes in relation to their on-site activities. Previous soundscape studies indicate three potential factors that can affect the user-soundscape relationship: *the level of social interaction of users' activities*, users' (auditory) *expectations* and users' *familiarity* with the space and with what they hear. To research how the factors interact while influencing users' evaluations of their soundscapes in relation to their activities, we integrate the concept of *affordance* (Gibson, 1977, 1979) as a conceptual framework for understanding the user-soundscape relationship, focusing on how people bring their soundscapes "into use" in their everyday life (Ingold, 2000), through their activities.

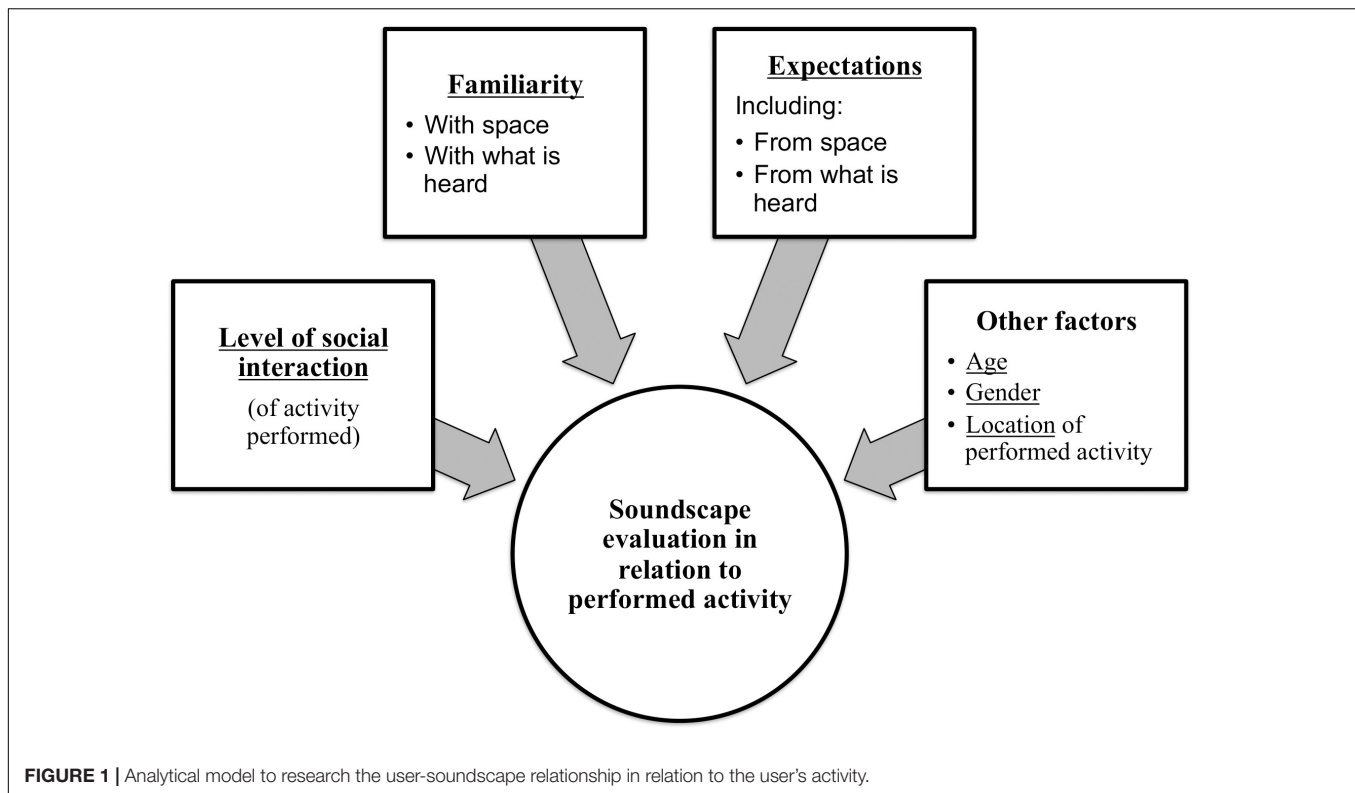We aim to answer the following two research questions:

(1) To what extent do the level of social interaction of users' activities, users' expectations, and users' familiarity (with the space and with what is heard) influence their soundscape evaluations in relation to their activities, and how are the factors associated?

(2) What are the possible reasons for those associations and how do users describe the factors in relation to their soundscapes?

The scientific work that we build on in this paper is detailed in Section "Background." The data collection and analysis methods are discussed in Section "Materials and Method" and the findings of the analysis are covered in Section "Results." We discuss the research and practice-related gaps that we address in detail in the following section and in the concluding discussion (see section "Discussion"); in the latter we also outline the three contributions of this study: empirical, methodological and policy and design-oriented.

# BACKGROUND

The evaluation of soundscapes is at the center of efforts of scientists from disciplines as diverse as psychology or anthropology, particularly of those working at the intersection of theoretical and applied research, as they aim to understand how users of various urban public and private spaces engage with and relate to what they hear, and how that influences the quality of their experience. In this section, we review the scientific literature key to the evaluation of soundscapes and develop the analytical model described below guiding the empirical research (**Figure 1**). First, we discuss studies exploring the role of activity in soundscape evaluations, including the specific effect of *level of social interaction of one's activities*. Then we review studies

---

[1]For a review of papers published in this project, check the project URL at: https://www.salford.ac.uk/research/sirc/research-groups/acoustics/psychoacoustics/positive-soundscapes-project [Accessed May 2nd 2018].

**FIGURE 1 |** Analytical model to research the user-soundscape relationship in relation to the user's activity.

researching the role of other factors influencing soundscape evaluations, like one's previous experience, as it relates to *expectation* and *familiarity*, on these evaluations (discussed below in detail and summarized in the analytical model proposed in **Figure 1**). Finally, we use the concept of *affordance* to understand how users of public spaces bring their soundscapes into use through their engagement with and activities performed in their public spaces (Ingold, 2000; Steenson and Rodger, 2015).

## Soundscape Evaluations and Activity

Scientific efforts have been made to determine soundscape descriptors and indicators that can help explain or predict users' soundscape evaluations (Nilsson et al., 2007; Jennings and Cain, 2013; Steenson and Rodger, 2015; Herranz-Pascual et al., 2017), with an eye on operationalizing this knowledge and implementing it in sound-related practices. The dominant approach studies evaluations in relation to sound/soundscape quality (see Schulte-Fortkamp and Fiebig, 2016) and integrates aspects of pleasantness (Raimbault, 2006; Axelsson et al., 2010; Can et al., 2016; Herranz-Pascual et al., 2017, *inter alios*) and quietness (Pheasant et al., 2008; Booi and van den Berg, 2012; Bloomfield, 2014; Aspuru et al., 2016), usually in contrast with annoyance (see e.g., Lercher and Schulte-Fortkamp, 2003; Andringa and Lanser, 2013).

While the role of users' activities as a variable potentially influencing their relationship with their soundscapes has been suggested before (Dubois, 2000; Lercher and Schulte-Fortkamp, 2003), the effective and explicit integration of activity in scientific research with a focus on urban public spaces is still in its

incipient, exploratory phase (Aletta et al., 2016b; Bild et al., 2016, 2018; Lavia et al., 2016; Steffens et al., 2017). Most of these research projects arise from more practice-oriented questions, either dealing with specific soundscape interventions with some form of behavioral control in mind (see Lavia et al., 2012, 2016), or emphasizing the role users' soundscapes play in relaxation or rehabilitation activities or in relation to auditory comfort, both indoors and outdoors (Mzali, 2002; Delepaut, 2009; Cerwén et al., 2016; Filipan et al., 2017). Consequently, many questions remain on how best to define and operationalize *activity* in empirical studies and what methods are suited for researching the relationship between soundscape evaluations and activities in an ecologically valid manner (Guastavino et al., 2005). For example, one laboratory study demonstrated that various soundscape recordings were evaluated as being appropriate[2] for different imagined activities by participants in a listening experiment (Nielbo et al., 2013); it, however, remains unclear how we can transfer the outcomes from research performed in a laboratory to research performed on-site. We address this issue by furthering the exploration of the aforementioned relationship with a focus on understanding the role activity plays in influencing public space users' soundscape evaluations. We base part of our inquiry on preliminary studies

---

[2]The term "appropriateness" has been used to measure users' evaluation of their soundscapes in relation to their settings, defined either as their performed or imagined activity – appropriateness for activity (Nielbo et al., 2013; Bild et al., 2018) or, more commonly, the geographical setting they find themselves in – appropriateness in/for "a place" (Lavia et al., 2012; Aletta et al., 2016a). In this paper, we focus on the former understanding of the term.

in the field, showing that *the level of social interaction of users' activities* has an influence on how users evaluate their soundscapes in relation to their activities (Bild et al., 2018) and a marginal effect on their soundscape descriptions. In other words, to what extent does whether users are alone or with others influence how they evaluate their soundscapes in relation to what they were doing (e.g., talking, reading, sunbathing)?

## Influencing Factor: Expectation

Bruce and Davies (2014) relate soundscape expectations to Truax's concept of "soundscape competence" (2001), referring to the ability of users of a space to interpret and make sense of what they hear, based on previous experience, and framing soundscape expectations for future situations. Filipan et al. (2017) suggest that the presence or absence of certain *expected* sounds in a context (like a park) can affect users' evaluations of their soundscapes in terms of, e.g., tranquility. Along similar lines, Bruce et al. (2009, p. 6) argued that a user's soundscape "becomes an issue when it does not conform to subjects 'perceived' sense of normality or interferes with information [...] transfer," thus not conforming to the users' expectations. The complexity of expectations in relation to one's experience has been explored extensively apropos music (see e.g., Huron, 2006) and only recently has it been researched explicitly in relation to soundscape (Bruce et al., 2009, 2015; Bruce and Davies, 2014). We build on the conclusions of the latter research avenue, particularly their preliminary findings on the effect of users' expectations from the space and what they hear on their soundscape evaluations, as well as what users refer to as "expected activities" within the space, as influenced by their soundscapes (Bruce and Davies, 2014).

## Influencing Factor: Familiarity

Familiarity is understood as "how usual or common a stimulus is in the subject's realm of experience" (Marcell et al., 2000, p. 834), referring to the previous experience of the user with their space, which includes their frequency of use of a space as well as activities performed in the space (Kogan et al., 2017). Particularly for the auditory domain, "familiarity" is one of the three factors that influence the "identifiability" of sounds along with "complexity" (Marcell et al., 2000) and "pleasantness," as well as one of the three features or perceptual attributes that Axelsson found to be most relevant for users' evaluations of their soundscapes (third after pleasantness and eventfulness – Axelsson et al., 2010). Axelsson found that variance in familiarity ratings tends to be low for urban respondents sharing a similar cultural framework, thus implying a limited applicability of the feature for design initiatives (Axelsson et al., 2010). We nonetheless consider that users' reported familiarity both with the space and with what they hear provides valuable insight into users' evaluations of their soundscapes in relation to their intended or performed activities; familiarity is essential in relation to aspects of expectations, and failure or success to meet them, as it relies on users' previous knowledge and experience.

## Soundscape and Affordance

Considering the activity-centered approach we take in this paper, we integrate the concept of (auditory) affordances in a public space context. In Gibson's formulation, affordances are defined as the qualities of an object or an environment that allow for the performance of an activity (Gibson, 1977). Turvey (1992, p. 174) describes an auditory affordance as a way to "provide a description of the environment that was directly relevant to behavior.". Affordances have been discussed and used previously in auditory research particularly in relation to music, in reference to what music can *afford* to a listener (see DeNora, 2000; Clarke, 2005; Reybrouck, 2012, *inter alios*). There have also been proposals and strategies for integrating the concept in soundscape research (Thibaud, 1998; Pecqueux, 2012; Nielbo et al., 2013; Nielbo, 2015; Steenson and Rodger, 2015) to more accurately address the complexities of user-soundscape relationships and articulate the role that users' soundscapes play in guiding or informing their public space experiences and uses. We follow Steenson and Rodger's reading of Gibson in relation to the auditory domain, suggesting that "auditory information is formed relationally, emerging with the situated activity of the agent" (2015, p. 181). We build on the work of Pecqueux (2012), who expands on the idea of affordances elicited by sounds in urban settings (p. 221) and demonstrates the relevance of an activity-centered strategy to researching the urban auditory experience, with implications for design practice. In our approach, we also extend on the idea of "actualization of affordances" (Kyttä, 2002; Stoffregen, 2003), that is, turning possibilities for action into actual activities, focusing on understanding how sounds are brought into use in an urban context (Steenson and Rodger, 2015). We articulate the notion that, by affording users' activities, users' soundscapes can enable or impede their activities.

## Proposed Analytical Model

**Figure 1** summarizes the various strands of soundscape research to understand the individual and interaction effect of three factors over users' soundscape evaluations in relation to their performed activities: (1) the level of the social interaction of users' performed activities (i.e., solitary vs. socially interactive), (2) expectation (including expectation from the space and from what is heard), and (3) familiarity (with a focus on familiarity with the space and with what is heard). The analytical model informs our mixed methods approach to the evaluation of soundscapes in relation to activity detailed in the next section.

## MATERIALS AND METHODS

To research to what extent the level of social interaction of users' activities, users' expectations, and users' familiarity (with the space and with what is heard) influence their soundscape evaluations in relation to their activities, and how these factors are associated, we combined quantitative, qualitative and spatial methods in the collection and analysis stages as part of a mixed methods approach (see Creswell and Plano Clark, 2011). Mixed methods approaches are common in soundscape research, as

the complexity of people's urban experiences cannot be fully grasped in mono-method studies (Bloomfield, 2014; Aletta et al., 2016a; Herranz-Pascual et al., 2017; Bild et al., 2018). They are conducive to a more nuanced, situated and integrated exploration of the relationship between users of public spaces and their soundscapes, in context (see also Knigge and Cope, 2006 with respect to the integration of qualitative and quantitative data).

In our research, we relied on a combination of on-site data collection methods, including self-completion questionnaires with randomly selected public space users, and non-participant observation of activities performed in the selected public spaces. The questionnaires included both soundscape evaluations/ratings as well as open-ended questions asking respondents to reflect on their ratings. We collected different types of data suited for both quantitative and qualitative analyses, ultimately contributing to a multi-layered understanding of users' on-site experience in relation to their activity as follows. The quantitative analysis allowed us to measure potential differences in soundscape ratings between public space users performing different activities and to test the role of various factors in influencing these ratings; the qualitative analysis offered a more nuanced understanding of users' ratings as well as an in-depth exploration of the reasons behind the aforementioned potential differences between user groups. The non-participant observation of activities was done to situate users' auditory experiences and their soundscape evaluations in a spatial and behavioral context. In the following sections, we first describe the data collection, including the research design, fieldwork locations and data collection methods, and then elaborate on the data analysis methods.

## Data Collection

We employed a mixed methods research design relying on parallel data gathering. Building on previous pilot studies (Steele et al., 2016; Bild et al., 2018), we combined field observations with on-site questionnaire data collection in a multi-sited field research. We collected data in three public spaces (two large urban public parks and one small urban "plein"/square) over the summer of 2016. 208 self-completion questionnaires were collected with Dutch public space users in similar weather conditions (sunny, warm, and dry), during two data collection sessions per space. Two types of data were collected at the same time: (1) using questionnaires, the ratings and open-ended responses on users' experiences, and, (2) using field observation, the patterns of occupancy of the public space by solitary and socially interactive users (including the spatial position of users who completed the questionnaire).

## Fieldwork Locations

The fieldwork was conducted in various areas of three different locations (**Figure 2**): two traditional urban parks (Oosterpak and Sarphatipark) and one smaller square-park hybrid (Frederiksplein). The spaces are located in central Amsterdam and were selected due to their heavy use for leisure purposes. They represent typical Dutch urban public spaces that can be split in smaller areas bordered by paths and greenery, and
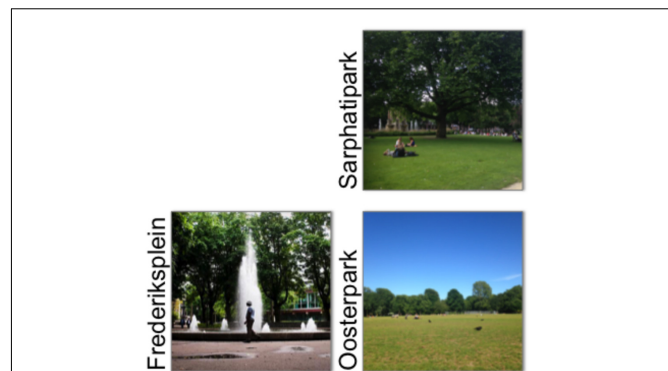


**FIGURE 2 |** Fieldwork locations. Photo credits: Frederiksplein: Het Parool[3]; Sarphatipark: authors; Oosterpark: authors.

are designed with diverse amenities encouraging mixed use and users (see **Table 1** below).

## Questionnaire Data Collection

### Questionnaire design

The aim of the questionnaire was to understand whether soundscapes were evaluated as affording users' activities in a public space context by researching users' soundscape evaluations in relation to their activities. Questionnaires used in previous research on soundscape evaluations tend to address experiences of spaces in relation to perceptions of pleasantness or eventfulness (e.g., Axelsson et al., 2010; Herranz-Pascual et al., 2017), rarely going in depth on the relationship between use of space and soundscape evaluation. These lines of questioning usually rely on semantic scales and seldom employ additional open-ended questions asking respondents to expand on their evaluations, effectively limiting their applicability in practice (Raimbault, 2006; Nielbo et al., 2013). Current standardized protocols (e.g., the "Soundscape Quality Protocol" – SSQP, Axelsson et al., 2010) might prove insufficient to collect insight useful for both urban researchers aiming to understand the user-soundscape relationship as well as city makers interested in developing spaces with sound in mind, as they would not offer

**TABLE 1 |** Fieldwork locations: description and amenities for observed areas.

| Location | Description | Amenities |
| --- | --- | --- |
| Oosterpark | Large urban park | Large green fields Benches Pond with waterfront green areas Paths Gray/built open area with benches and other sitting possibilities |
| Sarphatipark | Large urban park | Large green fields Benches Pond with waterfront green areas Paths Water fountain |
| Frederiksplein | Smaller-sized square-park hybrid Transition space (from center to adjoining neighborhood) | Gray/built open area Benches Paths Water fountain with benches surrounding it Tram tracks cutting through the space |

substantial insight into *what* in users' soundscapes is perceived as disrupting or suitable for their activities or purposes of use of space. We addressed these two challenges in the design of the questionnaire by focusing on users' soundscape evaluations in relation to their activity and by combining Likert items with open-ended questions in one questionnaire to understand how users reflect on the effect of their soundscapes on their activities and explain potential discrepancies in their evaluations (see **Table 2**).

Based on the analytical model outlined in **Figure 1**, we aimed to research the potential influence of three factors on users' soundscape evaluations in relation to activity, i.e., the level of social interaction of their activity, their familiarity (with what is heard and with the space), and their expectation, and whether these three factors interact for a stronger effect. Additionally, as indicated by literature, we also explored the potential effect of age and gender to influence auditory experiences and potentially the aforementioned evaluations (see e.g., Yang and Kang, 2005).

To understand whether their soundscapes afforded their on-site activities, we asked users to evaluate their soundscapes from three perspectives: in terms of disruption, stimulation and overall suitability; we afterwards asked for detailed explanations of their evaluations (see **Table 2** below). Stimulation is a common term used in relation to soundscapes and particularly in soundscape evaluation usually used as an adjective (Axelsson et al., 2010; Botteldooren et al., 2015), but we use it as an active verb ("to stimulate"). While some authors prefer "to disturb" (and "disturbance") to convey a similar message (see e.g., Lercher et al., 2016), we selected "to disrupt" as an antonym for "stimulate," due to its nature as a transitive verb as well

as its common use in relation to activity (e.g., Truax, 2001). We did not introduce the concept of "soundscape" in the questionnaire, as we wanted to ensure the statements were phrased in a "natural," everyday language, allowing respondents to focus on their experience rather than on relating to a new concept.

### Questionnaire data collection protocol

We approached park users who were usually seated (not in transit), and were willing to engage with the data collector and complete the questionnaire; the questionnaires were completed by native Dutch speakers. Park users were handed clipboards and pens, and were invited to fill out the questionnaires themselves. The data collector offered clarifications when needed. We gathered 188 questionnaires in the three fieldwork locations (Oosterpark: 81 questionnaires, Sarphatipark: 83, Frederiksplein: 24), as part of two data collection sessions per location (one in the weekend and one during the week).
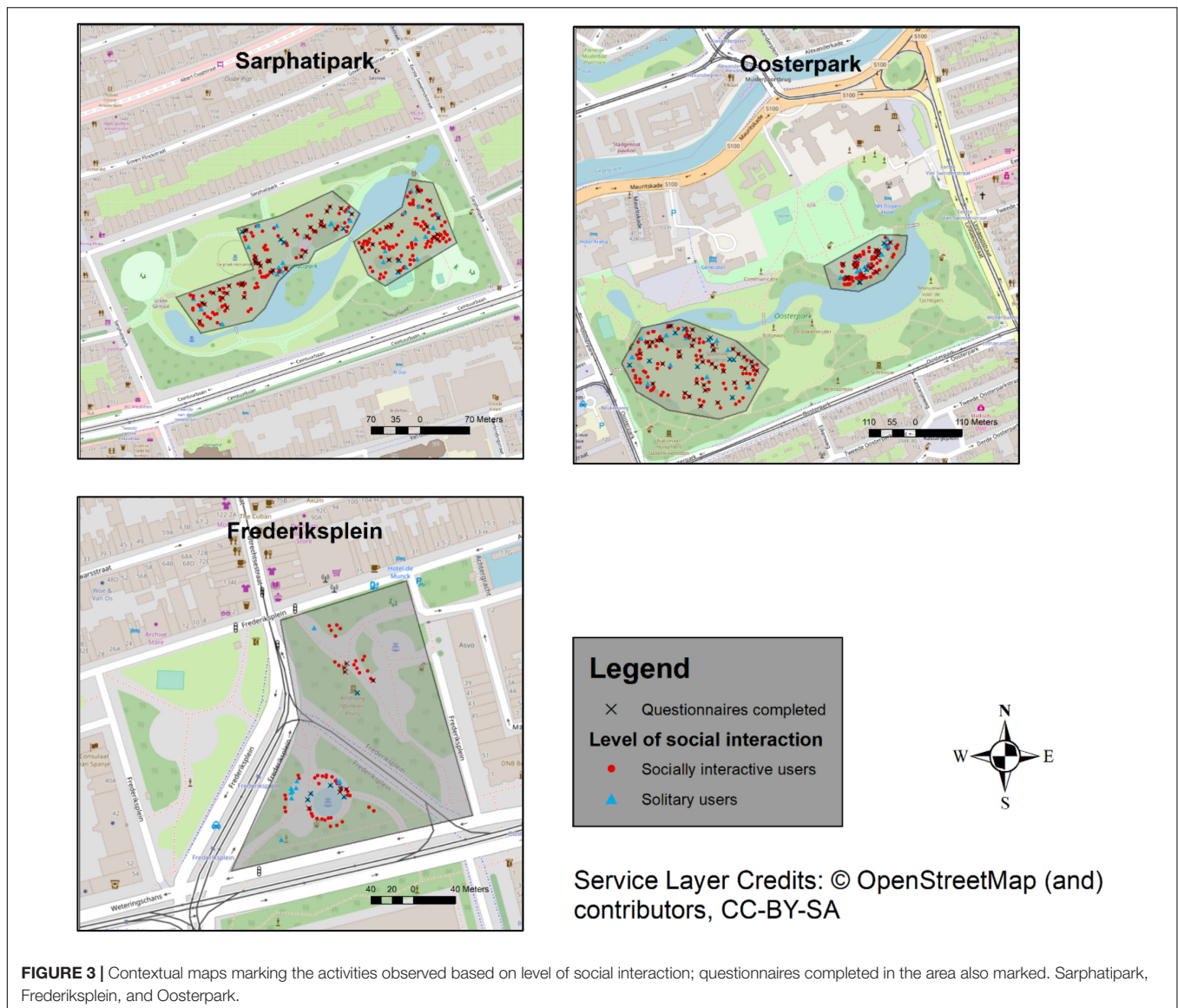
## Non-participant Observation

To situate the questionnaire data on users' soundscape evaluations in relation to their activities in a spatial and behavioral context, we also relied on systematic non-participant observation as a fieldwork method, more specifically, behavioral mapping (see e.g., Cosco et al., 2010; Goličnik and Thompson, 2010; Bild et al., 2018). Field observation (Aletta et al., 2016b; Lavia et al., 2016) has been increasingly integrated in urban soundscape research, particularly to document the effects of certain acoustic interventions on the ways in which people

**TABLE 2 |** Questions (in the order in which they were asked in the questionnaire).

| Variable | Question/statement (translation EN) | Question/statement (original NL) | Type of response |
|---|---|---|---|
| Activity (including level of social interaction) | Think back on the activities you perform in this park and **describe** them in as much detail as possible. | Denk terug aan uw activiteiten in dit park vandaag en **beschrijf** deze zo uitgebreid mogelijk. | Open-ended response. Coded for level of social interaction |
| Disruption | The performance of my activities was disrupted by what I heard. | Het uitvoeren van mijn activiteiten werd verstoord door wat ik hoorde. | Likert item: 1–5 (1, "I completely disagree" to 5, "I completely agree") |
| Stimulation | The performance of my activities was stimulated by what I heard. | Het uitvoeren van mijn activiteiten werd gestimuleerd door wat ik hoorde. | Likert item: 1-5 |
| Explanation of disruption or stimulation ratings | In what ways did what you hear disrupt or stimulate the activities you performed? | Op welke manieren werd het uitvoeren van uw activiteiten verstoord dan wel gestimuleerd door wat u hoorde? | Open-ended response |
| Suitability | What I heard was suitable for the activities that I performed. | Dat wat ik hoorde was toepasselijk voor de activiteiten die ik uitvoerde. | Likert item: 1–5 |
| Expectations | Did you have expectations about this park before you came here? If yes, what were they? | Had u verwachtingen over het park voordat u hier kwam? Zo ja, wat verwachtte u? | Open-ended response |
| Satisfaction of expectations | In what respect did what you hear match (or not) your expectations? | In hoeverre voldeed wat u hoorde aan uw verwachtingen? | Open-ended response |
| Familiarity with what is heard | I am familiar with what I heard during the activities I performed. | Ik was bekend met wat ik hoorde tijdens het uitvoeren van mijn activiteiten. | Likert item: 1–5 |
| Familiarity with space (i.e., frequency of use of space) | How often do you visit this park? | Hoe vaak bezoekt u dit park? | Categorical scale: 1–4 (from 1, "Once a week," to 4, "It is my first time here") |
| Age | What is your age? | Wat is uw leeftijd? | Continuous |

*Original Dutch and English translation.*

**FIGURE 3 |** Contextual maps marking the activities observed based on level of social interaction; questionnaires completed in the area also marked. Sarphatipark, Frederiksplein, and Oosterpark.

engage with and act in their public spaces[3]. Documenting public space use is crucial for on-site studies, as it shows how users relate to and behave in their physical (built) environments and how this relationship can further connect with their soundscape evaluations. By spatially mapping and situating the evaluations of users and their engagement with its amenities and with each other, we can explore how their physical environments and their soundscapes may interact to influence their urban experience in relation to their activities.

In this paper, using a behavioral mapping application[4], we gathered data on the level of social interaction of activities performed by public space users (individual, in pairs or in groups), in parallel with the collection of questionnaires, as part of hour-long sessions throughout the research period. The behavioral mapping resulted in a total of 665 distinct data points, referring to both individual users and users in groups in all three locations, in selected areas of each location. The 665 points include the 188 questionnaire respondents/public space users[5], and are marked on the resulting behavioral maps (**Figure 3** in section Contextual Maps").

## Data Analysis

To answer our two research questions, we analyzed the questionnaire data using a sequential approach, first statistically analyzing the responses to the closed-ended questions and

---

[3]For another application of this method as part of a mixed methods approach, see Steele et al. (2016).

[4]The method has been tested and implemented in two previous smaller scale studies (viz. Bild et al., 2018, for a detailed explanation of the behavioral mapping method and tools; see also Steele et al., 2016).

[5]We hereinafter refer to the public space users who completed our questionnaire as "respondents."

afterwards qualitatively analyzing the responses to the open-ended questions. The quantitative analysis served to establish potential patterns in the ways in which solitary and socially interactive respondents evaluate their soundscapes in relation to their activities and the role of, e.g., familiarity as a factor in influencing the evaluation. The qualitative analysis, for which we transcribed and combined the open-ended responses of the questionnaire from all three public spaces, helped to interpret the potential inter-group differences in evaluation and to provide richer, more nuanced knowledge on soundscapes as affordances for respondents' activities, including exploring the role of expectation as a further factor influencing the evaluation. We created contextual maps to situate users' questionnaire responses in a spatial and behavioral context.

## Quantitative Analysis

The variables used in the quantitative analysis are described in **Table 3** below.

The three dependent variables in the analysis – disruption, stimulation and suitability were measured on a 5-point ordinal scale and in our data were non-normally distributed, so we relied on two non-parametric tests for our analysis (Ruxton, 2006). First, using the Kruskal–Wallis test, we tested whether there are statistically significant differences between the categories of the independent variables on each of the three soundscape evaluations. Second, we applied the Mann–Whitney $U$ test to investigate whether soundscape evaluations differed significantly between activity types (solitary or socially interactive) according to frequency of use, familiarity with what is heard, location, age and gender. We considered relationships with $p < 0.05$ as statistically significant. We also discussed cases where $p < 0.1$ to indicate trends in the data, given the limited sample and number of variables we had at our disposal. The quantitative analysis was performed with the help of statistics software (SPSS version 19).

## Qualitative Analysis

We performed an in-depth analysis of responses to the open-ended questions that respondents provided when asked to explain how their soundscape stimulated or disrupted their activities (if at all), focusing on *what was disrupting/stimulating* (cause) and *what was disrupted/stimulated* (effect). We also analyzed how respondents articulated their (auditory) expectations and whether they were met during their time in the space. Our thematic coding approach was inspired by previous work on soundscape and place expectations (Bruce and Davies, 2014), focusing on respondents' expectations from the space itself, their auditory expectations (namely expected sounds), what they expected to experience in the space as well as expectations from others present in the space. We contrasted the answers of respondents performing solitary activities with those performing socially interactive activities.

## Contextual Maps

We visualized the data collected through behavioral mapping for the three fieldwork locations using GIS-based methods to situate the data on soundscape evaluation in a spatial setting. The resulting maps show the spatial distribution of questionnaires

**TABLE 3 |** Variables used for quantitative analysis.

| **Dependent variables** | |
|---|---|
| Disruption | 5-point Likert item |
| Stimulation | 5-point Likert item |
| Suitability | 5-point Likert item |
| **Independent variables** | |
| Level of social interaction | Binary variable ("solitary" and "socially interactive") |
| Familiarity with what is heard* | Ordinal variable ("low and medium familiarity," "high familiarity," "very high familiarity") |
| Frequency of use of public space | Ordinal variable ("this is my first visit," "a few times per year," "at least once a month," "at least once a week") |
| Location | Categorical variable; three distinct locations ("Sarphatipark," "Oosterpark," "Frederiksplein") |
| Age** | Binary variable: "35 or younger" and "older than 35" |
| Gender | Binary variable: "male," and "female" |

*The original five categories were collapsed in three for group comparison: "low and medium familiarity" (including "very low," "low," and "medium familiarity"), "high familiarity" and "very high familiarity." **The original continuous "age" variable was collapsed in two for group comparison.

and are accompanied by an overview of patterns of occupancy for each observed location, in relation to the level of social interaction of users' activities, illustrating the social interaction context within which the questionnaire responses were collected.

# RESULTS

## Contextual Maps

We begin with the analysis of the maps resulting from the behavioral mapping process (**Figure 3**) as they play a descriptive role, that is, to illustrate the larger context in which the questionnaires were filled out in terms of patterns of use based on the level of social interaction of the activities performed. The maps for each public space are an aggregation of the data collected during the two sessions per space and visualize the use of space exclusively in the areas where the behavioral mapping was carried out (marked with light gray in the resulting maps); the other areas have not been observed due to practical reasons, yet they were also consistently frequented by users.

The maps clearly show that socially interactive users are dominant in the space, throughout all three locations. The main observed physical factors influencing the distribution of use and subsequent concentration of users were: the surface materials (i.e., pavement or grass), presence/absence of shade (influenced by trees and other greenery), location and presence of conventional seating amenities (i.e., benches) or other elements that could used as seating amenities (e.g., other built structures), points of attraction (e.g., water fountains), proximity to bodies of water (i.e., ponds), and proximity to foot/bicycle paths.

The less dense, more spread out occupancy of the large open area in, e.g., the Western part of Oosterpark or all of Sarphatipark was influenced by the existence of conventional seating amenities

mainly along the foot/bicycle paths, with large open grass fields in between. We observed the clustering of users both in Eastern Oosterpark and throughout Frederiksplein. This could be due to the lack of grass where users could sit on and the dominance of various seating amenities (users, mostly socially interactive, also sat on the side of the fountain in Frederiksplein, and on round, elevated built structures in Oosterpark). The users closest to the body of water in Easter Oosterpark, largely performing socially interactive activities, were facing the water while sitting on grass, whereas solitary users mostly faced the water from a larger distance, while sitting on benches. The clear dominance and clustering of socially interactive users in the NE section of Sarphatipark was due to three separate birthday celebrations taking place at the same time, bringing together large groups of users.

The location of the completed questionnaires document in **Figure 3** demonstrates that the sample of users approached to complete our questionnaires is representative for the distribution of users in space in the timeframe and the locations where we conducted our research, with socially interactive users dominant across spaces, usually occupying the larger grass fields (generally in the sun), and solitary users equally distributed between the open fields and seating amenities closer to the paths (the latter generally in the shade).

## Quantitative Results: Statistical Analyses of Soundscape Evaluations

The sample distribution according to the main variables (**Table 4**) shows that, for the dependent variables, the vast majority of respondents (86%) evaluated their soundscapes as having low or very low disruption values, while no respondent evaluated them as being very disrupting. The sample was split rather evenly for stimulation ratings (low, medium and high stimulation), with around 30% of respondents each. The majority of respondents (67%) evaluated their soundscapes as highly or very highly suitable for their activities. Most of the respondents were participating in socially interactive activities. The sample was divided rather evenly also by frequency, with 47% visiting the locations at least once a month. The vast majority of respondents (90%) stated to be highly or very highly familiar with their soundscapes. 76% were 35 or younger and a slight majority of the sample identified as female.
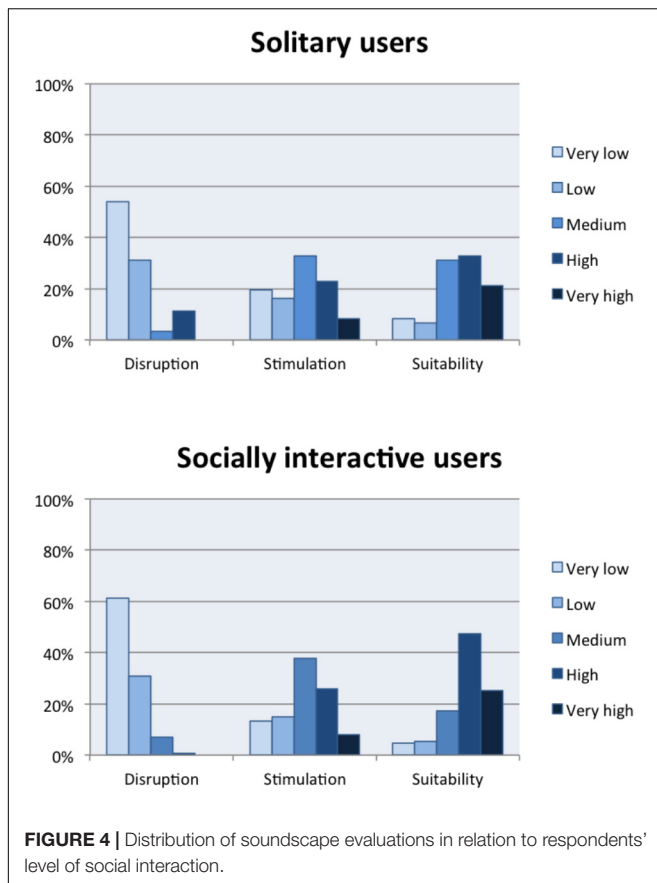
The distribution of soundscape ratings split by level of social interaction is presented in **Figure 4**. For disruption, a larger share of solitary respondents evaluated their soundscapes as having very high levels of disruption than socially interactive users; 11% of solitary respondents evaluated their soundscapes as highly or very highly disruptive, compared to 1% of socially interactive respondents. For stimulation, a larger share of solitary users evaluated their soundscape as having very low or low levels of stimulation: 36% compared to 28% of socially interactive users. Finally, for suitability, a smaller share of solitary respondents evaluated their soundscapes as highly or very highly suitable (54% of respondents compared to 72% of socially interactive respondents).

**TABLE 4** | Distribution of valid responses by variable used in quantitative analyses.

| Variable | Values | N (%) |
| --- | --- | --- |
| Level of social interaction | Socially interactive | 127 (67.6%) |
| | Solitary | 61 (32.4%) |
| Frequency of use of space | This is my first visit | 30 (16%) |
| | A few times per year | 70 (37.2%) |
| | At least once a month | 58 (30.9%) |
| | At least once a week | 30 (16%) |
| Familiarity with what is heard | Low and medium familiarity | 18 (9.6%) |
| | High familiarity | 57 (30.3%) |
| | Very high familiarity | 113 (60.1%) |
| Location | Sarphatipark | 83 (44.1%) |
| | Oosterpark | 81 (43.1%) |
| | Frederiksplein | 24 (12.8%) |
| Age | 35 or younger | 143 (76.1%) |
| | Older than 35 | 45 (23.9%) |
| Gender | Male | 85 (45.2%) |
| | Female | 103 (54.8%) |
| Disruption | Very low disruption | 111 (59%) |
| | Low disruption | 58 (30.9%) |
| | Medium disruption | 11 (5.9%) |
| | High disruption | 8 (4.3%) |
| | Very high disruption | 0 |
| Stimulation | Very low stimulation | 29 (15.4%) |
| | Low stimulation | 29 (15.4%) |
| | Medium stimulation | 68 (36.2%) |
| | High stimulation | 47 (25%) |
| | Very high stimulation | 15 (8%) |
| Suitability | Very low suitability | 11 (5.9%) |
| | Low suitability | 11 (5.9%) |
| | Medium suitability | 41 (21.8%) |
| | High suitability | 80 (42.6%) |
| | Very high suitability | 45 (23.9%) |

*(N = 188).*

The Kruskal–Wallis test was conducted to evaluate the differences between the categories of the five independent variables (level of social interaction, frequency of use of space, familiarity with what is heard, location and age) on the three soundscape evaluations (disruption, stimulation and suitability). The results (**Table 5**) showed that there was a significant difference in suitability rating for all independent variables, albeit a weak significance for level of social interaction; there was also a significant difference in disruption ratings between the three locations. These differences demonstrate the relevance of the independent variables in influencing the extent to which respondents' soundscapes are perceived to afford/be suitable for their on-site activities. Overall, the tests showed that the independent variables included in this research are related mainly with suitability ratings and minimally disruption ratings. The independent variables do not significantly relate with stimulation ratings. This suggests that "suitability" is the clearest construct for respondents to grasp, while "stimulation," and to an extent "disruption," are somewhat more challenging to assess.

**FIGURE 4 |** Distribution of soundscape evaluations in relation to respondents' level of social interaction.

**TABLE 5 |** Results for the Kruskal–Wallis test to compare the soundscape evaluations between categories of independent variables.

| Variables | Disruption | Stimulation | Suitability |
|---|---|---|---|
| Level of social interaction | 1.668 | 0.739 | 3.523* |
| Frequency of use of space | 2.270 | 5.849 | 13.066** |
| Familiarity with what is heard | 4.027 | 2.222 | 11.230** |
| Location | 17.897** | 2.543 | 9.474** |
| Age | 0.416 | 0.547 | 4.547** |

*Chi-square values reported. P-value significance: ** for p < 0.05, * for p < 0.1 (trend of significance). N = 188.*

To further understand the relationships identified above, we used the Mann–Whitney $U$ test to calculate whether soundscape ratings, grouped by the level of social interaction of respondents' activities, differ among categories of the independent variables (**Table 6**).

The Mann–Whitney $U$ test indicated that for those visiting the locations for the first time, socially interactive respondents have significantly higher stimulation ratings than solitary respondents ($U = 34.500$, $p < 0.05$), and higher suitability ratings, albeit with a weak significance ($U = 34.500$, $p < 0.05$). A possible explanation could be that the locations researched here are more geared toward group activities. Groups were especially dominant in those spaces on sunny days, usually engaged in various – likely audible – interactive activities throughout the

observed areas (as seen in the contextual maps in **Figure 3**). Also among respondents who visit at least weekly, evaluations of stimulation were higher – although weakly significant – for socially interactive respondents than for solitary respondents ($U = 74.000$, $p < 0.1$).

For respondents with high familiarity ratings, there is a weakly significant difference between socially interactive and solitary respondents, with the former having lower disruption ratings than the latter ($U = 197.000$, $p < 0.1$).

In the particular case of Sarphatipark, socially interactive respondents have significantly lower disruption and significantly higher suitability ratings than solitary respondents ($U = 447.000$, $p < 0.05$, and $U = 418.5000$, $p < 0.05$, respectively). This suggests that Sarphatipark is uniquely perceived as affording socially interactive activities rather than solitary ones in a significant manner, when compared to the other two locations.

For users older than 35, socially interactive respondents have lower disruption and higher stimulation ratings than solitary respondents, albeit weakly significant ($U = 189.000$, $p < 0.1$, and $U = 177.000$, $p < 0.1$, respectively). Finally, no significant differences between socially interactive users and solitary ones were found for males and females.

## Qualitative Results

The quantitative analysis partially confirmed our literature-driven expectations on the role of age, the level of social interaction of respondents' activities and of respondents' familiarity, both with the space and with what they hear, on their soundscape ratings in relation to their activity. The location in which the research was conducted was also identified as having an effect on soundscape ratings, particularly for Sarphatipark. The findings provided little detail on the respondents' experience that could guide, for example, design interventions, e.g., what they find disrupting or stimulating or what their expectations were from their space and their soundscapes, thus leaving much to speculation. To address this, we relied on qualitative insights from an in-depth analysis of respondents' explanations of their disruption and stimulation ratings, as well as their expectations, to better understand what specifically in their space and their soundscapes affords (or discourages) their activities. We first grouped the responses of all three spaces together, and categorized respondents' descriptions of their expectations according to: the type of space they were expecting to find, its amenities, what they expected to hear and how they expected others to use the space. Considering that Sarphatipark stood out in the quantitative analysis as a space evaluated as particularly affording of respondents' activities, we investigated whether the responses in the park differed from those in the other two fieldwork locations. However, no particular differences were observed, so below we report only on the aggregated data from the three spaces.

### Explanation of Disruption and Stimulation Ratings

Respondents described *how* their soundscapes disrupted and/or stimulated their activities, with a particular emphasis on *what* in their soundscapes they considered to be disrupting or stimulating.

**TABLE 6 |** Results of the Mann–Whitney $U$ test: comparison of soundscape evaluations between users according to level of social interaction.

| Variable | Values | Disr. SI (mean rank) | Disr. Sol (mean rank) | Disr. (sig.) Mann–Whitney $U$ | Stim. SI (mean rank) | Stim. Sol (mean rank) | Stim. (sig.) Mann–Whitney $U$ | Suitability SI (mean rank) | Suitability Sol (mean rank) | Suitability (sig) Mann–Whitney $U$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Frequency of use of space | This is my first visit | 14.39 | 19.14 | 55.000 | 17.50 | 8.93 | 34.500** | 17.50 | 8.93 | 34.500** |
| | A few times per year | 34.69 | 33.97 | 388.500 | 35.95 | 33.97 | 407.500 | 36.85 | 30.94 | 359.000 |
| | At least once a month | 28.13 | 31.59 | 354.500 | 32.26 | 25.30 | 306.000 | 30.37 | 28.17 | 372.000 |
| | At least once a week | 14.93 | 16.07 | 104.000 | 12.93 | 18.07 | 74.000* | 17.23 | 13.77 | 86.500 |
| Familiarity with what is heard | Low and medium familiarity | 9.58 | 9.33 | 35.000 | 9.92 | 8.67 | 31.000 | 10.29 | 7.92 | 26.500 |
| | High familiarity | 26.98 | 35.85 | 197.000* | 29.41 | 27.62 | 268.000 | 30.00 | 25.62 | 242.000 |
| | Very high familiarity | 55.18 | 60.07 | 1362.000 | 58.01 | 55.29 | 1419.000 | 60.96 | 50.30 | 1209.500 |
| Location | Sarphatipark | 39.10 | 51.15 | 447.000** | 44.15 | 35.23 | 494.500 | 45.36 | 31.43 | 418.500** |
| | Oosterpark | 41.64 | 39.79 | 708.000 | 40.12 | 42.66 | 695.500 | 41.88 | 39.34 | 695.500 |
| | Frederiksplein | 12.68 | 12.35 | 69.500 | 13.18 | 11.92 | 64.000 | 12.23 | 12.73 | 68.500 |
| Age | 35 or younger | 70.93 | 74.86 | 1916.500 | 72.02 | 71.94 | 2025.500 | 75.52 | 62.62 | 1662.000 |
| | Older than 35 | 20.22 | 25.91 | 189.000* | 26.30 | 19.55 | 177.000* | 25.83 | 20.05 | 188.000 |
| Gender | Male | 41.39 | 45.09 | 810.500 | 44.86 | 40.58 | 798.500 | 45.40 | 39.89 | 773.000 |
| | Female | 50.46 | 57.08 | 826.000 | 52.51 | 50.33 | 908.000 | 54.41 | 44.08 | 758.000 |

*Mean rank and Mann–Whitney U values reported (the higher the mean rank, the higher the rating). Dependent variables: disruption ("disr."), stimulation ("stim."), and suitability ("suit."). Level of social interaction: solitary ("Sol") and socially interactive ("SI"). P-value significance ("sig."): ** for p < 0.05, * for p < 0.1 (trend of significance). N = 188.*

### Disruption

The main source of disruption was, for both solitary and socially interactive respondents, the sounds of others in the space, especially the sounds of loud conversations and of children crying; surprisingly, the sounds of traffic (and public transportation) were mentioned only in passing as a source of disruption, the focus remaining on other public space users and their sound-producing activities. Solitary respondents also tended to cite more holistic reasons for their disturbance (e.g., "city sounds," "all sounds," "racket"[6]) than socially interactive respondents.

Both solitary and socially interactive respondents focused on the disturbing/distracting effect that some sounds had over their own activity: in the case of solitary respondents, what they heard disturbed their thought process or their ability to unwind, whereas for socially interactive respondents, their conversation was interrupted or they had to adjust their speaking levels to be able to understand each other.

### Stimulation

While for sources of disruption, there was quite some consensus on which sources are considered disrupting (see above) and a relatively small number of sounds were listed, there was a comparatively larger array of sources of stimulation mentioned by both categories of users. Socially interactive respondents stood out by listing comparatively more aspects of their auditory experience that they considered stimulating, including not only sounds but also using more holistic descriptions like "coziness" ("gezelligheid"). The sources of stimulation were, to relatively equal extents, nature-related sounds (i.e., fountain, birds, water, with socially interactive respondents putting an emphasis on the sound of wind through the leaves of trees) and human activity-related sounds.

Both solitary and socially interactive respondents focused on how what they heard stimulated the "atmosphere" in their space and the effect it had over users, particularly in relation to a relaxing[7] effect or to a "holiday feeling": "the buzz/murmur contributes to a pleasant atmosphere"[8]. Interestingly enough, solitary respondents focused particularly on how what they heard stimulated hypothetical conversations ("if I hear other people talk, it is also easier for me to talk"[9]) or doing what they wanted ("I'm stimulated to do what I like"[10]), e.g., fall asleep ("calming sounds allow me to fall asleep"[11]). Comparatively, socially interactive respondents further emphasized the importance of the presence of others for coziness and cheerfulness: "The fact that you can hear life around you makes it pleasant and cozy. In either way, it makes [this] pleasant and cozy".[12]

---

[6] Original in Dutch: "lawaai."

[7] Original in Dutch: "rustgevend."

[8] Original in Dutch: "het geroezemoes draagt bij aan gemoedelijke sfeer."

[9] Original in Dutch: "als ik andere mensen hoor praten, dan is het voor mij ook makkelijk om te praten."

[10] Original in Dutch: "gestimuleerd om met te doen wat ik leuk vind."

[11] Original in Dutch: "rustgevende geluiden laten me in slaap vallen."

[12] Original in Dutch: "dat je leven om je heen hoort, maakt het aangenaam en gezellig. In dit geval maakt het hier erg gemoedelijk en gezellig."

Solitary respondents focused more on the effect of what they heard had on their intended or current activities, whereas socially interactive respondents were more embedded in and engaged with their soundscapes, emphasizing not only the quiet dimension of their experience, but also the dynamism generated by the presence of others.

One socially interactive respondent (offering a low disruption rating and a high stimulation rating) summarized the complexity of their relationship with their soundscape: "music offers an atmosphere, so does the water and people. The tram is a bit disturbing but it is allowed here in the city"[13].

Not all respondents that offered explanations to their disruption or stimulation ratings identified particular sounds that affected their evaluation. Some respondents focused only on one or two disrupting sounds, stating that "the rest" is neither stimulating nor disruptive. Others stated that some sounds were "distracting," but that in general they were neither stimulated nor stimulated by what they heard; a sub-group of respondents stated that they were too focused on their activity to be aware of their soundscape: "I was very busy with my own activity so I was not very aware of the ambient sound"[14].

### Users' Expectations From Their On-Site Experiences

A count of occurrences showed that the majority of respondents reported that their expectations were met in all three spaces during their activities; however, only a slight majority of solitary respondents felt their expectations were met, compared to slightly over three quarters of socially interactive respondents. The subtle differences in expectations between solitary and socially interactive respondents indicate slightly different auditory experiences for those who use the public spaces alone or with others. As indicated by Bruce and Davies (2014), public space users tend to expect a limited number of sounds in an urban environment, especially for leisure-related uses and in relation to urban parks. Both categories of respondents expected the sound of fountain and water (due to two of the three public spaces being designed with large water fountains around which users tended to cluster, as shown in "Contextual Maps," and visualized in **Figure 3**) as well as "city sounds," However, socially interactive respondents also expected to hear the sounds of birds, which solitary respondents did not mention in their responses: "quiet environment with a fountain and birds"[15]. Furthermore, only socially interactive respondents stated they expected to hear the sounds of people and traffic-related sounds: "many people because of the nice weather. Tram + car also expected because we are close to the road. Oosterpark is not so big"[16].

Solitary respondents were more likely than socially interactive respondents to expect *quietness* first, with *crowdedness* mentioned second; the latter placed *crowdedness* first in their list of expectations, followed by *quietness* and, equally important,

*atmosphere* (whatever it entailed for respondents, usually in relation to *coziness*). Not surprisingly, in relation to the expected behavior of *others* in the public space, both groups of users expected the presence of others; however, while solitary respondents referred only marginally to the expected behavior of others, a large proportion of socially interactive respondents specifically mentioned they expected the presence of others when they decided to use the public spaces. Furthermore, they emphasized the expected level of interaction and dynamism of the activities that *others* would be performing: "crowdedness, many groups of people, young men playing football"[17].

Finally, in relation to expectations from the public spaces themselves, both categories of respondents stated that they expected a *city park* ("a park, just like any other park in Amsterdam"[18]), which comes with its assumptions in terms of patterns of use (shown in **Figure 3**) and, of course, audible sounds. This is particularly interesting for the case of Frederiksplein, not a traditional large urban park but rather a small urban square – park hybrid (a "plein").

Despite the variety in expectations, a majority of both solitary and socially interactive respondents stated their expectations were mostly or fully met, with some respondents explaining that their expectations were influenced by their previous uses of the park: "I was here before so I knew what I could expect"[19].

## DISCUSSION

This paper employed a mixed methods approach to study the user-soundscape relationship in a public space context, with an emphasis on users' activities; we further investigated how the level of social interaction of users' activities, individually or interacting with other factors, influence users' evaluations of their soundscapes, following the analytical model introduced earlier. We thus sought to demonstrate the relevance of considering the relationship between activity and soundscape evaluations when designing spaces for specific *uses* and interactions, rather than exclusively for generic goals like restoration. Through a mixed methods approach, we tested a number of factors that soundscape literature suggested to be likely to influence the user-soundscape relationship in a specific context, given the users' activities. We framed our research questions in relation to affordance theory, which helped to explain how public space users refer and evaluate the relationship between what they hear and what they do, i.e., do they evaluate their soundscapes as disruptive, stimulating or overall suitable for their activity.

We make three contributions to help address the challenge detailed in the introduction, and discuss each one below:

(1) A methodological contribution, adding to existing soundscape evaluation methodologies, reflecting at the same time on the limitations and ways of improving current methods,

---

[13]Original in Dutch: "muziek geeft sfeer, water er mensen ook. Tram verstoord beetje maar mocht erbij hier in de stad."

[14]Original in Dutch: "was teveel bezig met mijn eigen activiteit dat ik me weinig bewust was van het omgevings geluid."

[15]Original in Dutch: "rustige omgeving met een fontein en vogels."

[16]Original in Dutch: "veel mensen want het is mooi weer. Tram + auto ook verwacht omdat we vlakbij de weg zitten. Oosterpark is niet zo groot."

[17]Original in Dutch: "drukte, veel groepjes mensen, voetballende jongens."

[18]Original in Dutch: "een park, zoals elk ander park in A'dam."

[19]Original in Dutch: "ik ben hier vaker geweest dus wist wat ik kon verwachten."

(2) An empirical contribution, providing insight on how users explain their soundscape evaluations in relation to their activities and building on previous use of affordance theory to further the idea of a strong relationship between users' soundscape evaluations and their activities

(3) A policy and design-related contribution, offering additional insight on a transferable methodology/process that practitioners can employ in their work on the built environment.

## Methodological Contribution and Reflection

We used a mixed methods approach for a multi-layered analysis of the user-soundscape relationship in a public space setting, by (1) integrating users' activities as a key variable framing evaluations, (2) exploring the individual and interaction effect of additional factors on this relationship, (3) combining Likert items with open-ended responses where users can explain their soundscape ratings, and (4) combining questionnaires with on-site behavioral mapping to situate users' soundscape evaluations in a public space context.

The behavioral mapping was used to integrate a spatial dimension and to situate the data collected through questionnaires, not only in a physical environment, but also in a behavioral setting of *others* performing activities, that can offer additional insight into user evaluations and that cannot sufficiently be grasped via questionnaires alone. The quantitative method was used to collect categorical data on public space users' evaluations of their soundscapes in order to compare the ratings between users engaged in activities with different levels of social interaction, as well as across various factors that might influence the user-soundscape relationship. The qualitative analysis was used to offer more depth to the statistical findings; as the quantitative findings show a similar trend in soundscape ratings, the qualitative insight helped to understand the subtle differences in ratings and offer an interpretation of the findings. As people continue using the public spaces despite some (albeit low) level of reported disruption, only an in-depth approach could allow researchers and practitioners to understand, e.g., what are the sources of disruption and what makes users apparently accept them. Open-ended responses were encouraged through open-ended questions, which meant offering users the space to reflect on their experience and their subsequent evaluations. While analyzing such responses is time-consuming, it allows researchers and practitioners alike to make sure that they understand what the users of spaces are experiencing, focusing on and, ultimately, evaluating; simply asking users if they "like" what they hear in a space or if they find it "pleasant" is insufficient, as responses to such questions can potentially lead the data collector (designer, planner, researcher, etc.) to resort to a top-down interpretation of *what* the users evaluated. The knowledge collected through open-ended responses was thus essential in understanding what users focused on and referred to in their evaluation, as well as grasping the specific aspects in their experience that disrupted or stimulated their use of spaces, thus allowing for an exploration

of sounds and soundscapes as affordances for users' activities on site.

In this paper, we provide practitioners and researchers with an example of an insightful research *process* and with methods and tools to observe, ask and engage with users (actual or potential) of public spaces in relation to their multisensory experience. The methodology we put forward can also be used to research and document other aspects of the built environment, without being restricted to the auditory experience. We thus do not put forward a one-size-fits-all model, but rather a qualitative user-centered process that must be adapted to the specific and unique needs of each case, but that can provide a wealth of knowledge on *what* disrupts or stimulates users' activities in a public space. However, one minor limitation of this study method is that single Likert items were the variables analyzed (disruption, suitability, and stimulation); a future approach would be to substitute these with validated multi-item Likert scales as variables instead. For example, considering that suitability is indicated as the most useful/robust rating for users' soundscapes, it would be worth it in future studies to formulate a "suitability scale" based on multiple items, which may incorporate disruption and stimulation as well as other variables as the ones explored in this paper.

While time consuming and heavily reliant on users' willingness to participate, the methodology is nevertheless valuable for understanding what types of activities users' soundscapes and physical environments afford. A limitation of our mixed methods approach is that it questions and observes current users of mostly green parks, that are therefore less likely to have negative evaluations of their environment and their experience (as seen in the largely positive soundscape evaluations of both solitary and socially interactive respondents). Furthermore, asking users to reflect on what they hear through questionnaires encourages them to actively focus on their soundscapes, which results in responses that might not fully reflect their on-site everyday auditory experiences. This shows the need for further improving our methodology to elicit auditory knowledge in more creative, but systematic ways. The study described here is a first effort to research this topic on site and the questions raised by the results of our data can be used as opportunities for guiding or improving future research. For example, further attention could be paid to developing additional protocols to analyze the responses to the open-ended question on activity or on testing hypotheses on specific auditory affordances that soundscapes "create" for public space users, in terms directly relevant to users' activities or behavior. Furthermore, given the responses of, e.g., first-time solitary respondents, who evaluated their soundscapes as less suitable for their activities than first-time socially interactive respondents, an additional line of inquiry could be focused on their likeliness to return in the future to the public space in the future (or would prefer a different space). This could provide insight into whether there are aspects of their auditory experience in that particular location that have failed to meet the needs of various solitary respondents on multiple occasions. Finally, to better benefit from behavioral mapping as a method, more complex data could be collected on space users and uses, for example more

detailed insight on the type of social interaction they are engaged in, e.g., families with children, pairs, and more sophisticated spatial statistics could be employed to analyze the relationship between characteristics of the space, its patterns of use and users' soundscape evaluations.

## Empirical Contribution

The behavioral mapping showed the patterns of use of the three fieldwork locations by both solitary and socially interactive users during the research period, as a context in which the questionnaire respondents provided their soundscape evaluations. The quantitative analysis indicated that the level of social interaction of users' activities had an association with their suitability ratings, albeit weakly significant. It also showed that familiarity levels, both with what was heard and with the space (frequency of use of space) differed significantly for suitability ratings. The experience of first time visitors was of particular interest, as there was a significant difference between stimulation and suitability ratings for solitary and socially interactive respondents (with the latter being having higher ratings than the former for both ratings). For frequent users, a weakly significant association was shown in relation to stimulation ratings, with socially interactive users having higher stimulating evaluations than solitary users. Location also had a statistically significant association with disruption and suitability ratings, particularly for Sarphatipark, where solitary respondents reported significantly lower suitability and higher disruption ratings than socially interactive respondents. This difference in ratings also holds true for users older than 35 across locations for disruption and stimulation ratings, for whom the differences between solitary and socially interactive are weakly significant.

The qualitative analysis confirmed that solitary and socially interactive respondents differed slightly both in terms of sources of disruption and stimulation, as well as in the particular expectations (auditory and otherwise) from their experience. The sounds of people were considered as the main source of both disruption and stimulation for both groups; while conversations and the sounds of others in general were referred to as stimulating, *loud* conversations and children crying were disrupting. Surprisingly, the sounds of traffic were not mentioned as a main source of disruption; unsurprisingly, "natural" sounds were mentioned as a main source of stimulation (with only socially interactive respondents mentioning birds among stimulating sources). While solitary respondents were more likely to include holistic sounds (e.g., "city sounds") among sources of disruption, socially interactive respondents were more likely to include such sounds among sources of stimulation, thus affording their activities (e.g., "atmosphere," "buzz/murmur"). In terms of expectations, both solitary and socially interactive respondents reported that their expectations were largely met, which explains the soundscape ratings in relation to their activities reported on in the previous section, i.e., overall low disruption ratings and high stimulation and suitability ratings. Socially interactive respondents tended to focus not only on the presence of others, but also on their activities as well as the others' levels of social interaction.

They were also more likely to emphasize the importance of the general atmosphere/ambiance in their expectations, whereas solitary respondents focused on their expectations in relation to quietness. In terms of sounds expected, socially interactive respondents tended to expect a larger variety of "natural sounds" (including birds, wind in the trees, etc.), as well as more traffic and street-life related sounds (e.g., cars, tram, "the street"). The presence of others, not only as a source of disruption but rather as an *affordance* that both helps in the "creation of atmosphere" and encourages one's own engagement with the space is thus essential when discussing/addressing auditory concerns in relation to public space use.

## Policy/Design Implications

The policy and design implications are twofold, based on the empirical findings, as well as our methodological approach. On the one hand, the empirical insights demonstrate the added value of considering users' soundscapes in relation to their activities (with a focus on whether the activities performed are solitary or socially interactive) when considering new policy or design initiatives; it also showed the potential of including the analytical framework developed in the background section to help unpacking the complexity of the auditory experience. For example, tools like questionnaires commonly used by various practitioners should include activity questions in soundscape-related queries, such as asking users about their activity and whether they were by themselves or with others at the time of the completion of, for example, noise exposure surveys.

On the other hand, the methodological approach described in this paper, and the resulting research process described above can be used by policy makers and designers to gain contextual insights in users' experiences. For example, integrating open-ended questions in current questionnaires for soundscape evaluations can help with verifying the suitability or relevance of commonly (and uncritically) used terms like "annoyance" or "pleasantness" by accessing the everyday sound-related vocabulary of urbanites, which can in turn feed into and help adjust existing tools used by local, regional and national authorities.

Overall, in this paper, we provided both empirical and methodological insights that researchers and practitioners alike can adjust and employ in their own investigations of urban auditory complexity to contribute to the creation of spaces that afford a large array of activities. For future research and practice, it would be interesting to explore whether using richer descriptions or evaluations brings designers and policy makers to different kinds of interventions.

## ETHICS STATEMENT

This study was carried out in accordance with the guidelines outlined by the Ethics Committee (AIEC) of the University of Amsterdam. The participation in the study was voluntary and the responses were fully anonymized; subjects were informed that, if

they completed the questionnaires, their answers would be used in a scientific study. By completing the written questionnaires, they confirmed that they provided consent to participate in the research.

## AUTHOR CONTRIBUTIONS

EB performed the data collection and analyses and wrote the majority of the paper. KP, MC, and LB provided extensive comments and textual edits on previous versions of the manuscript. OR contributed to the quantitative analysis in this manuscript. The doctoral project this research is part of was initiated by MC, as part of INCAS, and further elaborated jointly by EB, KP, MC, and LB.

## REFERENCES

Aletta, F., Kang, J., and Axelsson, Ö. (2016a). Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landsc. Urban Plan.* 149, 65–74. doi: 10.1016/j.landurbplan.2016.02.001

Aletta, F., Lepore, F., Kostara-Konstantinou, E., Kang, J., and Astolfi, A. (2016b). An experimental study on the influence of soundscapes on people's behaviour in an open public space. *Appl. Sci.* 6:276. doi: 10.3390/app6100276

Andringa, T. C., and Lanser, J. J. L. (2013). How pleasant sounds promote and annoying sounds impede health: a cognitive approach. *Int. J. Environ. Res. Public Health* 10, 1439–1461. doi: 10.3390/ijerph10041439

Andringa, T. C., Weber, M., Payne, S. R., Krijnders, J. D., Dixon, M. N., Linden, R. V., et al. (2013). Positioning soundscape research and management. *J. Acoust. Soc. Am.* 134, 2739–2747. doi: 10.1121/1.4819248

Aspuru, I., Garcia, I., Bartalucci, C., Borchi, F., Carfagni, M., Governi, L., et al. (2016). LIFE+ 2010 QUADMAP project: a new methodology to select, analyze and manage quiet urban areas defined by the European Directive 2002/49/EC. *Noise Mapp.* 3, 120–129. doi: 10.1515/noise-2016-0009

Aspuru, I., Garcia, I., Herranz-Pascual, K., and Garcia-Borreguero, I. (2011). Understanding soundscape as a specific environmental experience: highlighting the importance of context relevance. *Proc. Mtgs. Acoust.* 14:015004. doi: 10.1121/1.3676685

Axelsson, Ö., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *J. Acoust. Soc. Am.* 128, 2836–2846. doi: 10.1121/1.3493436

Bild, E., Coler, M., Dubois, D., and Pfeffer, K. (2015). "A pilot experiment on effects of motor and cognitive activities on memories of soundscapes," in *Proceedings of the Euronoise*, Maastricht, 1529–1533.

Bild, E., Coler, M., Pfeffer, K., and Bertolini, L. (2016). Considering sound in planning and designing public spaces: a review of theory and applications and a proposed framework for integrating research and practice. *J. Plan. Lit.* 31, 419–434. doi: 10.1177/0885412216662001

Bild, E., Steele, D., Pfeffer, K., Bertolini, L., and Guastavino, C. (2018). "Activity as a mediator between users and their auditory environment in an urban pocket park: a case study of Parc du Portugal (Montreal, Canada)," in *Handbook of Research on Perception-Driven Approaches to Urban Assessment and Design*, eds F. Aletta and X. Jiao (Hershey, PA: IGI Global), 100–125. doi: 10.4018/978-1-5225-3637-6.ch005

Bloomfield, A. (2014). *Good Practice Guide on Quiet Areas*. EEA Technical Report No. 4/2014. Copenhagen: European Environment Agency.

Booi, H., and van den Berg, F. (2012). Quiet areas and the need for quietness in Amsterdam. *Int. J. Environ. Res. Public Health* 9, 1030–1050. doi: 10.3390/ijerph9041030

Botteldooren, D., Andringa, T., Aspuru, I., Brown, A. L., Dubois, D., Guastavino, C., et al. (2015). "From sonic environment to soundscape," in *Soundscape and the Built Environment*, eds J. Kang and B. Schulte-Fortkamp (Boca Raton, FL: CRC Press), 17–42. doi: 10.1201/b19145-3

## FUNDING

## ACKNOWLEDGMENTS

Bruce, N., Condie, J., Henshaw, V., and Payne, S. R. (2015). *Analysing Olfactory and Auditory Sensescapes in English Cities: Sensory Expectation and Urban Environmental Perception. Ambiances.* Available at: http://ambiances.revues.org/560

Bruce, N. S., and Davies, W. J. (2014). The effects of expectation on the perception of soundscapes. *Appl. Acoust.* 85, 1–11. doi: 10.1016/j.apacoust.2014.03.016

Bruce, N. S., Davies, W. J., and Adams, M. D. (2009). "Expectation as a factor in the perception of soundscapes," in *Proceedings of the Euronoise*, Edinburgh.

Can, A., Aumond, P., Michel, S., Coensel, B. D. E., Ribeiro, C., and Lavandier, C. (2016). "Comparison of noise indicators in an urban context," in *Proceedings of the INTER-NOISE and NOISE-CON Congress*, Vol. 253, (Reston, VA: Institute of Noise Control Engineering), 775–785.

Cerwén, G., Pedersen, E., and Pálsdóttir, A. M. (2016). The role of soundscape in nature-based rehabilitation: a patient perspective. *Int. J. Environ. Res. Public Health* 13:1229. doi: 10.3390/ijerph13121229

Clarke, E. F. (2005). *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195151947.001.0001

Cosco, N. G., Moore, R. C., and Islam, M. Z. (2010). Behavior mapping: a method for linking preschool physical activity and outdoor design. *Med. Sci. Sports Exerc.* 42, 513–519. doi: 10.1249/MSS.0b013e3181cea27a

Creswell, J. W., and Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research*, 2nd Edn. Thousand Oaks, CA: Sage Publications, Inc.

Delepaut, G. (2009). "Questionner: comment et pourquoi? Le choc des questions, le poids des mots!," in *Le Sentir et le Dire. Concepts et Méthodes en Psychologie et Linguistique Cognitives*, ed. D. Dubois (Paris: L'Harmattan), 157–188.

DeNora, T. (2000). *Music in Everyday Life*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511489433

Dubois, D. (2000). Categories as acts of meaning: the case of categories in olfaction and audition. *Cogn. Sci. Q.* 1, 35–68.

Filipan, K., Boes, M., De Coensel, B., Lavandier, C., Delaitre, P., Domitrović, H., et al. (2017). The personal viewpoint on the meaning of tranquility affects the appraisal of the urban park soundscape. *Appl. Sci.* 7:91. doi: 10.3390/app7010091

Gibson, J. J. (1977). "The theory of affordances," in *Perceiving, Acting, and Knowing*, eds R. E. Shaw and J. Bransford (Mahwah, NJ: Lawrence Erlbaum Associates), 67–82.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Goličnik, B., and Thompson, C. W. (2010). Emerging relationships between design and use of urban park spaces. *Landsc. Urban Plan.* 94, 38–53. doi: 10.1016/j.landurbplan.2009.07.016

Guastavino, C., Katz, B. F., Polack, J. D., Levitin, D. J., and Dubois, D. (2005). Ecological validity of soundscape reproduction. *Acta Acust. United Acust.* 91, 333–341.

Herranz-Pascual, K., García, I., Diez, I., Santander, A., and Aspuru, I. (2017). Analysis of field data to describe the effect of context (Acoustic and Non-Acoustic Factors) on urban soundscapes. *Appl. Sci.* 7:173. doi: 10.3390/app7020173

Huron, D. B. (2006). *Sweet Anticipation: Music and the Psychology of Expectation.* Cambridge: MIT Press.

Ingold, T. (2000). *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill.* Hove: Psychology Press, doi: 10.4324/9780203466025

Jennings, P., and Cain, R. (2013). A framework for improving urban soundscapes. *Appl. Acoust.* 74, 293–299. doi: 10.1016/j.apacoust.2011.12.003

Knigge, L., and Cope, M. (2006). Grounded visualization: integrating the analysis of qualitative and quantitative data through grounded theory and visualization. *Environ. Plan. A* 38, 2021–2037. doi: 10.1068/a37327

Kogan, P., Turra, B., Arenas, J. P., and Hinalaf, M. (2017). A comprehensive methodology for the multidimensional and synchronic data collecting in soundscape. *Sci. Total Environ.* 580, 1068–1077. doi: 10.1016/j.scitotenv.2016.12.061

Kyttä, M. (2002). Affordances of children's environments in the context of cities, small towns, suburbs and rural villages in Finland and Belarus. *J. Environ. Psychol.* 22, 109–123. doi: 10.1006/jevp.2001.0249

Lavia, L., Dixon, M., Witchel, H. J., and Goldsmith, M. (2016). "Applied soundscape practices," in *Soundscape and the Built Environment*, eds J. Kang and B. Schulte-Fortkamp (Boca Raton, FL: CRC Press), 243–301. doi: 10.1201/b19145-11

Lavia, L., Easteal, M., Close, D., Witchel, H., Axelsson, Ö., Ware, M., et al. (2012). "Sounding Brighton: practical approaches towards better soundscapes," in *Proceedings of the INTER-NOISE and NOISE-CON Congress*, New York, NY.

Lavia, L., Witchel, H. J., Aletta, F., Steffens, J., Fiebig, A., Kang, J., et al. (2018). "Non-participant observation methods for soundscape design and urban panning," in *Handbook of Research on Perception-Driven Approaches to Urban Assessment and Design*, eds F. Aletta and X. Jiao (Hershey, PA: IGI Global), 73–99. doi: 10.4018/978-1-5225-3637-6.ch004

Lercher, P., and Schulte-Fortkamp, B. (2003). "The relevance of soundscape research to the assessment of noise annoyance at the community level," in *Proceedings of the Eighth International Congress on Noise as a Public Health Problem*, Rotterdam, 225–231.

Lercher, P., Van Kamp, I., Von Lindern, E., and Botteldooren, D. (2016). "Perceived soundscapes and health-related quality of life, context, restoration, and personal characteristics," in *Soundscape and the Built Environment*, eds J. Kang and B. Schulte-Fortkamp (Boca Raton, FL: CRC Press), 89–132. doi: 10.1201/b19145-6

Marcell, M. M., Borella, D., Greene, M., Kerr, E., and Rogers, S. (2000). Confrontation naming of environmental sounds. *J. Clin. Exp. Neuropsychol.* 22, 830–864. doi: 10.1076/jcen.22.6.830.949

Mehta, V. (2014). Evaluating public space. *J. Urban Des.* 19, 53–88. doi: 10.1080/13574809.2013.854698

Mzali, M. (2002). *Perception de L'ambiance Sonore et Évaluation du Confort Acoustique des Trains.* Doctoral dissertation, Paris, Université Paris, 6.

Nielbo, F. L. (2015). "From soundscape to meaningscape," in *Proceedings of the EuroNoise*, Edinburgh.

Nielbo, F. L., Steele, D., and Guastavino, C. (2013). "Investigating soundscape affordances through activity appropriateness," in *Proceedings of Meetings on Acoustics ICA2013*, (London: ASA). doi: 10.1121/1.4800502

Nilsson, M., Botteldooren, D., and De Coensel, B. (2007). "Acoustic indicators of soundscape quality and noise annoyance in outdoor urban areas," in *Proceedings of the 19th International Congress on Acoustics*, Madrid.

Pecqueux, A. (2012). Les affordances des événements: des sons aux événements urbains. *Communications* 90, 215–227. doi: 10.3917/commu.090.0215

Pheasant, R. J., Horoshenkov, K., Watts, G., and Barret, B. T. (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

Raimbault, M. (2006). Qualitative judgements of urban soundscapes: questionning questionnaires and semantic scales. *Acta Acust. United Acust.* 92, 929–937.

Reybrouck, M. (2012). Musical sense-making and the concept of affordance: an ecosemiotic and experiential approach. *Biosemiotics* 5, 391–409. doi: 10.1007/s12304-012-9144-6

Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behav. Ecol.* 17, 688–690. doi: 10.1093/beheco/ark016

Schulte-Fortkamp, B., and Fiebig, A. (2016). "Impact of soundscape in terms of perception," in *Soundscape and the Built Environment*, eds J. Kang and B. Schulte-Fortkamp (Boca Raton, FL: CRC Press), 69–88. doi: 10.1201/b19145-5

Steele, D., Bild, E., Cynthia, T., Irene, M., Cubero, J., and Guastavino, C. (2016). "A comparison of soundscape evaluation methods in a large urban park," in *Proceedings of International Congress on Acoustics*, Montreal.

Steenson, C. J., and Rodger, M. W. (2015). Bringing sounds into use: thinking of sounds as materials and a sketch of auditory affordances. *Open Psychol. J.* 8, 174–182. doi: 10.2174/1874350101508010174

Steffens, J., Steele, D., and Guastavino, C. (2015). "New insights into soundscape evaluations using the experience sampling method," in *Proceedings of the Euronoise*, Maastricht, 1496–1500.

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. Am.* 141, 1414–1425. doi: 10.1121/1.4976627

Stoffregen, T. A. (2003). Affordances as properties of the animal-environment system. *Ecol. Psychol.* 15, 115–134. doi: 10.1207/S15326969ECO1502_2

Thibaud, J. P. (1998). "The acoustic embodiment of social practice," in *Proceedings of the Conference Stockholm, Hey Listen*, Stockholm.

Truax, B. (2001). *Acoustic Communication*, Vol. 1. Westport, CT: Greenwood Publishing Group.

Turvey, M. T. (1992). Affordances and prospective control: an outline of the ontology. *Ecol. Psychol.* 4, 173–187. doi: 10.1207/s15326969eco0403_3

Van Kempen, E., Devilee, J., Swart, W., and Van Kamp, I. (2014). Characterizing urban areas with good sound quality: development of a research protocol. *Noise Health* 16, 380–387. doi: 10.4103/1463-1741.144416

Yang, W., and Kang, J. (2005). Acoustic comfort evaluation in urban open public spaces. *Appl. Acoust.* 66, 211–229. doi: 10.1016/j.apacoust.2004.07.011

# Exploring the Validity of the Perceived Restorativeness Soundscape Scale: A Psycholinguistic Approach

*Sarah R. Payne[1]\* and Catherine Guastavino[2]*

[1] *The Urban Institute, Heriot-Watt University, Edinburgh, United Kingdom,* [2] *School of Information Studies and Centre for Interdisciplinary Research in Music Media and Technology, McGill University, Montreal, QC, Canada*

Soundscapes affect people's health and well-being and contribute to the perception of environments as restorative. This paper continues the validation process of a previously developed Perceived Restorativeness Soundscape Scale (PRSS). The study takes a novel methodological approach to explore the PRSS face and construct validity by examining the qualitative reasons for participants' numerical responses to the PRSS items. The structure and framing of items are first examined, to produce 44 items which are assessed on a seven-point Likert agreement scale, followed by a free format justification. Ten English speaking participants completed the PRSS interpretation questionnaire in two cafes in Montréal, Canada. Interpretation of participant free format responses led to six themes, which related to either the individual (personal attributes, personal outcomes), the environment (physical environment attributes, soundscape design) or an interaction of the two (behavior setting, normality, and typicality). The themes are discussed in relation to each Attention Restoration Theory (ART) component, namely Fascination, Being-Away, Compatibility, and Extent. The paper concludes by discussing the face and construct validity of the PRSS, as well as the wider methodological and theoretical implications for soundscape and attention restoration research, including the terminology importance in items measuring ART components and the value of all four components in assessing perceived restorativeness.

Keywords: soundscape, perceived restorativeness scale, perceived restorativeness soundscape scale, attention restoration theory, soundscape assessment, behavior setting, café

## INTRODUCTION

Soundscapes have the potential to enhance or damage our experience of a place and can have important consequences for people's behavior (e.g., Aletta et al., 2016b; Bild et al., 2016) performance (Clark and Sörqvist, 2012), health and well-being (Stansfeld et al., 2005; World Health Organisation, 2011; Van Kamp et al., 2015). To help design supportive, sustainable environments, soundscape assessment tools are necessary to understand individuals' experiences. An important evaluation criterion for people's experience of some places and its soundscape, is the level of psychological restoration that users may achieve from visiting the place (Gidlöf-Gunnarsson and Öhrström, 2007; Payne, 2008). One form of psychological restoration is attention restoration which refers to individuals' need to recover from attentional fatigue (drained cognitive resources from

directed attention) and reflect upon daily or life issues (Kaplan and Kaplan, 1989; Herzog et al., 1997). Restorative environments enable individual users to experience high levels of attention restoration. Assessments of an environment's potential to provide attention restoration can be made using scales assessing the extent an environment is perceived as having the qualities, or components that are theoretically considered important for restoration. Scales such as the Perceived Restorativeness Scale (PRS; Hartig et al., 1997) and the Perceived Restorative Component scale (PRC; Laumann et al., 2001) are commonly used in studies which only present visual cues. To help understand and design soundscapes which enhance restoration, these previous measures of perceived restorativeness were adapted to create a tool to specifically assess the perceived restorativeness of the *soundscape* (Payne, 2013). An important part of creating new tools is to test their reliability and validity. Reliability and partial concurrent validation (getting similar results to existing scales) have previously been demonstrated for the Perceived Restorativeness Soundscape Scale (PRSS) (Payne, 2013). However, public comprehension of scale items was unclear and this affects its face validity (does it measure what it is supposed?) and construct validity (does it measure the underlying construct?). Therefore, the aim of this paper is to further examine the validity of PRSS items, through a psycholinguistic analysis of participants' free format descriptions which justify their numerical PRSS item ratings.

As reported in Payne (2013) the (PRSS) was developed from the Perceived Restorativeness Scale (PRS; Hartig et al., 1997) and the Perceived Restorative Component scale (PRC) (Laumann et al., 2001). These assessment tools that measure the perceived qualities of an environment in terms of the presence of four theoretical components considered necessary to create a restorative environment and experience. Fascination, Being-Away, Compatibility, and Extent, are the four Attention Restoration Theory (ART) components considered necessary for an environment to be restorative (Kaplan and Kaplan, 1989; Kaplan, 1995). Fascination is a description of involuntary, effortless attention. It is the ability of a stimulus to have attention-holding properties, either without the individual needing to direct attention to focus upon the stimulus, or by inhibiting other stimuli from gaining attention. Being-Away involves a physical or conceptual shift away from the present situation or problems, to a different environment or way of thinking, allowing tired cognitive structures to rest while activating others. Compatibility is the matching of the environment's affordances to the individual's needs and inclinations. The environment needs to be responsive enough to an individual's planned behavior and for the individual to have aims that fit the environment's demands. A high match between the individual and the environment results in the individual using little directed attention as few differences need to be resolved, thus providing opportunities for restoration. An environment with Extent is one that is "rich enough and coherent enough so that it constitutes a whole other world" (Kaplan, 1995, p. 173). Extent has two subcomponents, Coherence and Scope (Kaplan and Kaplan, 1989). Coherence relates to how elements in the environment connect, with their structure and organization

combining to make sense (a coherent whole). Scope relates to the scale of the environment (imagined or physical) and quantity of its attributes that the individual is sufficiently engaged.

The PRS and PRC are component-measuring scales which examine the attributes of the person-environment interaction to help determine what makes an environment, or specifically the soundscape in the case of the PRSS, potentially restorative. Understanding the person-environment relationships through these components will enable designers to consider people's perception and behavior in context and elements of the environment that could be enhanced or removed to improve restoration. Taking the individual contextualized approach is in line with the soundscape definition set by the International Organization for Standardization (ISO, 2014, p. 1): "acoustic environments as perceived or experienced, and/or understood by a person or people, in context."[1]

The PRSS measures the perceived level of these four ART components in relation to an environment's sound through a number of items, each designed to measure one of the four components (Payne, 2013). Developed largely by replacing the word "place" in PRS and PRC items with "sonic environment,"[2] the PRSS successfully differentiated between soundscapes from different types of environment, the same type of environment, and within the same place (Payne, 2010, 2013). Similar to perceived restorativeness environment scale findings and measures of restorative outcomes for environments (Hartig et al., 1997; Laumann et al., 2001; Herzog et al., 2003; Kahn Jr et al., 2008), the more "natural" the soundscape the more restorative the soundscape was perceived to be Payne (2013). These results in part support concurrent validity of the PRSS as a measure of perceived restorativeness. However, still to be determined, is its face validity which evaluates if it is measuring what we think it is, the restorativeness of *soundscapes* (rather than say visual elements), and its construct validity which evaluates if it's measuring the underlying construct, such as Fascination.

To enhance construct validity, the PRSS and original PRS and PRC items use words relating to the theoretical attention restoration components and their definition. This results in words reflecting the researcher's interpretation and understanding of the relevant concepts, rather than words that the public would normally recognize and use to evaluate a soundscape or place. In turn, the public who are unfamiliar with the concepts being explored, find items in the scale strange and difficult to interpret (Payne, 2013). The wording of items may be particularly problematic for a soundscape scale as people are not used to discussing sounds to the same degree as visual aspects, and in comparison have a limited vocabulary (Dubois, 2000; Guastavino, 2006, 2007; Davies et al., 2013). Therefore, any restorative soundscape measuring tool needs to have simple, comprehensive language that is easy for a respondent to understand.

---

[1]See Brown et al. (2011) for a review of previously used definitions and a description of ISO Working Group 54 - Assessment of soundscape quality.
[2]The term sonic environment was used rather than soundscape as initial comments from laypeople at the time was that the former was easier to comprehend.

Examination of the grammatical structure and framing of items within existing perceived restorativeness environment or soundscape scales (Hartig et al., 1997; Laumann et al., 2001; Purcell et al., 2001; Payne, 2013) also highlight a number of inconsistencies. Differences occur between item composition depending on the ART component being assessed, as well as within and between items developed by different authors. Namely, differences exist in (i) the presence or absence of personal pronouns (e.g., I, me), (ii) the location (holistic framing) or elements (specific framing) under discussion (e.g., soundscape vs. sounds), and (iii) the terms used to describe each theoretical ART component through adjectives qualifying the environment (e.g., fascinating) or verbs refering to individual actions (e.g., discover). Each of these are problematic for the face and construct validity of the tool, as differences in item structure and framing of assessment tools can influence respondent's ratings (Bentler et al., 1971; Scott and Canter, 1997). For example, individual items may not be interpreted in the intended manner (low face validity) and items grouped together will therefore not represent their associated ART component but another aspect instead (low construct validity). If public responses are influenced by these psycholinguistic differences, it affects how the PRSS results should be interpreted.

The aim of this paper is to explore public comprehension and interpretation of PRSS items to explore item face and construct validity in assessing the perceived restorativeness of soundscapes. Initially, the paper examines the vocabulary, grammar, and framing of items used within perceived restorativeness scales to develop a "PRSS interpretation questionnaire." Innovatively, to determine the interpretation of PRSS items this study examines the free-format description used by participants to justify their numerical PRSS ratings, rather than conducting numerical analyses on the provided ratings. Face and construct validity cannot be definitively answered through this psycholinguistic approach as results are, for example, not being tested against comparable previously validated measures of the same concept. However, indications of face validity are expected in terms of participant responses being dominated by reference to sounds, rather than visual features. Whilst construct validity should be indicated from participant responses predominantly referring to terms used to describe their designated ART component, and potentially mention ART outcomes (recover and reflect).

## METHODOLOGY

### Psycholinguistic Analysis of PRS, PRC, and PRSS Items

In a previous study, public respondents previously raised comprehension issues with some PRSS items which had been developed from PRS and PRC items (Payne, 2013). Therefore, this study examined the linguistics of PRS, PRC, and PRSS items. The linguistic examination identified a number of deviations by the PRSS away from the original words used by PRS and PRC items. For example, some of the key theoretical words such as "Fascinating" did not appear in the PRSS, which could reduce the PRSS construct validity. The choice of nouns and

adjectives used within items are important as they represent the operationalization of each theoretical ART component and should be influential in respondents' ratings. Therefore, the key nouns and adjectives used should be comprehensible and are vital for construct validity. This issue is not restricted to differences between the perceived restorativeness soundscape scale (PRSS) and environment scales (PRS, PRC). Differences also exist in the descriptive words used in items to assess the same ART component *between* perceived restorativeness environment scales by different authors, as well as differences existing *within* each authors scale of perceived restorativeness. For example, some Fascination items refer to an interpretation of the *content* of the environment, using an adjective ("I find this place *fascinating*"; Hartig et al., 1997) while others explicitly refer to a *process*, using the infinitive verb ("There is much *to explore* and *discover* here," Hartig et al., 1997). This infers subtle differences in the conceptual processes that the item is measuring and the manner in which the individual interacts with the environment. Both may be important for defining and measuring the concept, or they may be a by-product of the item development through the chosen language (e.g., Swedish, German, French, or English) and word composition, without a full consideration of the implications for the concept measurement. From *what* is being said and *how* it is being said, psycholinguistic analysis can be used to derive inferences about how people process and conceptualize sensory experiences (Dubois, 2000). Examination of the linguistics spontaneously used by participants to justify their numerical responses will determine item comprehension and interpretation in relation to the underlying theoretical component being measured. Specifically, the analysis of the use of personal pronouns can be used to infer different conceptualizations at varying levels of subjectivity. For example, the use of singular first-person pronouns ("I", "me") refers to idiosyncratic experiences rather than shared knowledge, the use of collective pronouns ("we", "us") refers to negotiated meaning as collective knowledge, and the absence of personal pronouns (e.g., "*it*") refer to consensual knowledge conceptualized as objective "facts."

In instructions for completion of the perceived restorativeness environment scales, participants are asked to consider the statements in relation to how much it applies to their experience, through the use of the pronoun "my." However, when participants are completing questionnaires with numerous items, at times participants may not thoroughly read and take on board all parts of the instructions, thus the emphasis on *their* experience can be missed if it is only referred too in an opening instruction. In other validated scales, where the individual's perspective is required, the items all begin with the words "*I…*" to emphasize the individual experience (e.g., Warwick-Edinburgh Mental Well-being Scale; Tennent et al., 2007). In contrast, there are variations in the use of personal pronouns ("I," "my") within and between the sets of PRS and PRC items designed to measure each ART component. Seventy-three percent of the 64 original items examined included a pronoun. Items that include personal pronouns infer that the interaction between the individual and the environment is important for perceiving the restorative qualities of a soundscape. Without the inclusion of a personal

pronoun, items could be agreed with *in principle* but does not necessarily mean the individual thinks the soundscape provides restorative qualities for *themselves*. Although this is only a subtle difference, to understand variations in responses from different groups of people, it is important to know exactly how the item is being interpreted and if the personal element is involved in the given rating. Examination of the presence or absence of personal pronouns in individuals' perceived restorativeness responses will indicate the importance of the interaction between the individual and the environment for each ART component.

Finally, as highlighted earlier, the framing of items is influential over participant interpretation and responses (Bentler et al., 1971). PRS, PRC, and PRSS items have been framed in three different ways; holistic, specific, both holistic, and specific. Most items refer to the holistic environment, namely the "soundscape," "place," or "setting," such as "*I find this sonic environment appealing.*" Some items refer to individual or specific elements within an environment, namely the "sounds," "things," or "objects," such as "*When I hear these sounds I feel free from work, routine and responsibilities.*" Occasionally items refer to both the holistic environment and specific elements, such as "*Hearing these sounds hinders what I would want to do in this place.*" Variation in item framing across the different scales may result in their face validity differing. Furthermore, construct validity may suffer if items are framed differently across each component, for example framing Fascination items holistically and Being-Away items specifically, without any theoretical justification for this variation. Further issues arise if items assessing the same component are framed both holistically and specifically, as rating outcomes become harder to interpret. For example, if the result was a low perceived restorativeness rating, to improve the soundscape should an individual sound be removed/altered or is it the combination of sounds that is detrimental? If the framing of the item causes individual respondents to interpret and answer differently, the unsystematic variation in the framing of PRSS items makes it impossible to redesign a soundscape based on the results. Examination of responses to "identical" paired items, either framed holistically or specifically, will identify if both sets of items are easy to interpret, if both are important for evaluating a component, and if interpretation of prior results should be reviewed as item framing has caused variations in responses.

## Development of a PRSS Interpretation Questionnaire

To examine the public understanding of items evaluating the perceived restorativeness of the soundscape an "interpretation questionnaire" was developed, consisting of items for participants to numerically respond too and provide qualitative justifications for those responses, from which their interpretation of the item can be inferred. Development of the questionnaire was directly based upon a large number of PRS and PRC items as well as PRSS items. This was due to the above observations of differences in PRS, PRC, and PRSS item structure and composition, and the importance of, and consistent use of, PRS and PRC by researchers. In total there were 64 potential perceived restorativeness items from three different scales ($n = 23$, Hartig

et al., 1997; $n = 22$, Laumann et al., 2001; $n = 19$, Payne, 2013). To make a feasible questionnaire for participants, the list was reduced to 22 items, which were based upon 35 of the original items (**Table 1**, column 2). The reduction was achieved by noting similarities in items and removing items that included words referring to a sensory modality (such as "I see"), or ambiguous items using homonyms [words with two meanings; such as "Everything here seems to have a proper place," with "place" meaning a location within the place/environment, rather than the place (environment) itself]. To avoid affecting construct validity, items using similar words but representing different ART components were also removed (Compatibility: "This sonic environment *fits* with my personal preferences"; Extent: "The sounds I am hearing *fit together* quite naturally with this place"). Additionally, some items were included to ensure a balance between the different types of item compositions used in the original scales for each ART component, such as the use of adjectives or infinitive verbs.

A series of adaptations to the original items (**Table 2**) were necessary to reflect the psycholinguistic issues identified earlier, namely including personal pronouns when they were missing, and addressing differences in item framing. Therefore, all items, except for Extent items were adapted to include a personal pronoun. Pronouns were not included into Extent items due to its definition which refers to the environment more than the interaction between the environment and an individual. Additionally, there was a major absence of personal pronouns in existing PRS or PRC Extent items; only six of 16 PRS/PRC Extent items include personal pronouns, with none in the Extent items by Laumann et al. (2001). To convert PRS and PRC items to soundscape items, the words "place," "here," "setting," and "surroundings" were changed to "soundscape." The previously used "sonic environment" in PRSS items was also converted to "soundscape" in line with the new ISO definition (ISO, 2014), and a definition was provided to participants. This initially generated 15 holistically framed items. Specific framed items were then generated by changing the word "soundscape" to "sounds," before reversing their sentence order to avoid a repetitive feeling for participants, whilst keeping the same item meaning (**Table 2**). Six items (*refuge, adapt, coherent, clearly organized, spacious, whole world*) however kept the same sentence structure or they became incomprehensive. One Extent item (*belong*) was framed in both a holistic and specific way to avoid a nonsensical item. All items, except two (*obligations* and *concentration*), were positively framed, with high agreement relating to high perceived restorativeness.

The final PRSS interpretation questionnaire consisted of 44 items (**Table 1**, column Holistic framing and Specific framing). These were 22 paired items which were similar except for their framing being holistic (*soundscape*) or specific (*sounds*), and all included personal pronouns, except for the Extent items.

## Environment

The PRSS interpretation questionnaire and subsequent interviews were conducted in two downtown cafes in Montréal, Canada. An indoor environment was necessary due to weather

**TABLE 1 |** Relationship between the original PRSS, PRS and PRC items and their adapted versions for the PRSS interpretation questionnaire.

| Keyword | Original PRSS, PRS or PRC item | Holistic framing | Specific framing |
|---|---|---|---|
| **FASCINATION** | | | |
| Curiosity | This place awakens *my* curiosity[a] | This soundscape awakens *my* curiosity | *My* curiosity is awoken by these sounds |
| Discover | There is much to explore and discover here[a]; There is plenty to discover here[b] | There is plenty for *me* to discover in this soundscape | There are plenty of sounds for *me* to discover |
| Fascinating | This place is fascinating[a] | *I* find this soundscape fascinating | These sounds, *I* find fascinating |
| Interest | Following what is going on here really holds *my* interest[a] | Following what is going on in this soundscape really holds *my* interest | *My* interest is really held by following what is going on with these sounds |
| **BEING-AWAY** | | | |
| Break | Spending time here gives *me* a break from *my* day-to-day routine[a]; Listening to these sounds gives *me* a break from *my* day-to-day listening experience[c] | Spending time in this soundscape gives *me* a break from *my* day-to-day routine | *I* get a break from *my* day-to-day routine from spending time with these sounds |
| Concentration | I experience few demands for concentration when *I* am here[a] | This soundscape demands *my* concentration | *My* concentration is demanded by these sounds |
| Demands | This is a place to get away from things that usually demand *my* attention[a] I experience few demands for concentration when *I* am here[a] | *I* experience few attentional demands by this soundscape | From these sounds, *I* experience few attentional demands |
| Free from | When *I* am here *I* feel free from work and routine[b]; When *I* am here *I* do not need to think of *my* responsibilities[b]; When I hear these sounds *I* feel free from work, routine, and responsibilities[c] | When *I* am in this soundscape *I* feel free from work and/or responsibilities | *I* feel free from work and/or responsibilities when *I* am with these sounds |
| Obligations | *I* am away from *my* obligations[b]; When *I* am here *I* do not need to think of *my* responsibilities[b] | When *I* am in this soundscape *I* need to think of *my* obligations | *I* need to think of *my* obligations when *I* am with these sounds |
| Refuge | This place is a refuge from unwanted distractions[a]; This sonic environment is a refuge from unwanted distractions[c] | This soundscape is a refuge for *me* from unwanted distractions | These sounds are a refuge for *me* from unwanted distractions |
| **COMPATIBILITY** | | | |
| Accordance | There is an accordance between what *I* like to do and these surroundings[b] | There is an accordance between what *I* like to do and this soundscape | There is an accordance between these sounds and what *I* like to do |
| Adapt | *I* rapidly adapt to this setting[b]; *I* rapidly get used to hearing this type of sonic environment[c] | *I* rapidly adapt to this soundscape | *I* rapidly adapt to these sounds |
| Do what want | It is easy to do what *I* want here[a] | It is easy to do what *I* want while *I* am in this soundscape | While *I* am with these sounds, it is easy to do what *I* want |
| Fit | Being here fits with *my* personal inclinations[a]; This sonic environment fits with *my* personal preferences[c] | Being in this soundscape fits with *my* personal inclinations | My personal inclinations fits with being with these sounds |
| **EXTENT—COHERENCE** | | | |
| Belong | The existing elements belong here[b]; All the sounds *I'm* hearing belong here[c] | The existing sounds belong to this soundscape | |
| Fit together | The things and activities *I* see here seem to fit together quite naturally[b] | The sounds fit together to form a coherent soundscape | |
| Coherent | The surroundings are coherent[b] All the sounds merge to form a coherent sonic environment[c] | This soundscape is coherent | These sounds are coherent |
| Organized | It is easy to see how things here are organized[a] | This soundscape is clearly organized | The sounds are clearly organized |
| Order | There is a clear order in the physical arrangement of this place[a] | There is a clear order in the physical arrangement of this soundscape | The physical arrangement of these sounds has a clear order |
| **EXTENT—SCOPE** | | | |
| Exploration | This place is large enough to allow exploration in many directions[a] | This soundscape is large enough to allow exploration in many directions | There are plenty of sounds to allow exploration in many directions |

*(Continued)*

**TABLE 1 |** Continued

| Keyword | Original PRSS, PRS or PRC item | Holistic framing | Specific framing |
|---------|-------------------------------|------------------|------------------|
| Limitless | It seems like this place goes on forever[a]; The sonic environment suggests the size of this place is limitless[c] | It seems like the extent of this soundscape is limitless | The extent of these sounds seems limitless |
| Spacious | *I* experience this place as very spacious[a] | This soundscape feels very spacious | These sounds feel very spacious |
| Whole world | This place has the quality of being a whole world to itself[a] | This soundscape has the quality of being a whole world to itself | These sounds have the quality of being a whole world to themselves |

*Personal pronouns are in italics. Author of original item: [a](Hartig et al., 1997), PRS; [b](Laumann et al., 2001), PRC; [c](Payne, 2013), PRSS.*

**TABLE 2 |** Example of item development for the PRSS interpretation questionnaire.

| | | |
|---|---|---|
| 1. | Original PRS or PRC item | This place is fascinating |
| 2. | Conversion to PRSS item | This soundscape is fascinating |
| 3. | Addition of personal pronoun (holistic framing) | I find this soundscape fascinating |
| 4. | Conversion to PRSS item with specific framing | I find these sounds fascinating |
| 5. | Reverse sentence structure to avoid repetitive feel | These sounds, I find fascinating |

conditions and cafés are frequented for restoration as well as occasional work, thus providing the potential to show the validity of scale items in an environment that may be restorative to some and not others. This helps test the breadth of the scale comprehension, rather than testing it in a traditional restorative environment, such as a quiet, outdoor green space. To be valuable the PRSS should be comprehensible for studies indoors and outdoors, thus although subtle result differences may arise from using an indoor environment, this study helps extend the range of environments used in restoration studies. Additionally, two cafés were utilized to test the scale across multiple conditions and to avoid results being dependent on the interpretation of items in relation to the specific conditions of one environment. The ability of the PRSS to differentiate within one given context is particularly necessary if it is to be helpful in designing restorative soundscapes, and be of value to restorative environment research which is progressing beyond outdoor natural environments.

The two cafés are located across the road from each other, by offices and a university campus, and were distinctly different (**Figure 1**). Café A had expansive windows on the outer "wall," resulting in little need for artificial lighting, and the adjacent busy road and pavement was clearly visible. Overall, it had a rustic theme, basic chairs and tables, as well as a service counter at the entrance displaying food. Café B was enclosed by a small internal wall to separate the café from the surrounding thoroughfare to apartments and a small shopping complex. This café relied on artificial lighting and had considerably fewer customers during interviews than Café A. Overall, it had a modern luxurious theme, large cushioned chairs or stalls at a variety of table types, and an open plan kitchen on one side. Both cafés had a television on with no sound, and pre-recorded

music or a radio station played from the array of speakers. Acoustic measurements were not taken as this study is interested in the interpretation of the items, rather than documenting and assessing the perceived restorativeness of the soundscapes in these two cafés.

## Participants

Ten English speaking participants, aged 20–47 years (median = 25–34 years, 70% female) were recruited via public forums. Two participants had slight hearing issues (undiagnosed tinnitus; right ear hears less treble), but both said it did not knowingly affect their results. In general, participants reported being fairly sensitive to noise ($\bar{x} = 5.5$, s.d. = 1.27, on a 7 point scale) and very aware of sounds ($\bar{x} = 5.9$, s.d. = 1.45, on a 7 point scale). On average participants visited a café weekly, thus it was a familiar setting. Participants visited cafés for multiple reasons, including for food and drink, or for work, but their main reason was for socializing ($n = 8$).

This study was conducted in accordance with the recommendations of the British Psychological Society and the protocol was approved by the Research Ethics Board II at McGill University. All participants gave written informed consent in accordance with the Declaration of Helsinki.

## Measures

The PRSS interpretation questionnaire consisted of 44 items, which were presented in a random order to each participant. All items were rated on a seven point Likert scale from completely disagree (1) to completely agree (7). Each item was followed by a space to provide the "*reason for your chosen response.*" This paper only explores the reasons for responses rather than the numerical assessment, as the sample size is small and the aim is the interpretation of the items, not the assessment of these café soundscapes.

## Procedure

Participants were recruited via an advertisement for a study on the experience and evaluation of urban places. Questionnaire and interviews were completed on weekdays between 10 a.m. and 12 p.m. ($n = 3$) and 3 and 6 p.m. ($n = 7$). Half of the participants participated in café A, half in café B. Information sheets entitled "The evaluation of soundscapes within urban places and evaluating a soundscape assessment tool" were provided. This included a soundscape definition; "A soundscape is the collection of sounds and subsequent ambience that can be heard within

**FIGURE 1 |** Café A (**Top Two**) and Café B (**Bottom Two**).

a particular location. It is a holistic aspect, whereby everything together is larger than the sum of its parts. Thus the sum or collection of sounds is more than each individual sound."

Participants were met within the café, bought a hot drink, invited to review the information sheet again followed by the completion of consent forms. They were then asked to consider the soundscape and sounds for 30 s before listing perceived sounds. The PRSS interpretation questionnaire was then completed. Participants could ask questions at any point and were to underline questions they particularly struggled answering or understanding. A semi-structured interview and demographic questionnaire followed before debriefing. Only free format written responses are analyzed in this paper. Participation lasted around an hour and was recorded on Dictaphones and transcribed. Participants received $10 and a hot drink for taking part.

## Analysis

Participants' written justifications for their numerical ratings were analyzed using the method of constant comparison (Glaser, 1965). Using Nvivo and Excel software Author 1 coded all data (participant responses) and compared notes with Author 2 who separately coded half the data. Coding of individual responses was not mutually exclusive as the 440 responses were coded multiple times on occasions. Both authors produced similar codes with slight variance in the terminology used to name the coding. Discussions and further constant comparison of the data occurred until authors were confident of their interpretation of the data. All components had 80 potential participant responses,

except Being-Away which had 120 potential responses, thus percentages, rather than occurrences, are used to compare across components.

## RESULTS AND DISCUSSION

The results of participants' justifications for their PRSS ratings are presented and discussed below. First, the themes developed from the authors' interpretation of the data is presented. This is followed by a detailed explanation of each theme using data examples, and discusses the implications of the results in relation to the ART components, PRSS validity, and related literature. Participant quotes begin with their numerical response (1–7), followed by their descriptive justification, with the following brackets stating the item keyword (**Table 1**) and if framed holistically (soundscape) or specifically (sounds).

## Interpretation Themes

The qualitative justifications of participants' numerical responses are depicted by six themes (**Table 3**). Two of the themes relate to the Individual (Personal Attributes, Personal Outcomes), two related to the Environment (Physical Environment Attributes, Soundscape Design), and two are an Interaction of environment and individual perspectives (Behavior Setting, Normality, and Typicality). These themes and their sub-themes are discussed in turn below, followed by a comparison of responses from holistic and specific framed items. Thirteen per cent of item responses were not interpreted as belonging to one of these themes due to; (i) no answer was provided at all ($n = 7$); (ii)

**TABLE 3 |** Percentage (and frequency) of responses for each ART component per theme.

| | | Fascination (n = 80)[a] | | Being-away (n = 120) | | Compatibility (n = 80) | | Extent: coherence (n = 80) | | Extent: scope (n = 80) | | Total number of codes | Percent of all responses (n = 440) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Individual | Personal attributes | 24% | (19) | 28% | (34) | 48% | (38) | 8% | (6) | 13% | (11) | 108 | 25% |
| | Personal outcomes | 49% | (39) | 41% | (49) | 63% | (50) | 6% | (5) | 13% | (10) | 153 | 35% |
| Environment | Physical environment attributes | 26% | (21) | 12% | (14) | 26% | (21) | 31% | (25) | 46% | (37) | 118 | 27% |
| | Soundscape design | 3% | (2) | 2% | (2) | 1% | (1) | 40% | (32) | 25% | (20) | 57 | 13% |
| Interaction | Behavior setting | 3% | (2) | 8% | (10) | 30% | (24) | 5% | (4) | 9% | (7) | 47 | 11% |
| | Normality/typicality | 18% | (14) | 3% | (4) | 8% | (6) | 18% | (14) | 5% | (4) | 42 | 10% |
| | | | | | | | | | | | | 525 | 121[b] |

[a]The number of potential participant responses for this component. [b] Interpretation of item responses into themes were not mutually exclusive, hence total percentage >100.

no explanation of the numerical value was given ($n = 23$); (iii) the response did not provide an explanation ("not really"; $n = 23$); (iv) or it related to the study task ("Yes I'm doing the survey"; $n = 6$). There were significant differences across all the ART components and the three overarching theme categories of Individual, Environment and Interaction ($\chi^2 = 141.8$, df = 8, $p < 0.001$). There were more Fascination, Being-Away, and Compatibility item responses themed as Individual than statistically expected and less Environment responses than expected. Compatibility items also had slightly more responses themed as Interaction than expected. In contrast to the other ART components, Extent Scope and Coherence responses were themed more often as relating to the Environment and less about Individual themes than statistically expected.

## Personal Attributes

Personal attributes were referred to in a quarter of participant responses (25%; $n = 108$) with four different subthemes. These related to participants noting their: (i) preferences for certain types of sounds or experiences; (ii) responses may vary depending on their mood, desire, cognitive ability, or activity; (iii) conscious changes in their perception; and (iv) unconscious perceptual changes.

In nearly half of the Compatibility items and a quarter of the Fascination and Being-Away items participants referred to themselves as an important factor in their response rating (**Table 3**). This emphasizes the importance of individuals' assessment that the soundscape has the restorative qualities of these ART components. They are making judgements about the restorative nature of the soundscape for themselves and not for others, again emphasized by their higher use of personal pronouns in Compatibility and Being-Away responses (see section Framing of Items With Personal Pronouns). Thus, the PRSS can be a measure to compare individual differences in perceived restorativeness across settings as well as collating information from a number of people to monitor trends in soundscapes' perceived restorativeness across different groups of people.

## Preference

Individual preferences for sounds and activities ($n = 42$) were frequently mentioned from Compatibility items ($n = 25/42$) and sometimes Being-Away items ($n = 12/42$). Phrasing of the Compatibility items *accordance* and *fit* encourage participants to reflect upon what soundscapes and activities they like in general. Participant responses showed successful contemplation around whether the café soundscape matched those preferences ["6 I prefer this sort of soundscape to one that is too quiet or too loud, like a library or a club/bar" (*fits soundscape*)] or did not match ["2 No, I generally prefer quiet time away from people, unless it's people I've chosen to be with." (*accordance soundscape*)]. Previous studies have found preferred environments tend to also be perceived as restorative environments, with particularly high Compatibility PRS ratings for favorite places (Korpela and Hartig, 1996). This study's results also suggest high compatibility scores in the PRSS may also relate to favorite soundscapes. Examination of the phrasing and participant responses suggest the relationship between preferred and restorative environments may partly be an artifact of the measuring tool, as half of the Compatibility items are indeed measuring preference, thus associations with restorativeness are to be expected. Indeed part of the Compatibility definition relates to preference as it refers to "one's inclinations" and "fitting to what one would like to do" (Kaplan, 1995, p. 173). Given preference is a common assessment in soundscape studies, further consideration of the relationship between preferred and restorative soundscapes could be explored.

Only one response to the accordance item did not refer to sounds, suggesting the comprehension and face validity was good. In contrast, 6 out of 10 responses to the *fits* item did not directly refer to sounds or soundscapes suggesting it was not assessing the compatibility of the soundscape. Instead, other features of the environment were referred too, such as "5 I like cafes" (*fits sounds*) and "1 If I were to spend time and money in a cafe environment, I would choose one with more character and windows" (*fits soundscape*). Given the wording within the fits item of "personal inclinations" has previously been noted as confusing by participants (Payne, 2013), adaptations to improve or remove this item seem necessary.

## Depends on Mood, Desire, Cognitive Ability, or Activity

In this study, participants were sitting in the café and considering the restorativeness of the soundscape due to the task, rather than purposefully having chosen to come to this environment to restore. The consequence of this artificial arrangement meant it made it trickier for participants to provide a fair rating, resulting in many middle numerical ratings of four; "4 Potentially—possibly. But normally it'd be a soundscape I turn off a bit from, in order to concentrate on something else." (*discover soundscape*). People need restoration for different reasons at different times, and different soundscapes may support these needs or hinder it and may vary depending on the specifics of the scenario that caused the need for restoration. Therefore perhaps unsurprisingly this resulted in participants frequently responding with the statement "it depends" ($n = 49$) particularly for Being-Away ($n = 23$) and Compatibility items ($n = 14$). For example "4 Depends on my level of distractability" (*Attentional demands soundscape*) and "5 Well, I've got things on my mind at the moment, so yes. But at other times, I'd disagree more with this statement" (*Obligations sounds*). Knowing how a person is feeling, their level and type of fatigue (if any) prior to doing ratings would help understand the reasoning behind responses and how this varies the perceived restorativeness qualities.

The environment the study was conducted in, cafés, are also multi-purpose environments with the potential to both work or relax, which may have exaggerated the "it depends" issue; "4 In this soundscape, if I were here to relax, I would feel relaxed. If I were here to work I would probably feel as stressed as my mindset was." (*free from soundscape*). Unlike other studies where participants are either purposively fatigued beforehand, such as partaking in a lecture (Laumann et al., 2001) or asked to imagine a scenario where they are fatigued and need restoration (Staats et al., 2003), this study did not provide any such situation. Perhaps if a scenario had been provided the "it depends" variation would have been reduced. However, these responses highlight that both Being-Away and Compatibility items are encouraging the individual to consider the restorative qualities of a soundscape at a particular point in time and it may be useful to understand that context fully to understand the reported level of the perceived restorative qualities of the soundscape. This is highlighted by the response to a Being-Away item where the relative differences between the previously experienced and current soundscape being assessed is important in understanding the response: "3 It certainly is if I'm coming off the street; not so much if I'm coming from my home" (*refuge soundscape*). These relative differences are particularly important to consider when making planning decisions as the relative difference between previously exposed soundscapes and the soundscape under investigation may be important for defining the soundscape as restorative.

## Conscious Perceptual Changes

Responses from eight of the Fascination items raise interesting points about how environmental assessments can change over a period of time; "2 After acclimatizing to this environment the sounds begin to feel uniform" (*discover soundscape*); "5

only initially → once they are identified, I'd rather they disappear" (*curiosity sounds*). In this study participants were in the soundscape for around 10 min before they started assessing the soundscape which took around 30 min. Thus, they had prolonged "exposure" to the soundscape which meant their initial assessments could change over time. This allows for the addition of new sounds which may increase fascination, but in this instance, the fascination actually waned over time. In contrast, most laboratory soundscape studies present stimuli for 15, 20, or 30 s (e.g., Carles et al., 1999; Dubois, 2000; Guastavino, 2007; Axelsson et al., 2010) with only a few including longer recordings of 5 min (e.g., Guastavino et al., 2005) before requesting an evaluation of the sound. The results from this study question the suitability of brief exposures to stimuli to make restorative assessment judgements and perhaps other environmental assessment criteria.

## Unconscious Perceptual Changes

On an equal number of occasions to the conscious perceptual changes ($n = 14$), participants referred to unconsciously directed perceptual changes ($n = 14$), where "2 no, elements seem to ebb in and out of importance within the soundscape" (*clear order soundscape*). These unconscious variations appeared more in the response to Extent items ($n = 7/14$) suggesting stimuli variations help define the extent of a soundscape; "5 Yes, but now that I'm hearing the outside a little more, that world has expanded to include the street" (*whole world sounds*).

# Personal Outcomes

The likely personal outcomes from experiencing the soundscape was the most dominant participant response ($n = 153$). These included subthemes of attention, other cognitive aspects, behavioral actions, and emotions. The PRSS was designed to measure the perceived restorative qualities of the soundscape and thus the likelihood of an individual being psychologically restored, particularly in relation to directed attention, after experiencing a given soundscape. Two specific types of restorative outcomes identified are the ability to recover and reflect (Kaplan and Kaplan, 1989; Herzog et al., 1997). Thus, responses should and did display participants' consideration of what may happen from experiencing the soundscape, with particular reference to attention, recovering, and reflecting. Indeed, over half of the responses to Compatibility items (63%) and nearly half of the Fascination (49%) and Being-Away (41%) item responses mentioned "personal outcomes" from experiencing the soundscape (**Table 3**). In contrast, Extent items hardly referred to outcomes. Personal Outcomes was the most coded theme, which supports the overall aim of the scale measuring what is likely to happen from experiencing the soundscape, and in part supports its content validity (fair representation of the topic).

## Attention

Over half of the comments regarding outcomes from the experience were attention related ($n = 81/153$), thus they partially support the role of the PRSS in measuring the attentional qualities of the soundscape (face validity). The

PRSS is specifically designed to measure the degree to which directed and involuntary attention is likely to be activated, particularly through the ART component Fascination, which is defined as involuntary effortless attention (Kaplan and Kaplan, 1989; Kaplan, 1995). The majority of Fascination items coded as personal outcomes referred to attention in some form (*n* = 29/39), with some responses directly using the word attention; "2; I don't find it particularly attention grabbing at all" (*fascinating soundscape*). The ability to ignore or tune in and out of attending to the sounds was occasionally referred to, such as; "3; I'm generally curious of the origin of sounds/how they shift etc. But there is a uniformity that is also easy to tune out" (*interest sounds*). Participants' "search" for stimuli that evoked involuntary attention was also associated with a level of interest in the sounds and soundscapes; "3; Fascinating in their transparency and interaction, but I am bored of it and look forward to other soundscapes" (*fascinating sounds*); "1; none of the sounds hold my interest, just my attention most of the time" (*interest soundscape*). Therefore, although the cafe sound/scape did not necessarily invoke positive involuntary attention for some participants, their responses suggest these PRSS Fascination items have construct validity in the sense that they assess the extent of specific types of attention being activated. However, their responses also highlight that involuntary attention may be produced by unwanted and undesirable stimuli which would not be restorative. This is in line with early critiques that negatively evaluated nature, like snakes, can induce involuntary attention, and thus "Fascination" alone cannot result in restoration (Ulrich et al., 1991). Thus, positive associated words are important to include in items measuring Fascination, such as "interest." Additionally, in line with the positively framed word "Fascination," originally chosen by the Kaplans to represent involuntary attention (Kaplan and Kaplan, 1989), its definition should always include a positive word, such as "desirable." This explicit emphasis would assist in the development of valid items for measuring Fascination; for example "There is plenty for me to discover in this soundscape" should become "There is plenty *I want* to discover in this soundscape." Indeed, researchers have previously noted that when measuring Fascination, three dimensions should be emphasized, namely pleasantness, intensity (amount of effort), and functionality (recover and/or reflect), and in part have been proposed to differentiate between Hard and Soft Fascination (Hartig et al., 1997; Herzog et al., 1997).

For responses relating to outcomes from the experience, Attention was the second most coded personal outcome for Being-Away items (*n* = 29/49) and for Compatibility items (*n* = 18/50). Half of the personal outcome codes for Extent-Scope were for attention but there were very few of them (*n* = 4). Given there is a Being-Away item on "*attentional demands*" it is hardly surprising attentional aspects were often referred too. However, as with Fascination items, participant comments suggest involuntary attention being invoked, but not in a positive way for this soundscape; "1, The sounds themselves are unwanted distractions!" and "1, These are unwanted distractions and I crave the refuge of my own company, space and voice" (*refuge sounds*). Involuntary attention is generally discussed in terms of

positive attributes in relation to restoration, however, in both the Fascination and Being-Away responses, participants make it clear that their attention at times is being demanded by the sounds, whether they like it or not. This recalls that although Fascination may often be defined as "involuntary, effortless attention," involuntary attention is not equal to Fascination. Unlike visual perception, where an individual can generally choose the direction of their gaze and what they want to look at, audio perception is harder to control and sounds have to be continually filtered and processed to ignore some auditory streams and focus on others (Moore, 2003). This implies another important word in the definition of Fascination, alongside desirable, is "effortless," which was a key aspect of work by James from which the ART evolved (Kaplan and Kaplan, 1989; Kaplan, 1995). Reemphasising effortless attention, and perhaps the distinction between Hard and Soft Fascination (Herzog et al., 1997), may also place a greater importance on the component Extent—Coherence as a coherent environment would aid effortless attention. Indeed a close relationship between Fascination and Extent has previously been hypothesized, albeit in the opposite direction; a fascinating environment would contribute to a sense of extent (Hartig et al., 1997). Additionally, Extent may therefore be important in differentiating between environments or soundscapes with Hard and Soft Fascination if they vary in degrees of intensity (effort) needed. This is important as Coherence, along with Scope, is sometimes not considered in some restorative environment studies (e.g., Nordh et al., 2009; Lindal and Hartig, 2013) but yet may still be an important component for restoration—particularly restorative soundscapes.

As with many psychological processes, it is hard to study the natural conditions of what is occurring and the influence on the individual, as focussing on the topic causes the individual to think or behave differently. The very nature of the PRSS requests participants to focus on and consider the soundscape thereby activating directed attention; sounds that participants would otherwise have been able to "tune out," may now take prominence and "demand attention." Additionally, the process of active listening, which is closer to *musical listening* than the more usual *everyday listening* (Gaver, 1993a,b), clearly influenced some participants' responses. For example, one participant provided a high Fascination rating due to the "interesting" study *process* of active listening and engaging *directed* attention rather than *involuntary* attention, and not because of interesting sounds; "6, Yes, they are ordinary, but it is interesting to be attentive" (*fascinating sounds*). Therefore, although the discussion of attention by participants helps validate PRSS items, it should be remembered that participants' attentional responses during PRSS completion may be different to usual.

## Other Cognitive Aspects
Related to attention, participants also noted their ability to concentrate, be focussed, and productive, when in this type of soundscape; "7, Absolutely. I need sounds to keep me focussed" (*refuge sounds*). For others, they referred to the degree the soundscape allowed their thoughts to wander ["2, Not especially, I think my thoughts could drift off" (*concentration sounds*)] or

even a wandering exploration of the sounds ["6, If it is busy it allows for my ears to wander; however if it is slow then I tend to keep to my own thoughts" (*exploration sounds*)]. This connects to one of the main (but often neglected in research) outcomes from a restorative environment—reflection (Kaplan and Kaplan, 1989; Herzog et al., 1997). Two Being-Away items, *obligations* and *free from* particularly led to statements about "thinking" and "reflecting"; "5, They did torrent through my head as I waited (obligations soundscape)"; "2, Reminiscent of studying at cafes during undergrad" (free from sounds). However, some questioned whether it was the sounds/cape that was causing this or if the holistic environment caused it instead "4, Sitting in a cafe usually causes me to reflect on my obligations but I would not say the sounds make me" (obligations sounds). This raises face validity issues as it highlights some concerns over the ability for people to answer questions specifically about the soundscape without influence from other sensory stimuli. The soundscape definition includes "in context" (ISO, 2014), thus other sensory stimuli should be included in soundscape assessments. This is in line with current multisensory research showing the interaction between sensory modalities, with sensory stimuli presented in one modality impacting our sensory experiences of another modality (Bayne and Spence, 2015) including the impact of sound on visual landscape assessments (Carles et al., 1999).

Overall, Being-Away items particularly mentioned these other cognitive aspects ($n = 34/49$). This meant that along with the comments relating to Attention, half of all the participant responses to Being-Away items were about cognitive outcomes ($n = 61/120$). The words used in the items such as "think" (*obligations*) also help direct participants to consider reflective aspects. These participant responses support the face validity of the PRSS items in evaluating the soundscapes' potential for providing attention restoration, however some responses question if the focus on soundscapes is appropriate for assessing the involvement of Being-Away in providing attention restoration outcomes.

### Action

The behavior or activity participants would do because of the soundscape was mentioned in 41 responses (27% of personal outcome responses). Largely, these included participants' ability to do their desired activity, such as reading, talking, and working. Importantly, participants mentioned the action of relaxing which is often associated with recovery (six from Compatibility items, three from Being-Away items), sometimes as the result of other behavioral actions; "5, Socializing and work are both pleasant and relaxing to do here. If occasionally it's distracting" (*accordance soundscape*). Participants had split views on their ability to relax in this soundscape and the task and particular environment may have prevented the word "relax" from being mentioned more frequently by participants. Reference to relaxation and discussing whether it was possible or not, helps validate that the PRSS items were activating restorativeness assessments. This supports the PRSS validity as an instrument to assess the soundscape qualities as those that could produce restorative outcomes of relaxing and recovery (see cognition above).

Overall, when Compatibility items produced responses relating to personal outcomes, they were more frequently about Actions ($n = 25/50$), and Compatibility items also had more action responses than any other ART component. Given the definition of Compatibility, being a match between the environment's affordances and the individuals' needs and planned behavior (Kaplan and Kaplan, 1989; Kaplan, 1995), it is positive that so many desired actions are mentioned in the responses to justify their ratings. This is in addition to the consequences of those actions, such as relaxing, and the impact on their emotions and attention, which were also frequently mentioned in Compatibility items. The diverse spread of responses across three of the four personal outcome subthemes (attention, emotive, and action) suggest that Compatibility items are an important component for the scale and for ART. It draws on all aspects of the theory, with the activation of attention depending on the environment's match to the individual's need and the potential for restorative outcomes.

### Emotive

Participants at times described the valence of the experience, with Compatibility items responsible for nearly half of the emotive comments ($n = 16/34$). The pleasantness, annoyance, and comfort of the sounds were particularly mentioned; "5 Yes, but the more I listen, the more I'm becoming irritated" (*adapt soundscape*); "7, I am very comfortable in these sounds (*fit soundscape*)." These responses suggest that the soundscape matching the participants' desired emotional mood is important for the individual. Thus emotional aspects are being partly assessed with the PRSS, and support the identified relationship between preferred environments and restorative environments (Korpela and Hartig, 1996) as also discussed earlier. Together the results suggest restorative environments, as assessed by the PRS, PRC, and PRSS, are influenced by emotional responses, however, emotions still play a small role compared to attention and other cognitive outcomes.

## Physical Environment Attributes

A quarter of participant responses included references to specific physical environment attributes within the environment, such as describing sounds by their sources ($n = 91$), the size of the environment ($n = 12$), or visual elements ($n = 4$). There were also references to present sounds changing over time ($n = 20$).

As there are numerous individual sounds that can be listed, this response theme was the second most frequent to occur, as participants only needed to refer to one sound or visual attribute to be coded here. As found with previous linguistic observations of soundscape work, people tended to describe sound sources rather than sounds (Dubois, 2000). Sounds listed by participants included building services (e.g., ventilation), café related objects (e.g., coffee machine), entertainment systems (e.g., music), people (e.g., talking), external street sounds (e.g., traffic) and a few referring to an "ambience." Considered sounds will of course vary depending on the environment and the soundscape being assessed, but the presence of this theme highlights that people did consider the individual sound sources that comprise the soundscape. The reference to sounds

changing over time emphasized the Extent of the soundscape to participants [n=11 Extent item responses; "6 It is continually shifting in terms of the composition and nature of the sounds" (*limitless soundscape*)], and its long-term ability for Fascination to remain [*n* = 6 Fascination item responses; "5 lots of new sounds being introduced" (*discover soundscape*)]. Together, the listing of sounds in the environment and their variation overtime, help validate the PRSS, as the variation of sounds between the two café environments and within them at different times was affecting the perceived restorativeness rating of the soundscape. The prolonged exposure to the stimuli (real world soundscape) in this study also provided the opportunity to note the variations in sound sources, and this prolonged exposure has helped the rating of Extent items and some Fascination items. This again suggests the importance for longer stimuli exposure times in laboratory studies to ensure realistic ratings are provided. Extent items are excluded in a number of online and laboratory studies (e.g., Nordh et al., 2009; Lindal and Hartig, 2013) as prior studies have found Extent results do not compare well with the other components, thus questioning the importance of Extent in restoration, while others have critiqued the items for being unrepresentative of the definition (Pals et al., 2009). However, participant responses in this study suggest Extent items are good measures of the perceived restorativeness of the physical environment attributes. It may just be that participants need longer exposure periods to provide valid ratings for Extent, and this has not been noted due to the tendency to use shorter stimuli exposure in lab settings.

Visual attributes such as "blank walls," "café menu," and "no windows" were infrequently noted by participants as well as the size of the environment; "1 it feels neither spacious nor crowded I like the enclosed café area which feels private while I'm inside it and prevents people accidently (or not) wandering through on the way elsewhere" (*spacious soundscape*). Although it only occurred on a few occasions, references to non-sound related elements highlights difficulties in translating some perceived restorativeness environment scale items to become sound specific (for the PRSS) rather than a focus on all elements of the environment, as with the PRS and PRC. Therefore, there is still further work to increase face validity of some PRSS and to ensure the items are framed in a way that enables the assessor to focus only on the soundscape if the intention is to assess the perceived restorativeness of the soundscape.

Physical Environment Attributes were particularly mentioned by items assessing Extent (31% by Extent-Coherence and 46% by Extent-Scope items). Thus, what the soundscape is perceived to comprise, and the size of the environment has a strong influence over Extent ratings. The composition of the soundscape (discussed in relation to which sounds are present) and the size of the environment clearly relate to the definition of Extent Coherence (structure and organization) and Extent Scope (scale of environment) (Kaplan and Kaplan, 1989). This suggests that the PRSS items are a good measure of the concept, albeit sometimes currently with an influence from non-sound related elements.

## Soundscape Design

Responses from Extent items also predominated the noted theme of Soundscape Design, where the design of the soundscape may have been intentional or not (91% of these responses were Extent items). These largely related to instances of the location of sounds ["3 not really, most sounds seem to emanate from one localized area" (*exploration sounds*)], distances between sound sources ["5 the room 'feels' big, sounds are coming from different distances from one another…" (*exploration soundscape*)] or the composition of the sounds and the environment, ["7 absolutely, the decor and the music go well together and the sounds of people passing are hardly noticeable" (*fit together sound to soundscape*); "2 not organized, at all -> they occur independently of any plan" (*organized sounds*)]. References to sounds being in the foreground and background were also made (*n* = 9); "6 foreground/background clearly defined" (*order sounds*). Although this can depend on the individual perceiver rather than the soundscape design, participants referred to it as if it was an objective description of the acoustic environment rather than having the potential to vary by perceptual differences "1 no the sounds clash although the radio is dominant" (*coherent soundscape*).

Soundscape design is an important area of growing interest (Andringa et al., 2013; Kang et al., 2016) to help reduce the negative effects of environmental noise (World Health Organisation, 2011) and potentially consider the positive impacts soundscapes can also have (Davies et al., 2009). Therefore, it is valuable to note the design of the soundscape is considered within the Extent item responses. The Extent items seem to particularly assess the perceived restorativeness of the physical environment attributes with little influence in the potential variation that may arise between individuals (e.g., little reference to personal attributes and outcomes from Extent items, and little use of personal pronouns). In a future study, it would be interesting to examine statistical responses to Extent items from a large number of people's assessment of the same soundscape. If there was little variation, then these items could be used by independent evaluators to help with assessing and designing restorative soundscapes, without the need for large-scale surveys. The words used within the Extent items also has similarities to words frequently used to assess soundscapes in other studies. For example, ART research uses the words "coherent," "order," and "spacious," while soundscape assessment research uses the words "congruence," "organized," "harmonious," "nearby/far," and "open" (e.g., Carles et al., 1999; Raimbault et al., 2003; Ge and Hokao, 2005; Axelsson et al., 2010).

## Behavior Setting

Behavior settings is the interplay between behavior episodes (goal-directed actions), social inputs, and environmental force units (combination of distinct environmental inputs) (Barker, 1965); behavior settings are the physical environment where standing patterns of behavior occur independent of individuals' perception (Schoggen, 1989). In short, a setting where a series of known activities and behaviors would be conducted. The theme of behavior settings emerged in participants' responses as they often referred to "this type of place" and "in an

environment such as this," or quite simply "6 yup café" (*fit together sound to soundscape*) (*n* = 30). Similar to personal outcomes, activities that occurred in the café were mentioned but responses were coded here when they particularly referred to the activity being in this setting, such as "4 I often associate these environments with work/studying and/or planning things in my life - however I do associate it with socializing as well" (*free from soundscapes*). There tended to be a focus on the environment overall rather than a particular consideration of the sounds or soundscapes which only occurred a few time (*n* = 8/30); "3 I feel this type of sound is associated for me with the type of space - cafe - which I do not usually think of as spacious" (*spacious sounds*). It is the matching of both the activity and the physical environment that explains why half of the responses coded in this theme are from Compatibility items, in line with its definition. Therefore the wording of the PRSS Compatibility items successfully induce people to consider both the environment and intended activities, however the intention of the PRSS is supposed to be on the soundscape's affordances, rather than the general environment, questioning its face validity. The soundscape definition refers to the context of the perceiver (ISO, 2014), thus research is increasingly focussing on the activity of the soundscape assessor (Aletta et al., 2016b; Kang et al., 2016). Thus, although Compatibility item responses do question if the soundscape was focussed on, the inclusion of activity focussed items (via Compatibility) will still be important for assessing perceived restorativeness of soundscapes.

On a number of occasions participants ratings were based upon comparisons of this type of soundscape, a café soundscape, with other cafés' or other environments' soundscapes (*n* = 9). For example, "1 I prefer a quieter place to read without distraction and don't drink coffee too often, mainly when the weather is cold" (*accordance soundscape*) and "6 I prefer this sort of soundscape to one that is too quiet or too loud, like a library or a club/bar" (*fit soundscape*). On other occasions participants contrasted the café *environment* to other cafes or environments with no reference to the sound (*n* = 6); "1 If I were to spend time and money in a cafe environment, I would choose one with more character and windows" (*fit soundscape*). These comparisons highlight the choices people usually make regarding where they go to do certain activities, feel certain things, and to have certain outcomes resulting in choosing one behavior setting over another or choices within a type of behavior setting. This supports prior findings that PRSS is sensitive enough to differentiate soundscapes between environments, such as rural, urban park, and city center, and within the same environment type (Payne, 2013). However, questions remain as to whether participants can truly consider the restorativeness of the soundscape without all other aspects of the behavior setting influencing their ratings. As behavior settings of the same type, say café, will produce similar sounds as there will be similar activities and objects in each place, this is understandable, but the interplay of these aspects should be acknowledged when reporting PRSS results. Therefore, it may only be valuable to compare PRSS ratings from the same behavior setting rather than across behavior settings, to avoid non-soundscape aspects strongly influencing the comparative results.

## Normality, Typicality, Expected, and Familiarity

Ten per cent of coded responses referred to the normality of the sounds or soundscapes, and the typicality of them for a café, thus they were sounds they expected to hear there, and that it was a familiar soundscape or environment (*n* = 42). A third of these responses (*n* = 14) came from Fascination items and another third from Extent-Coherence items (**Table 3**). Two sub themes emerged, with the normality, typicality, expectedness or familiarity referring to the behavior setting (n=29/42) or referring to individual physical environment attributes (*n* = 13/42). All of the behavior setting subtheme responses were in addition to the main Behavior Setting theme (i.e., mutually exclusive), as were eight of the physical environment attributes sub theme responses in addition to the main Physical Environment Attributes theme. Despite the alignment with other themes, conceptually this theme was interesting to discuss and remain separately. A previous study has also identified familiarity as one of three basic dimensions in soundscape perception (along with Pleasantness and Eventfulness) (Axelsson et al., 2010).

The expectedness of sounds for this study's behavior setting, a café, justified a number of Extent-Coherence items' ratings (*n* = 14/42); "6 sounds I would expect to hear at a café" (*coherent sounds*); "7 Very typical sounds for a cafe. Music, coffee, conversations" (*belong sounds to soundscape*). Thereby the interpretation of coherence was partly about the relationship between the sound and the behavior setting. This may also overlap with a consideration of the Soundscape Design theme which was also predominantly interpreted from Extent-Coherence items.

Fascination was defined earlier as the ability of a stimulus to have attention-holding properties, either without the individual needing to direct attention to focus upon the stimulus, or by inhibiting other stimuli from gaining attention (Kaplan and Kaplan, 1989; Kaplan, 1995). The normality, typicality, expectedness, familiarity of the sounds and behavior setting contributed to participants rating of the stimulus holding their attention. In this instance, it generally resulted in negative ratings ["2 Nothing out of the ordinary happening" (*interest sounds*); "2 I find expected, typical" (*fascinating sounds*)] apart from the positive novel listening experience caused by the study task itself. These negative ratings again suggest the interpretation of Fascination items as relating to *desirable* effortless attention holding stimuli. This is understandable given the wording of most of the Fascination items (*fascinating, curious, interest*) but should be emphasized in the main definition. Familiarity may also be an important aspect to include again in future soundscape studies (see Steffens et al., 2017 for further investigation of the effect of familiarity on soundscape assessments).

## Framing of Items as Holistic and Specific

Comparison of item responses to items framed in relation to the sounds (specific), or related to the soundscape (holistic; **Table 4**) found no significant differences in the frequency with which they were interpreted as part of a theme ($\chi^2$ = 1.81, df = 5, *p* = 0.88). This is in agreement with comparisons of the numerical ratings

| | | Framing | |
|---|---|---|---|
| | | **Sounds** | **Soundscape** |
| Theme | Personal attributes | 65 | 56 |
| | Personal outcomes | 87 | 88 |
| | Physical environment attributes | 65 | 56 |
| | Soundscape design | 24 | 30 |
| | Behavior setting | 23 | 26 |
| | Normality/typicality | 17 | 16 |

of each set of matched specific and holistic items, that showed little variation (median difference of 1, with 41% of identical responses) (Payne and Guastavino, 2013). This suggests that the framing of the question in holistic or specific terms did not have a strong influence on participants' interpretations.

## Framing of Items With Personal Pronouns

There was little variation between participant's individual responses to the use of personal pronouns in holistic or specific framed question, with some having none, and two participants mentioning personal pronouns four times more in holistic than specific framed questions. There was a much larger variation across participants though, with one individual only using personal pronouns four out of the 44 potential responses, whilst the other participants used them between 15 and 30 times. Of greater importance was the variance in the use of personal pronouns across ART component item responses. Personal pronouns were used in two thirds of the responses for Compatibility items (66%), and over half of the time for Being-Away items (58%). This is in line with the high level of Compatibility and Being-Away item responses interpreted as relating to Personal Attributes, with the *individual* being an important aspect of the assessment process. Personal pronouns were used 41% of the time for Fascination items, which is surprising given a high number of responses relating to personal outcomes. Instead participants tended to say "it's interesting" or "it's boring," perhaps assuming that other people, like themselves, would also rate things similarly (consensual knowledge). Extent Coherence and Extent Scope item responses only included personal pronouns 28 and 38% of the time, respectively. This emphasizes again participants interpretation of the Extent items as less about the individual's assessment of the soundscape, and more of a consensual knowledge conceptualized as an "objective" assessment of the physical environment attributes. Given personal pronouns were excluded from the PRSS Extent items, this may have influenced the results, however, as some participants did respond using personal pronouns occasionally, this suggests the lack of personal pronouns did not completely direct how an individual should respond to the question.

## Sound or Vision Leading Responses

Face validity of the items is generally supported, in this respect, as only on two occasions did participants specifically use visual terms in their responses ("see"), compared to the multiple times participants used acoustic terms ("hear," "listen," or "eavesdrop"; $n = 12, 13, 2$, respectively). Participants also considered the predominance of the sounds as either "foreground/background" or "tuning in" on particular sounds ($n = 10, 8$, respectively). This suggests the focus on acoustics rather than visual features was consciously adhered too, however other sensory aspects may have unconsciously affected participants ratings, particularly when behavior setting aspects predominated responses.

## Study Limitations

This study only had a small sample size as the focus was on the qualitative descriptions people used to provide reasons for their numerical ratings, rather than gathering a sample size sufficient for statistical testing. Authors were satisfied data saturation was reached as no new codes were being generated with the addition of the last few participants. Providing a "fatigue" scenario to participants may have helped set the situation a little better and made it easier to answer some of the questions. However, the lack of a scenario also aided the results being generalizable to a variety of situations as the responses highlighted how participants felt they would perceive the soundscape depending on a variety of situations, and thus the construct validity of the items across a variety of situations. These insights may have been lost if a fixed "fatigue scenario" had been provided to participants. The lack of personal pronouns in the Extent items for the PRSS interpretation questionnaire may have resulted in the strong emphasis on the physical environment without a consideration of the individual, compared to the other items. These were excluded due to the definitions focus on the environment rather than the individual's interpretation of the environment. Participants also seemed to interpret the items designed to measure Extent in this way too, however it is unknown if the inclusion of personal pronouns would have resulted in an individual perspective or if the concept of Extent does and should only relate to the "objective" physical characteristics of the environment. Ideally to assess the differences in the framing of the items holistically or specifically, the words "sounds" and "soundscapes" should have been a straight switch, but the order of the sentence was also reversed. Unfortunately, this resulted in some awkwardly read items which may have slightly affected comparisons between the holistic and specific framed items. The items were designed in this way to try and avoid a strong feeling of repetition for the participants. Half of the participants however still noticed the similarity of items "5, see no. 17" (*curiosity soundscape*) at least on one occasion. This suggests that regardless of the structure of the item, participants would have responded similarly anyway. Finally the study was conducted in one context, an indoor environment, thus differences may arise if conducted in a different environment such as outdoors. Such differences are expected to be minor, but further research could check this and determine if the interpreted themes remain consistent across each ART component in different environments.

## CONCLUSION

Through this qualitative study, which investigated the construction and interpretation of PRS, PRC and PRSS

items, advancements in understanding the face and construct validity of the PRSS have occurred. In addition, theoretical and further methodological implications have arisen from the findings which are summarized below.

## PRSS Face and Construct Validity

The PRSS was originally adapted from PRS and PRC scales which focus on all aspects of the environment rather than one sensory aspect, although the PRS and PRC have largely been used to rate elements in visual images. The PRSS has previously been tested in experimental and real world conditions where both visual and acoustical information was present (Payne, 2013; Evensen et al., 2016), as was the case in this study. Examination of participant responses suggest at times, participants were considering other information than just the sounds, although acoustic terms were used more frequently than visual terms. This highlights the difficulty in constructing a subjective measure for a singular sense when multiple sensory stimuli is available, particularly as evidence suggests one sense is strongly influenced by other sensory information (Bayne and Spence, 2015). This brings into question the value and validity of the PRSS when used in real world environments and potentially of other sound specific subjective measures. However, the PRSS still has value in laboratory settings where sensory stimuli can be systematically manipulated and the perceived restorativeness of different sounds and soundscapes can be monitored, including in interaction with other sensory stimuli. For example, the PRSS can differentiate between soundscapes within the same environment type, such as urban parks (Payne, 2010) and cafes as suggested in this study. This means that under controlled laboratory conditions where all other visual and contextual information remains the same, the PRSS could be a useful tool for helping designers to determine whether the addition of certain sounds, such as a fountain into a café, would be beneficial in creating an environment with greater perceived restorative qualities.

During the development of the PRSS interpretation questionnaire used in this study, differences in the vocabulary, grammar, and framing of PRSS, PRS, and PRC items were noted both within and between ART components. From this, two sets of items were developed, one set framed specifically (about sounds) and one set framed holistically (about soundscape). Results indicated that participant responses did not differ numerically or thematically between paired specific and holistic items.

Six themes were interpreted from participant justifications of numerical responses to PRSS items. Two related to the individual (personal attributes, personal outcomes), two related to the environment (physical environment attributes, soundscape design), and two were an interaction between individuals and the environment (behavior setting, normality/typicality). This mix of individual and environment themes is in line with ART which discusses both restorative environments and restorative experiences (Kaplan, 1995). Therefore the PRSS items appear to be engaging participants to think about all the necessary aspects for measuring perceived restorativeness, thus supporting construct validity. In addition, this study identified that respondents interpreted items measuring the different ART

components in thematically different ways; ART component responses varied in the extent to which the individual, the environment, or the interaction between the individual and its environment was emphasized. This has implications for studies which choose to only use items that measure some of the ART components such as Fascination and Being-Away. In such studies, the environmental aspects and the interaction between the individual and environment may not be included as much in the perceived restorativeness rating, which may reduce the full understanding of a soundscape's restorative qualities. However, participants freely referred to the two main theoretical outcomes from restoration, recovery of attentional fatigue and reflection, which again supports the construct validity of the scale.

## Methodological Implications for Soundscape and Restoration Research

A number of wider methodological issues were raised from this study. First, many studies ask participants to rate a soundscape after a brief exposure time, lasting a few seconds or minutes. This study suggested that longer periods of exposure to a soundscape (around 40 min) can influence soundscape assessments, in particular for the ART components Fascination and Extent. Future studies should review suitable exposure times to ensure a fair assessment of all evaluative criteria. Secondly, for restoration research setting a fatigue scenario (and perhaps measuring baseline fatigue levels,) is important to avoid many of the "it depends" responses provided in this study. In this study the lack of a fatigue scenario was useful in highlighting the range of potential reasons people may use to respond to perceived restorativeness soundscape assessments, but fatigue scenarios are necessary for studies aiming to produce a restorativeness soundscape value. The type of fatigue scenario used should however be carefully considered, particularly if the environment can be used for a variety of activities. Indeed some of the responses in this study suggest different soundscapes may have different restorativeness values depending on the individual's type of attentional fatigue (such as work related or personal life issues). Thirdly, this study found assessing involuntary attention (Fascination) via self-reporting subjective statements problematic when the study task involves directing attention to the soundscape. Future studies may need to explore other means of assessing involuntary attention of soundscapes, such as through electroencephalogram scans (EEG), as an equivalent to the eye tracking studies starting to be used to assess Fascination in visual studies (Berto et al., 2008; Nordh et al., 2013).

## Theoretical Implications for Attention Restoration

Broader ART implications also arose from the research. Minor adjustments or reemphasis to the definitions of Fascination and Compatibility are suggested to emphasize characteristics that are assumed from the interpretation of the current definitions or how they are currently measured. The positive quality of Fascination always needs noting, alongside an emphasis of the effortlessness of involuntary attention, as sounds can direct attention involuntarily, but sometimes in

a draining and undesirable way (e.g., erratic banging from a neighboring construction site). Using explicit definitions will improve the accuracy of tools designed to measure the defined concept. A relationship between Compatibility and Preference often found in restorative environment research was also highlighted in these soundscape assessment responses, due to the words used in the Compatibility items. Examination of the statistical analysis of the relationship between Compatibility and preference scores in other studies is necessary to decide if there is a need to measure and assess *both* preference and compatibility in restorative soundscape research if they are highly related.

Compatibility was highlighted as an important ART component as more than any other component it led participants to specifically focus on the personal outcomes from experiencing the soundscape, including the two main outcomes said to derive from restorative experiences and environments—recovery and reflect. Extent was also identified to be particularly important for the perceived restorativeness of soundscapes and was particularly affected by the "objective" physical environment attributes rather than individual experiences. Extent is often neglected in restoration research, but this study suggests it may be particularly important for the restorativeness of soundscapes and key for considering the implications of soundscape design and the rest of the physical environment.

Restorative soundscapes are created through a combination of the physical environment and individuals' interpretation of that soundscape as restorative. This research suggests all four ART components are important to ensure soundscapes can be designed to create a potentially restorative environment and that people have a restorative experience, with each component contributing to understanding the environment, the individual, or a mixture of the two. To confirm these theoretical implications, further investigation into them would be necessary

via examination of responses to the original PRS, rather than soundscape specific ones to ensure the implications related to broader theoretical aspects rather than sensory specific issues. Such work was conducted at the same time of this study but is not fully analyzed. Finally, as discussed by Aletta et al. (2016a) the relationship between restorativeness and other soundscape descriptors, such as pleasantness-eventfulness (Axelsson et al., 2010) and appropriateness (Brown et al., 2011) could be explored further to monitor any overlap.

## AUTHOR CONTRIBUTIONS

## FUNDING

## ACKNOWLEDGMENTS

## REFERENCES

Aletta, F., Kang, J., and Axelsson, O. (2016a). Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landscape Urban Plan.* 149, 65–74. doi: 10.1016/j.landurbplan.2016.02.001

Aletta, F., Lepore, F., Kostara-Konstantinou, E., Kang, J., and Astolfi, A. (2016b). An experimental study on the influence of soundscapes on people's behaviour in an open public space. *Appl. Sci.* 6:276. doi: 10.3390/app6100276

Andringa, T. C., Weber, M., Payne, S. R., Krijnders, J. D., Dixon, M. N., Lindon, R., et al. (2013). Positioning soundscape reserach and management. *J. Acoust. Soc. Am.* 134, 2739–47. doi: 10.1121/1.4819248

Axelsson, O., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *J. Acoust. Soc. Am.* 128, 2836–46. doi: 10.1121/1.3493436

Barker, R. G. (1965). Explorations in ecological psychology. *Am. Psychol.* 20, 1–14.

Bayne, T., and Spence, C. (2015). "Multisensory perception," in *The Oxford Handbook of Philosophy of Perception,* ed M. Matthen (Oxford: Oxford University Press), 603–620.

Bentler, P. M., Jackson, D. N., and Messick, S. (1971). Identification of content and style:a two-dimensionalinterpretation of acquiescence. *Psychol. Bull.* 76, 186–204. doi: 10.1037/h0031474

Berto, R., Massaccesi, S., and Pasini, M. (2008). Do eye movements measured across high and low fascination photos differ? Addressing Kaplan's Fascination

hypothesis. *J. Environ. Psychol.* 28, 185–191. doi: 10.1016/j.jenvp.2007. 11.004

Bild, E., Steele, D., Tarlao, C., Guastavino, C., and Coler, M. (2016). "Sharing music in public spaces: social insights from the Musikiosk project (Montreal, CA)," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Hamburg), Vol. 253, 3657–3666.

Brown, A. L., Kang, J. A., and Gjestland, T. (2011). Towards standardization in soundscape preference assessment. *Appl. Acoust.* 72, 387–392. doi: 10.1016/j.apacoust.2011.01.001

Carles, J. L., Lopez-Barrio, I., and de Lucio, J. V. (1999). Sound influence on landscape values. *Landscape Urban Plan.* 43, 191–200. doi: 10.1016/S0169-2046(98)00112-1

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

Davies, W., Adams, M. D., Bruce, N., Cain, R., Jennings, P., Poxon, J., et al. (2009). "The positive soundscape project: a synthesis of results from many disciplines," in *Inter-Noise* (Ottawa, ON), 379.

Davies, W. J., Adams, M. D., Bruce, N. S., Cain, R., Carlyle, A., Cusack, P., et al. (2013). Perception of soundscapes: an interdisciplinary approach. *Appl. Acoust.* 74, 224–231. doi: 10.1016/j.apacoust.2012.05.010

Dubois, D. (2000). Categories as acts of meaning: the case of categories in olfaction and audition. *Cogn. Sci. Quart.* 1, 33–66. Available online

at: https://www.researchgate.net/publication/228707245_Categories_as_acts_of_meaning_The_case_of_categories_in_Olfaction_and_audition

Evensen, K. H., Raanaas, R. K., and Fyhri, A. (2016). Soundscape and perceived suitability for recreation in an urban designated quiet zone. *Urban Forestry Urban Green.* 20, 243–248. doi: 10.1016/j.ufug.2016.09.003

Gaver, W. W. (1993a). How do we hear in the world? Explorations in ecological acoustics. *Ecol. Psychol.* 5, 285–313.

Gaver, W. W. (1993b). What in the world do we hear? An ecological approach to auditory event perception. *Ecol. Psychol.* 5, 1–29.

Ge, J., and Hokao, K. (2005). Applying the methods of image evaluation and spatial analysis to study the sound environment of urban street areas. *J. Environ. Psychol.* 25, 455–466. doi: 10.1016/j.jenvp.2005.10.003

Gidlöf-Gunnarsson, A., and Öhrström, E. (2007). Noise and well-being in urban residential environments: the potential role of perceived availability to nearby green areas. *Landscape Urban Plan.* 83, 115–126. doi: 10.1016/j.landurbplan.2007.03.003

Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Soc. Prob.* 12, 436–445. doi: 10.2307/798843

Guastavino, C. (2006). The ideal urban soundscape: Investigating the sound quality of French cities. *Acta Acust. United Acust.* 92, 945–951. Available online at: http://www.ingentaconnect.com/contentone/dav/aaua/2006/00000092/00000006/art00013

Guastavino, C. (2007). Categorization of environmental sounds. *Can. J. Exp. Psychol. Revue Can. De Psychol. Exp.* 61, 54–63. doi: 10.1037/cjep2007006

Guastavino, C., Katz, B. F. G., Polack, J. D., Levitin, D. J., and Dubois, D. (2005). Ecological validity of soundscape reproduction. *Acta Acust. United Acust.* 91, 333–341. Available online at: http://www.ingentaconnect.com/contentone/dav/aaua/2005/00000091/00000002/art00015

Hartig, T., Korpela, K. M., Evans, G. W., and Gärling, T. (1997). A measure of restorative quality in environments. *Scand. Hous. Plan. Res.* 14, 175–194. doi: 10.1080/02815739708730435

Herzog, T. R., Black, A. M., Fountaine, K. A., and Knotts, D. J. (1997). Reflection and attentional recovery as distinctive benefits of restorative environments. *J. Environ. Psychol.* 17, 165–170. doi: 10.1006/jevp.1997.0051

Herzog, T. R., Maguire, C. P., and Nebel, M. B. (2003). Assessing the restorative components of environments. *J. Environ. Psychol.* 23, 159–170. doi: 10.1016/S0272-4944(02)00113-5

ISO (2014). *12913-1:2014 Acoustics–Soundscape–Part 1: Definition and Conceptual Framework.* ISO.

Kahn Jr, P. H., Friedman, B., B, G., Hagman, J., Severson, R. L., Freier, N. G., et al. (2008). A plasma display window? - The shifting baseline problem in a technologically mediated natural world. *J. Environ. Psychol.* 28, 192–199. doi: 10.1016/j.jenvp.2007.10.008

Kang, J., Aletta, F., Gjestland, T. T., Brown, L. A., Botteldooren, D., Schulte-Fortkamp, B., et al. (2016). Ten questions on the soundscapes of the built environment. *Build. Environ.* 108, 284–294. doi: 10.1016/j.buildenv.2016.08.011

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

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

Korpela, K. M., and Hartig, T. (1996). Restorative qualities of favourite places. *J. Environ. Psychol.* 16, 221–233. doi: 10.1006/jevp.1996.0018

Laumann, K., Gärling, T., and Stormark, K. M. (2001). Rating scale measures of restorative components of environments. *J. Environ. Psychol.* 21, 31–44. doi: 10.1006/jevp.2000.0179

Lindal, P. J., and Hartig, T. (2013). Architectural variation, building height, and the restorative quality of urban residential streetscapes. *J. Environ. Psychol.* 33, 26–36. doi: 10.1016/j.jenvp.2012.09.003

Moore, B. C. J. (2003). *An Introduction to the Psychology of Hearing.* San Diego, CA: Academic Press.

Nordh, H., Hagerhall, C. M., and Holmqvist, K. (2013). Nordh_tracking restorative components patterns in eye movements as a consequence of a restorative

rating task. *Landscape Res.* 38, 101–116. doi: 10.1080/01426397.2012.691468

Nordh, H., Hartig, T., Hagerhall, C. M., and Fry, G. (2009). Components of small urban parks that predict the possibility for restoration. *Urban Forest. Urban Green.* 8, 225–235. doi: 10.1016/j.ufug.2009.06.003

Pals, R., Steg, L., Siero, F., and van der Zee, K. I. (2009). Development of the PRCQ: a measure of perceived restorative characteristics of zoo attractions. *J. Environ. Psychol.* 29, 441–449. doi: 10.1016/j.jenvp.2009.08.005

Payne, S. R. (2008). "Are perceived soundscapes within urban parks restorative?," in *Proceedings of Acoustics '08* (Paris).

Payne, S. R. (2010). "Urban park soundscapes and their perceived restorativeness," in *Proceedings of the Institute of Acoustics and Belgium Acoustical Society* (Ghent), Vol. 32, 264–271.

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

Payne, S. R., and Guastavino, C. (2013). "Measuring the perceived restorativeness of soundscapes: is it about the sounds, the person, or the environment?," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Innsbruck), Vol. 247, 2223–2229.

Purcell, T., Peron, E., and Berto, R. (2001). Why do preferences differ between scene types? *Environ. Behav.* 33, 93–106. doi: 10.1177/00139160121972882

Raimbault, M., Lavandier, C., and Berengier, M. (2003). Ambient sound assessment of urban environments: field studies in two French cities. *Appl. Acoust.* 64, 1241–1256. doi: 10.1016/S0003-682X(03)00061-6

Schoggen, P. (1989). *Behaviour Settings: A Review and Extension of Roger G Barker's Ecological Psychology.* Stanford, CA: Stanford University Press.

Scott, M. J., and Canter, D. V. (1997). Picture of place? A multiple sorting study of landscape. *J. Environ. Psychol.* 17, 263–281.

Staats, H., Kieviet, A., and Hartig, T. (2003). Where to recover from attentional fatigue: An expectancy-value analysis of environmental preference. *J. Environ. Psychol.* 23, 147–157. doi: 10.1016/S0272-4944(02)00112-3

Stansfeld, S. A., Berglund, B., Clark, C., López-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

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. Am.* 141, 1414–1425. doi: 10.1121/1.4976627

Tennent, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., et al. (2007). The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): development and UK validation. *Health Qual. Life Outcomes* 5:63. doi: 10.1186/1477-7525-5-63

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., and Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *J. Environ. Psychol.* 11, 201–230. doi: 10.1016/S0272-4944(05)80184-7

Van Kamp, I., Klæboe, R., Brown, A. L., and Lercher, P. (2015). "Soundscapes, human restoration and quality of life," in *Soundscape and the Built Environment* (Boca Raton, FL: CRC Press), 43–68.

World Health Organisation (2011). *Burden of Environmental Noise: Quantification of Healthy Life Years Lost in Europe.* Copenhagen: World Health Organisation.

# Soundscape Assessment of Aircraft Height and Size

Gianluca Memoli[1]*, Giles Hamilton-Fletcher[1] and Steve Mitchell[2]

[1] School of Engineering and Informatics, University of Sussex, Brighton, United Kingdom, [2] Environmental Resources Management, London, United Kingdom

It is accepted knowledge that, for a given equivalent sound pressure level, sounds produced by planes are worse received from local communities than other sources related to transportation. Very little is known on the reasons for this special status, including any interactions that non-acoustical factors may have in listener assessments. Here we focus on one of such factors, the multisensory aspect of aircraft events. We propose a method to assess the visual impact of perceived aircraft height and size, beyond the objective increase in sound pressure level for a plane flying lower than another. We utilize a soundscape approach, based on acoustical indicators (dBs, $L_{A,max}$, background sound pressure level) and social surveys: a combination of postal questionnaires (related to long-term exposure) and field interviews (related to the contextual perception), complementing well-established questions with others designed to capture new multisensory relationships. For the first time, we report how the perceived visual height of airplanes can be established using a combination of visual size, airplane size, reading distance, and airplane distance. Visual and acoustic assessments are complemented and contextualized by additional questions probing the subjective, objective, and descriptive assessments made by observers as well as how changes in airplane height over time may have influenced these perceptions. The flexibility of the proposed method allows a comparison of how participant reporting can vary across live viewing and memory recall conditions, allowing an examination of listeners' acoustic memory and expectations. The compresence of different assessment methods allows a comparison between the "objective" and the "perceptual" sphere and helps underscore the multisensory nature of observers' perceptual and emotive evaluations. In this study, we discuss pro and cons of our method, as assessed during a community survey conducted in the summer 2017 around Gatwick airport, and compare the different assessments of the community perception.

Keywords: soundscapes, aircraft, height perception, size perception, multisensory perception, questionnaire design, survey development, interviews

## INTRODUCTION

It is well-accepted that, for a given sound pressure level (SPL), aircraft are perceived by local communities to be more annoying than other transportation sources (WHO, 2009). This special status of aircraft-generated sounds has been evolving with time, so that recent studies identified an ongoing increase in sensitivity to aircraft sounds in communities: for the same sound-pressure level, these studies record a larger percentage of annoyed respondents than, say, 10 years ago

(Guski et al., 2017). The reasons for this increase are still unclear: part of the research community attributes this to the "rate of change" in the number of aircraft movements (MVA-Consultancy, 2007) and in the composition of aircraft fleets (Janssen et al., 2011; Guski, 2017), while others report a general change in the attitude toward planes and an increase in the weighting of non-acoustical factors (Bartels et al., 2015).

Recent estimates attribute 66–75% of the variation in recorded perception to non-acoustical factors (Guski, 1999; Arras et al., 2003; Nillson et al., 2007). However, while factors like demographics, occupation, self-reported sensitivity, feeling of being in control are broadly covered in the literature, aspects such as visual perceptions, expectations, and judgments regarding these sound sources are rarely covered.

In this context, different airports, in the United Kingdom (Redeborn and Lake, 2016) and elsewhere (Schreckenberg et al., 2016; Hiroe et al., 2017), have recorded in their local communities evidence of a specific non-acoustical factor, usually worded as "planes are flying lower than before." As reported in Gatwick's Independent Arrivals Review (Redeborn and Lake, 2016), this perception often finds no correspondence in objective data, which show only negligible changes in the height distribution of arriving aircraft, in their average arriving paths or in the measured sound pressure levels.

To the soundscape scientist, this apparent discrepancy between objective and subjective heights suggests a combined effect of visual and acoustic factors in the perception of residents under arrivals routes. Similar cross-modal interaction on acoustic judgements has been highlighted in the context of quiet areas (Pheasant et al., 2007) but, to the authors' best knowledge, has not been properly investigated for aircraft sounds so far. This study is a first attempt to address this aspect of community perception.

Here we propose a method, based on the combination of measurements and social surveys, to address questions like "is aircraft height perceived by individuals reasonably accurately?" and "is there a correlation between aircraft size and height perception?" In a context where it is not clear what causes the reported effect on perception, we propose to run simultaneously one measurement campaign and two social surveys: the first, based on postal questionnaires, 30–40 min long and oriented to long term perception, and the second, based on 15-min face-to-face interviews and focused on assessing perception contextually to the planes passing during the interview. We discuss the design of the two social surveys and their interplay, highlighting how they offer two different but complementary windows on the perception of local communities.

Finally, we discuss pro and cons of the method following a preliminary test on about 200 residents around Gatwick in the summer of 2017.

## MATERIALS AND METHODS

### From Research Hypotheses to Survey Design

According to Frankfort-Nachmias et al. (2015), the design of a social survey requires at least one question (i.e., "is there a non-acoustic impact of aircraft height and size on acoustic perception?") and one hypothesis. At the start of this work, we had two.

The first hypothesis, suggested in Gatwick's Independent Arrivals review (Redeborn and Lake, 2016), attributes the perceived effect to the changing fleet makeup, with larger, but similar proportioned planes being increasingly used over time: an argument used by other studies to explain an increased awareness toward plane-originated sounds (Guski, 2017). This suggests that observers may believe the planes to be closer due to their larger visual size during observation and, potentially, due to a potential contribution on the acoustic side (i.e., larger aircraft may appear even bigger due to increased SPL). This hypothesis is mainly visual and can be assessed by a survey containing appropriate questions on height and size only and by a thorough analysis of aircraft movements and physical dimensions (e.g., from radar tracks).

The second hypothesis, proposed by the authors, was inspired by a well-known report into soundscape research (Payne et al., 2009), which highlighted the multisensory character of what are normally labeled simply as "auditory" experiences. The "soundscape approach" suggests evaluating the interaction between the sounds, the visual size, and the spatial height of passing planes.

If such a multisensory interaction between vision, perception, and interpretation of aircraft sounds exists, this should not be balanced: there is in fact a stronger tendency to favor visual information on acoustic stimuli, rather than the reverse (Posner et al., 1976; Bregman, 1990). In this context, the intrinsic difficulty of judging the height of a passing plane would generate an ambiguity, which is resolved by an increased reliance on alternate senses. For testing this second hypothesis, height-specific questions needed to be accompanied by sound perception ones, like those in the standardized surveys (Fields et al., 2001).

Aircraft sounds, however, can be experienced both indoors and outdoors. Height effects on perception can come from long-term memory (e.g., an opinion built on the repeated passage of lower aircrafts) or short-term judgements (e.g., the occasional passage of an outlier aircraft, sedimented in the memory). To remove these ambiguities, in this study we use in parallel two different interaction modalities: a 40 min long questionnaire, focused on long-term perceptions, and a 15-min questionnaire, targeting short-term judgements. Inspired by the high response rate (60%) recently achieved near Narita (Hiroe et al., 2017), we decided to deliver the 40-min questionnaires by post and the 15-min one during semi-structured interviews. The postal questionnaire was designed to be completed by the participants unassisted and indoors. The semi-structured interviews, designed to be run with a researcher, were targeted to participants outdoors and included a component of "plane spotting," which was used to assess perceptual judgements "there and then."

We designed the two surveys to be interconnected, so that some key questions were repeated, in view of a future comparison. As an example, while exposure outdoor was primarily assessed by interviews, the postal survey also contained two key questions related to aircraft perception outdoors. When possible, we maintained the ICBEN 11-point numeric scale in the
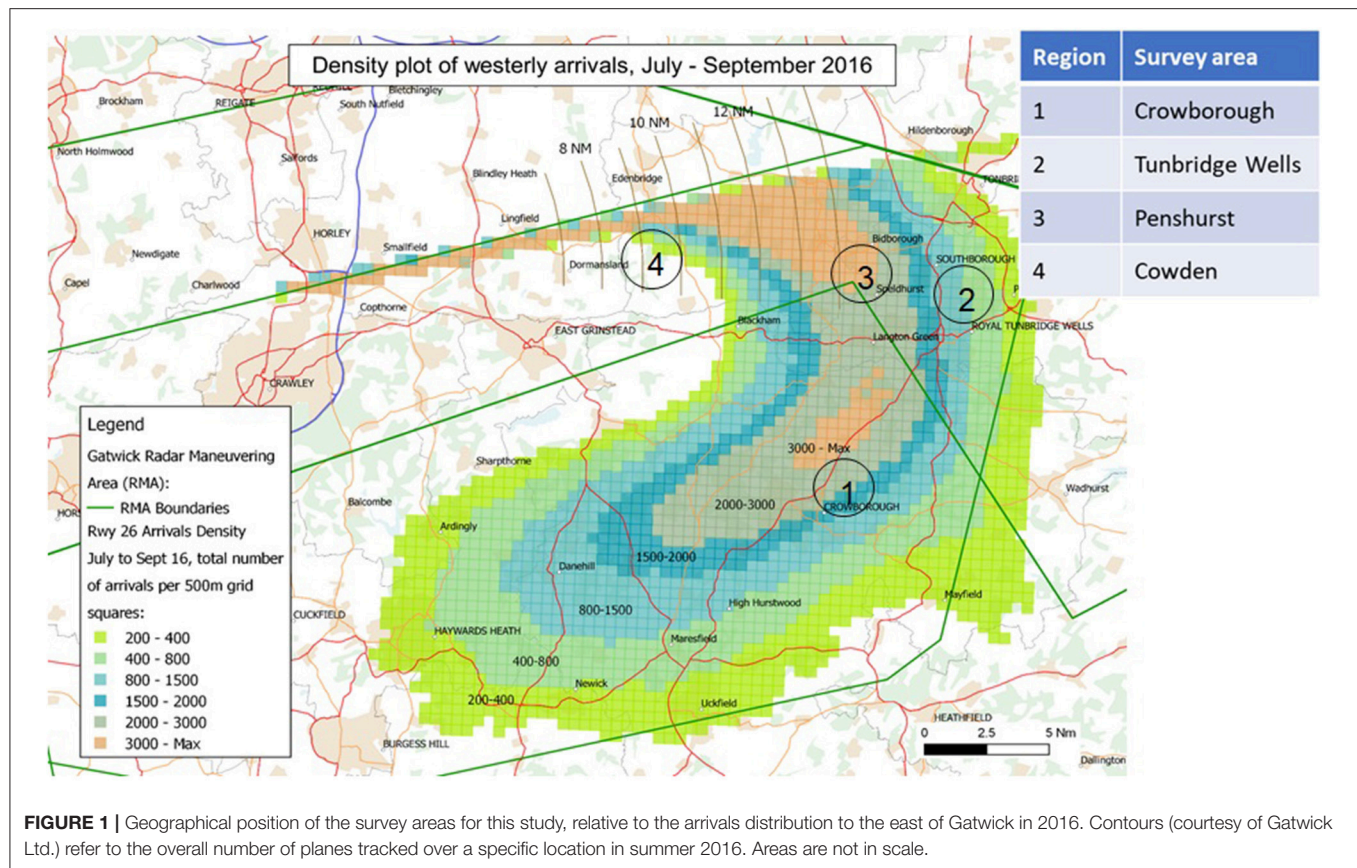
**FIGURE 1 |** Geographical position of the survey areas for this study, relative to the arrivals distribution to the east of Gatwick in 2016. Contours (courtesy of Gatwick Ltd.) refer to the overall number of planes tracked over a specific location in summer 2016. Areas are not in scale.

postal questionnaires and the 5-point ICBEN verbal scale in the interviews (Fields et al., 2001). A similar choice was taken near Narita (Hiroe et al., 2017) and the two scales were compared using recent guidelines (Brink et al., 2016).

Finally, the two social surveys were designed to be assisted by a measurement campaign, also to be ran in parallel, with the goal of assessing the acoustic climate in the selected survey areas, but also of associating acoustic indicators like $L_{A,max}$ and *SEL* (WHO, 2009) to the planes observed during the field interviews. Measurements of plane trajectories (to assess visual distances[1] and real heights) could be done in post-processing, linking the exact time of the passage with the data from flight-tracking apps like *FlightRadar24* or *CASPER*.

## Characterization of the Survey Areas

We tested our method in the summer of 2017, when the number of flights reaches its peak. In the period 28/8–30/9, we focused on three locations to the east of Gatwick airport, along the main arrival path ("westerly arrivals," see **Figure 1**): Crowborough, Penshurst and the center of Tunbridge Wells. Each of these three areas was characterized by a different average aircraft altitude over the ground level (as measured by Gatwick using radar tracks) and contained about 300 households. **Figure 1** also shows

the site of Cowden, which was used as a control, with 200 households.

For the purposes of this study, we will assume that the height distribution of the planes passing over each survey areas is very close to a Gaussian[2]. This hypothesis defines the first statistical parameter with which to characterize each area i.e., the mean height, which corresponds to the height of the most frequently observed plane. As second descriptive parameter, instead of using the standard deviation, we used the height of the lowest plane (defined as the 1st percentile of the height distribution). Having received from Gatwick the numerical height distributions relative to summer 2016 for the different locations (Helios, 2016), we therefore characterized each of the survey areas with two parameters: the height of the "most frequent" plane and that of the "lowest" plane (see **Table 1**).

In terms of population, while Cowden and Penshurst are small villages surrounded by countryside, Crowborough and Tunbridge Well are more urbanized areas. Simply walking through the areas shows that most of the residents live in detached or semi-detached houses. According to the most recent census (Office for National Statistics, 2011), the overall population living in the selected villages and towns could be stratified as follows:

---

[1]In this study, "visual distance" is the distance between the eye of the observer and a passing plane, along the line connecting the two. For clarity, the closest distance to the observer was considered, both during the interviews and in post-processing.

[2]An assumption very close to the real facts, as shown from the distributions acquired by radar tracks for the summers in the period 2011-2016 .

| Location | Lowest plane/ft. | Most frequent plane/ft. |
|---|---|---|
| Crowborough | 2390 | 5257 |
| Tunbridge Wells | 2590 | 4683 |
| Penshurst | 2019 | 4236 |
| Cowden | 1748 | 3818 |

- Age according to census: 18–24 (8%), 25–34 (18%), 35–44 (18%), 45–54 (19%), 55–64 (15%), 65, and over (23%).
- Gender according to census: males (51%), females (49%).

In terms of exposure to aircraft sounds, the selected areas are at least 18 km away from the local airport: a distance much larger than the ones typically surveyed in other studies (MVA-Consultancy, 2007; Civil Aviation Authority, 2017) and beyond the lowest contour (57 dBA $L_{Aeq, 16h}$ by day) of the local noise map (Environmental Research Consultancy Department, 2017). It was therefore necessary to assess acoustical indicators by direct measurements.

Gatwick airport contributed to this study by deploying a mobile acoustic monitor in each of the 4 survey areas. The monitors (Larson Davis, type 870) were mounted inside a weatherproof metal cabinet and connected to an outdoor microphone located at about 4.0 m from the ground (ISO 1996-2, 2017). The monitors were programmed to record all noise events, but those with $L_{Aeq} \geq$ 55 dBA (and lasting at least 10 s) were correlated automatically with details of the aircraft and its flight path using a Noise Track Keeping (NTK) system. Values of $L_{A,max}$ were acquired using a Slow (1 s) time constant.

In addition, a calibrated class I spectrum analyzer (Norsonic 121) was present during most of our field interviews, with its 1/2" microphone mounted on a tripod at 1.5 m from the ground (ISO 1996-2, 2017). These measurements were aimed at planes with $L_{A,max} <$ 55 dBA, for which (we thought) the visual component (i.e., the aircraft height, size, and visual distance) could distinguish planes characterized by the same acoustics. Here, the assignment of $L_{A,max}$ to a specific airplane was performed in post-processing, by synchronizing the measurement with the radar tracks as reported by CASPER (Casper, 2017).

We did not apply any correction for ground reflections (ISO 1996-2, 2017) to the Norsonic measurements, because most of the time the tripod with the microphone was on soft ground (grass), all the interviews were taken in the same (favorable) weather conditions, our acoustic sources were very far from the microphone, and we only used the $L_{A,max}$ of events as they happened.

Our measurements showed that, in each of the areas, plane sounds contributed with an estimated[3] value of $L_{DEN}$ between 47 and 50 dBA, while background sounds (i.e., as given by the level

that was overcome 90% of the time, or $L_{90}$) were between 35 and 37 dBA. In summary, all the survey areas were subject to the same exposure to aircraft sounds, in terms of average energy levels.

# THE POSTAL SURVEY
## Recruitment
A package was sent to randomly selected residents in each survey area (50% of the households), including a pre-paid return envelope and three items (an introductory letter, a consent form and the postal questionnaire), anonymized with a unique ID, in the format "Y-XXXX" where "Y" identified the survey area and "XXXX" is a random number.

The consent form was based on a template produced by the Sciences & Technology Cross-Schools Research Ethics Committee at Sussex and explained how returning the questionnaire was considered an "explicit" act of consent to take part in the study (European Commission, 2011) and to treat the answers anonymously, unless further consent was given (e.g., volunteering for a follow-up, see below). It also detailed how data would be stored and reported instructions on how to withdraw participation.

As a novelty compared to previous studies, we provided an additional mechanism, at the end of the postal questionnaire, aimed at recruiting a small control set of participants. Postal responders could volunteer also to be interviewed (by appointment), in their garden or in a park nearby, thus providing an immediate check between the two interaction modalities (i.e., the postal and the face-to-face interviews).
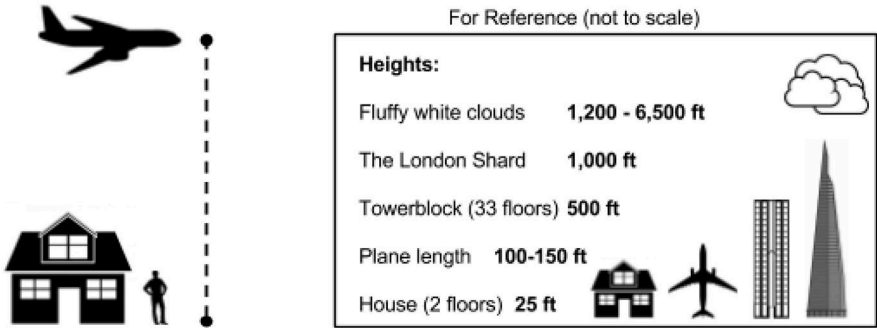
## Questionnaire Design
The postal questionnaire consisted in 80 questions: a combination of the well-established, key questions from technical specification ISO/TS 15666:2003 (Fields et al., 2001; ISO/TS 15666:2003, 2017) and of a set of custom questions, specific to assessing long-term perception of aircraft height/size (see below). The postal questionnaire used in this study can be found attached as Annex 1 and a detailed description of its sections has been added to the **Supplementary Material S1**.

Whether by postal questionnaires, filled at home (Janssen et al., 2011; Hiroe et al., 2017), interviews by telephone (Schreckenberg et al., 2016) or in-person appointments (MVA-Consultancy, 2007; Civil Aviation Authority, 2017), the surveys based on ISO/TS 15666:2003 measure the impact of unwanted sounds on perception in terms of the single parameter "annoyance," evaluated over long periods and at home (ISO/TS 15666:2003, 2017). They share a variant of the same question ("Thinking about the last 12 months, when you are at home, how much does noise from [planes, traffic, rail] bother, disturb, or annoy you?") and their results are quantitatively assessed using either a 5-point verbal scale ("not-at-all" to "extremely"), for use with verbal questions, or an 11-point numerical scale (0–10), for use in written questions (Fields et al., 2001).

There is additional difficulty in adding height-specific questions to such a survey, as the exact nature of forming expectations around height may be informed via visual inspection or auditory influences, and the mere fact of asking

---

[3]We only had measurements for 21 days during the peak summer period, so our values of $L_{DEN}$ are not yearly averages.

**FIGURE 2** | Quantitative assessment of the perceived height of planes in the postal questionnaire.

participants to evaluate the acoustic environment may alter their attention and listening strategy[4] (Truax, 2001). Unwanted effects were mitigated by allowing neutral or positive responses even for what are usually defined "unwanted sounds" (i.e., "noise") in standard questions (Fields et al., 2001). When possible, we also maintained the same wording and positional sequence of questions (Abe et al., 2006). We decided, however, to stick to the traditional single dimension of "annoyance" (which has a negative connotation in itself), even if more recent studies demonstrate that a multi-dimensional analysis may be more appropriate (Schreckenberg et al., 2017).

## Height Scale

Since the postal questionnaire refers to the memory of the respondent, it is not possible to compare directly a perceptive judgement with the real height of a passing plane: the comparison can only be done with statistical quantities. As shown in **Figure 2**, we decided to introduce two perceived quantities– i.e., the "average plane" and the "lowest plane"–without further instructions for the respondents. Nevertheless, as discussed in section Result and Discussion, this apparently free choice linked very clearly to a specific perception of the participants. In the postal questionnaire, we assess height in two ways:

- Quantitatively, asking the respondent a numerical judgement on the height of the "average" plane and the "lowest" plane flying over his/her home (questions C1 and C2 in **Figure 2**).

[4]e.g. "I didn't notice the plane, but now that you mention it, it is annoying". As in quantum mechanics, the act of measuring (perception) influences the result.

- Qualitatively, asking the participant a perceptual judgement on the average/lowest plane flying over his/her home (question C8). We also ask whether the height of the lowest/average plane had changed compared to 1 year or 5 years ago (question C9 in **Figure 3**).

As shown in **Figure 2**, during the initial testing phase for the postal questionnaire, we realized that height assessment required some visual reference, either in the memory of the observer (e.g., famous local landmarks like the Shard or a tower block) or, better, something that could be found on the scene. We initially thought of the clouds but discarded the idea once we saw that their potential height range (1,200–6,500 ft.) is weather-dependent. We then realized that the only object always on the scene is the plane itself, so we added one to the graphical scale. Equally important in **Figure 2** is the presence of a dotted vertical line, to resolve any potential ambiguity between "visual distance" (i.e., the distance between the observer and the passing plane, which may be at an angle) and "height" (which may not be close to the observer).

## Size Scale

**Figure 4** shows the graphical scale that accompanies questions on size (C5 and C6) in the postal questionnaire, with the instructions to use it and the wording of the relative questions.

For assessing size, we wanted a method that could be used with as little guidance as possible and that could be valid for different visual distances. Eventually, we took inspiration from astronomy, where the size of a far star is assessed measuring its image on the eyepiece of the telescope, and devised a method
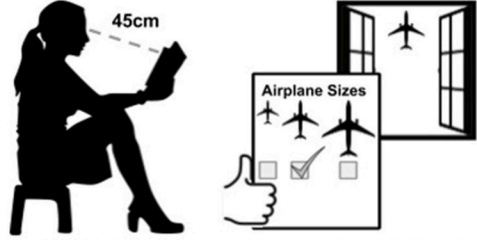
FIGURE 3 | Qualitative assessment of the perceived height of planes in the postal questionnaire.



FIGURE 4 | Qualitative assessment of the perceived size of planes in the postal questionnaire. This chart allows for a quantitative assessment when the distance between the eye and the chart is known.

based on the visual angle i.e., the amount of space that an image will subtend on the retina (Swearer, 2011). For a fixed object size, the visual angle depends on the distance between the object and the observer (i.e., the visual distance), so that larger distances lead to smaller visual angles. Similarly, for a fixed visual distance, larger objects lead to larger visual angles.

**TABLE 2 |** Visual distance at which different planes enter a new size class (in 1000s of feet).

**Visual Plane Size (45 cm away)**

|  | 0.1 cm | 0.5 cm | 1 cm | 1.5 cm | 2 cm | 2.5 cm | 3 cm | 3.5 cm | 4 cm | 4.5 cm | 5 cm |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Class | A | B | C | D | E | F | G | H | I | J | K |
| A319 | 50.0 | 10.0 | 5.0 | 3.3 | 2.5 | 2.0 | 1.7 | 1.4 | 1.2 | 1.1 | 1.0 |
| A320 | 55.5 | 11.1 | 5.5 | 3.7 | 2.8 | 2.2 | 1.8 | 1.6 | 1.4 | 1.2 | 1.1 |
| 737–800 | 58.3 | 11.7 | 5.8 | 3.9 | 2.9 | 2.3 | 1.9 | 1.7 | 1.5 | 1.3 | 1.2 |
| A321 | 65.7 | 13.1 | 6.6 | 4.4 | 3.3 | 2.6 | 2.2 | 1.9 | 1.6 | 1.5 | 1.3 |
| A330 | 94.0 | 18.8 | 9.4 | 6.3 | 4.7 | 3.8 | 3.1 | 2.7 | 2.3 | 2.1 | 1.9 |
| 777–200 | 94.0 | 18.8 | 9.4 | 6.3 | 4.7 | 3.8 | 3.1 | 2.7 | 2.4 | 2.1 | 1.9 |
| 787 | 92.7 | 18.5 | 9.3 | 6.2 | 4.6 | 3.7 | 3.1 | 2.6 | 2.3 | 2.1 | 1.9 |

*According to this chart, an A320 flying 2,200 feet above the observer is seen as class F i.e., the same size as an A330 flying at a visual distance of 3,700 ft. Visual distance at which the plane is seen under the same angle as at 45 cm.*

This method, which appears qualitative, becomes quantitative when the distance between the eye and the reference is known. We therefore put at normal reading distance (45 cm) the silhouettes of an A330[5], scaled at sizes between 0.1 and 5 cm (see **Figure 4**) and asked the participant to select the one that appeared closest to either the average or the lowest plane. This assessment, together with the visual distance between the observer and the passing plane (that can be evaluated from flight tracks), gives a "perceived plane size," which can then be compared with the true size (from flight tracks).

**Table 2** shows a practical reference for size assessment, based on the last plane in **Figure 4** being 5 cm long. As an example of using **Table 2**, an A320 flying at 5,200 ft. just above the observer (visual distance is 5,200 ft.) should be seen as "size C" (row: A320, column: the closest class to 5,200 ft.), while should be perceived as "size D" when flying at 3,800 ft.

The uncertainty related to this method depends mainly on the distance between the reference chart and the eye of the observer. Short-sighted participants, for instance, would tend to keep the reference chart further away. Equally, as confirmed later observing participants during the interviews, long-sighted participants tend to keep it closer. During the testing phase of the postal questionnaire, we estimated an uncertainty of ±5 cm, which introduces an uncertainty of approximately one step in the perception scale (i.e., a correct judge of size, holding the visual chart at 40 cm instead than 45 would judge the planes to be one size larger).

The difficulty in making independent size and height judgements is demonstrated by the effect known as "the moon illusion." It is in fact undisputed that the moon over the horizon appears to be larger than the moon high in the sky (Hershenson, 1989a). This difference in the perception of the size of the moon is illusory: while the perceived size is different at different elevations above the horizon, the physical stimulus that is produced by the light reflected from the moon (i.e., the visual angle at the eye of the viewer) does not change. If a similar effect applies to planes, the perceived size should get larger as the plane gets closer to the horizon (i.e., as the angle to the observer increases).

## Results and Discussion

For this study, we will only report the results concerning the perception of height and size and their relationship with noise measurements and annoyance. Further details can be found in a public report on the Gatwick website (Memoli et al., 2018).

### Demographics

In the selected areas, we collected 112 postal questionnaires (20% response rate). The sample was stratified as follows:

- Self-declared age (postal): 18–24 (1%), 25–34 (3%), 35–44 (12%), 45–54 (15%), 55–64 (31%), 65, and over (38%).
- Self-declared occupation: full-time employed (27%), part-time employed (9%), retired (43%), home/carer (8%), other (12%), prefer not to say (2%).
- Self-declared type of home: detached house/bungalow (55%), semi-detached house/bungalow (32%), other (13%).

According to the age distribution, even if the sample was small, it was representative of the demographics in the area–as assessed by Office for National Statistics (2011). A good part of the postal respondents was over 55, while the younger side of the age distribution (i.e., 18–24) was much less represented. This was either due to the request, at the start of the postal questionnaire, of selecting "the person who spends most time at home" as representative of the household or to a concentration of aged residents in the specific survey areas.

### Perception of Height and Size

**Figure 5** reports a comparison between the measured heights of the "most frequent" plane (i.e., from **Table 1**) and the perceived heights of the "average" plane, as reported by the postal respondents in questions C2 and C6 (see **Supplementary Data Sheet 1**). In looking at these results, it is worth remembering that the wording of the relative questions (see e.g., **Figure 2**) does not define what the "average" and

---

[5]The A330 was chosen as reference since it is the closest in size to the mean of all planes that arrive at Gatwick, and thus should produce the least amount of error.

**FIGURE 5 |** Height **(A)** and size **(B)** of the "average" plane as reported in the received postal questionnaires. Error bars in **(A)** refer to one standard deviation, while letters in **(B)** refer to the size categories described in Figure 4. The "mean" line in **(B)** is mainly a guide to the eye, treating all the survey areas like one single sample.
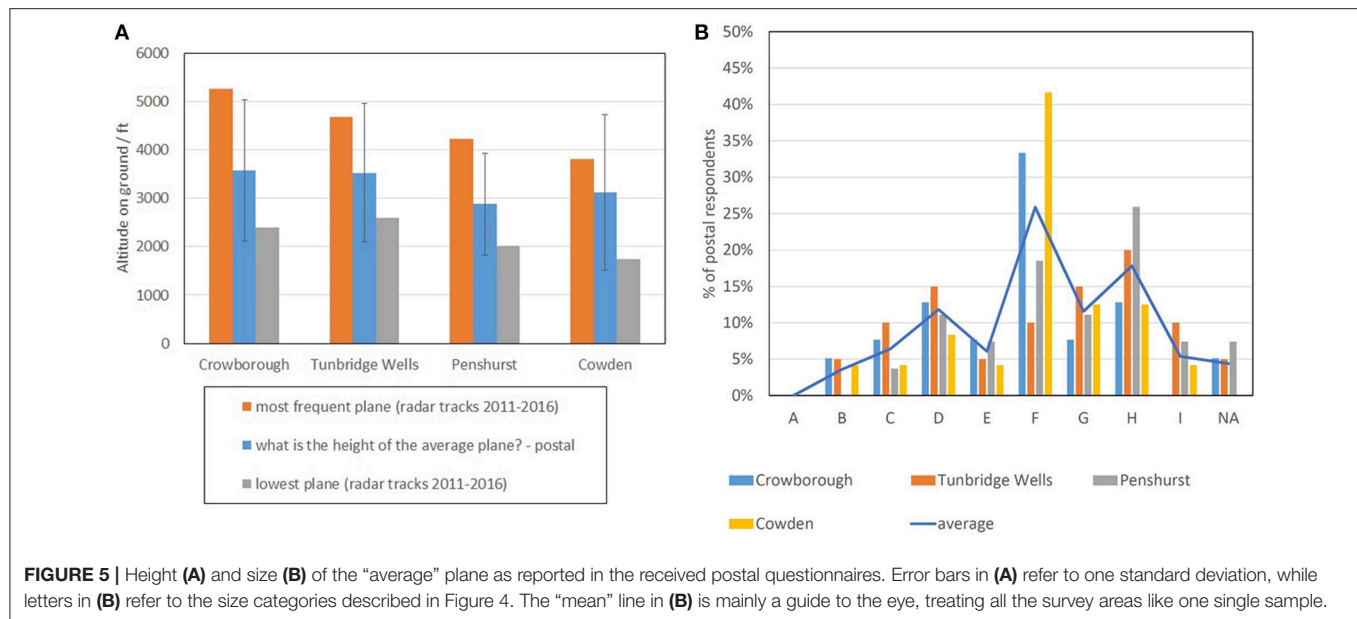
the "lowest" plane are: these are categories assigned by the respondents according to their perceptions.

Respondents reported a perceived height that was typically lower than the one determined by radar tracks (**Table 1**). Most of the postal respondents, for instance, (under)estimated the height of the "lowest" plane within 400 ft, while (under)estimating the height of the "most frequent" plane by 900–1,500 ft (see **Figure 5A**). The fact that the height of lowest plane is so accurately reported highlights its strong presence in the memory of the respondents.

Similarly, most of the respondents reported the correct size class for the lowest plane but perceived the "most frequent" plane to be at least one size larger. According to its size, the "most frequent" plane should in fact be seen in the range C of **Figure 4**, but only 15% of the respondents judged the "average plane" to be in this class (i.e., first peak from left in **Figure 5B**). The other respondents reported a size for the "average plane" at least two classes higher.

A plausible reason for this discrepancy (in terms of height and size of the "most frequent" plane) is labeling the postal sample as more prone to negative comments (Janssen et al., 2011). In support to this conclusion, we noted that 22 of the 112 postal respondents (20%) declared to have filed at least one complaint to the airport. These represent about 50% of the highly annoyed in our sample (i.e., a total of 44 out of 112 respondents reported a score $\geq$ 7 to the annoyance question D3 in the part regarding "planes") and 48% of the ones who reported sleep disturbance (i.e., a total of 46 out of 112 respondents scored $\geq$ 7 to question D3 in the part for "sleep disturbance"). With the expected percentage of those complaining ranging from 2% (Avery, 1982) to 19% (Van Wiechen et al., 2003) of the highly annoyed ones, this is a much larger value than what reported in other studies (Maziul et al., 2005). This hypothesis was further tested in the field studies, which typically offer a different window into community perceptions.

## FIELD INTERVIEWS

As described above, we decided to run two surveys in parallel to probe both long-term and short-term perceptions. Investigations on outliers or on the correlations between acoustical and visual indicators were only possible by commenting on the planes as they passed over the observer. Running two surveys simultaneously also allows the researchers to maximize community involvement (e.g., picking the age groups or group of respondents not fully represented by the postal survey returns) and, at the same time, build up their own impression of the local reality. In hindsight, we also noticed that sending a postal questionnaire improves the chances of being well-received when visiting for unannounced interviews[6], just like conducting interviews increases the response rate of postal studies.

To minimize impact on the participants' life, we designed our interviews to last no longer than a successful marketing or fundraising interaction, i.e., 15-min (Market Research Society, 2014). Advantageously, 15-min should also be sufficient to establish a perceptual acoustic judgement, according to recent models of acoustic perception (De Coensel and Botteldooren, 2008) and to some experimental studies on planes (Breugelmans et al et al., 2017) and other traffic sources (Memoli et al., 2008; Memoli and Licitra, 2012).

We assigned to the field interviews also the role of looking at planes "there and then." This was achieved by what we called "plane spotting": as soon as a plane appeared in the field of view of the interviewee, the flow of the interview was interrupted, and the interviewer delivered a set of targeted questions related to that specific plane ("single-plane questions").

---

[6]Randomly selected residents who have already received the postal questionnaire know already what is happening, while others may have heard by word-of-mouth.

## Recruitment

The field interviews in this study occurred unannounced, to avoid the establishment of prejudices that could affect short-term judgements. Consistently, we decided to recruit participants not by appointment, but meeting them on their doorstep or in a local park and to run the interviews in a semi-structured way, to leave more space for free comments and to create a friendlier atmosphere between the researcher and the participant.

In September 2017, the research team visited each survey areas at various times of the day, at least once during the week and once during the weekend. Once in a location, the team split: one was fixed near the noise meter and the other knocked at the doors in a specific road. Then the noise meter was moved in another road and the roles were inverted. Every time one of the researchers encountered a person willing to be interviewed, he/she would start reading the ethics form (see **Supplementary Data Sheet 2**). In doing so, he/she would formally invite the potential interviewee to be part of the study, would explain our procedure of data storage, would mention how to cancel the responses at any time and would ask for an explicit consent. Following advice from the Ethics Committee at Sussex, we registered consent either by getting a signature or by recording a pre-prepared sentence.

The researcher would then follow the flow suggested by the pre-prepared questionnaire, interrupting it as soon as a plane could be spotted in the sky. In our design, in fact, the goal for each interview was to acquire the interviewee's opinion on at least one passing plane, while the interaction lasted[7].

## Questionnaire for the Semi-structured Interviews

The guide questionnaire (see **Supplementary Data Sheet 2**) is like the one used in the postal survey, plus something specific. It has questions on:

- demographics (age, gender, type of home, employment status, local to the area);
- non-acoustical parameters ("feeling in control," presence of sound insulation at home, sensitivity to unwanted sounds);
- annoyance at home and sleep disturbance (a direct link to the postal questionnaire);
- changes in the number/height/loudness of planes in the past 24 h;
- a section assessing "when do you feel a plane flies over you," assessed in two questions, like in the postal case.

The key differences with the postal survey are:

- The scales used. Since interactions were verbal, we used in the interviews only 5-points verbal scales throughout (Fields et al., 2001).
- An additional question in the ice-breaking section (i.e., at the start of the guide questionnaire). We asked whether the interviewee had heard about our study. This allows the researchers to identify potential external influences on the
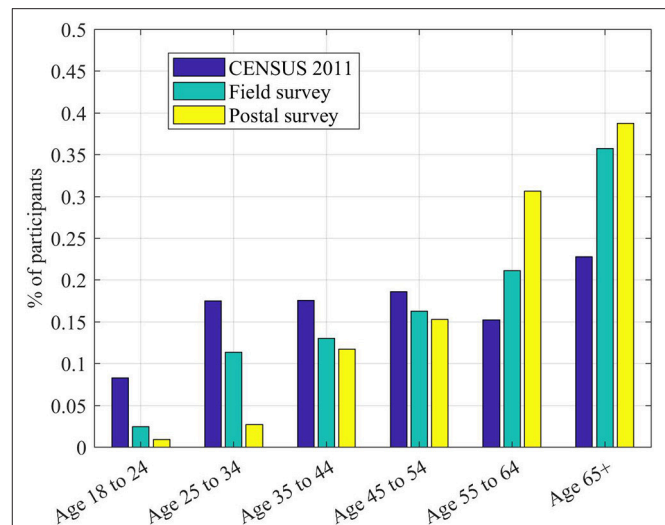


**FIGURE 6** | Demographics of the postal and in-person samples, compared to the latest census available for the selected survey areas (Office for National Statistics, 2011). Data reported in this study refer to the Parishes and Wards databases within CENSUS 2011.

interviewee but also, more simply, the interviewees who had already filled in the postal questionnaire.
- The role of outliers was assessed only in the field interviews, interrogating the participant on "extremely noticeable planes" (Questions 8, 9, and 10) and on which of their activities they felt aircraft sounds impacted most.
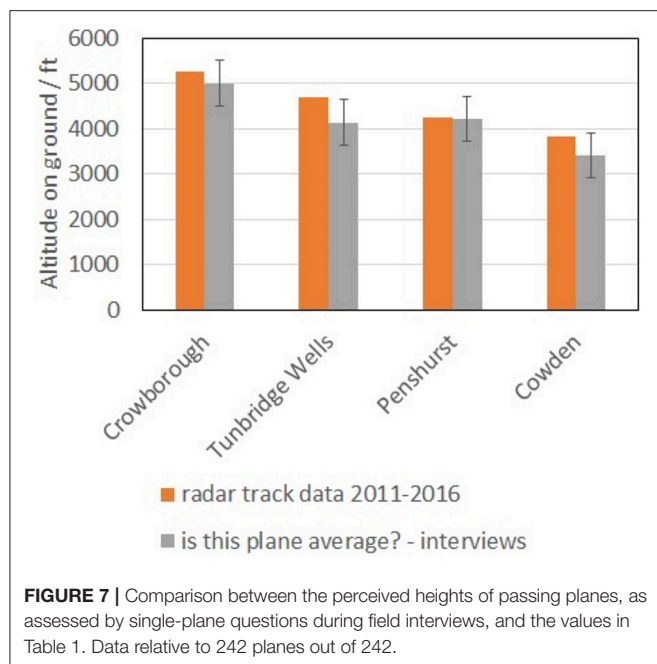
Whenever a plane passed on sight, however, the interviewer would pass to a "single-plane" questionnaire (inset of the field questionnaire, as shown in **Supplementary Data Sheet 2**). This part contained questions on the absolute assessment of height/size of the specific plane, but also an assessment of short-term annoyance. The single plane questions also covered by how much the observed aircraft was far from the "average plane." The reference scales for height (**Figure 2**) and size (**Figure 4**) were handed to the participant, so that the researchers could check the appropriate reading distance was used (**Supplementary Data Sheet 3**).

## Results and Discussion

As in the case of the postal survey, in this work we focus on the perception of height and size as determined during the semi-structured interviews.

### Demographics

In this part of the study, we collected 123 field interviews, observing 242 planes. The questions probing the demographics of the participants (**Figure 6**), their occupational status and the type of home gave results very similar to the ones in the postal questionnaire. It is worth noting that, while we did not have a direct question on whether the participant worked at the airport, this was part of the conversation: only in one case (i.e., a pilot) the participant declared to be directly related to Gatwick.

---

[7]Preliminary tests, conducted with students before going in the field, showed that 15 minutes allowed a maximum of three planes to be observed for each participant.

## Perception of Height

**Figure 7** reports a comparison between the perceived height of the "average plane," as determined during interviews, and the height of the "most frequent" plane, from **Table 1**. The perceived values in **Figure 7** were determined by selecting the planes that interviewees labeled as of "average height" and finding the mean and the standard deviation (error bar in **Figure 7**) of their distribution. This process defines the "average plane." **Figure 7** shows that, in this survey, the "average" plane corresponded, according to our reported answers, to the "most frequent" plane. Also, given the relatively small value of the standard deviation, it can be concluded that interviewees distinguished well when a plane was "average."

Figure 8 shows a comparison between the perceived changes from the "average" plane, as assessed during interviews, and the real changes in height (as determined by radar tracks). Results show that, except for Cowden, interviewees also distinguished well changes from the "average plane": when planes were higher, they were perceived as higher. Equally, when planes were lower, they were perceived as lower. Particularly interesting is the case of Crowborough, where the planes fly higher than the others and with a wider spread.

Conversely, when asked a numerical judgement on the height of the "average" plane, the interviewees (**Figure 9**) tended to underestimate it, like the postal respondents (**Figure 5A**), by about 1,200–1,500 ft (i.e., 350–450 m). As discussed earlier, this is potentially not surprising, given the absence of references on the line of sight between the observer and the plane: it may simply show that the references we used on paper were not sufficient.

**Figures 5–8** answer the question "is aircraft height perceived by individuals reasonably accurately," showing evidence that residents well-know the height of the most frequent plane (i.e., where most of the planes should be in the sky), but also that their
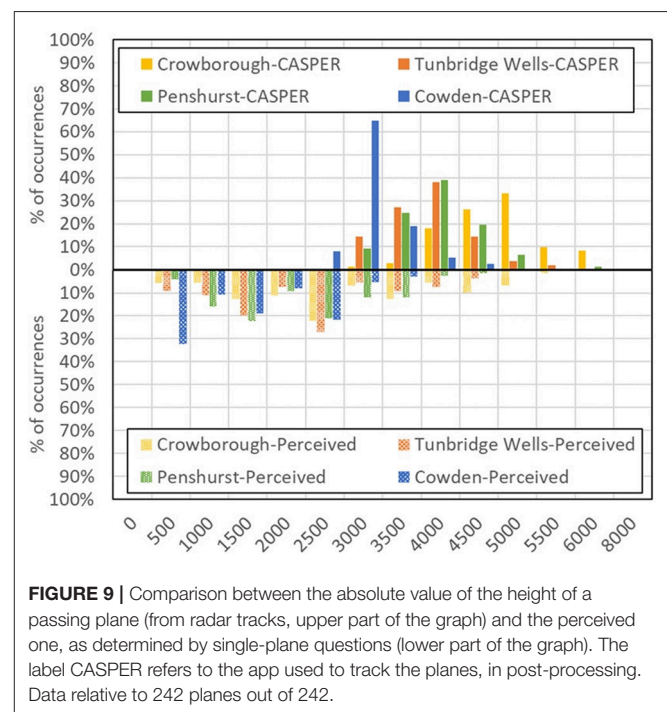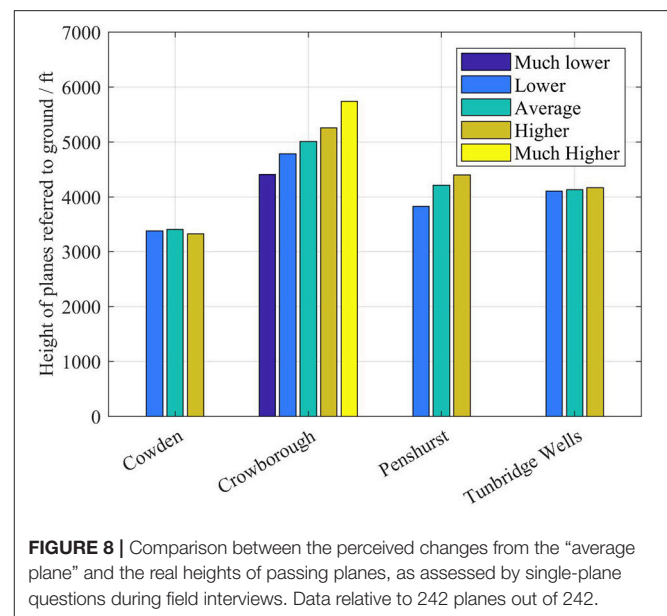
absolute estimate of the height of the most frequent plane is not accurate.

Interestingly, the real heights of "most frequent plane" and of the "lowest plane" were within one standard deviation from the perceived height of the "average plane" (this is particularly clear in **Figure 5A**). There is therefore evidence that, in the process of averaging the height distribution in their memory, postal respondents may have weighted the lowest planes more than the highest ones.

**FIGURE 10 |** Mean annoyance for different qualitative judgements on the height of a plane, as assessed during "single-plane questions" (242 planes out of 242).



**FIGURE 11 |** Comparison between the absolute value of the size of a passing plane (from radar tracks, upper part of the graph) and the perceived one, as determined by single-plane questions (lower part of the graph). The label CASPER refers to the app used to determine the actual size of the planes, in post-processing.

**Figures 5–8** also suggest that, since the participants to our study were sensitive to planes not flying like the "average plane" (with a sensitivity that depends on the location, as shown in **Figure 8**), it is the changes from the average that may trigger negative perceptions and annoyance.

A further evidence in this direction comes from **Figure 10**, where the mean annoyance (European Environmental Agency, 2010; Guski, 2017) has been calculated relatively to the qualitative judgements on plane height, for each location. **Figure 10** shows that, at least for the locations of Penshurst and Cowden, the mean annoyance increases as the planes are perceived to be lower than the "average plane." The absence of a trend for Crowborough and Tunbridge Wells confirms that a larger sample would need to be analyzed, before drawing definite conclusions.

This finding, however, goes in the direction proposed by a recent study (Filipan et al., 2017), where the authors have found that the perception of tranquil areas in the city parks of Antwerp is mostly affected by the sounds that visitors are not expecting to hear. Changes from the expected may be the cause underpinning annoyance.

## Perception of Size

If height tends to be underestimated, both surveys confirm that participants tend to overestimate the size of passing planes: as shown in **Figure 11** (relative to single-plane observations), they were reported to be up to two classes larger (i.e., up to twice as large). Due to the uncertainty on the reading distance discussed earlier, however, this effect may well be within the limits of the method.

We did not observe any correlation between the error in assessing size (EAS, defined as the ratio between the perceived size and the actual size of a passing plane and therefore reported in arbitrary units or a.u.) and the actual size of a plane ($r =$
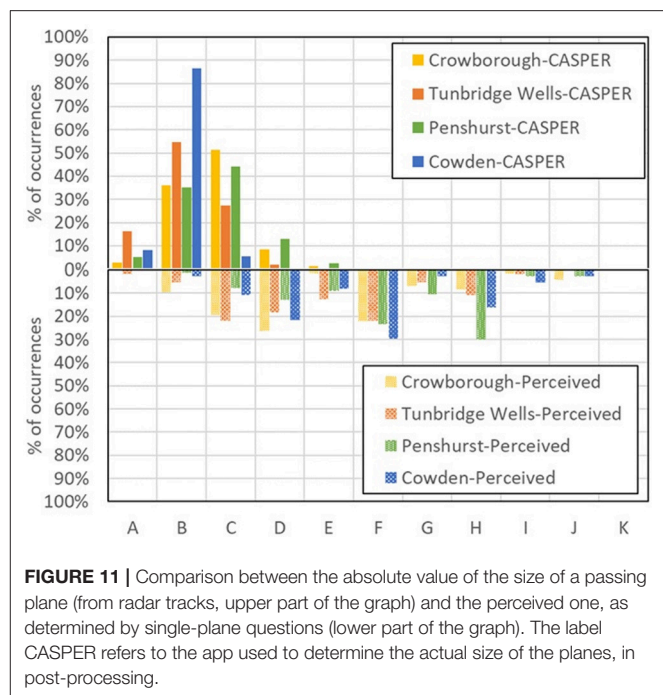
$-0.15$, $p = 0.07$). We found instead a correlation between EAS and the visual distance ($r = 0.66$, $p < 0.001$): it is much easier to get the size wrong for planes further away i.e., the size-distance invariance hypothesis fails at large distances, like in the moon illusion (Hershenson, 1989b). Unfortunately, our results do not show a clear trend that could be linked to one of the existing theories for the size-distance paradox (see **Supplementary Figure 1**).

## Comparison With Acoustic Indicators

As mentioned earlier (section Characterization of the survey areas), a measurement survey run in parallel to the social surveys: one of its aims was to assign a value of $L_{A,max}$ to each passing plane captured during the field interviews. In this part of the study, we only use 144 of the 242 available plane events i.e., those where our tracking procedure managed to assign a unique value of $L_{max}$ and were therefore clearly unaffected by other acoustic sources in the background. On these planes we run a preliminary analysis, based on the Pearson correlation test (using MATLAB R18), which did not show any correlation between the error in assessing aircraft height (EAH, defined as the difference between the real height of the plane, as obtained by radar tracks, and the perceived one, as reported during the interviews, with negative values corresponding to underestimation) and the objective variables. Specifically, assuming $p \leq 0.05$ as significance level, we found no correlation between EAH and the real height ($r = -0.22$, $p = 0.08$), the size of the plane ($r = 0.045$, $p = 0.56$), the visual distance ($r = 0.16$, $p = 0.06$) or the peak noise level during an aircraft pass-by ($r = -0.11$, $p = 0.178$). Recent studies, however, suggest that the Pearson test may not be sufficient while analyzing sparse data (Liu et al., 2012).

In the case of EAH vs. $L_{A,max}$ (**Figure 12A**), in fact, while the results are clearly sparse ($SD$ : 6 $dB$ for $L_{A,max}$ and 1,000 ft for EAH), most of them can be found in the central region of the graph. This statement is confirmed by **Figure 12B**, which reports the number of data points in a grid spaced 500 ft vertically and 2 dB horizontally (the pace of the grid reflects the categories in the questionnaire and the measurement uncertainty).

This finding suggests a linear regression $y = a + b \cdot x$ based on the chi-square merit function (Press et al., 1992):

$$\chi^2 (a,b) = \sum_{i=1}^{N} \left( \frac{y_i - a - b \cdot x_i}{\sigma_i} \right)^2 \tag{1}$$

where $x_i$ is the i-th value of $L_{A,max}$, $y_i$ the corresponding value of EAH and $\sigma_i$ is the "weighted uncertainty" on the value $EAH_i$, obtained from the initial uncertainty ($\varepsilon_i = 500$ ft, from the questionnaires) in order to weight some regions of **Figure 12A** more than others (see below). This method gives $a \pm \sigma_a$ and $b \pm \sigma_b$ where (Press et al., 1992):

$$a = \frac{\sum_i \frac{x_i^2}{\sigma_i^2} \cdot \sum_i \frac{y_i}{\sigma_i^2} - \sum_i \frac{x_i}{\sigma_i^2} \cdot \sum_i \frac{x_i y_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2} \cdot \sum_i \frac{x_i^2}{\sigma_i^2} - \left( \sum_i \frac{x_i}{\sigma_i^2} \right)^2};$$

$$\sigma_a^2 = \frac{\sum_i \frac{x_i^2}{\sigma_i^2} \cdot}{\sum_i \frac{1}{\sigma_i^2} \cdot \sum_i \frac{x_i^2}{\sigma_i^2} - \left( \sum_i \frac{x_i}{\sigma_i^2} \right)^2} \tag{2}$$

$$b = \frac{\sum_i \frac{1}{\sigma_i^2} \cdot \sum_i \frac{x_i y_i}{\sigma_i^2} - \sum_i \frac{x_i}{\sigma_i^2} \cdot \sum_i \frac{y_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2} \cdot \sum_i \frac{x_i^2}{\sigma_i^2} - \left( \sum_i \frac{x_i}{\sigma_i^2} \right)^2};$$

$$\sigma_b^2 = \frac{\sum_i \frac{1}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2} \cdot \sum_i \frac{x_i^2}{\sigma_i^2} - \left( \sum_i \frac{x_i}{\sigma_i^2} \right)^2} \tag{3}$$

In this study, the weighted uncertainties $\sigma_i$ were assigned to $y_i$ by taking the initial value $\varepsilon_i = 500$ ft (which is equal for all the points) and dividing it by the number of occurrences in the region that contains $y_i$. Therefore, if $(x_1, y_1)$ and $(x_2, y_2)$ are all the points contained in the same region $\Xi$ of the 2D histogram in **Figure 12B**, they both get $\sigma_1 = \sigma_2 = 250$ ft; if $(x_3, y_3)$ is the only point in region $\Phi$, its uncertainty remains $\sigma_3 = 500$ ft.

This approach corresponds to looking for a regression that does not depend on other parameters, where the single data points have a weight related to their statistical significance (i.e., if a larger number of people gave a similar answer, that answer counts more than others). Using all the data (144 points) and the weights $1/\sigma_i$, minimizing the chi-square functions leads to $a_1 = 0 \pm 100$ ft and $b_1 = -26 \pm 3$ ft/dBA (see **Figure 12C**). This fit suggests that the louder the plane, the larger the value of EAH. Its "goodness of fit," however, is barely acceptable: MATLAB *fitnlm* function gives in fact ($r = -0.149, p = 0.07$).

We therefore applied a form of subset selection (Miller, 2002), focusing on the center of **Figure 12A** and neglecting all data with

$\sigma_i \geq 250$ ft. In this way only 72 data points (of the 144 available) are used in the fit, but the linear regression is much stronger ($r = -0.407, p < 0.001$), with $a_2 = 16,200 \pm 600$ ft and $b_2 = -300 \pm 10$ ft/dBA in the region $54 \leq L_{A,max} \leq 64$ dBA (see **Figure 12D**).

To clarify the potential impact of our findings, we will use the fitting line in **Figure 12D** and consider a plane flying on day 1 over Crowborough at 4,200 ft., with $L_{A,max} = 57$ dBA. Following the vertical at 57 dBA, we encounter the guiding line joining our data at $-900$ ft, so this plane will be perceived to be flying at 3,300 ft., with $L_{A,max} = 57$ dBA. If the same plane, on day 2, overflies Crowborough at 3,400 ft, its emission as a point source[8] will increase to $L_{A,max} = 58.8$ dBA. Joining the vertical at 58.8 dBA with the red dotted line gives an increase in the EAH, which becomes $\approx -1,400$ ft. The second day, then, this plane would be perceived to fly at 2,000 ft. The plane would be flying lower, by 800 ft., but would be perceived to fly much lower, by $\sim$2,200 ft.

No other correlation was found for EAH, even when the subset selection method was applied to the other variables. If confirmed over a larger sample (e.g., including the 98 plane events not used in this study, as their $L_{A,max}$ was affected by non-aircraft sources), these results may give a new insight into the perceptual mechanism causing annoyance due to unwanted plane sounds to rise much quicker (due to changes in perceived height) than the one corresponding to other traffic sources.
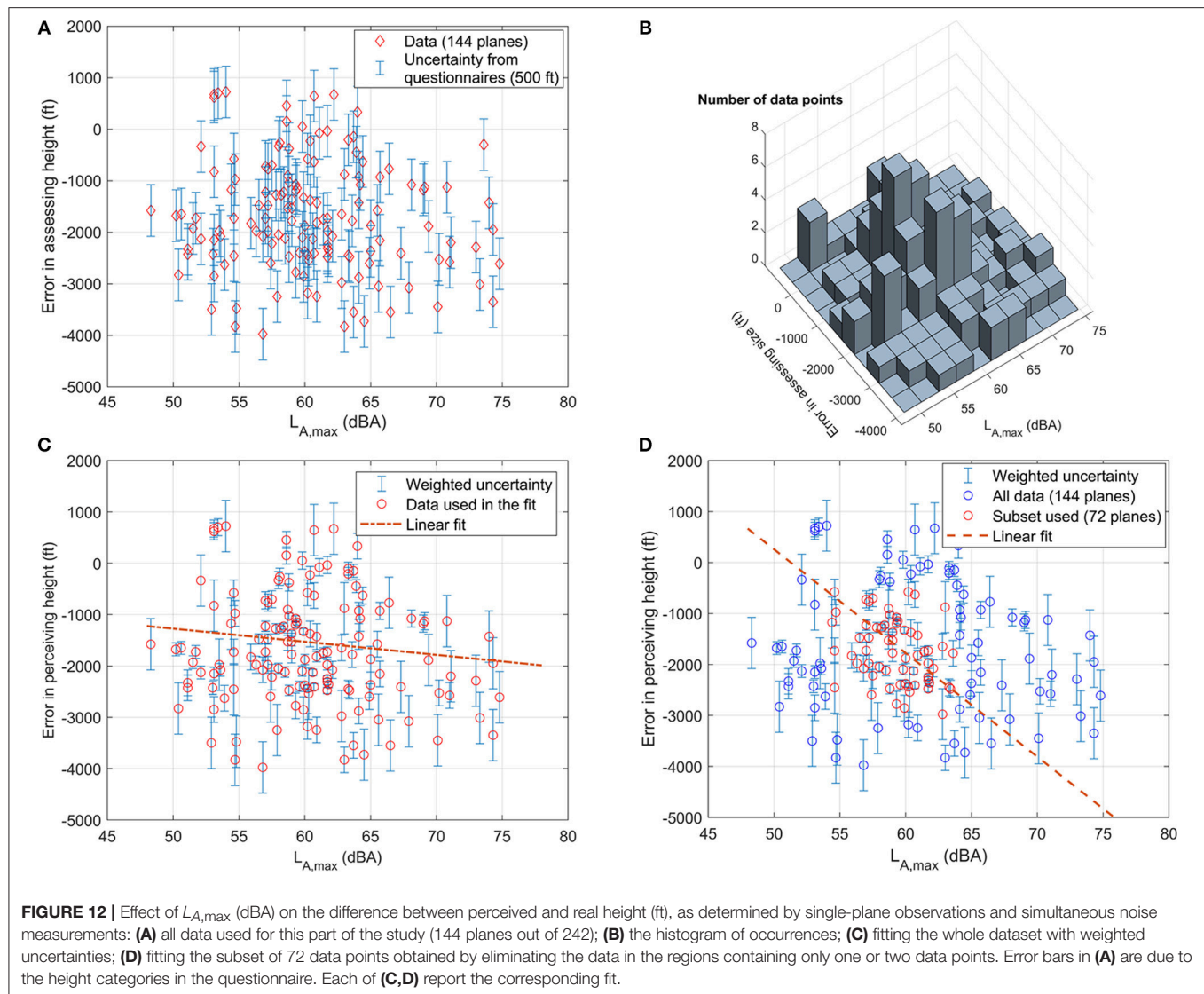
In this study, we could not detect any effect of $L_{A,max}$ on the ratio between perceived and actual size (EAS): as shown in **Figure 13A**, EAS does not depend on $L_{A,max}$ (i.e., it stays constant for different values of $L_{A,max}$). This conclusion remained similar ($r = -0.061, p = 0.47$) even when the subset selection method was applied: as shown in **Figure 13B**, most of the data clearly align with a horizontal line. Since there is an effect of plane peak emission on perceived height, but not on size, it is reasonable to think that there is no correlation between perceived size and height. This result, if confirmed by a larger sample, may give a negative answer to the question "is there a correlation between aircraft size and height perception?".

## OVERALL DISCUSSION

In the previous sections, we have presented the results of testing our method on a selection of 4 survey areas around Gatwick airport. For what concerns the qualitative and the quantitative assessment of height and size, both the postal survey and field survey gave the same result: the main advantage of running two types of survey simultaneously was in reinforcing the confidence in the overall message, even with a limited sample. This consideration is valid in general for studies involving multiple types of social surveys (Bartels et al., 2015; Hiroe et al., 2017).

In some cases, however, our distinct types of survey disagreed: this offered different points of view on the same population and may help inferring the mechanisms

---

[8]i.e. assuming a decrease in $L_{A,max}$ of 6 dBA for each doubling of the distance, which was confirmed by the measurements.

**FIGURE 12** | Effect of $L_{A,max}$ (dBA) on the difference between perceived and real height (ft), as determined by single-plane observations and simultaneous noise measurements: **(A)** all data used for this part of the study (144 planes out of 242); **(B)** the histogram of occurrences; **(C)** fitting the whole dataset with weighted uncertainties; **(D)** fitting the subset of 72 data points obtained by eliminating the data in the regions containing only one or two data points. Error bars in **(A)** are due to the height categories in the questionnaire. Each of **(C,D)** report the corresponding fit.
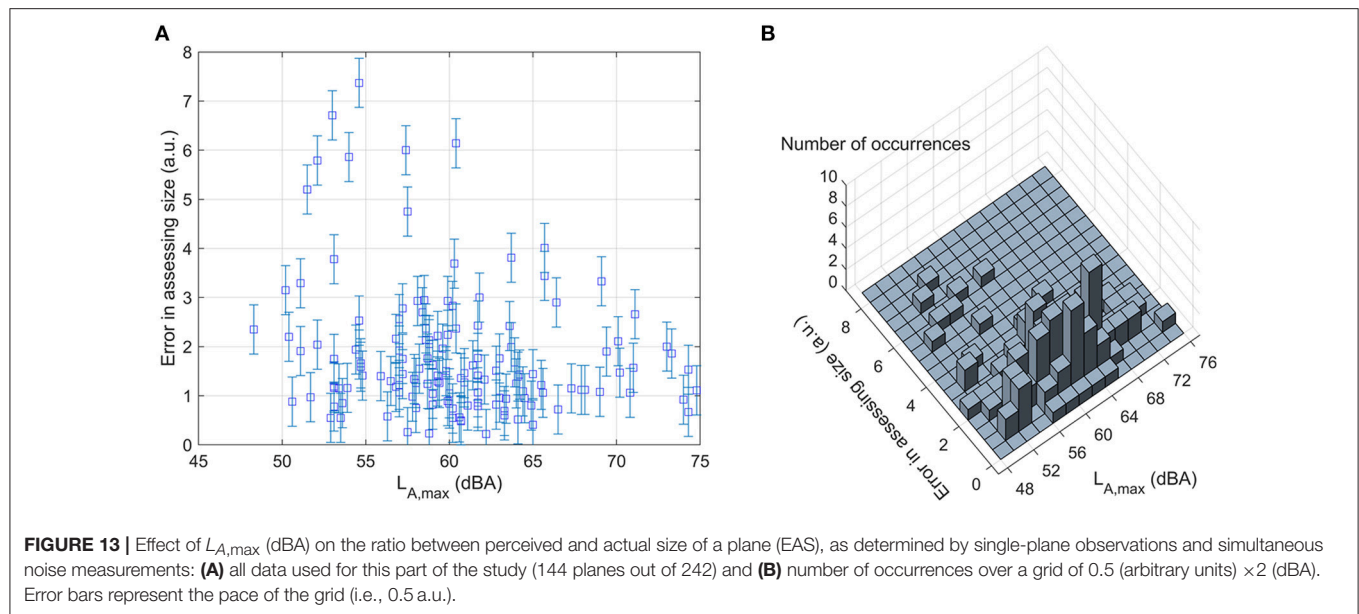
underpinning perception in the sampled residents (e.g., whether a perceptual judgement is due to short-term or long-term memories). In our method we put in place a control mechanism to investigate these cases, where postal respondents could volunteer to be interviewed too, but the number of volunteers was eventually very limited (13 over 112). Future studies will need a mechanism to maximize this control sample.

Our proposed method includes 15-min interviews: an absolute minimum in the literature of face-to-face surveys–e.g., (The HYENA Consortium, 2009; Schreckenberg et al., 2016; Civil Aviation Authority, 2017; Hiroe et al., 2017). This choice is extremely convenient and was welcomed by the participants, who only interacted with the researcher for a limited amount of time, but the planned duration was an informed guess, based on previous field studies (e.g., Memoli et al., 2008). A proper psychoacoustic analysis will be needed, before the interview time can be optimized.

In this study, the participants were extremely good at determining where most of the planes should be in the sky (i.e., as the "average" plane was easily identified with the "most frequent" one) but underestimated significantly aircraft height from the ground. We identified a relationship between the noise produced by a plane and this perception error, but nothing similar could be found on the size. We also highlighted a special role of the "lower" planes in the memory of the residents, suggesting a prevalent role of the outliers in perception-based judgement, Our conclusions, however, are limited in their significance by the size of the respondents/interviewees samples: even if the demographics is similar to the local census, future studies will need to be benchmarked on much larger samples.

Our findings support the second of our initial research hypotheses: in absence of clear references, when it can be very difficult to evaluate the absolute height and size of planes passing by, our brain counts on cross-modal interactions between

**FIGURE 13** | Effect of $L_{A,max}$ (dBA) on the ratio between perceived and actual size of a plane (EAS), as determined by single-plane observations and simultaneous noise measurements: **(A)** all data used for this part of the study (144 planes out of 242) and **(B)** number of occurrences over a grid of 0.5 (arbitrary units) ×2 (dBA). Error bars represent the pace of the grid (i.e., 0.5 a.u.).

audio and visual stimuli, leading to potentially erroneous judgements on height. The fact that we could not find a correlation between size and height, however, goes against the first hypothesis (i.e., that planes were perceived to be lower because of fleet changes). This may suggest that our brain prefers auditory stimuli to additional visual cues, not only in signal detection (Frassinetti et al., 2002), but also in assessing planes. Specific experiments may be needed for a definite conclusion.

## CONCLUSIONS

In this study, we designed a survey method to assess two specific non-acoustical factors in the soundscape perception of residents under the routes of arriving aircraft: the height and the size of arriving planes. The hypothesis of a multisensory interaction between visual and acoustical factors led us to complement existing standardized surveys with specific questions. To our best knowledge, this approach, used in the past for soundscape assessment, has not been applied to aircrafts before.

The ambiguity on whether height effects on perception were due to long-term memory or short-term judgements, and the desire to maximize the involvement of residents, led us to design two different interaction modalities, to be run in parallel: a 40-min long questionnaire and a 15-min interview. The first, delivered by post, was designed to be completed by the participants unassisted, presumably indoors. The second was designed to be run with a researcher, who would recruit the participant either on their doorstep or in a local park, for interviewing him/her outdoors. Interviews also included a component of "plane spotting," which was used to assess perceptual judgements "there and then."

Our "double-survey" method, assisted by acoustic measurements and aircraft tracking, was tested in 4 locations around Gatwick airport in the summer of 2017, involving a total of ∼200 participants.

When the two surveys arrived at a similar result, the outcome message was reinforced. In this way, we found evidence that:

a. Participants living below arriving aircraft could correctly describe the "average plane" i.e., the most frequent aircraft in their area.
b. Qualitatively, participants were very good at accurately perceiving how a passing aircraft was different from the "average plane": in height, size, and distance from where they lived.
c. Quantitatively, most participants underestimated the height of a specific aircraft—including the "average" one—by between 1,200 and 1,500 ft and overestimated its size by as much as twice.
d. For the same height, louder planes are perceived as lower, but not as larger.
e. Planes which are different from the "average plane" (i.e., "outliers") are the ones affecting perception, generating annoyance.

These observations, if confirmed in other studies or with a larger sample, may underpin the differences between the perception of arriving aircraft and the annoyance judgements on other sources of noise (i.e., unwanted sounds). Assessing the visual variations in the height of arriving planes may become one of the key non-acoustical factors in surveys oriented to arriving aircraft. The fact that outliers seem to play a key role in the perception of overflown residents, even more than the absolute height of the "most frequent plane," may have a significant impact on aircraft movement strategies in the future.

The fact that the two parallel surveys captured the impressions of two complementary parts of the population, if confirmed in other studies, may affect the way we determine community perception in the future: running two types of samples, supported by measurements, may become the new standard.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Data Protection Act 1998 and the General Data Protection Regulation (GDPR). The protocol was approved on 23/08/2017 by the Sciences & Technology Cross-Schools Research Ethics Committee at the University of Sussex, under project reference number ER/GM330/1. All subjects gave written or recorded informed consent, in accordance with the Declaration of Helsinki.

## AUTHOR CONTRIBUTIONS

GM led the study, conducted data analysis and wrote the paper. GH-F participated to the survey design and to the field survey. SM chaired the steering committee at Gatwick: he designed the scoping document and provided information on the context.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02492/full#supplementary-material

## REFERENCES

Abe, K., Ozawa, K., Suzuki, Y., and Sone, T. (2006). Comparison of the effects of verbal versus visual information about sound sources on the perception of environmental sounds. *Acta Acustica* 92, 51–60. Available online at: https://www.ingentaconnect.com/content/dav/aaua/2006/00000092/00000001/art00007

Arras, F., Massacci, G., and Pittaluga, P. (2003). *Soundscape Perception in Cagliari Italy.* Naples: Proceedings of Euronoise.

Avery, G. (1982). Comparison of telephone complaints and survey measures of noise annoyance. *J. Sound Vib.* 82, 215–225. doi: 10.1016/0022-460X(82)90530-2

Bartels, S., Márki, F., and Müller, U. (2015). The influence of acoustical and non-acoustical factors on short-term annoyance due to aircraft noise in the field—The COSMA study. *Sci. Tot.Environ.* 538, 834–843. doi: 10.1016/j.scitotenv.2015.08.064

Bregman, A. (1990). *Auditory Scene Analysis: The perceptual Organization of Sound.* Available online at: https://www.researchgate.net/publication200045100_Auditory_Scene_Analysis_The_Perceptual_Organization_of_Sound (Accessed October 2017).

Breugelmans et al, 2017 Breugelmans, O., Houthuijs, D., van Poll, R., Hajema, K., Hogenhuis, R., et al. (2017). *Predicting Aircraft Noise Annoyance: Exploring Noise Metrics Other Than Lden.* Zurich.

Brink, M., Schreckenberg, D., Vienneau, D., Cajochen, C., Wunderli, J. M., Probst-Hensch, N., et al. (2016). Effects of scale, question location, order of response alternatives, and season on self-reported noise annoyance using ICBEN scales: a field experiment. *Int. J. Environ. Res. Public Health* 13:E1163. doi: 10.3390/ijerph13111163

Casper, B. V. (2017). *CASPER - Live Aircraft Tracking.* Available online at: http://flighttracking.casper.aero/lgw/ (Accessed October 2018).

Civil Aviation Authority (2017). *CAP 1506: Survey of Noise Attitudes 2014: Aircraft.* Available online at: https://publicapps.caa.co.uk/docs/33/CAP%201506%20FEB17.pdf (Accessed October 2017).

De Coensel, B., and Botteldooren, D. (2008). *Modeling Auditory Attention Focusing in Multisource Environments.* Paris, France. doi: 10.1121/1.2934070

Environmental Research and Consultancy Department-Civil Aviation Authority (2017). *Noise Exposure Contours for Gatwick Airport 2015,* London: Department for Transport.

European Commission (2011). *Opinion 15/2011 on the Definition of Consent.* [Online] Available online at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf (Accessed October 2017).

European Environmental Agency (2010). *Good Practice Guide on Noise Exposure and Potential Health Effects.* Copenhagen: EAA.

Fields, J. M., De Jong, R. G., Gjestland, T., Flindell, I. H., Job, R. F. S., Kurra, S., et al. (2001). Standardized general-purpose noise reaction questionnaire for community noise surveys: research and recommendation. *Int. J. Sound Vib.* 242, 641–679. doi: 10.1006/jsvi.2000.3384

Filipan, K., Boes, M., De Coensel, B., Lavandier, C., Delaitre, P., Domitrović, H., et al. (2017). The personal viewpoint on the meaning of tranquility affects the appraisal of the urban park soundscape. *Appl. Sci.* 7, 1–16. doi: 10.3390/app7010091

Frankfort-Nachmias, C., Nachmias, D., and DeWaard, J. (2015). *Research Methods in the Social Sciences.* New York, NY: Macmillan Education.

Frassinetti, F., Bolognini, N., and Ládavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.* 147, 332–343. doi: 10.1007/s00221-002-1262-y

Guski, R. (1999). Personal and social variables as co-determinants of noise annoyance. *Noise Health* 1, 45–56.

Guski, R. (2017). "The increase of aircraft noise annoyance in communities. Causes and consequences." in *Proceedings of 12th ICBEN Congress on Noise as a Public Health Problem* (Zurich).

Guski, R., Schreckenberg, D., and Schuemer, R. (2017). WHO environmental noise guidelines for the european region: a systematic review on environmental noise and annoyance. *Int. J. Environ. Res. Public Health* 14:E1539. doi: 10.3390/ijerph14121539

Helios (2016). *Report for the Implementation of the Imm-05 Recommendation: CDA to 7,000 ft*, Gatwick Airport.

Hershenson, M. (1989a). "Moon illusion as anomaly," in *The Moon Illusion.* ed M. Hershenson (Hillsdale, NJ; Hove; London: Taylor and Francis).

Hershenson, M. (1989b). "That most puzzling illusion," in *The Moon Illusion,* ed M. Hershenson.Taylor and Francis.

Hiroe, M., Makino, K., Ogata, S., and Suzuki, S. (2017). "A questionnaire survey on health effects of aircraft noise for residents living in the vicinity of narita international airport: the results of physical and mental health effects," in *Proceedings of the 12th ICBEN Congress on Noise as a Public Health Problem* (Zurich).

ISO 1996-2 (2017). *Acoustics – Description, Measurement and Assessment of Environmental Noise – Part 2: Determination of Sound Pressure Levels.* Geneva: International Organization for Standardization.

ISO/TS 15666:2003 (2017). *ISO/TS 15666:2003 Acoustics – Assessment of Noise Annoyance by Means of Social and Socio-acoustic Surveys.* Available online at: https://www.iso.org/standard/28630.html.

Janssen, S., Vos, H., van Kempen, E. E., Breugelmans, O. R., Miedema, H. M., et al. (2011). Trends in aircraft noise annoyance: the role of study and sample characteristics. *J. Acoust. Soc. Am.* 129, 1953–1962. doi: 10.1121/1.3533739

Liu, Y., Nelson, P. I., and Yang, S.-S. (2012). An omnibus lack of fit test in logistic regression with sparse data. *Stat. Methods Appl.* 21, 437–452. doi: 10.1007/s10260-012-0197-0

Market Research Society (2014). *MRS Code of Conduct.* Available at: https://www.mrs.org.uk/standards/code_of_conduct.

Maziul, M., Job, R., and Vogt, J. (2005). Complaint data as an index of annoyance-theoretical and methodological issues. *Noise Health* 7, 17–27. doi: 10.4103/1463-1741.31628

Memoli, G., Bloomfield, A., and Dixon, M. (2008). Soundscape characterization in selected areas of central london," in *Proceedings of Acoustics'08* (Paris).

Memoli, G., Hamilton-Fletcher, G., Ceraudo, F., Ratkevicius, L., and Brookes, M. (2018). *Imm-15: Perception of Aircraft Height and Noise.* Brighton: University of Sussex.

Memoli, G., and Licitra, G. (2012). "From noise to annoyance mapping." in *Noise Mapping in the EU: Models and Procedures* (Boca Raton, FL: CRC Press), 371.

Miller, A. (2002). *Subset Selection in Regression. 2nd Edn.* Boca Raton, FL: ChapmanandHall/CRC.

MVA-Consultancy (2007). *Attitudes to noise from aviation sources in England (ANASE): Final Report for Department for Transport.* Available Online at: http://webarchive.nationalarchives.gov.uk/20081022220159/http://www.dft.gov.uk/pgr/aviation/environmentalissues/Anase/.

Nillson, M., Botteldooren, D., and De Coensel, B. (2007). *Acoustic Indicators of Soundscape Quality and Noise Annoyance in Outdoor Urban Areas.*

Office for National Statistics (2011). *CENSUS 2011.* Available Online at: https://www.ons.gov.uk/census/2011census.

Payne, S., Davies, W., and Adams, M. (2009). *DEFRA: Research into the Practical and Policy Applications of Soundscape Concepts and Techniques in Urban Areas (NANR 200).* Available Online at: randd.defra.gov.uk/Document.aspx?Document=NO0217_8424_FRP.pdf

Pheasant, R., Horoshenkov, K., Watts, G., and Barrett, B. (2007). 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

Posner, M., Nissen, M., and Klein, R. (1976). Visual dominance: an information-processing account of its origins and significance. *Psychol. Rev.* 83, 157–171. doi: 10.1037/0033-295X.83.2.157

Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. (1992). "15.2 Fitting Data to a Straight Line." in *Numerical Recipes in C-The Art of Scientific Computing.* Cambridge: Press Syndicate of the University of Cambridge, 661–666.

Redeborn, B., and Lake, G. (2016). *Gatwick Airport Independent Arrivals Review: Report and Recommendations.* Available Online at: https://www.eastsussex.gov.uk/media/4761/independent-arrivals-review-report-jan2016.pdf.

Schreckenberg, D., Belke, C., Faulbaum, F., Guski, R., Mohler, U., Spilski, J., et al. (2016). Effects of aircraft noise on annoyance and sleep disturbances before and after expansion frankfurt airport - results of the NORAH study, WP1 'annoyance and quality of life'," in *Proceedings of Internoise'16* (Hamburg).

Schreckenberg, D., Belke, C., Spilski, J., and Guski, R. (2017). *First Results of the Development of a Multiple-item Annoyance Scale (MIAS).* Zurich.

Swearer, J. (2011). "Visual angle," in *Encyclopedia of Clinical Neuropsychology.* New York, NY: Springer. 2626–2627.

The HYENA Consortium (2009). Annoyance due to aircraft noise has increased over the years - Results of the HYENA study. *Environ. Int.* 35, 1169–1176. doi: 10.1016/j.envint.2009.07.012

Truax, B. (2001). *Acoustic Communication.* Westport, CT: Ablex Publishing.

Van Wiechen, C., Houthuijs, D., and Heisterkamp, S. (2003). *Complaining About Aircraft Noise.* Rotterdam: State-of-the-art around Amsterdam Airport Schiphol.

WHO (2009). *Night Noise Guidelines for Europe.* Available Online at: http://www.euro.who.int/__data/assets/pdf_file/0017/43316/E92845.pdf.

Check for
updates

# Use of Creative Writing to Develop a Semantic Differential Tool for Assessing Soundscapes

David Welch[1]*, Daniel Shepherd[2], Kim Dirks[1], Mei Yen Tan[1,3] and Gavin Coad[1]

[1] School of Population Health, The University of Auckland, Auckland, New Zealand, [2] Department of Psychology, Auckland University of Technology, Auckland, New Zealand, [3] Bay Audiology Ltd., Auckland, New Zealand

Exploring our understanding of soundscapes to understand why and how sound impacts people is important. The aim of this study was to develop a short quantitative questionnaire that would use terms generated by creative writers to assess people's experiences of a soundscape. This process may provide different items for the questionnaire and thus, potentially, different dimensions or fuller definitions of dimensions that have already been identified. In the preliminary phase, a group of people identifying themselves as good writers listened to recordings of natural, traffic, and human sound environments and wrote about their impressions and responses to each. Qualitative analysis was used to extract themes from the writing. These themes were identified by key words, and scalar items were developed to form a short 17-item questionnaire. The questionnaire was administered to 228 people in Auckland City, New Zealand, with participants recruited from city streets and in a central-city park. Respondents were comfortable to use the questionnaire. Factor analysis revealed patterns of responding with five dimensions: Calming, Protecting, Hectic, Belonging, and Stability. There were correspondences between these and others previously reported in the literature, as well as differences. The use of items derived from creative writing provided interesting insights into the soundscape, including spirituality, the sense of time passing, and physical wellbeing. The park soundscape was measurably better than the street soundscapes on all dimensions, and streets with less vehicular traffic tended to be experienced as more Calming and Protecting, and less Hectic. This implies that there is validity in the scales generated. In future, it would be valuable to test the questionnaire in more varied environments, to add greater variability to the soundscapes.

Keywords: soundscape, questionnaire, qualitative methodology, quantitative methodology, psychometric

## INTRODUCTION

Sound has been shown to impact on people's physical and mental health (Basner et al., 2014), as has the loss of the access to sound in severe or profound hearing loss (Guitar et al., 2013). Nonetheless, many people appear to lack awareness about the importance of sound and hearing so that troubling noise is widespread in our society (Welch et al., 2013), and noise-limiting or hearing-health programs are fraught with difficulty (e.g., Reddy et al., 2012). The concept of the

soundscape may be a useful way to understand and thus communicate with other people about the effects of perceived sound in order to improve our societies' sound environments (Schafer, 1977; Andringa et al., 2013; Steele et al., 2015). The aim of this study was to improve our understanding of it and to add to the development of an instrument to measure it quickly and effectively.

One way of accessing a person's representation of the world based on their sensory experience is through language (Raimbault, 2006), though other approaches [e.g., comparison to music (Botteldooren et al., 2006)] have been considered. Qualitative approaches have the capacity to delve deeply into people's narratives for meaningful descriptions of what they perceive. Compared to quantitative data, qualitative data are rich. However, it can be more difficult to make comparisons between qualitative measures and can be harder to obtain quick and accurate responses at a population level, especially when seeking responses from less educated or literate people. Another issue with qualitative descriptors of soundscapes is that they may be limited by a person's vocabulary and ability to express themselves using language. Given that most people understand more words than they will actually use (Laufer, 1998), providing people with a set of descriptors which can be rated may allow them to report on experiences for which their active vocabulary would be insufficient but for which their passive vocabulary compensates.

A tool that has been used in the context of a soundscape is the semantic differential scale (Osgood et al., 1957; Kang and Zhang, 2010; Cain et al., 2013). The approach takes the form of a set of adjectives, and requires the respondent to select a number between two poles of a continuum (e.g., pleasant vs. unpleasant). An advantage of this method is that the same subjective attributes may be compared between different locations quantitatively. Some terms are described as 'denotative,' or referring to aspects of the sounds being experienced (e.g., fast/slow); and others are described as 'connotative,' or referring to a person's response to the sounds (e.g., calming/agitating). Parallel terms, more suited to soundscape research, have been used, with 'descriptive' for denotative and 'affective' for connotative. The challenge is to find terms that are easy to understand but which also allow a respondent to express the subtleties of their experience of a soundscape (Raimbault, 2006).

The semantic differential method presents opposing soundscape descriptors on a scale which is considered to be unidimensional (Osgood et al., 1957). On the other hand, previous research into the cognitive representations of soundscapes and their descriptors suggests that there may be heterogeneity in the interpretation of the lexical items used and thus of the determinants of respondents' choices (Dubois et al., 2006). In other words, while the semantic differential may be a useful quantitative method for the analysis of experiential factors, it must be borne in mind that it cannot represent an objective or absolute measurement of a soundscape attribute. A superficial appearance of consensus which may occur is that respondents will use a set of terms presented to them, but there may be variation in the meaning of those terms for each person. The process of developing the semantic markers is thus crucial in providing respondents with acceptable and clear responses,

and the introduction of different markers may potentially provide the opportunity to present new ways of perceiving the world.

Factor analysis, and the related principle components analysis, has been used extensively with semantic differential scales in the soundscape literature (e.g., Kang and Zhang, 2010). Factor Analysis combines a statistical approach with subjective judgment (Tabachnick and Fidell, 1996). It aims to simplify people's responses to many semantic differential scales by identifying the underlying perceptual/emotional dimensions (factors) that influence the original responses. To do this, it measures correlations between responses to different items, and where several items correlate to a reasonably high degree, a factor is generated. The subjective exercise is the 'naming' of factors based on aspects of the items contributing most to each. The naming exercise relies heavily on the choice of word made in forming the semantic differential scales. Furthermore, in reducing several items to one name that encompasses all of their meanings, it relies on the minds and vocabularies of the researchers to capture the commonality appropriately. Factor analysis cannot, of course, look outside the set of original items and the responses to them so it searches for correspondences within a closed set and cannot be used to comment on the extent to which a particular approach has captured the true variance in people's thoughts.

As such, the process of deciding upon the original set of items is crucial and a range of approaches has been used. For example, one approach has been review of the literature on sound descriptors and rendering down of a larger list into twelve items researchers perceived to be most appropriate for the task along with pilot testing (Berglund and Nilsson, 2006). Another approach used a list of 116 items that were based originally on terms extracted from interviews about photographs, with reference to sound-relevant terms and consensus from a group of experienced listeners (Axelsson et al., 2010). Others have used a combination of terms derived from literature and items decided upon by the researchers as relevant to the environment being studied (Kang and Zhang, 2010). These approaches are well-considered and have generated quite similar sets of items, each of which has face-validity as a potential descriptor of a soundscape. A class of approach that has been applied to gather data relevant to the soundscape is interviewing with qualitative analyses (Liu and Kang, 2016), but it has not been reported as a preliminary stage in the consideration of items for semantic differential scales. An advantage of using such an approach would be that a reduced battery of questionnaire items could be used. These items would be based on the themes identified in the qualitative research, and would therefore provide a good structure for the soundscape while also reducing the length of the questionnaire.

On the basis of this, it is desirable to establish a set of dimensions that people tend to use generally when making judgments about a soundscape. Some progress has been made in this direction (Davies and Murphy, 2012). Furthermore, a theoretical basis, rooted in evolutionary psychology, has been proposed to explain why these dimensions might be common for people across cultures (Andringa and Lanser, 2013; van

den Bosch et al., 2018). Essentially, this theory suggests that the environment might be perceived in terms of two factors: whether it is pleasant for the organism, and whether much is happening. The two concepts may be regarded as orthogonal in that an environment can be rich (pleasant and eventful), dangerous (unpleasant and eventful), calming (pleasant and uneventful), or boring (unpleasant and uneventful). Two dimensions, Pleasantness (emotional valence) and Eventfulness (vibrancy), have been identified in several soundscape studies (De Coensel and Botteldooren, 2006), and it has been suggested that these might be seen as two basic dimensions of soundscapes (Davies et al., 2013; Aletta et al., 2016). These dimensions may be seen to reflect the basic dimensions of human mood, as expressed in earlier research (Russell, 1980).

The dimensionality of the soundscape may be more complex, and has varied across studies. There are many possible reasons for this, including differences in the sound environments and the methods used to collect responses. A four-dimensional model derived from factor analysis: "Relaxation," "Communication," "Spatiality," and "Dynamics" has been developed to account for urban soundscapes (Kang and Zhang, 2010). Research using only affective (i.e., connotative) semantic differential attributes (e.g., "pleasant" and "calm") and not descriptive (e.g., "loud" and "sharp") found three components: "Pleasantness," "Eventfulness," and "Familiarity" (Axelsson et al., 2010). In other research, two principal components: "Calmness" and "Vibrancy," which may be seen to parallel Pleasantness and Eventfulness, were identified (Cain et al., 2013). Other work has identified the concept of "Restorativeness," the sense that a soundscape helps people to recover from tiredness or malaise (Payne, 2013). Furthermore, the concept of "Appropriateness" (a sense that the soundscape is right for the place in which it is experienced) has been considered as an aspect of sound environments which should be considered in terms of soundscapes (Axelsson, 2015). Each of these dimensions has been shown to have some reliability, and yet they vary and differ between studies. The differences may arise partly due to variations in the sound environments or stimuli used in different studies, but they may also depend upon observers' ability to express their perceptual experiences. The more varied the response options that can be provided, the more detail may be understood about human soundscapes.

It is likely that there is commonality in the human experience of soundscapes (Brown et al., 2011), so it may be possible to generate a short and quantitative measure to capture this. A key issue is the need for a good set of terms to allow people's responses to the soundscape to be captured, since if a concept is missing, there will be no way to detect its absence. Our approach had three main stages:

(1) We asked people with high active vocabularies and an interest in written expression to write about their responses to three different sound environments.
(2) We analyzed these writings for themes that were present in the responses.
(3) We used these themes in a short questionnaire that we administered to a small sample of people in Auckland City.

The rationale for choosing literate people was based on the principle described above that people with limited active vocabularies will typically have larger passive vocabularies. Since people may be induced to draw upon their passive vocabulary when prompted, and a semantic differential questionnaire is essentially a set of prompts, the approach seemed reasonable. We ran exploratory analyses on the results to see whether the approach had produced potentially useful data. In particular, we were interested to see whether members of the general public could use the questionnaire to describe their perception of the sound environment and their responses to it quickly and easily.

## PHASE 1: QUALITATIVE STUDY

There were two phases to the research. Phase 1 involved recruiting literate people with an interest in descriptive writing and/or sounds. These participants wrote about their perceptions and responses to three different sound environments, and their writing was analyzed using thematic analysis (Braun and Clarke, 2006). A questionnaire was developed based on the themes identified. Phase 2 was a piloting of the questionnaire in a sample of people in real sound environments in Auckland City, New Zealand. The research was approved by the University of Auckland Human Participants Ethics Committee: Approval number 8150.

## Materials and Methods
### Participants
Twenty-five adult participants aged 20–38 years (Mean = 25.04, $SD = 4.71$) participated; 52% were male ($n = 13$). Recruitment was through advertisement in the form of posters, electronic flyers, and social media. It was desirable to attract participants who would be willing and able to provide rich written descriptions of their responses to different sound environments, so advertising was targeted to students in creative-writing courses at the university. All participants had hearing thresholds of better than 20 dBHL in their better ear for all tested frequencies.

### Procedure
Three recordings were selected on the basis that they represented sound environments dominated by sounds of nature, humanity, and technology. These classes of environment have previously been shown to produce differences in the types of descriptor used for the soundscapes arising from them (e.g., Axelsson et al., 2010). They were purchased from a database of environmental sound recordings at www.shockwavesound.com. The three soundscape recordings were in 5.1 surround sound AC3 (Dolby digital) file format, and brief descriptions of each are as follows:

(1) Traffic: Road traffic noise recorded at a town junction. Cars, mopeds, motorcycles, and occasional buses accelerating past in all four directions with some distant voices.
(2) Human: Crowded pedestrian street in town. People walking by in all directions, distant sound of children playing.

(3) Nature: Light surf with small birds chirping and tweeting to the front and rear.

The original recordings were looped to extend the presentation duration using Audacity® 2.0.0. After this processing, the participants were presented with recordings of sound environments that lasted between 19 min 27 s and 26 min 41 s. The recordings were crossfaded over 3 s to avoid sudden changes. The presentation order was randomized across participants, and soundscape assessment conducted in a sound attenuating chamber 2.21 m wide and 2.48 m long.

A Sony 6.1 surround speaker system consisting of left (L), centre (C), right (R), left surround (Ls), right surround (Rs), centre back (Cb) speakers and a subwoofer (Sub) was used. The speaker system was treated as a 5.1 surround system, and no input was received at the Cb speaker for the 5.1 soundscape recordings file format. All speakers were facing the listener and mounted on adjustable stands, with the exception of the subwoofer. The speakers were amplified with a Sony digital audio/video (Model STR-DG500 6.1 Channel) amplifier.

The 6.1 surround system was set up as follows:

(1) The C speaker was positioned straight ahead of the listener at 0° azimuth.
(2) The L and R speakers were positioned at each corner of the front of the booth, approximately 45° left and right, respectively, to the horizontal. The speakers were raised slightly above ear level.
(3) The Ls and Rs speakers were positioned at each corner of the back of the booth, approximately 45° left and right, respectively, to the midline. The speakers were aligned at ear level.
(4) The non-functioning Cb speaker was positioned directly behind the listener.
(5) The Subwoofer was positioned at the front between speakers C and L.

A calibration spot approximately 150 cm from each of the L, R, Ls, and Rs speakers to the middle of the room was marked with masking tape. A comfortable chair on which participants were seated was positioned over the calibrated spot, and a large table was situated in front of the chair where the amplifier and a laptop were placed.

Sound recordings were delivered through the surround sound speakers using VLC media player on a Macbook. The coupling of the laptop with the amplifier was carried out with a Creative Labs Sound Blaster THX® TruStudio Pro external USB soundcard and an optical audio cable.

Output levels of the three sound environment recordings were calibrated using a Brüel and Kjær Hand-Held Analyzer Sound Level Meter (Type 2250) with a $1/2$ inch microphone. The sound level meter was mounted on a Manfrotto 804RC2 tripod at participants' ear level when seated over the calibrated spot.

The average sound pressure level (SPL) of traffic sounds was set to 75 dBA (LAeq, 4 min). We based this on an estimate which indicated that the average SPL of traffic noise taken from major Australian cities ranges between 55 and 75 dB (Austroad Facts, 2000). The upper limit of this range was taken because the

intersection was very busy and this level sounded appropriate to the researchers. On the same basis of the researchers' subjective experience of the sound (what "sounded right"), the average SPL of human sounds was set to 65 dBA (LAeq, 4 min), and the average SPL of nature sounds was set to 55 dBA (LAeq, 4 min).

Each participant was seated and briefed about the context of the sound environments before commencement of each recording. While listening to each recording, participants were instructed to write about their soundscape experience. Participants were given the option of manually writing their responses with pen and paper or typing on a laptop, but all preferred the latter. A blank Microsoft Word document was created headed with an open-ended question:

*"Please describe the soundscape you have just heard, and the feelings, emotions, and impressions it may have evoked in you (for example, positive or negative reactions you may have)"*

Participants were instructed to write as freely as possible in response to the question. For each of the soundscape recordings, participants were informed that the minimum writing time was 8 min. However, they were encouraged to write as much as they could, and allowed as long as they required. A count-up timer was set up in the top right-hand corner of the laptop screen to notify participants when 8 min had passed.

During the experiment, the researcher waited outside the booth in order not to interfere with the soundscape experience and to preserve the anonymity of participants' writings. Participants were asked to leave mobile phones outside the booth. The lights of the sound-proofed booth were dimmed during the experiment.

## Qualitative Analyses

Each participant wrote in response to each of the three sound recordings. Participants' subjective writings in response to the open-ended question were analyzed using NVivo Software. A thematic analysis of the writings was conducted, and a set of themes and concepts within the data was identified. These categories were organized in a hierarchical manner, illustrating the emergence of more specific themes from general concepts. Coding was conducted by authors MT and DW, who worked both independently and together in order to propose and clarify themes, and achieve consensus.

A thematic analysis approach was used (Braun and Clarke, 2006). The coders read the writing and described themes that they felt underlay each passage. A 'passage' is not clearly defined, but is described by the coder in the process of analysis according to their understanding of meaning, and quoted as appropriate (see below). Furthermore, a given passage may potentially be coded as expressing multiple themes, and a theme can be expressed many times or just once: the frequency is not relevant since no sampling frame or specific, *a priori* definitions are used. The analysis seeks to discover a hierarchical structure whereby the themes expressed can be described. The hierarchy is a system of general themes and subthemes that allows the coders to perceive a pattern and to extract elements of meaning. It is thus a subjective approach, and uses the coder's mind as the lens for understanding the themes underlying what is written. It is possible to approach the data with

a pre-conceived theory, and look for themes that are relevant to the components of that theory. In this case though, themes were allowed to emerge from the data without an explicit theoretical stance. However, both coders were aware of themes/descriptors that had been used in previous research into the soundscape, so this may have influenced our thinking. Themes were labeled based on what the coders believed to be the soundscape feature that underlay the writing.

More specifically, at the highest level, we classified responses into those which were descriptive of qualities of the sound environments, and those which were relevant to the response generated in the person while experiencing the soundscapes. In the qualities of the sound environment were responses that reflected: temporal qualities of the sound, which contained sub-themes related to (1) pace (leisurely versus fast) and (2) patterning (with concepts like rhythm or predictable patterns versus irregular or unpredictable sounds); (3) the overall level of the sounds; (4) the extent to which the sound environment was described as clear versus blurred or disorderly; (5) the complexity of the sound environment; (6) the spatial qualities, including sub-themes relating to vastness as opposed to congestion; (7) the sense of tonality or harmony versus discordancy or harshness; and (8) the stability as opposed to variability of the sound environment. The responses to the soundscape were classified into three general areas: health, physical responses, and responses of the *psyche*. In this latter category, we drew on its usage in reference to cognition, as well as the concept of the spirit or soul. Health responses included themes relating to (9) wellbeing, with ideas like wholesomeness versus a sense of affliction; (10) stimulation or arousal versus hypnosis; and (11) stress, including distress and anxiety versus a sense of relaxation. Physical responses included themes relating to (12) safety versus feeling threatened and fearful; and (13) comfort including ideas like contentedness versus having a desire to escape. The responses of the psyche were divided into those which were either cognitive or soulful. Within the cognitive set of themes were those related to (14) cognitive load or burden as opposed to a feeling of being refreshed; and the sense of (15) familiarity or usualness versus novelty. The soulful themes included feelings of (16) connection to the soundscape, and (17) a spiritual sense of being uplifted versus being oppressed. In these different themes, there were statements that supported positive and negative aspects, and the ideas that were expressed helped to develop anchor points for the scales generated from each theme. The numbering in the foregoing text are to allow the reader to see the eventual themes that emerged and were included as scales in the questionnaire; these are explained more fully with supporting quotes in the "Results" section below.

## Results

Themes fell into two general classes: themes relating to the perceptions of the sounds themselves, and themes about the feelings and impressions that were evoked by the soundscapes. The distinction was not always clear, but we presented the themes according to this. For example, the sound of an internal combustion engine presented at a high level may be perceived as loud, and this may make a person feel disturbed. In our analysis,

the component of the report, 'loud,' was treated as a report about the qualities of the sound and the component, 'disturbing,' was treated as a report about the person's deeper feelings and emotions. We acknowledge that 'loudness' is a perceptual quality, and may contain the sense of being disturbing.

### Qualities of the Sound

The impression of loudness was identified as a theme, especially in the exposure to traffic sounds:

> "The blood that runs in the city's veins is harsh and loud. . ."
> "Lots of loud noise, motor noises are not the most relaxing sounds- especially motorbikes. Constant noise- there may be quieter moments, but there is always background noise, and the quieter moments do not last long."

It can be seen from the second quote that the sense of loudness/quietness was, as might be expected, seen as a continuum from loud through quieter stimuli. This quantitative aspect seemed to be present generally and accorded with our use of bipolar scales in the design of the questionnaire.

A sense of pace, particularly speed was perceived in the sounds. Again, this was referred to as if it were a continuum, and manifested as a sense of the temporal combined with the emotional. In other words, it suggested that the soundscape included a sense of the passage of time and that this was intertwined with a need to act in a manner consistent with the temporal imperative. For example, urgency is conveyed in these quotes from the traffic and pedestrian sound environments, respectively:

> "The brakes stop abruptly, signifying that time is short, and nobody has time to spare in their busy schedule. Nobody has time to spare, everyone minding their own business."
> "A sense of urgency followed by a wave of panic fills the air. . . Everything is moving so fast in this town, like someone or something is coming."
> "Hustle. And. Bustle. Not in the good way. Someone get me out of here."

In the last quote, the implication is that speed, 'hustle and bustle,' is sometimes a pleasant thing but that it can also be unpleasant as in the case of the busy traffic. At the other end of this continuum was a sense of leisureliness and the gradual nature of processes in the sound environments associated with a change in the pace of time:

> "Time slows down to an almost standstill."
> "The water's course over the stone will erode it. The stone fades just as we do, just a little slower. One must appreciate the beauty that must all fade away. One feels, too, that the sea is hidden behind a verdant curtain; tall trees at the border of the garden perhaps, or simply thick growth on shorter flora. One catches glimpses of the fast passing of waves and the slow passing of stones through this curtain, just as one does of the world, of life."

The idea that the sound environment provides cues about the slow erosion of stone due to the action of water provides a compelling sense of how a sound environment may evoke the quality of slowness. And the relationship between the soundscape and time was not necessarily straightforward; variations in the

speed of different aspects of it could be part of the reason why a sound environment was pleasant:

*"These sounds of nature to me are so peaceful and different. They change but stay the same. The surf is always rushing but it's the intensity that changes. The birds tweet but the rhythm and speed changes."*

A sense of the clarity of sounds seems to reflect the apparent signal-to-noise ratio for interesting sounds in each environment. In some soundscapes, sounds were distinct with a clear source:

*"It is not difficult to separate the sounds of the ocean versus the bird calls, but if I close my eyes the sounds start to merge into an overall panorama of peaceful noise that is just very pleasant to listen to."*

Whereas in others they were not:

*"People's footsteps and voices are drowned out by the constant hum of traffic."*
*"I can hear many voices. A hairball of voices. A clogged pipe system of voices. An imperfect spaghetti bowl of vocal chords tied together and spiraling inefficiently. A sound, assassinated."*

Interestingly, in the first quote, the clarity is present but the participant was happy to allow the sounds to merge. In the second quote, the sense is that the human sounds are overwhelmed and lost in the noise from the traffic.

Part of the descriptions of the sound environments seemed to refer to their complexity or lack thereof, and neither was intrinsically good or bad; sometimes the complexity seemed a violent tangle:

*". . . there are more people speaking at once and several other background noises competing against each other for attention."*
*"There's too many things happening (like different people's conversation) and it gets distracting"*

On the other hand, for some, the lack of complexity in the natural soundscape could be seen negatively:

*"I feel as if I would be easily bored as there aren't many new sounds (just birds and waves crashing around)."*

There was an awareness of spaciousness associated with some sound environments:

*"There is also a sense of a large expanse of the ocean, the beach (perhaps) and because there are birds there would be places that they can fly off away to."*

While others conveyed a sense of crowding, proximity, and congestion:

*"Sounds busy and congested. Felt a bit tight and restricting at first, almost stressful initially."*
*"There is a distant clanging of cutlery, babies crying. . . everything that exists in a densely populated space."*

A tonality was perceived. In some sound environments it was harmonious:

*"The chorus the sea sings as the wind encourages its wave to crash. What other melody can compare to that?"*

Whereas others were discordant, jangling, or harsh:

*"There are a range of voices of different pitches that I can hear. The higher pitched voices – children and women – seem easier to pick out as they move around. But occasionally a man's voice stands out. Sounds such as babies crying are suddenly quite startling."*

A sound environment could be stable and unchanging or varied and changing. Stability did not seem to refer to the individual acoustic components, but rather that there was a constancy to the various components of the sound environment:

*"The surf is always rushing but it's the intensity that changes. The birds tweet but the rhythm and speed changes. You feel like you could sit for hours and never tire of hearing the same sounds over and over."*

A pattern was observed in some soundscapes which had predictability:

*"The ocean waves are rhythmic and predictable and quickly become part of a soothing background."*
*"I can hear a low thunderous rumble almost continuously in the background, which seems to stay at about the same volume throughout. At times this rumble seems to almost pulsate and feel sort of rhythmic."*

But others were irregular and unpredictable:

*"Sounds such as babies crying are suddenly quite startling and immediately noticeable, as are short claps"*

## Feelings and Emotions

Like the sound qualities above, the internal feelings that people expressed as resulting from the sound environments generally followed a pattern of having two poles with intermediate states.

Stimulation was experienced as a result of perceiving the soundscapes. At one end of the spectrum was the effect of arousing people. This could be pleasant, invigorating and exciting:

*"I enjoy myself. It's not every day I get to go to such a busy and exciting place. The clatter of shoes, the banter of people, the merchants having welcoming and, sometimes sly, smiles, it's to be an eventful afternoon."*

Or else the level of stimulation could seem too much:

*"I can sense urgency in the air. My heart is starting to race. [. . . ] Why can't I relax? I need to breathe."*

At the other pole from arousal was a sense of feeling soothed or calmed by the sound:

*"I like the sound of the ocean waves. The repetitive white noise has a kind of calming, hypnotic effect that could put me to sleep at night."*

There was a theme reflecting a perception of connection to the environment and the things in it:

*"But this is no kind of loneliness, for there is the connection with the greener beings."*
*"There is synchrony between the birds and waves. They sing to me with love. Each wave, though far, seems to lap playfully at my feet like a playing child, wanting me to come and join. Welcoming. Appreciating. I have nothing more to think. My body unwinds and settles into this natural rhythm."*

*"I also feel almost a sense of belonging – I am most familiar with the sounds of a busy inner city and foot traffic. I feel like I'm in my comfort zone and I know where I'm going."*

And on the other hand, the soundscapes could produce a sense of alienation:

*"I feel isolated from them. This is their everyday [. . . ] I feel invisible, lost, lonely even. It's as though they are alien to me."*
*"I feel squashed and I don't feel quashed. Just. . . removed. I am observing, remember, and I am watching people at the game of going."*
*"There is also a sense of isolation even though it sounds like there are people around me and I hear voices and people slamming car doors. I know that I am not alone because there are people driving the vehicles and people walking on the street – but it seems everyone in the scene is busy focusing on their own lives and their own actions and almost ignoring me. . . This makes me feel rather alone and isolated"*

The sound environments caused people to feel stressed:

*"It's only when I've stopped and listened to it that I realize how harsh and stress-inducing it is. Seems to be a tiring environment to be in – I'm really craving for some quiet time in a park or at home from all the noise."*

Or to feel relaxed

*"I feel at rest, worrisome thoughts I may once have had are long forgotten, and I pause to enjoy the sound of nature."*
*"It feels great to be listening to this. I can literally feel my body relax. . . my muscles being less tensed, my mind slowing down in thoughts. I feel like I just find somewhere to lie down and rest, maybe read a book, enjoy the breeze, hear the birds sing. Ahh. . . it'll be such a wonderful experience."*

A sense of familiarity is evoked by some soundscapes:

*"I feel that this is a very normal, everyday environment to be in."*
*"The whole thing is busy, bristling with noise and bustling with the familiar sounds of modern life."*

And this may either put people at ease:

*"I feel like I'm in my comfort zone and I know where I'm going."*

Or it may seem unpleasant, dull and boring:

*"The racket is almost unbearable. Although all too familiar."*
*"We are all following a pattern designed by something larger than ourselves, all moving, busy ants picking up a little lump of the bigger sugar pile, picking it up, carrying it and dropping it somewhere, only to be picked up and moved again by a fellow ant. I want out. This is not me."*

The sense of a Cognitive Load or burden placed on or removed from the mind by experiencing the soundscapes was suggesting. Sometimes this load was heavy and crushing:

*"Too many things, too many noises surround me it's hard to hear your own thoughts."*
*"I feel smothered by the constant hustle and bustle. Mind feels saturated with thought trying to take in everything but unable to hear my own thoughts."*

Whereas the removal of the load could be refreshing:

*"I can feel my mind coming alive, as if a blanket of responsibility that has been smothering me has been removed."*
*"There is nothing to think, nothing to clutter my mind with. I do not yearn to think either. I leave my thoughts on my bed and come out a free soul."*

Another theme was that of the sense of safety that was experienced while listening to some environments and the sense of danger or threat induced by others. The sense of safety was associated with the natural environment:

*"I feel a sense of control – I can move close to the birds or the waves and interact with it if I want to and only if I want to. Nothing in this environment is going to move in a way that may threaten my safety. I don't have to be on my guard the whole time."*
*"This makes me feel safe, I am not enclosed or locked up and I can control my actions and walk away if anything threatening occurs; there is an escape route. Also I feel safe because there are lots of bird calls. I guess this means that there is nothing overly threatening in my environment currently – if there was the birds would fly away or sound some bird alarm. They sound pretty contented and going about their lives so there must not be much to fear around."*

In contrast, the sound environment that was dominated by human pedestrians produced mixed responses with respect to people's feeling of safety. Sometimes safety was enhanced:

*"I feel almost slightly calm and almost like I'm waiting or walking at a leisurely pace. Nothing in this environment is threatening to me at all. . . Also if something bad occurs or something threatening occurs, I think my chances that people will come to my physical aid is high. I hear voices of both men and women of all ages – someone will be able to help me. Knowing that help is readily available also gives me a sense of peace."*

Or even a sense of it being so safe that it is dull:

*"Generally this soundscape seems mundane and everyday. Sounds I am familiar with and not threatened by. Not quite peaceful, but certainly not annoying."*

Whereas sometimes there could be a mixed view of threats and excitement:

*"On one hand, I feel anxious, I have associated crowds to danger and theft. I try my best to avoid them and the busyness of town while on the other hand, the busyness can become very exciting and positive, giving me a sense of adventure and disarray, a break from the boring routines I've inadvertently put in place in my life."*

And the same environment may be perceived as a threat:

*"I find this environment quite loud and feel a bit nervous as to what is happening. The voices do not appear friendly and I feel unsafe. The motorbikes especially evoke a sense of fear and I don't like them."*

On the other hand, the sound environment that was dominated by traffic was perceived only as threatening:

*"I don't feel safe at all – in fact I feel rather threatened. Passing traffic sounds are really close I'd much rather be a bit further away from them. . . the environment seems threatening and dangerous."*
*"I find this environment quite loud and feel a bit nervous as to what is happening. The voices do not appear friendly and I feel unsafe.*

*The motorbikes especially evoke a sense of fear and I don't like them. I think there are lulls in the traffic which relax me, although at times I feel an accident could be imminent."*

Responses suggested that contemplation of sound environments could awaken spiritual feelings in people. The natural sounds were associated with spiritual uplifting:

*"I want to find discover new things, and learn about the answers of life. I feel like I have so many questions and very little in the way of answers so far. Some questions I cannot even express, but I have a feeling of curiosity and hope that that feeling will take me somewhere should I act upon it."*

*"So carefree, so full of spirit the chirping reverberates through the surrounding and penetrates the darkness. So happy, so light-hearted, bringing a sense of purity and innocence to this place."*

While the other soundscapes could be spiritually deadening:

*"The whole drab affair is soul crushing when exposed to it for so long – need a change."*

*"Why is there such a profound hate, hate I did not know I possessed. But yet it is there, it etches deep into me, scraping at my heart and pulling out ghosts which have been safely buried away."*

The sense of physical wellbeing was enhanced by some sound environments:

*"I feel healthy – I'm awake early enough to hear the birds. [. . .] I'm breathing fresh, unpolluted, virgin air."*

Whereas a sense of affliction was caused by others:

*"I cannot get through, for there are too many people. The wait is giving me lines, a tight forehead, and I feel tired, very tired and a little short of breath."*

A sense of comfort and contentedness was associated with some soundscapes:

*"Overall it's pretty warm and cozy. . . "*

*"I feel relaxed and a lot less on edge. The waves almost seem to lull me to sleep and the bird sounds are comforting."*

Whereas the traffic-dominated environment produced discomfort and the desire to escape:

*"The air is dirty and I'm not comfortable. I feel like I'm heading for another long, restless and monotonous day in the office."*

*"I can feel myself trying to leave my own body. Withdrawing just to escape the screams and roars."*

*"I look to escape. I do not want to be here. I want us to be away from the city. Why did we create this? Is this necessary? Can't it go back to the way it was before?"*

In summary, the qualitative analysis generated eight themes related to the quality of the sound environment and nine themes related to the deeper internal response to it (**Table 1**). These themes, and the polar terms used to capture our understanding of each theme in semantic differential scales were used in the questionnaire. However, this process was not straightforward because (as pointed out earlier) the distinction between the perception of a sound and the emotions it evokes are not straightforward. To attempt to address this, and in order to

**TABLE 1 |** Themes and terms used on each pole of the scale in the questionnaire.

| Theme | Lower pole | Upper pole |
|---|---|---|
| **Qualities of the sound** | | |
| Level | Very soft | Very loud |
| Pace | Leisurely | Fast |
| Clarity | Clear/distinct | Unclear/blurred/disorderly |
| Complexity | Simple sounds | Complex sounds |
| Space | Spacious/liberating/vast | Congested/claustrophobic/enclosed |
| Tone | Harmonious/melodious | Discordant/harsh |
| Stability | Dynamic/changing/up-and-down | Monotonous/in the same manner/flat |
| Pattern | Rhythmic/predictable | Irregular/random |
| **Feelings and emotions** | | |
| Stimulation | Soothing/hypnotic | Arousing |
| Connection | A sense of belonging | A sense of alienation |
| Stress | Relaxation/tranquility/peace | Stress/anxiety annoyance/anger |
| Familiarity | Familiar/usual | Novel/unusual |
| Cognitive Load | Refreshed/rejuvenated | Distracted/mentally overloaded |
| Safety | Safe/a sense of control | Threatened/fearful |
| Spirit | Uplifted/meditative/transcendent | Oppressed/depressed |
| Wellbeing | Healthy/wholesome | Affliction/infirmity |
| Comfort | Contented/comfortable | Desire to escape/uncomfortable |

provide a sense of the meaning intended for each scale, sets of semantic markers were used as seemed best subjectively. This meant that we used terms which were not necessarily in simple opposition to each other, and which potentially described different aspects of a theme. For example, the theme related to 'Space' was labeled 'spacious/liberating/vast' at one end of the scale, and 'congested/claustrophobic/enclosed' at the other. The meanings of the terms used, and their overlap captured as best we were able our understanding of the meaning behind the themes identified.

# PHASE 2: QUANTITATIVE STUDY

## Materials and Methods

The themes identified in Phase 1 were adapted to items in a questionnaire (**Table 1** and **Supplementary Material**). There were 17 different themes identified by the qualitative analysis. Of these, two (Stimulation and Familiarity) appeared to have a multidimensional structure (see Qualitative Results), and therefore two more items were introduced to the questionnaire to allow this to be captured, but these were little-used by participants and were dropped from analyses and not presented in the quantitative part of the Results. The semantic-differential items were introduced by a short passage of text reading: "*Please listen to the sounds around you and rate the <u>sound</u> environment and*

*your response(s) toward it by circling a number (1–6) on the following scales. If the scale is irrelevant to you, tick the 'not applicable' box beside it."*

We did not attempt to limit or differentiate the descriptive (denotative) and affective (connotative) items.

Typically, a seven-point bipolar rating scale has been used in semantic differential scale research. This allows for a range of responses while also allowing a respondent to adopt the midpoint of a scale as a 'null' option. It is important to include a null option where a particular soundscape may simply not cause a particular response in a person, but the null option is also a danger in that it allows a respondent to opt out of making a decision about the sound environment or their response to it and thus reduces the value of data collected. Recognizing the competing issues, we used a six-point rating scale with the null point presented as a separate tick box labeled 'N/A' for each scale. We believed that this would tend to preserve the usefulness of data while still allowing the null option in a position that required respondents to make a definite choice to select it.

Scales were generated with as much information for respondents as possible. Each scale had a heading that reflected a theme about the soundscape identified in the writing during the first phase, and the poles of each scale were anchored with at least one term which we agreed would capture that extreme of the scale in question. If possible, multiple terms were used on the principle that if the meaning of each term has variability in the mind of a respondent, the areas of meaning which overlapped between terms would specify the concept we were asking about more precisely (**Table 1**).

As the aim was to create a questionnaire that could be distributed and completed by members of the public, focus was placed on reducing the number of items where possible. Preliminary versions of the questionnaire were trialed among a small group of people to assess issues like clarity, readability and ambiguity. An iterative process allowed for refinement of the questionnaire.

Adults were stopped in the street and asked if they would be willing to answer the questionnaire. Of these, 228 agreed and their data are presented here. This was done in four different locations within the central city: a park ($N = 12$), a quiet shopping street with mixed pedestrian and light vehicular use ($N = 50$), a busy main street with a mixture of pedestrian and vehicular traffic ($N = 48$), and a street heavily used by buses and other vehicles with fewer pedestrians ($N = 59$). The sites were selected because we were interested to test whether the questionnaire would provide different responses in quite subtly varying sound environments (i.e., the different types of street), and in a qualitatively different environment (the park).

In preparation for the analysis (SPSS v25), a check and reorganization of the variables was carried out. Questionnaire items were presented in random order to encourage respondents to read each question and provide thoughtful answers. To align the direction of responses with the label used for each theme, reverse coding was performed prior to analysis. This was for the items: Clarity, Space, Tone, Pattern, Connection, Familiarity, Safety, Spirit, Well-being, and Comfort. Follow-up items: 'Type of Arousal' which sought to operationalize the difference between feeling aroused in the sense of being excited and aroused in the sense of being overwhelmed; and 'Feeling about familiar sounds' which tried to operationalize the difference between comfortably familiar sounds and boring sounds were dropped from this analysis because they were conditional on prior items and this made interpretation difficult, as well as resulting in many missing data points.

A preliminary principle components analysis was conducted.

Principle axis factoring with obliminal rotation was used to assess the remaining 19 soundscape items. Obliminal rotation was preferred because it allows factors to be non-orthogonal (i.e., correlated), and there is no *a priori* reason to assume that soundscape factors would be orthogonal. Using the standard approach (i.e., the Kaiser criterion) of selecting factors with eigenvalues > 1, five factors were generated.

Internal reliability was assessed using Cronbach's alpha. Since factor scores are based on all the items in the dataset, these values were generated based only on the raw scores for items with loadings >0.3 on each factor.

Mean factor scores for each of the five factors were compared between those responding in the park and in the different types of street using ANOVA. Distributions of scores within each area were checked for violations of normality and homoscedasticity assumptions and found to be acceptable. *Post hoc* least-significant difference tests were conducted: no attempt was made to control for 'Type-1' errors on the basis that the study was essentially exploratory, and therefore such errors would be less important than Type-2 errors which are more prevalent when using controlled testing.

## Results

The principle components analysis showed that there were five principle components with eigenvalues greater than one. The loadings of the 17 items on the first two components are plotted in **Figure 1**.

The sample was found to be suitable for factor analysis on the basis of a Kaiser–Meyer–Olkin score of 0.85 and a significant Bartlett's Score. Five factors had eigenvalues greater than one (**Figure 2**). These explained 62% of the variance: Factor 1 explained 37%, and Factors 2–5 explained 9–6%, respectively.

The rotated solution with five factors explained 47% of the variance based on the extraction sums-of-squares loadings. Absolute factor loadings of greater than 0.3 were used to characterize each factor (**Table 2**).

Stimulation, Stress, and Cognitive Load loaded negatively and Space, Tone, Pattern, and Spirit loaded positively on a factor that was labeled 'Calming.' A person who found a soundscape 'Calming' would tend toward the descriptors: 'soothing/hypnotic, spacious/liberating/vast, harmonious/melodious, rhythmic/predictable, tranquility/peaceful, refreshed/rejuvenated, and uplifted/meditative/transcendent.' Cronbach's alpha score for this scale was 0.81.

**FIGURE 1 |** Loadings of the 17 items on each of the first two principle components.
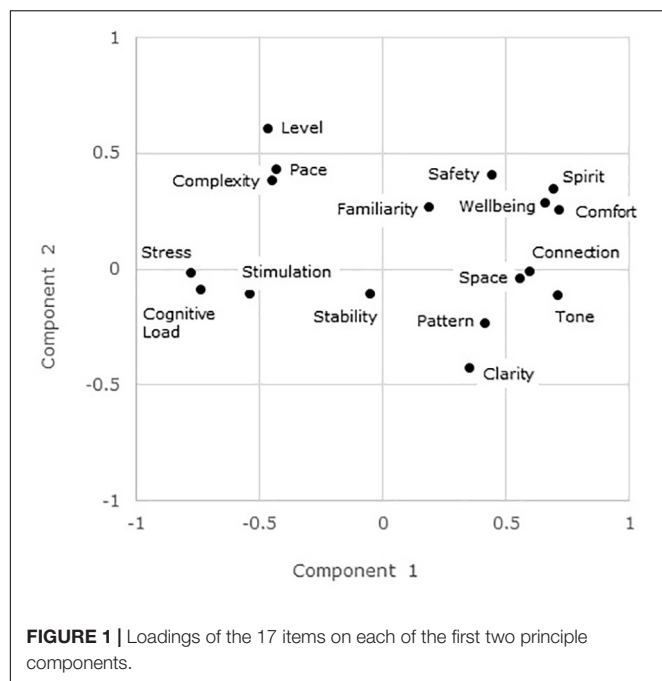
Safety, Spirit, Wellbeing, and Comfort loaded positively on a factor labeled 'Protecting.' A person who felt a sense of being 'Protected' in the soundscape would tend to use the descriptors: 'safe/a sense of control, uplifted/meditative/transcendent,

healthy/wholesome, and contented/comfortable.' Cronbach's alpha score for this scale was 0.78.

Level and Pace loaded positively and Clarity negatively on a factor that was labeled 'Hectic,' capturing as it did, loud, quickly changing, and unclear sounds. A person who found a soundscape 'Hectic' would tend toward the descriptors: 'very loud, fast, and unclear/blurred/disorderly.' Cronbach's alpha score for this scale was 0.60.

Connection and Familiarity loaded together and positively on a factor labeled 'Belonging.' A person who felt a sense of belonging to the soundscape would tend to use the descriptors: 'a sense of belonging, and familiar/usual.' Cronbach's alpha score for this scale was 0.39.

Stability alone had a high loading on the fifth factor, and thus this factor was labeled 'Stability.' A person who scored high on this scale would have used the descriptors 'monotonous/in the same manner/flat' to describe the soundscape.

Complexity did not produce sufficiently large loadings on any of the factors to be considered in the naming of factors, suggesting that its loading was distributed rather evenly across the factors. Considering cross-loading, only Spirit loaded > 0.3 on more than one factor: it was represented in both the Calming and Protecting factors.

Oblique factor analysis allows factors to correlate. In most cases, correlations were small (<0.3), however, Calming and Protected correlated moderately ($r = 0.47$), as did Calming and Hectic ($r = -0.43$). Given the nature of these factors, and the observation that Spirit loaded highly on both Calming



**FIGURE 2 |** Scree plot showing eigenvalues obtained for factors obtained.
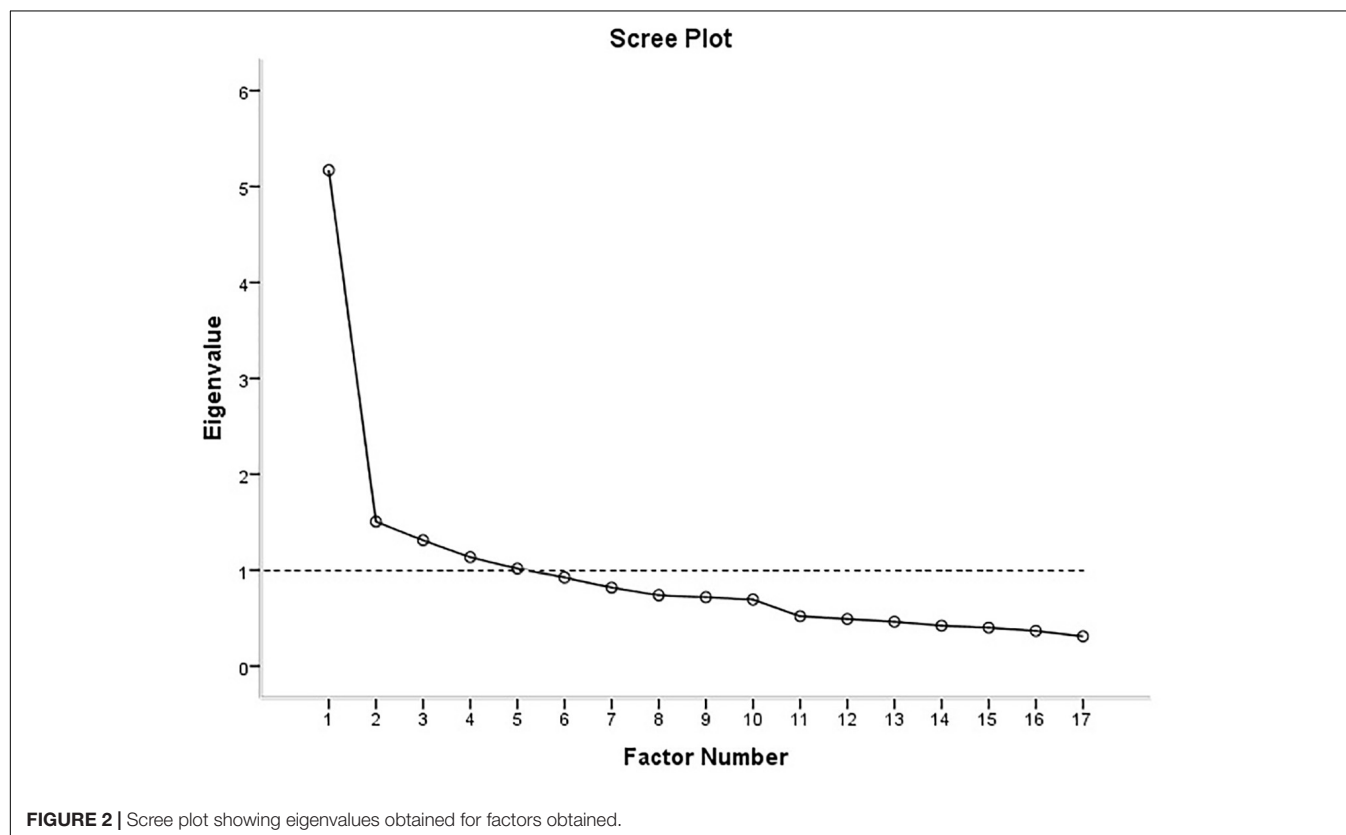
**TABLE 2** | Factor loadings for each of the 17 soundscape items.

| | Factor names | | | | |
|---|---|---|---|---|---|
| | **Calming** | **Protecting** | **Hectic** | **Belonging** | **Stability** |
| Level | | | 0.777 | | |
| Pace | | | 0.528 | | |
| Clarity | | | −0.439 | | |
| Complexity | | | | | |
| Stimulation | −0.553 | | | | |
| Space | 0.451 | | | | |
| Tone | 0.715 | | | | |
| Stability | | | | | 0.706 |
| Pattern | 0.324 | | | | |
| Connection | | | | 0.638 | |
| Stress | −0.616 | | | | |
| Familiarity | | | | 0.489 | |
| Cognitive Load | −0.697 | | | | |
| Safety | | 0.728 | | | |
| Spirit | 0.485 | 0.516 | | | |
| Wellbeing | | 0.642 | | | |
| Comfort | | 0.566 | | | |

*Only loadings of greater than 0.3 are shown to facilitate the visualization of the factors.*

and Protecting, the correlations and their directions are not unexpected.

Factor Scores were generated and the scores were compared between those who completed the questionnaire on city streets and those who completed it in a park (**Figure 3**). Negative scores indicate that, on average, people experienced the opposite of the factor name, and positive or negative scores further from zero reflect the degree that each factor was experienced.

Analyses were conducted to compare the responses across the four environments (**Figure 3**). This used General Linear Modeling with the four types of area (street dominated by vehicles, mixed, street dominated by pedestrians, and park) as the independent variable and the five Factors as dependent variables. The Calming [$F(3,168) = 17.25$, $p < 0.001$], Protecting [$F(3,165) = 9.09$, $p < 0.001$], and Hectic [$F(3,168) = 11.14$, $p < 0.001$] factors were clearly significantly different between areas. Belonging [$F(3,168) = 2.56$, $p = 0.056$] and Stability [$F(3,168) = 1.86$, $p = 0.14$] differed more marginally. The direction of effects was consistent: Calming, Protecting, Belonging, and Stability were higher, and Hectic was lower for the park than for the street environments. *Post hoc* testing (Least Significant Differences) showed that all the soundscape factors differed between the park environment and at least some of the street environments. Calming was highest in the park and was also higher in the pedestrian-dominated street than in either the mixed or vehicle street types. Protecting was higher in the park than all the street types and was also higher in the pedestrian-dominated than the vehicle-dominated street. Hectic was lower in the park than any of the streets and was also lower in the pedestrian-dominated street than in either the mixed or vehicle-dominated street. Belonging was higher in the park than any of the streets, which did not differ between each other.

Stability was higher in the park than in any of the streets apart from the pedestrian, and no other differences were observed (**Figure 3**).

## DISCUSSION

We exposed people who self-identified as expressive writers to different sound environments and asked them to write about their reactions. The writings were then subjected to a qualitative, thematic analysis and the themes which emerged were used as the basis for a seventeen-item questionnaire. This was administered to passers-by in Auckland City. The items generated by the thematic analysis could be rendered down to five factors which underlay the responses made by people to them: Calming, Protecting, Hectic, Belonging, and Stability.

The themes extracted from the expressive writing part of the research (i.e., Phase 1) were broadly consistent with the concepts used in other similar studies. For example, the set of items used for urban soundscapes by Kang and Zhang (2010) and since used in other studies included equivalent concepts to the themes of level, pace, complexity, tone, stress, and wellbeing. On the other hand, our writers came up with other themes that did not appear in that set (**Table 1**). Comparison with other, much larger sets of terms (Axelsson et al., 2010) shows similar parallels and discrepancies. The process of seeking to capture the elements of the soundscape is not straightforward, and the use of different approaches for this crucial first step is important. Overall, after rotation, our five factors explained 47% of the variance in the data we collected. This is similar to the 53% reported in the other study that used a similar approach of combining descriptive and affective items to describe urban spaces (Kang and Zhang, 2010).

Our questionnaire was based on the themes identified in the qualitative writing phase of the research. It was useable by the general public and showed patterns in the results consistent with previous research suggesting that parks would differ from urban streets in being calmer, more protecting and less hectic (Carrus et al., 2017). Our questionnaire was not, however, very useful for discriminating between soundscapes associated with city streets that had differing degrees of heavy vehicular traffic use. Saying this, differences were observed in some factors between a pedestrian-dominated street environment (with light vehicular use) and streets that had either mixed or predominately vehicular usage. The pedestrian-dominated environment was more Calming and less Hectic than the others. The capacity of the questionnaire to differentiate between these environments, provides some support for the validity of the measure. Though it is not possible to validate psychometric scales like these absolutely, the capacity to detect statistical differences associated with qualitatively different environments supports the idea that the scales had validity.

We assessed the internal reliability of the scales, and these were generally at acceptable levels (approximately 0.8) for scales with several items: scales with few items will tend to show lower measures of internal reliability, so lower values in these do not imply poor reliability. Use of the scales in similar urban environments (i.e., the different types of street) showed only
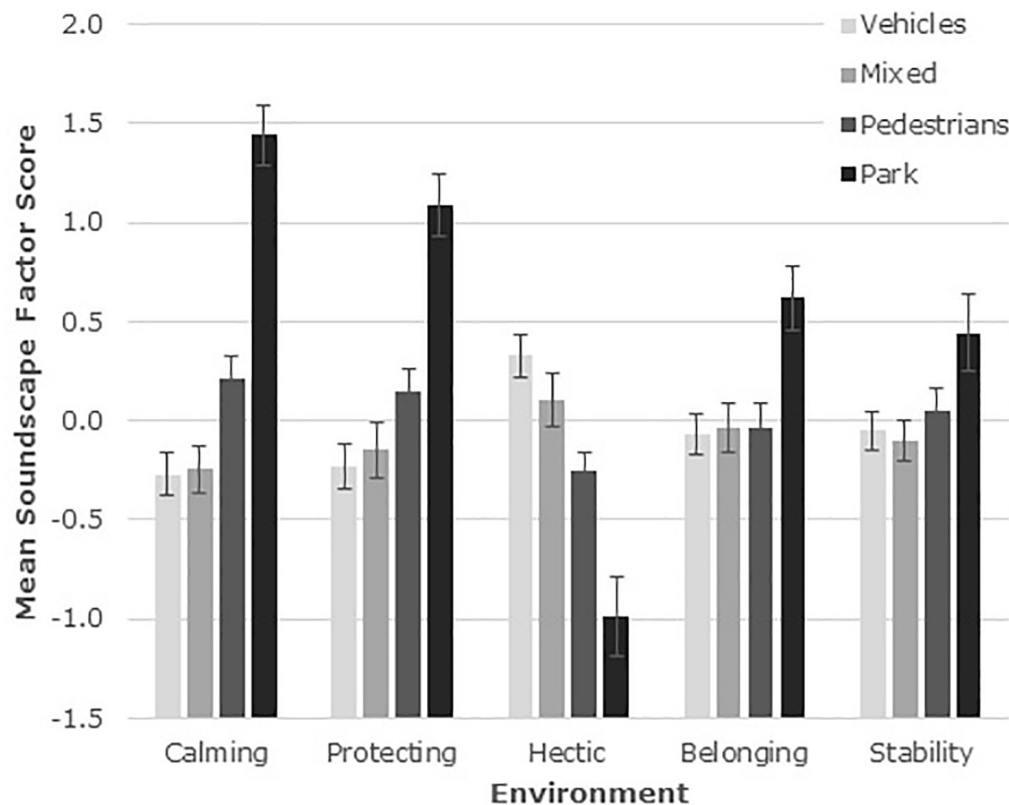
**FIGURE 3 |** Mean Soundscape Factor Scores for people responding in the four city environments. Scores on each of the five factors are represented separately. Error bars represent one standard error of the mean.

slightly different mean scores, which implies that there is some reliability in responses given the similar environments. It would also be interesting to test reliability over time in the same location, however, this would be difficult given that factors influencing the soundscape might change, so careful characterization of the environments would be important to allow any variation to be clearly indicated as either due to changing environment or to unreliability of the measure.

The generalizability of the results is questionable because our sample was quite small, and depended on voluntary involvement of passers-by in public streets rather than carefully conducted random sampling. Saying this, we were not seeking to provide a definitive set of data about the soundscape in Auckland City. Rather, we were seeking to test whether people could and would respond to the questionnaire, and if so whether there was some meaningful structure to the way they responded. We believe there was and so are comfortable that the research supports the approach as a way of generating data.

Soundscape research such as ours seeks to quantify a subjective judgment. As researchers, we hope that this is possible because there is an element of consistency in the sound environments, and since people are all from the same species, it would be likely that there would be commonality in the factors that drive us to experience different feelings (Andringa and Lanser, 2013). The lack of consistency in subjective judgments depends on

many factors. We propose that there are multiple loosely coupled systems in operation to explain individual responses to questions about the soundscape:

(1) The set of concepts and emotions that each person associates with each lexical item will vary according to their experiences and idiolect: a word will not necessarily capture the same meaning for everyone. Furthermore, for a person to express an emotion via a closed-set questionnaire such as we used in Phase 2 of this study, there must be some item present that would capture the meanings that person perceives; if there is not, then that internal state would go unreported.

(2) The emotional response to a given sound environment will differ depending on a person's understanding of the situation, their previous experiences, personality, and current state of mind. Furthermore, the pallet of affective states that different people experience may vary according to their experiences and psychophysiological make-up.

(3) Definitions of the soundscape explicitly accept interactions with other sensory information. The sound environment is highly integrated with other information in our minds. While the field has acknowledged this, it still presents an unquantified element in the coupling between the sound environment and the soundscape.

(4) The emotional states that people experience in response to sounds may be hard to distinguish from their experience of the sound itself. The concepts of sound in the sense of physical pressure fluctuations and sound as a percept may be intertwined in the mind. In other words, people know how they feel and what the world is like, but they do not necessarily separate these two sets of concepts. An example of this is the Calmness factor, which combined denotative (Space, Tone) and connotative (Stimulation, Stress, Cognitive Load, Spirit) items. It would be convenient if these were separate in the minds of people, but our results demonstrate that they are not cleanly separated. Other factors appeared to display organization along denotative and connotative lines: Protecting and Belonging appeared connotative, and Hectic and Stability were apparently more denotative in terms of the highly loaded items (**Table 2**). Even in these cases though, our qualitative analyses had already revealed that, for example, 'pace,' which was loaded into the Hectic factor, seems to reflect not only the denotative temporal quality of the sound environment, but also reflects a connotative adaptation of the internal state of listeners in response to this.

Together, these four issues combine to reduce a researcher's capacity to gain a full understanding of a person's perception of the sound environment. Accepting this requires us to put aside some of our tightly organized, analytical thinking at one level while maintaining it very carefully at another.

We identified five factors on the basis of eigenvalues. In factor analysis, there is no strict rule for deciding on the factor structure, and a structure with less factors is generally preferable on the basis that it can be imagined as a space (if there are three or less dimensions), or even drawn. Some soundscape studies have identified factor structures on the basis of eigenvalues and then dropped factors which the authors feel do not contribute much to the understanding of the data (e.g., Axelsson et al., 2010). This is a perfectly acceptable practice. We chose to preserve even the fifth factor (Stability) which had a strong loading from only one item. This might be regarded as improper on the basis that factor analysis is valuable because it reduces the number of dimensions below the original, and this is the basis for the Kaiser criterion that factors should have eigenvalues greater than one to be regarded as efficient. Nonetheless, we felt that it was justified. Firstly, the item 'Stability' did not load much on any other factor and it was identified as relevant in the writing. Secondly, in principle, the soundscape is still poorly understood so we felt that any contributing factor should not be neglected until the provision of evidence to the contrary. Thirdly, it must be remembered that in factor analysis the relative 'strengths' of the factors is somewhat arbitrary. The unfactorized data may be envisaged as an N-dimensional cloud, where N is the number of items in the questionnaire. Commonality in the alignment of underlying meanings of the items in the minds of respondents would tend to reduce the dimensionality of the cloud due to the tendency for correlations between responses to those items, and thus N can be reduced while losing only slight variations in the cloud's dimensions. However, we do not know what the true dimensionality of the soundscape is and our choice of items is thus rather arbitrary. The finding that only one item substantially loaded on the Stability factor, and that the factor explained 6% of the variance in the items used does not tell us that it is unimportant. It only tells us that Stability does not relate much to the other items we have chosen.

Factor analysis therefore allows the grouping of items which are originally separate. It provides a simplification of data but the process of naming the factors also adds to the understanding of the underlying influences on the data. The factor we labeled 'Calming' implies that soundscapes that were harmonious, following a pattern, and providing a sense of spaciousness were associated with people feeling soothed and tranquil; and that this was rejuvenating and enabled spiritual transcendence. This picture is helpful in that it seems to follow from descriptive features to affective states and then goes beyond simple emotions into higher aspects of our being. The implication is that soundscapes can influence us very deeply, and this dimension is consistent with the dimension 'Calmness' (Cain et al., 2013).

Similarly, the factor we named 'Protecting' captured the idea that soundscapes in which people felt safe provided contentment and in such soundscapes people felt both physically wholesome and spiritually uplifted. Again, the depth of the concepts drawn from the writing produced items which allowed respondents to express the deep impact of feelings beyond simple emotions and provides a picture of the unfolding influences of higher-level cognition. Usefully, this level of responding was accessible from passers-by in the street who took only moments to reflect. The idea that the acoustic environment feeling protected is important for people, and theoretical work has been done in this direction (van den Bosch et al., 2016). The emergence of a factor that relates directly to this suggests that the approach used may provide a useful model for soundscape research into improving the soundscape via interventions.

We used the term 'Hectic' to label another factor because it captures the idea of loud and low fidelity environments causing people to feel hurried and pressured temporally. The relationship of time to soundscapes has been considered previously (e.g., Kang and Zhang, 2010). Time is a physical dimension within which we have no control. Nonetheless, as people, we feel that our relationship with time varies, speeding up and slowing down depending on the conditions and our state of mind. The soundscape appears to contribute to this, and the associations described by this factor provide some sense of how. More understanding of how and why the soundscape contributes to this would be important.

The other two factors we identified had high loadings from only two and one item, respectively. 'Belonging' combined the idea that a person could feel familiar with a soundscape and that this would be associated with a feeling of belonging to it. Interestingly, everyone surveyed must have been familiar with the soundscapes dominated by traffic and other pedestrians, but nonetheless the responses did not reflect this: rather the sense of both familiarity and belonging was greater in the park. Connection and Familiarity loaded together and positively on a factor labeled 'Belonging' and this may partly correspond to

the factor labeled 'Familiarity' in previous research (Axelsson et al., 2010). We named the fifth factor 'Stability' and it was associated with a less varying soundscape, which people observed in the park more than in the vehicle-dominated streets. As we argued above, this factor might explain some important aspect of the soundscape for which we have no good theoretical understanding. Of course, it may relate to the concept of 'Eventfulness' (Axelsson et al., 2010) or 'Vibrancy' (Cain et al., 2013), which have also been identified as being important both from a theoretical perspective (Andringa and Lanser, 2013; van den Bosch et al., 2018). The structure of the factor analysis here may be driven partly by the lack of an interestingly eventful soundscape in the areas where we administered the questionnaire: streets and a park, none of which have much vibrancy or many events occurring. Future research using the questionnaire in areas with more interesting and relevant sounds might produce a different structure.

The research has other limitations. To us, the most significant caveat is that during the creative writing phase we used sound exposures in the absence of other (visual, olfactory, etc.) stimulation which may have influenced the experience of the soundscapes. We wanted the participants to focus on the sound so that their writing would capture those aspects of the soundscape for us to use in the development of the questionnaire. We thought that adding other sensory information alongside the sound, or asking participants to conduct the writing exercise in the real world would have provided distractions from the acoustic aspects of the environment to which we very much wanted them to attend. It is possible that the themes may have been broader had the writing been conducted in multisensory environments, but the positive aspect of running the study the way we did was that we found few references to non-auditory aspects of the virtual environments in the writing. Nonetheless, the seventeen themes identified in the qualitative phase of the research may have underestimated the potential themes in soundscapes, and thus is it worth considering whether more may be valuable.

The sound environments where we administered the questionnaires were reasonably similar, apart from the park, and it may be interesting to test the questionnaire in a more widely varying set of environments. Furthermore, the particular range of four environments we used might have introduced patterns to the data that could have led to the factor structure being different from what it would have been if we had included other environments. Future research in which we increase the number of environments may well alter the factor structure observed. Finally, we did not assess the extent to which the questionnaire would detect changes in an environment, and an interesting area for future research is to administer it in a longitudinal manner throughout a period of change in a sound environment such as a redevelopment of an area of the city.

Some excellent theory and research has moved us toward a unifying theory to explain the various findings from soundscape research. Good theory can help direct research and allow more specific hypotheses to be tested. We regard our present work as largely exploratory and aimed at stimulating ideas about possible directions for growth in existing theories. We have tried to demonstrate that there are possibly more complexities to the

soundscape than are captured by our two-dimensional models and to remind researchers that the items used to generate factors are crucial, because they dictate the entire conceptual space which then provides the components of the theory. With a different theoretical stance, the descriptors used in questionnaires would change, and thus the apparent factors that emerge from them would be different. An example of another way of considering the interplay between our senses, cognition, and emotions is that of a valuation-based process wherein we evaluate environments based on a complex internal model that would weigh up the survival benefits of a given environment and take into account factors such as the opposing principles of competition and social support from other people (Mercado-Doménech et al., 2017). By thinking more about the ways that people feel due to their experience of a soundscape, why they would feel this, and crucially, how they describe those feelings, we may move toward a more complex model and assessment of the soundscape.

We did not include previously used scales alongside our one for comparison, and this would be interesting to do in a future study. Nonetheless, it is interesting and useful to speculate about the possible correspondences between factors identified in different studies. We have mentioned above that Calming and Belonging seem to correspond, at least in part, with factors identified in earlier work. The factor we labeled 'Protecting' might correspond to earlier-identified 'Pleasantness,' and 'Hectic' might correspond to 'Eventfulness.' Stability could perhaps represent the other pole of factors that have been labeled 'Excitingness' or 'Vibrancy.' Published research with semantic differential scales and in similar sound environments (outdoor, urban) to those we used produced a pattern of responses that was somewhat similar to ours (Kang and Zhang, 2010). The earlier research identified four factors: relaxation, communication, spatiality, and dynamics. Our factor label 'Calming' sounds similar to the earlier 'Relaxation,' though our version did not include large loadings from our items 'comfort' or 'level' which appear to correspond to the items 'comfort-discomfort' and 'quiet-noisy' in the earlier study, while the other items in the relaxation factor did not have equivalents in our questionnaire. This may imply that the similarity in the factor name is rather superficial. The second and third factors in the earlier study were labeled 'Communication' and 'Spatiality,' and neither the names nor the items that load on these factors appear to correspond to factors we observed. On the other hand, the earlier study's fourth factor 'Dynamics' might possibly correspond to our 'Hectic' in that the scales 'hard-soft' and 'fast-slow,' which loaded highly on it might be similar to the experiences captured by our 'level,' 'pace' and 'clarity' items. It has been suggested previously that Kang and Zhang's 'calming' and 'dynamics' factors may correspond to the commonly reported two-dimensional soundscape structure (Davies and Murphy, 2012). If so, then the possible similarities with our work may support the notion of these two dimensions.

We encourage caution with respect to the factor structure we have described; the research was done to develop and field-test a questionnaire, and in this respect we feel it was successful.

However, our data were captured in a limited set of sound environments, and this would limit the scope of the factors that we could possibly identify, while potentially introducing spurious correlations that might have led to apparent factors that would not be present in more representative datasets. As more research is conducted and theory generated, greater understanding of the range of aspects of the soundscape will emerge.

## CONCLUSION

We believe that the approach of bringing creativity to the initial set of soundscape-related items was useful. It provided both similarities and differences with previous research, and the questionnaire was workable in principle, with measureable differences in soundscapes between different urban sound environments. It allowed ordinary people stopped in the street to provide complex and deep responses about the impact of the soundscape on themselves, and it might be useful for those such as acousticians, architects, and planners trying to influence soundscapes. From a scientific perspective, some of the aspects of soundscapes that are suggested by the research may open up interesting directions for more research and the development of theory. We hope to develop the techniques used

here further, and to test the questionnaire in more differing environments.

## AUTHOR CONTRIBUTIONS

DW wrote the first draft of the paper, contributed to the design of the research and the qualitative analyses, was responsible for the quantitative data analyses, and contributed to the interpretation of the data. DS and KD contributed to the writing of the paper and the interpretations of the data. MT contributed to the writing of the paper and the interpretations of the data, collected part of the data, and contributed to the design of the research and the qualitative analyses. GC contributed to the writing of the paper and the interpretations of the data and collected most of the quantitative data.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02698/full#supplementary-material

**TABLE S1 |** Soundscape questionnaire.

## REFERENCES

Aletta, F., Kang, J., and Axelsson, O. (2016). Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landsc. Urban Plan.* 149, 65–74. doi: 10.1016/j.landurbplan.2016.02.001

Andringa, T. C., and Lanser, J. J. L. (2013). How pleasant sounds promote and annoying sounds impede health: a cognitive approach. *Int. J. Environ. Res. Public Health* 10, 1439–1461. doi: 10.3390/ijerph10041439

Andringa, T. C., Weber, M., Payne, S. R., Krijnders, J. D., Dixon, M. N., van der Linden, R., et al. (2013). Positioning soundscape research and management. *J. Acoust. Soc. Am.* 134, 2739–2747. doi: 10.1121/1.4819248

Axelsson, O. (2015). How to measure soundscape quality. *Paper presented at the Euronoise 2015*, Maastricht.

Axelsson, O., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *J. Acoust. Soc. Am.* 128, 2836–2846. doi: 10.1121/1.3493436

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S., et al. (2014). Auditory and non-auditory effects of noise on health. *Lancet* 383, 1325–1332. doi: 10.1016/s0140-6736(13)61613-x

Berglund, B., and Nilsson, M. E. (2006). On a tool for measuring soundscape quality in urban residential areas. *Acta Acust. United Acust.* 92, 938–944.

Botteldooren, D., De Coensel, B., and De Muer, T. (2006). The temporal structure of urban soundscapes. *J. Sound Vib.* 292, 105–123. doi: 10.1016/j.jsv.2005.07.026

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Brown, A. L., Kang, J. A., and Gjestland, T. (2011). Towards standardization in soundscape preference assessment. *Appl. Acoust.* 72, 387–392. doi: 10.1016/j.apacoust.2011.01.001

Cain, R., Jennings, P., and Poxon, J. (2013). The development and application of the emotional dimensions of a soundscape. *Appl. Acoust.* 74, 232–239. doi: 10.1016/j.apacoust.2011.11.006

Carrus, G., Scopelliti, M., Panno, A., Lafortezza, R., Colangelo, G., Pirchio, S., et al. (2017). A different way to stay in touch with 'Urban Nature': the perceived restorative qualities of botanical gardens. *Front. Psychol.* 8:914. doi: 10.3389/fpsyg.2017.00914

Davies, W. J., Adams, M. D., Bruce, N. S., Cain, R., Carlyle, A., Cusack, P., et al. (2013). Perception of soundscapes: an interdisciplinary approach. *Appl. Acoust.* 74, 224–231. doi: 10.1016/j.apacoust.2012.05.010

Davies, W. J., and Murphy, J. E. (2012). Reproducibility of soundscape dimensions. *Paper Presented at the Internoise Conference*, New York, NY.

De Coensel, B., and Botteldooren, D. (2006). The quiet rural soundscape and how to characterize it. *Acta Acust. United Acust.* 92, 887–897.

Dubois, D., Guastavino, C., and Raimbault, M. (2006). A cognitive approach to urban soundscapes: using verbal data to access everyday life auditory categories. *Acta Acust. United Acust.* 92, 865–874.

Guitar, K., Giles, E., Raymond, B., and Welch, D. (2013). Health effects of cochlear implants. *N. Z. Med. J.* 126, 9–26.

Kang, J., and Zhang, M. (2010). Semantic differential analysis of the soundscape in urban open public spaces. *Build. Environ.* 45, 150–157. doi: 10.1016/j.buildenv.2009.05.014

Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Appl. Linguist.* 19, 255–271. doi: 10.1093/applin/19.2.255

Liu, F. F., and Kang, J. (2016). A grounded theory approach to the subjective understanding of urban soundscape in Sheffield. *Cities* 50, 28–39. doi: 10.1016/j.cities.2015.08.002

Mercado-Doménech, S. J., Carrus, G., Terán-Álvarez-Del-Rey, A., and Pirchio, S. (2017). Valuation theory: an environmental, developmental and evolutionary psychological approach. Implications for the field of environmental education. *J. Educ. Cult. Psychol. Stud.* 16, 77–97. doi: 10.7358/ecps-2017-016-merc

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

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

Raimbault, M. (2006). Qualitative judgements of urban soundscapes: questionning questionnaires and semantic scales. *Acta Acust. United Acust.* 92, 929–937.

Reddy, R., Welch, D., Thorne, P., and Ameratunga, S. (2012). Hearing protection use in manufacturing workers: a qualitative study. *Noise Health* 14, 202–209. doi: 10.4103/1463-1741.99896

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714

Schafer, R. M. (1977). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester: Destiny Books.

Steele, D., Steffens, J., and Guastavino, C. (2015). The role of activity in urban soundscape evaluation. *Paper Presented at the Euronoise, 10th European Congress*, Maastricht.

Tabachnick, B. G., and Fidell, L. S. (1996). *Using Multivariate Statistics*, 3rd Edn. New York, NY: Harper Collins.

van den Bosch, K. A., Vlaskamp, C., Andringa, T. C., Post, W. J., and Ruijssenaars, W. (2016). Examining relationships between staff attributions of soundscapes and core affect in people with severe or profound intellectual and visual disabilities. *J. Intellect. Dev. Disabil.* 41, 21–30. doi: 10.3109/13668250.2015.1114077

van den Bosch, K. A.-M., Welch, D., and Andringa, T. C. (2018). The evolution of soundscape appraisal through enactive cognition. *Front. Psychol.* 9:1129. doi: 10.3389/fpsyg.2018.01129

Welch, D., Shepherd, D., Dirks, K. N., McBride, D., and Marsh, S. (2013). Road traffic noise and health-related quality of life: a cross-sectional study. *Noise Health* 15, 224–230. doi: 10.4103/1463-1741.113513

# 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:** info@frontiersin.org | +41 21 510 17 00

**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