Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

Effects of contexts in urban residential areas on the pleasantness and appropriateness of natural sounds

Joo Young HONG[1], Bhan LAM[2*], Zhen-Ting ONG[2], Kenneth OOI[2], Woon-Seng GAN[2], Jian KANG[3], Samuel YEONG[4], Irene LEE[4], and Sze-Tiong TAN[4]

[1] Architecture and Sustainable Design, Singapore University of Technology and Design, 487372, Singapore.

[2] School of Electrical & Electronic Engineering, Nanyang Technological University, 639798, Singapore.

[3] UCL Institute for Environmental Design and Engineering, The Bartlett, University College London, Central House, 14 Upper Woburn Place, London WC1H 0NN, United Kingdom.

[4] Building & Research Institute, Housing & Development Board, 738973, Singapore.

*Send correspondence to: Bhan Lam (bhanlam@ntu.edu.sg)

50 Nanyang Ave, S2-B4a-03, Digital Signal Processing Lab,

School of Electrical & Electronic Engineering,

Nanyang Technological University

Singapore 639798, Phone: +65 6790 6901

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

Abstract

Before introducing natural sounds to potentially improve the soundscape quality, it is important to understand how key contextual factors (i.e. expected activities and audio-visual congruency) affect the soundscape in a given location. In this study, the perception of eight natural sounds (i.e. 4 birdsongs, 4 water sounds) at five urban recreational areas under the constant influence of road traffic was explored subjectively under three laboratory settings: visual-only, audio-only, and audio-visual. Firstly, expected socio-recreational activities of each location were determined in the visual-only setting. Subsequently, participants assessed the pleasantness and appropriateness of the soundscape at each site, for each of the eight natural sounds augmented to the same road traffic noise, in both audio-only and audio-visual settings. Interestingly, it was found that the expected activities in each location did not significantly affect natural sound perception, whereas audio-visual congruency of the locations significantly affected the pleasantness and appropriateness of the natural sounds. Particularly, the pleasantness and appropriateness decreased for water sounds when water features were not visually present. In contrast, perception with birdsongs was unaffected by their visibility likely due to the presence of vegetation. Hence, audio-visual coherence is central to the perception of natural sounds in outdoor spaces.

1. Introduction

As urban residential buildings are increasingly built closer to transportation infrastructure, acoustic environmental quality in urban residential areas has become a critical factor for improving urban sustainability [1–4]. Traditional environmental noise control approaches concentrate on the abatement of noise levels. However, many studies have reported that reduction in noise levels does not necessarily equate to an improvement in perceived acoustic comfort [3,5]. In this sense, the notion of soundscape has emerged as a new paradigm by emphasizing the importance of human perception of the acoustic environment for urban sound management and planning [5].

The soundscape design approach primarily focuses on how to improve perceived acoustic quality in a space. Aside from noise mitigation measures, introducing sounds of preference to a noisy environment is one representative soundscape design approach to enhance soundscape quality based on auditory masking phenomena [5,6]. Pleasant sounds, in theory, have the potential to reduce the perception of noise by energetically masking the noise or diverting the listener's attention to the pleasant sounds [7]. In this context, over the past decade, many studies have provided strong evidence for introducing pleasant natural sounds such as birdsongs [8–12] and water sounds [13–19] to reduce the perceived loudness of existing noise sources and to increase the pleasantness of a soundscape.

Most of these studies have primarily focused on evaluating the pleasantness of natural sounds and investigated the acoustic characteristics of natural sounds as key soundscape design factors. Desirable natural sound levels corresponding to background noise levels have previously been investigated as important acoustic design factors. For instance, several studies found that sound levels of natural sounds similar or 3 dB lower than ambient noise levels were evaluated as most desirable [6,18,20]. Spectro-temporal characteristics of natural sounds have also been found to affect the perceived pleasantness of soundscapes as key acoustic factors [6,8,9,21]. Some studies found that water sounds with high-frequency content and high temporal variability tended to be judged as more pleasant [14,20]. It has also been found that the dissimilarity in temporal characteristics between the target noise and natural sounds could improve the soundscape quality by diverting attention away from the noise [18].

However, to our best knowledge, few studies appear to have assessed perceptions of natural sounds added to an existing environment, whether or not they are subjectively judged to be appropriate in a given context. According to ISO 12913-1, the soundscape is defined as the "acoustic environment as perceived or experienced and/or understood by a person or people, in context" [22]. As emphasized in the definition, context plays a critical role in the perceptual construct of soundscapes through their auditory sensations and interpretations, as well as the listeners' "responses to the acoustic environment" [22,23]. The context, as described in ISO 12913-1, can be represented by the people-place-activity framework, whereby the interactions between the people, the place, and its activities are considered as important factors influencing a soundscape [24–26].

It has been established that primary functions and socio-recreational activities in a given space play a key role in soundscape assessment regarding the perceived appropriateness of the sound environment [27]. To describe the people-place-activity context, some researchers have adopted the concept of a 'sociotope', defined as "the commonly perceived direct use values of a place by a specific culture or group" [28,29], to explore the people-place-activity interaction in soundscapes [30–32]. These studies demonstrated significant relationships between the appropriateness of soundscapes and the socio-recreational activities in various urban spaces. Particularly, Lavia et al. [33] found a specific set of appropriate sound sources in a place that corresponded to a specified set of suitable social and recreational activities in that place. For instance, the sociotope of a city park was closely related to sound sources of nature (e.g., birdsong, wind in trees, etc.), while the sociotope of a beach area was associated with sound sources such as sea waves, seagulls, people talking, or music.

The visual environment is also one of the critical contexts that affect soundscapes [34–38]. In particular, congruency between the acoustic and visual environment is known to modulate audio-visual interactions, which strongly affect the appraisal of both soundscape [34,39] and landscape [40,41]. Notably, sound source visibility has been identified as a critical aspect of the audio-visual congruency, but the effect of sound source visibility on noise annoyance has yielded inconclusive observations; Some have reported that a visible road traffic noise source increased subjective annoyance [42], whereas others have reported that the perceived annoyance of road traffic noise reduced when a road was visible [43,44].

Nowadays, high-rise residential precincts are designed to host a multitude of social-recreational activities with multifunctional structures and facilities. In landscape guides for such precincts [45–47], outdoor areas are usually classified into active and passive activity zones based on their intended functions of space. Active zones provide facilities and structures for active activities, such as children's playgrounds and exercise equipment for the elderly.

On the other hand, passive zones provide facilities and structures such as benches, pavilions, shelters, and community gardening spaces to promote passive activities such as sitting, resting, and socializing. Since urban outdoor residential areas play multi-faceted roles, their soundscapes should be designed by considering their contexts in a given space such as their associated socio-recreational activities and visual environments [5].

Therefore, this study aims to investigate the influence of context on soundscape intervention by natural sound augmentation. Specifically, the effects of expected human activities in a place and sound source visibility are explored as critical contextual factors in outdoor residential areas. Two research questions are addressed in this study: (1) Do the expected socio-recreational activities affect the perception of birdsongs and water sounds in noisy outdoor residential areas? (2) Does the visibility of sound sources affect the perception of birdsongs and water sounds in noisy outdoor residential areas? To address these questions, laboratory experiments were conducted to evaluate soundscapes with varying audio and visual components in outdoor residential areas.

The experimental design and procedures are addressed in Section 2. The results of the experiments related to the expected outdoor activities in urban residential areas are analyzed in Sections 3.1. The effects of audio and visual components on pleasantness and appropriateness of soundscape are examined in Sections 3.2 and 3.3 respectively. The results associated with the first research question on the relationship between the expected outdoor activities and the soundscape attributes are addressed in Section 4.1. Subsequently, the results pertaining to the second research question about the effect of the visibility of sound sources are analyzed in Section 4.2. Lastly, the findings of this study are discussed with the limitations and implications of this study in Section 4.3.

2. Method

2.1 Visual stimuli

As 80% of Singapore's population resides in public housing [48], urban residential outdoor spaces are amongst the most prevalent high-utility public areas. Therefore, this study focuses on outdoor locations in high-rise public housing estates in Singapore. As shown in Fig.1, a total of five locations were selected for the laboratory experiment, of which two were active zones and three were passive zones. Table 1 presents landscape components and facilities at the selected location. Both the active zones, A_1 and $\widehat{\mathrm{A}}_2$, had a children's playground. A_1 was located in a neighborhood park, whereas $\widehat{\mathrm{A}}_2$ was adjacent to a minor road. For the passive zones, P_1 was on the walkway adjacent to a waterway in the neighborhood park, P_2 was a rooftop garden adjacent to residential buildings with sitting areas, and $\widehat{P}_3$ was an open space next to a minor road populated with sitting benches. Minor roads were visible only in $\widehat{A}_2$ and $\widehat{P}_3$. The hat accent '⌒' signifies that a location had visible roads, as can be seen from Fig. 1.

Figure 1. The five selected outdoor residential locations for the experiment, with corresponding equirectangular panoramic photos from spherical videos taken at the respective locations. Location names starting with 'A' and 'P' denote active and passive zones, respectively, and '⌒' denotes a location with a visible minor road. (source: Google Maps)

A spherical panoramic camera (Garmin VIRB 360 Action Camera, USA) mounted at a height of 1.6 m from the ground was used to capture an omnidirectional video at each location (4K 30-FPS resolution with a bit-rate of 80 Mbps). The omnidirectional videos were recorded in the absence of human activities to prevent the introduction of bias regarding the expected activities in each location. The recorded videos were post-processed (Adobe Premiere Pro CC 2017) into spherical projections for playback in a virtual reality head-mounted display (VR HMD).

Table 1. Visible landscape components at the selected locations

Visual components	Locations				
	A_1	$\hat{A}_2$	P_1	P_2	$\hat{P}_3$
Waterway			○		
Vegetation	○	○	○	○	○
Playground	○	○			
Sitting area				○	○
Minor road		○			○

2.2 Acoustic stimuli

Since road traffic is one of the most pervasive urban noise sources, the road traffic noise of an expressway (2 × 4 lanes) in Singapore was recorded at a distance of 40 m from the closest lane using an ambisonic microphone (Core Sound Tetramic, USA) via a digital recorder (Zoom F8, Japan). A 10-s audio sample of the road traffic noise was excerpted from the recording to serve as the road traffic noise stimulus, and its A-weighted equivalent sound pressure level (SPL) was 65.2 dB. The A-weighting used for this stimulus, as well as all other acoustic stimuli for the experiment, was according to the ANSI S1.42 standard.

For the natural sounds, birdsongs and water sounds were selected as previous studies have characterized birdsongs [8,9,35] and water sounds [13–16] as pleasant sounds for improving soundscape quality. Recordings of four bird species found in Singapore were taken from the Macaulay Library at the Cornell Lab of Ornithology (ML47282551, ML176161, ML85119) and the archive of the Korean Broadcasting System (sparrow call). The four birdsongs (i.e. labeled B1 to B4) were selected based on differing temporal structures, as they were identified as critical perceptual characteristics in a prior study on birdsong selection for auditory masking [12]. The perceptual properties of water sounds can be characterized by their temporal variation (soft-variable/steady-state) and spectral envelope (high/low sharpness) [14]. Thus, four audio clips of water sounds (i.e. labeled W1 to W4) were selected to account for variations in their temporal and spectral characteristics. The four recordings of water sounds were obtained from Sonniss Limited. Hence, eight recordings of natural sounds (four bird and four water) were selected for the experiment.

The 10-second audio samples of the eight natural sounds were excerpted from the recordings and respectively mixed with the traffic noise stimulus via pointwise addition to generate an additional eight stimuli for the laboratory experiment. The A-weighted 10-s equivalent sound pressure level $L_{Aeq,10\mathrm{s}}$ of the traffic noise (T) was calibrated to 65 dB. On the other hand, the $L_{Aeq,10\mathrm{s}}$ of the eight natural sounds was set to 62 dB because a previous study found that setting natural sound levels to 3 dB less than the ambient traffic noise level was most preferable in the range tested [17,18]. The headphone playback of each acoustic stimulus was calibrated using a head and torso simulator (Brüel & Kjær 4128-C, Denmark).

Acoustic parameters describing the loudness, spectral contents, and temporal structure of the 8 stimuli were calculated as displayed in Table 2. Psychoacoustic parameters of the audio stimuli, specifically the *loudness* and *sharpness,* were calculated using the ArtemiS software package (HEAD acoustics GmbH, Germany). The time-varying *loudness* of the audio stimuli was calculated in accordance with ISO 532-1:2017 [49]. *Sharpness*, a sensation value of the timbre caused by high-frequency content in sound, was calculated according to DIN 45692 [50]. Additionally, the differences between the 10th and 90th percentage exceedance levels ($L_{A10} - L_{A90}$) were calculated to describe the overall variation of the equivalent sound levels over time [51,52].

The temporal patterns of *loudness* varied across the four water sounds, as shown in Fig. 2. Regarding the temporal variations, W1 and W3 can be characterized as steady-state sounds due to the lower value of $L_{A10} - L_{A90}$, whereas W2 and W4 can be characterized as soft-variable sounds with higher $L_{A10} - L_{A90}$ values. Regarding the spectral characteristics, the *sharpness* values of W3 (2.14 acum) and W4 (2.20 acum) were relatively higher than those of W1 (1.69 acum) and W2 (1.73 acum).

Table 2. Acoustic parameters of the audio stimuli used for this study. The $L_{Aeq,10s}$ of traffic noise (T) was set to 65 dB, whereas the $L_{Aeq,10s}$ of the water sounds (W1 to W4) and birdsongs (B1 to B4) were set to 62 dB.

Audio Stimuli	Description	*Loudness* [sone]	*Sharpness* [acum]	$L_{A10} - L_{A90}$ [dB]
T	Expressway	8.86	1.29	1.21
W1	Waterfall	41.90	1.69	0.63
W2	Water flow	43.50	1.73	2.67
W3	Fountain 1	33.90	2.14	0.88
W4	Fountain 2	33.50	2.20	2.03
B1	Banded woodpecker	19.60	2.17	8.82
B2	Sparrow	15.70	2.61	18.97
B3	Cerulean Warbler	28.80	3.32	35.99
B4	Pine Siskin	27.20	2.62	7.25

Regarding the temporal structure of the birdsongs, B1 and B4 can be characterized as quasi-steady due to continuous chirping, whereas B2 and B3 can be considered as intermittent sounds according to the time-domain plots in Fig. 2. This is because the $L_{A10} - L_{A90}$ values of B1 (8.82 dB) and B4 (7.25 dB) were much lower than those of B2 (18.97 dB) and B3 (35.99 dB).

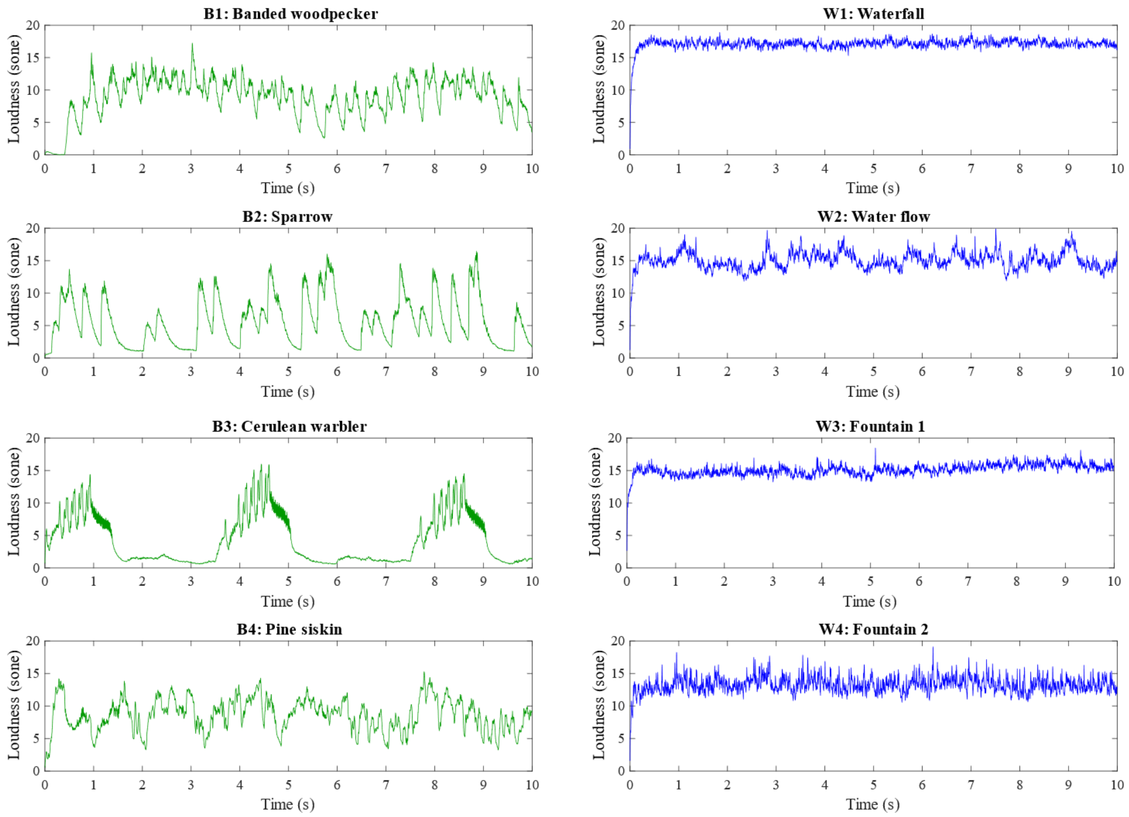

Figure 2. Loudness as a function of time for the selected birdsongs (B1 to B4) and water sounds (W1 to W4).

2.3 Experimental design and VR reproduction settings

In this study, a repeated-measures (RM) design, also known as a within-subjects design, was employed, which provides greater statistical power by controlling for differences between subjects. There were three sessions in the laboratory experiment: a visual-only, an audio-only, and a bimodal audio-visual session. The visual-only session was designed to determine the expected socio-recreational activities of the five locations based on the appropriateness of 12 pre-determined social and recreational outdoor activities excerpted from a previous study [32]. These activities are listed in Table 3. Participants were asked to evaluate the appropriateness of the 12 activities to each of the locations presented to them (in random order) through a VR HMD (Pimax 4K, China). The appropriateness was judged on a 7-point scale ('Entirely inappropriate', 'Mostly inappropriate', 'Somewhat inappropriate', 'Neither appropriate nor inappropriate', 'Somewhat appropriate', 'Mostly appropriate', and 'Entirely appropriate ') with the following question: "*To what extent do you think the location is suitable for each of the following activities?*"

The audio-only session used the nine 10-s audio stimuli, which were specifically the traffic noise stimulus (T) alone, and the 8 stimuli generated via pointwise addition of the natural sounds to the traffic noise stimulus (T+B1, T+B2, T+B3, T+B4, T+W1, T+W2, T+W3, and T+W4). The audio stimuli were presented to the participants through headphones (Beyerdynamic Custom One Pro, Germany) driven by a soundcard (Creative SoundBlaster E5, Singapore). Participants were asked to rate the pleasantness for each stimulus using a 7-point scale (i.e. 1: not at all pleasant and 7: extremely pleasant). Appropriateness of the audio stimuli was not assessed in the audio-only session because the concept of appropriateness is defined based on the context of the location and would thus necessitate a corresponding visual stimulus for a meaningful measurement.

Therefore, the bimodal audio-visual session evaluated both the appropriateness and pleasantness of combinations of artificially-generated visual and acoustic environments. A total of 45 audio-visual stimuli were generated (9 audio stimuli from audio-only session × 5 videos of the locations in the visual-only session). The participants were asked to assess the appropriateness and pleasantness of the soundscape in the given audio-visual stimuli using a 7-point scale (i.e. 1: not at all appropriate/pleasant, and 7: extremely appropriate/pleasant).

To provide an immersive and realistic audio-visual experience [53,54], omnidirectional videos of the five locations were presented via a VR HMD integrated with head-tracked first-order ambisonics (FOA) binaural rendering. This presentation method, also known as FOA-tracked binaural reproduction, enables a 3-degrees-of-freedom (3DoF) audio-visual experience. The Facebook Spatial Workstation Virtual Studio Technology (VST) plugin [55] for Reaper (version 5.4, USA) was employed to render the FOA-tracked binaural tracks. Positions of the traffic noise (T) and the natural sounds (W1 to W4; and B1 to B4) were rendered in the frontal direction of the participants to avoid spatial unmasking effects. In addition, the audio stimuli were equalized through inverse filtering with the measured headphone transfer function (HPTF) to avoid any changes to the frequency characteristics of stimuli due to the headphones used.

Table 3. List of the 12 socio-recreational activities in urban outdoor residential areas used for the visual-only session of the study

No.	**Socio-recreational activities**
1	Experiencing peace and quiet in general
2	Gardening/food-growing
3	Nature appreciation
4	Walking, jogging or running
5	Walking the dog
6	Using personal mobility devices
7	Children’s play
8	Informal outdoor games
9	People-watching
10	Socialising/conversing/chatting
11	Using electronic devices (e.g., smartphone)
12	Spending time with friends or family

2.4 Participants

A priori statistical power analysis was conducted to calculate the required minimum sample size for the within-subject design to achieve 80% power using G*Power 3.1 [56]. The power analysis suggested that 21 participants were needed to detect a medium effect: $f = 0.25$, $\alpha = 0.05$, and $(1 - \beta) = 0.80$. There were 50 participants (16 males and 34 females) in the experiment, which was more than twice the required number, thus indicating that this study had a probability of at least 80% to detect an effect that exists with a p-value of less than 0.05 in the statistical test.

The age distribution of the participants ranged from 19 to 26 years ($\mu_{\text{age}} = 21.4$, $\sigma_{age} = 1.9$). All the participants were locals who were familiar with the context of the chosen open spaces in the residential areas of Singapore. Before the experiment, a hearing test was conducted with an audiometer (Interacoustics AD629, Denmark) on all participants, and it was confirmed that all participants had normal hearing for all the tested frequencies (mean threshold of hearing <15dB at 0.125, 0.5, 1, 2, 3, 4, 6, and 8 kHz).

2.5 Procedure

In compliance with ethical procedures, formal ethical approval (IRB-2017-07-025) to conduct this experiment was obtained. The participants were informed about this study via written information, and written consent was obtained from all the participants. The audio-visual stimuli were presented to the participants in random order through a VR HMD and headphones. After experiencing each visual-only or audio-visual stimulus, the participants took off their VR HMD and completed the questionnaire. The participants were allowed to replay each stimulus as many times as required. The audio-only and visual-only session lasted between 10-15 minutes, and the audio-visual session lasted approximately 30-40 minutes. A mandatory break

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

time of at least 15 minutes was imposed between the sessions to relieve boredom and fatigue [57].

2.6 Statistical analyses

In the visual-only session, principal component analysis (PCA) was conducted to extract the main components of the socio-recreational activities in urban outdoor residential areas based on the subjective suitability ratings of the activities in each of the 5 locations.

For the audio-only and audio-visual sessions, two-way repeated-measures analysis of variance (RM ANOVA) tests were conducted to investigate the within-subjects effects of the audio stimuli, locations, and interaction between the audio stimuli and locations in the perceived appropriateness and pleasantness of soundscapes. The assumption of sphericity for the dataset was tested using Mauchly's test of sphericity. When the assumption of sphericity was violated, the Greenhouse–Geisser correction was applied and the corrected degrees of freedom of the F-distribution and p-values were reported.

Post hoc comparisons were conducted with Bonferroni correction. Partial eta squared (η_p^2) values were reported as an effect size measure. In addition, simple effect analyses were performed when the interaction effects were significant because the main effect is only meaningful when there is no interaction between the two independent variables. All statistical analyses were conducted using the statistical software package SPSS (version 23.0, IBM, USA).

3. Results

This section analyses the subjective responses obtained from the three sessions. In Section 3.1, the subjective responses obtained from the visual-only session are analyzed to characterize the expected socio-recreational activities in the selected locations. Using the subjective responses of the audio-only and audio-visual sessions, Section 3.2 examines how the visual and auditory components in the locations affect the perceived pleasantness of their soundscapes. In Section 3.3, the effects of the audio and visual stimuli on the appropriateness of soundscape are investigated based on the results in the audio-visual session.

3.1 Principal components of outdoor activities in urban residential areas

To identify the main components of outdoor activities in urban residential areas, PCA was conducted based on the responses of the visual-only session. Varimax rotation was applied to extract orthogonal components. As shown in Table 4, three components with eigenvalues larger than 1 were obtained. The Kaiser–Mayer–Olkin (KMO) measure of the sampling adequacy was 0.78 and Bartlett's test of sphericity was also significant (χ^2 (66) = 995.88, and $p < 0.001$), which indicates that the data set is appropriate for PCA. Components 1, 2, and 3 explained 34.7%, 13.3%, 10.6% of the variance in the data set, respectively.

Component 1, interpreted to represent *relaxation*, was highly associated with relaxation activities: ‘experiencing peace and quiet in general’, ‘Gardening/food-growing’ and ‘Walking, jogging or running’. Component 2, described as *outdoor play*, had higher component loadings with play activities: ‘Children’s play’ and ‘Informal outdoor games’. Component 3, characterized as a *social gathering*, was highly related to social activities: ‘socializing/conversing/chatting’ and ‘Spending time with friends or family’.

To characterize the five locations with respect to the extracted components, the component scores were calculated using the regression method. Fig.3 shows the mean component scores for each component, plotted as a function of the main components for each location. For active zones, A_1 (children’s playground in park) exhibited positive *outdoor play* and *relaxation* scores but negative *social gathering* scores (see Fig 3(a)), whereas $\widehat{A}_2$ (children’s playground

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

near road) exhibited highly positive scores for *outdoor play* and negative scores for both *relaxation* and *social gathering* (see Fig. 3(b)).

Table 4. Rotated component matrices of the PCA using subjective responses for the 12 socio-recreational activities (numbers in parentheses represent explained variance)

Socio-Recreational activities	Component (Explained variance, %)		
	1: *Relaxation* (34.70)	2: *Outdoor play* (13.29)	3: *Social gathering* (10.63)
Experiencing peace and quiet in general	**0.78**	-0.25	0.30
Gardening/food-growing	**0.77**	0.05	-0.09
Nature appreciation	**0.75**	0.31	0.12
Walking, jogging or running	**0.63**	0.52	-0.05
Walking the dog	**0.55**	0.50	0.07
Using personal mobility devices	**0.49**	0.38	0.14
Children's play	0.00	**0.78**	0.17
Informal outdoor games	0.21	**0.73**	0.02
People-watching	0.09	**0.56**	0.36
Socialising/conversing/chatting	0.04	0.16	**0.77**
Using electronic devices (e.g., smart phone)	0.00	0.02	**0.76**
Spending time with friends or family	0.32	0.34	**0.52**

The walkway in a neighborhood park P_1 was dominated by the *relaxation* component, as shown in Fig. 3(c). The rooftop garden P_2 showed neutral component scores for both *relaxation* and *social gathering*, but a negative mean score for *outdoor play* as seen in Fig. 3(d). Unsurprisingly, $\hat{P}_3$, which comprised of tables and benches, showed higher *social gathering* scores than the other locations, as shown in Fig. 3(e).

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

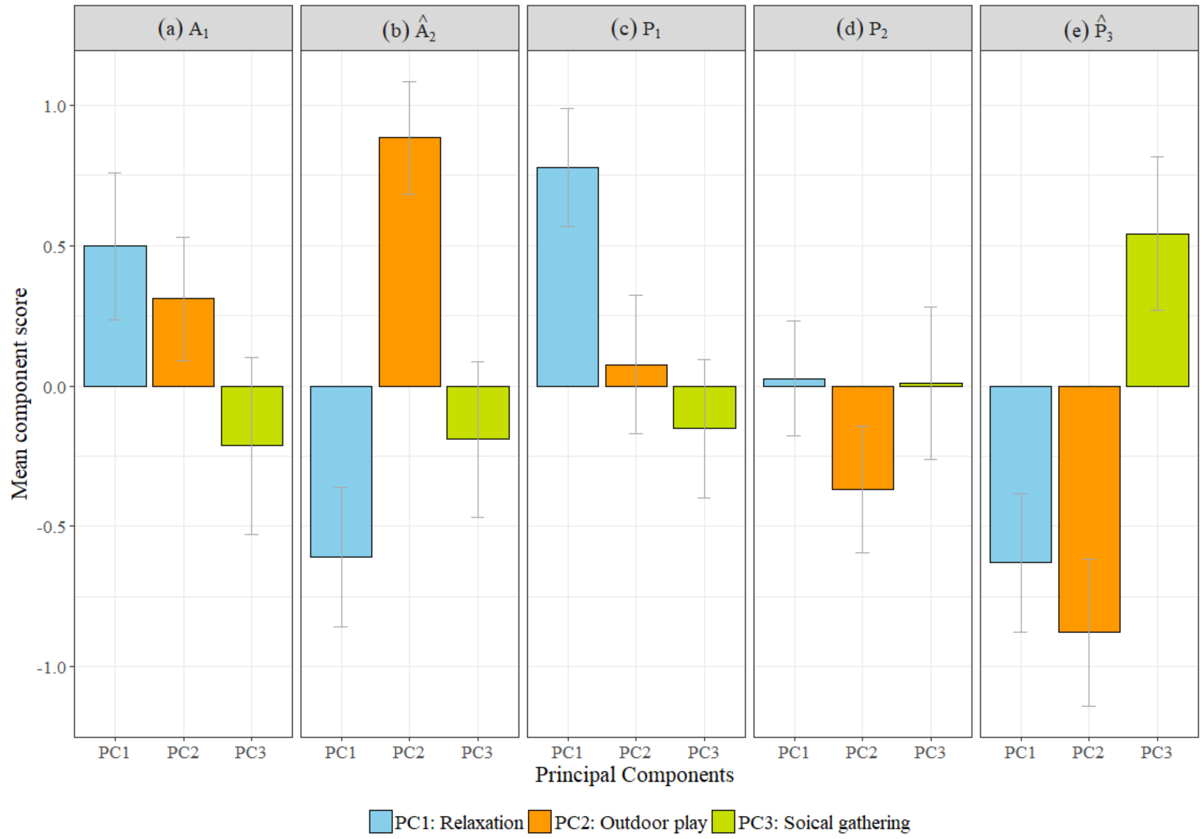

Figure 3. Mean principal component (PC) scores for the socio-recreational activities across the five locations denoted by the subplot labels. The abscissa indicates principal components for the socio-recreational activities; PC1, PC2, and PC3 are *relaxation, outdoor play, and social gathering*, respectively. The error bars indicate 95% confidence intervals.

3.2 Pleasantness of natural sounds

A two-way RM ANOVA was conducted on the results of the audio-only and audio-visual sessions to investigate the main effects of location and audio stimulus on the rated pleasantness. T he results for the audio-only session were treated as results for an additional location (with no video) in the audio-visual experiment. In other words, there were six levels for the independent variable "location", corresponding to the five locations in the audio-visual session and the additional set of results from the audio-only session. The independent variable 'audio stimulus' consisted of nine levels represented by each acoustic stimulus (i.e. T, T+B1 to T+B4, and T+W1 to T+W4) used for the audio-only and audio-visual sessions. The results showed that the main effects of locations [$F(3.28, 160.75) = 2.92, \eta_p^2 = 0.06,\ p = 0.031$] and audio stimuli [$F(4.26, 208.63) = 19.70, \eta_p^2 = 0.29,\ p < 0.001$] on pleasantness were significant. The interaction effect between locations and audio stimuli was also significant [$F(17.20, 843.02) = 4.79, \eta_p^2 = 0.09,\ p < 0.001$]. Hence, the simple effects of the locations and audio stimuli on the pleasantness of the soundscape are also analyzed in Sections 3.2.1 and 3.2.2, respectively.

3.2.1 Effect of location on pleasantness

The results of the F-tests for simple effects of locations on the pleasantness for each acoustic stimulus are summarized in Table 5. The significance of the *F*-values was tested at a 0.006 (0.05/9) significance level to compensate for the inflation of the family-wise error rate. The simple effects of locations were significant for T, T+W1, and T+W3, whereas no significant simple effects were found for the other stimulus types.

Post hoc tests were conducted to find significant differences in terms of the locations for each audio stimulus. Mean pleasantness scores of the acoustic stimuli are plotted as a function of the locations in Fig. 4. As shown in Fig. 4(a), the mean pleasantness score of the traffic noise T at location $\widehat{\mathrm{P}}_3$ was significantly higher than that at A_1. No differences were found in the pleasantness scores among the birdsongs across the five locations as presented in Figs. 4(b-e). However, for water sounds, statistically significant differences were observed in T+W1 and T+W3 – steady-state water sounds combined with traffic noise – between the audio-only and audio-visual sessions as shown in Figs. 4(f) and 4(h), respectively. Specifically, the pleasantness scores for T+W1 (P_2 and $\widehat{\mathrm{P}}_3$) and T+W3 (A_1, $\widehat{\mathrm{A}}_2$, and P_2) were significantly lower in the audio-visual sessions than in their respective audio-only sessions.

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

Table 5. Summary of the RM ANOVA showing the simple effects of the locations on the pleasantness of soundscape in each acoustic stimulus. Audio stimuli T, B, and W designate traffic noise, birdsong, and water sounds, respectively; the '+' sign denotes a pointwise addition of stimuli.

Audio Stimuli	Factor	df_1	df_2	F	p	η_p^2
T	Location†	3.98	195.09	4.84	0.001	0.09
T+B1	Location	5.00	245.00	3.08	0.010	0.06
T+B2	Location†	4.19	205.41	3.36	0.010	0.06
T+B3	Location	5.00	245.00	3.26	0.007	0.06
T+B4	Location†	3.74	183.29	2.63	0.040	0.05
T+W1	Location†	4.13	202.37	5.54	<0.001	0.10
T+W2	Location†	3.84	188.24	3.13	0.017	0.06
T+W3	Location†	3.81	186.47	8.41	<0.001	0.15
T+W4	Location	5.00	245.00	3.59	0.006	0.07

† Assumption of sphericity was violated, and Greenhouse–Geisser correction was applied.

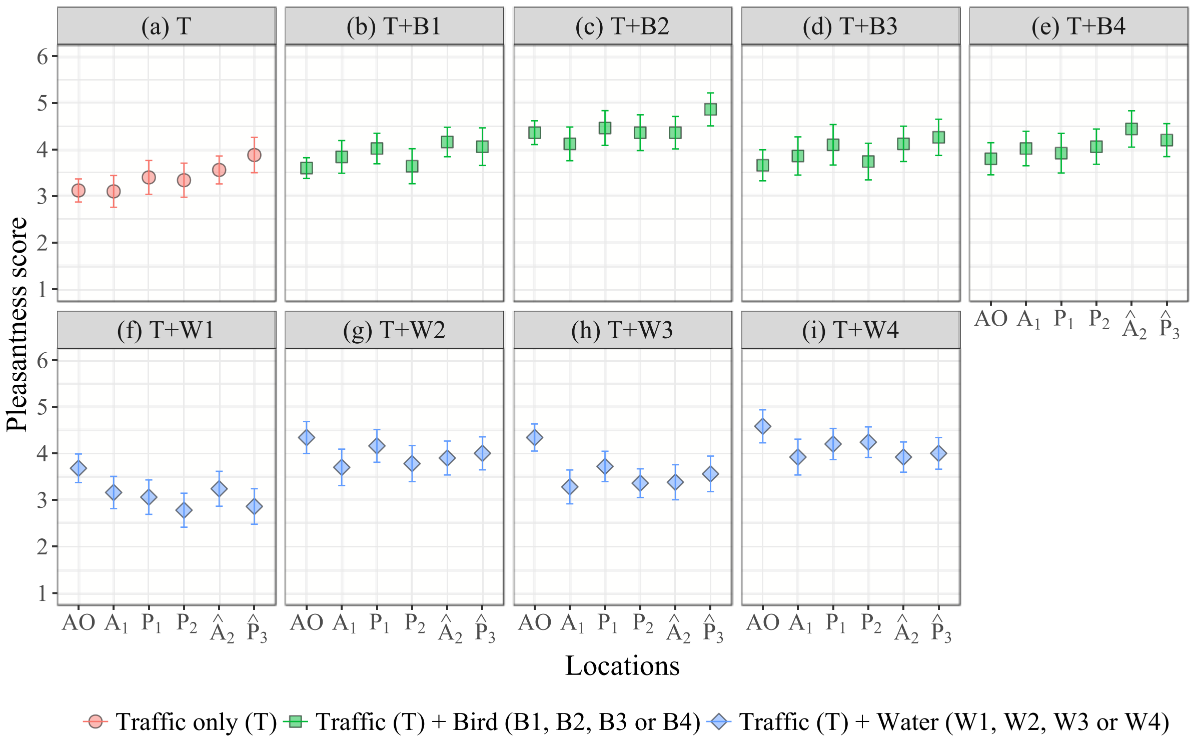

Figure 4. Mean pleasantness scores as a function of locations for all nine audio stimuli. The subplot labels denote the stimuli type, and '+' denotes a pointwise addition of stimuli. The abscissa indicates the locations. 'AO' indicates audio-only condition without visual stimuli and the error bars indicate 95% confidence intervals.

3.2.2 Effect of audio stimulus on pleasantness

The mean pleasantness scores for the nine stimuli are plotted across the locations in Fig. 5. The simple effects of audio stimuli on pleasantness in each location were tested at a 0.008 (0.05/6) significance level as shown in Table 6. The results show that the simple effects of audio stimuli on pleasantness were statistically significant in all the locations.

Post hoc tests were conducted to examine the effect of the audio stimuli in each location. Amongst the birdsongs in the audio-only session, only the sparrow chirp T+B2 when added to the traffic noise enhanced the pleasantness as compared to the traffic noise T alone ($p < 0.001$), as can be seen from Fig. 5(a). On the other hand, the addition of water sounds W2, W3 and W4

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

to the traffic noise (to make T+W2, T+W3, and T+W4 respectively) significantly increased the pleasantness (p < 0.001), whereas there was no significant increase due to the addition of W1 (to make T+W1).

Table 6. Summary of the RM ANOVAs: Simple effect of audio stimulus on the rated pleasantness of soundscape in each location. ‘AO’ indicates an audio-only condition without visual stimuli.

Location	Factor	df_1	df_2	F	p	η_p^2
A_1	Audio Stimuli†	4.89	239.63	9.00	<0.001	0.16
P_1	Audio Stimuli†	5.39	264.24	10.63	<0.001	0.18
P_2	Audio Stimuli†	5.16	252.78	11.78	<0.001	0.19
$\widehat{A}_2$	Audio Stimuli†	4.91	240.43	8.12	<0.001	0.14
$\widehat{P}_3$	Audio Stimuli†	5.39	264.28	12.18	<0.001	0.20
AO	Audio Stimuli†	5.80	284.04	13.55	<0.001	0.22

†Assumption of sphericity was violated, and Greenhouse–Geisser correction was applied.
^Road traffic was visible.

In the audio-visual session, similar mean pleasantness scores for the birdsongs were found with those in the audio-only session. Amongst the birdsongs, the excerpt of the sparrow call when combined with the traffic noise (T+B2) significantly increased the pleasantness across all five locations as compared to the traffic noise alone. Meanwhile, the excerpt corresponding to the banded woodpecker (B1) did not enhance the pleasantness at any location, as can be observed in Figs. 5(b-f). For B3 and B4, improvements in pleasantness were only found at location A_1 compared to T only, as shown in Fig. 5(b).

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

Interestingly, the pleasantness as a result of adding water sounds in the audio-visual session dramatically differed from the audio-only session. Most of the water sounds (i.e. W1, W2, and W3) did not enhance the pleasantness of soundscape in the five locations as presented in Figs. 5(b-f). Only W4, a soft-variable fountain sound, significantly improved the pleasantness at locations A_1, P_1, and P_2 ($p < 0.001$), while the effects of W4 were not significant at the locations $\hat{A}_2$ and $\hat{P}_3$. These results demonstrate that the judged pleasantness of water sounds largely depends on the visual context as compared to that of birdsongs.

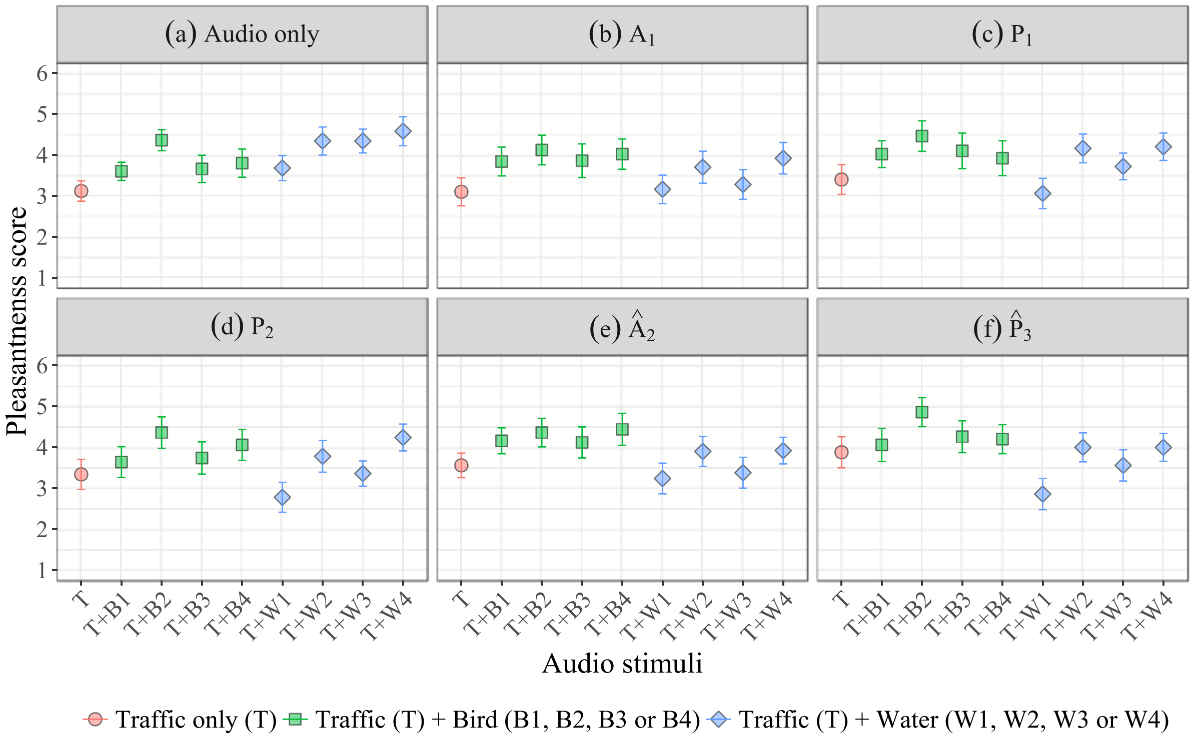

Figure 5. Mean pleasantness scores as a function of audio stimulus across the audio-only session and five locations in the audio-visual session. The locations are denoted by the subplot labels. The abscissa indicates audio stimuli, where T, B, and W designate traffic noise, birdsong, and water sounds, respectively; '+' denotes a pointwise addition of stimuli. The error bars indicate 95% confidence intervals

3.3 Appropriateness of natural sounds

A two-way RM ANOVA was conducted to examine the statistically significant mean differences in the appropriateness of soundscape according to the locations and audio stimuli using the subjective responses from the audio-visual session. Therefore, two independent variables (the locations and audio stimuli) were described by five and nine different levels, respectively.

The results showed that the main effects of locations [$F(3.03, 148.35) = 21.36$, $\eta_p^2 = 0.30$, $p < 0.001$] and audio stimuli [$F(3.80, 186.19) = 107.24$, $\eta_p^2 = 0.69$, $p < 0.001$] were significant. The interaction (locations × audio stimuli) was also significant [$F(13.54, 663.62) = 19.51$, $\eta_p^2 = 0.29$, $p < 0.001$]. Thus, we elaborate on the simple effects of the locations and the audio stimuli on the appropriateness of the soundscape in Sections 3.3.1 and 3.3.2 respectively.

3.3.1 Effect of location on appropriateness

To analyze the simple effects of the locations on the appropriateness of the soundscape, sets of one-way RM ANOVA were conducted by stimulus. The significance of the locations in each stimulus was tested at a 0.006 (0.05/9) significance level considering the inflation of the family-wise error rate. The simple effects of the locations were significant across all audio stimuli, as shown in Table 7.

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

Table 7. Summary of the RM ANOVA showing the simple effect of locations on the rated appropriateness of soundscape for each acoustic stimulus. Audio stimuli T, B, and W denote traffic noise, birdsong, and water sounds, respectively.

Stimuli	Factor	df_1	df_2	F	p	η_p^2
T	Location†	2.95	144.56	63.08	<0.001	0.56
T+B1	Location	4.00	196.00	20.51	<0.001	0.30
T+B2	Location†	3.19	156.44	17.69	<0.001	0.27
T+B3	Location†	3.31	162.14	10.57	<0.001	0.18
T+B4	Location†	3.41	166.94	13.42	<0.001	0.22
T+W1	Location	4.00	196.00	4.07	<0.001	0.08
T+W2	Location†	2.89	141.46	18.00	<0.001	0.27
T+W3	Location	4.00	196.00	6.56	<0.001	0.12
T+W4	Location†	2.52	123.35	16.28	<0.001	0.25

†Assumption of sphericity was violated, and Greenhouse–Geisser correction was applied

Figure 6 shows the mean appropriateness scores of the locations for the nine audio stimuli. Post hoc tests showed that the appropriateness scores of the traffic noise alone T at locations $\widehat{\mathrm{A}}_2$ and $\widehat{\mathrm{P}}_3$ were significantly higher than those at the other three locations ($p < 0.001$), as shown in Fig. 6(a). Similarly, the participants evaluated that the birdsongs when combined with traffic noise (T+B1 to T+B4) were more appropriate at locations $\widehat{\mathrm{A}}_2$ and $\widehat{\mathrm{P}}_3$ than at the other locations ($p < 0.001$), as shown in Figs. 6(b-e). These results demonstrate that the visibility of road traffic (i.e. $\widehat{\mathrm{A}}_2$ and $\widehat{\mathrm{P}}_3$) enhances the appropriateness of soundscape at locations usually exposed to traffic noise.

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

Interestingly, the appropriateness of the water sounds differed across all water audio stimuli (i.e. W1 to W4). The overall effects of traffic with steady-state water sounds T+W1 and T+W3 were not significant, as shown in Figs. 6(f) and (h), respectively. However, the appropriateness of soft-variable water sounds, T+W2 and T+W4, were significantly greater at P_1, an area beside a waterway, than those at the other locations ($p < 0.001$), as shown in Figs. 6(g) and (i).

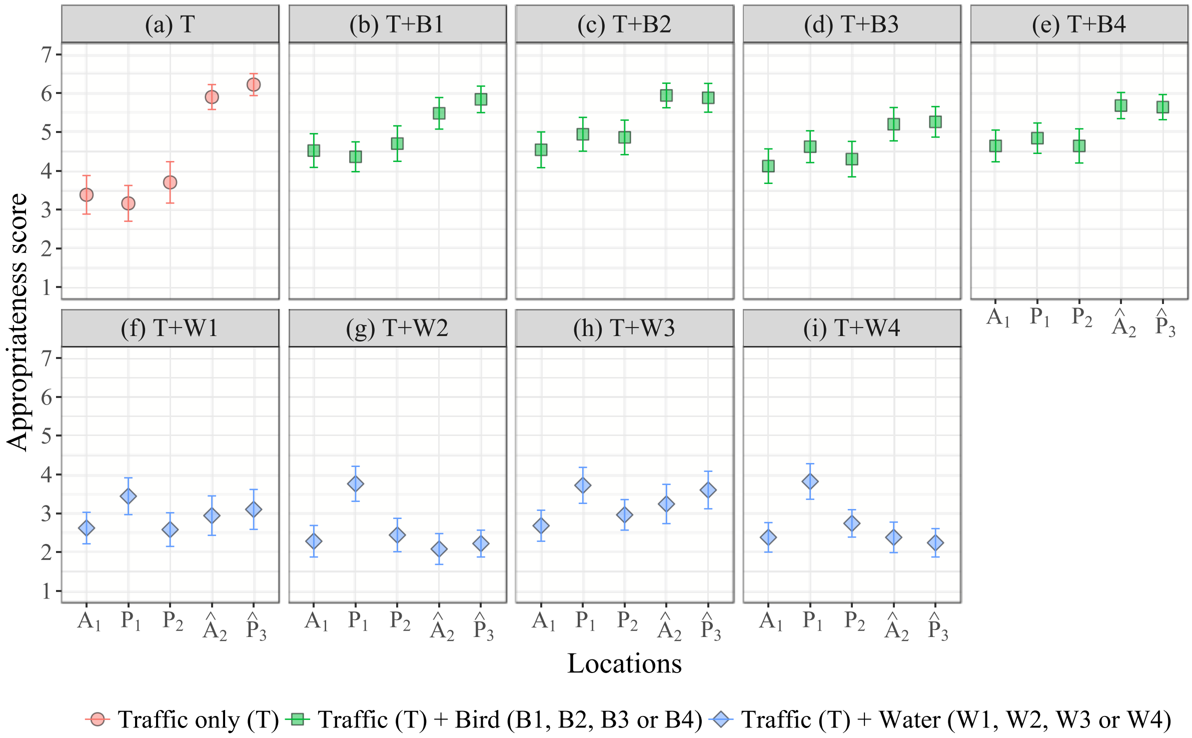

Figure 6. Mean appropriateness scores as a function of the locations across the nine audio stimuli denoted by the subplot labels. T, B, and W denote traffic noise, birdsong, and water sounds, respectively; '+' denotes a pointwise addtion of stimuli. The error bars indicate 95% confidence intervals.

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

3.3.2 Effect of audio stimuli on appropriateness

A series of one-way RM ANOVA were conducted for each audio stimuli to examine the simple effects of audio stimuli on the appropriates of the soundscape. Due to the inflation of the family-wise error rate for simple effect analyses, the significance of audio stimuli in each location was tested at a 0.01 (0.05/5) significance level. As shown in Table 8, The simple effects of the audio stimuli were significant across the five locations.

Table 8. Summary of the RM ANOVAs: Simple effects of audio stimuli on the appropriateness of soundscape in each location.

Location	Factor	df_1	df_2	F	p	η_p^2
A_1	Audio Stimuli†	4.19	205.12	40.32	<0.001	0.45
P_1	Audio Stimuli†	5.92	290.28	18.54	<0.001	0.27
P_2	Audio Stimuli†	4.50	220.57	32.54	<0.001	0.40
$\widehat{A}_2$	Audio Stimuli†	3.88	190.13	80.79	<0.001	0.62
$\widehat{P}_3$	Audio Stimuli†	4.47	219.11	96.04	<0.001	0.66

†Assumption of sphericity was violated, and Greenhouse–Geisser correction was applied

⌒Minor road was visible

The mean appropriateness scores for the nine audio stimuli as a function of the five locations are depicted in Figs. 7(a-e). Post hoc tests revealed that the mean appropriateness scores of the nine audio stimuli significantly differed across the five locations. Amongst the birdsongs, there were no significant differences observed across the locations.

Comparing the water sounds, the appropriateness of T+W3 (traffic with stream sound) was significantly higher than that of T+W2 (traffic with steady-state fountain sound) $(p < 0.01)$

at both $\widehat{A}_2$ and $\widehat{P}_3$, as shown in Figs. 7(d) and 7(e), respectively. Meanwhile, there were no significant differences among the other water sounds across the five locations.

Regarding the types of sound sources, there were also significant differences in the appropriateness between the three different audio stimuli types (traffic, birdsongs and water sounds) and their differences varied across five locations. In comparison, birdsongs were evaluated as more appropriate than water sounds at locations A_1, P_2, $\widehat{A}_2$ and $\widehat{P}_3$. Only at location P_1 was there no significant difference in the appropriateness between T+B1 (traffic with woodpecker) and T+W2 to T+W4, as shown in Fig. 7(b).

As shown in Figs 7(a-c), there were significant differences in the appropriateness scores between the traffic-only stimulus (T) and traffic combined with birdsongs (i.e. T+B1 to T+B4) at locations A_1, P_1 and P_2. Meanwhile, there were no significant differences in appropriateness scores between T and T+B1 to T+B4 at both locations $\widehat{A}_2$ and $\widehat{P}_3$, where the road traffic was visible, as shown in Figs. 7(d) and (e), respectively.

Figs. 7(a-c) also show that there were no siginificant differences in the appropriateness rating score at locations A1, P1, and P2 between T and T+W1 to T+W4. However, the appropriateness of T was significantly higher than traffic combined with water sounds (T+W1 to T+W4) at locations $\widehat{A}_2$ and $\widehat{P}_3$, as illustrated in Figs. 7(d) and (e).

Accepted version before copy editing
Published article: https://doi.org/10.1016/j.scs.2020.102475

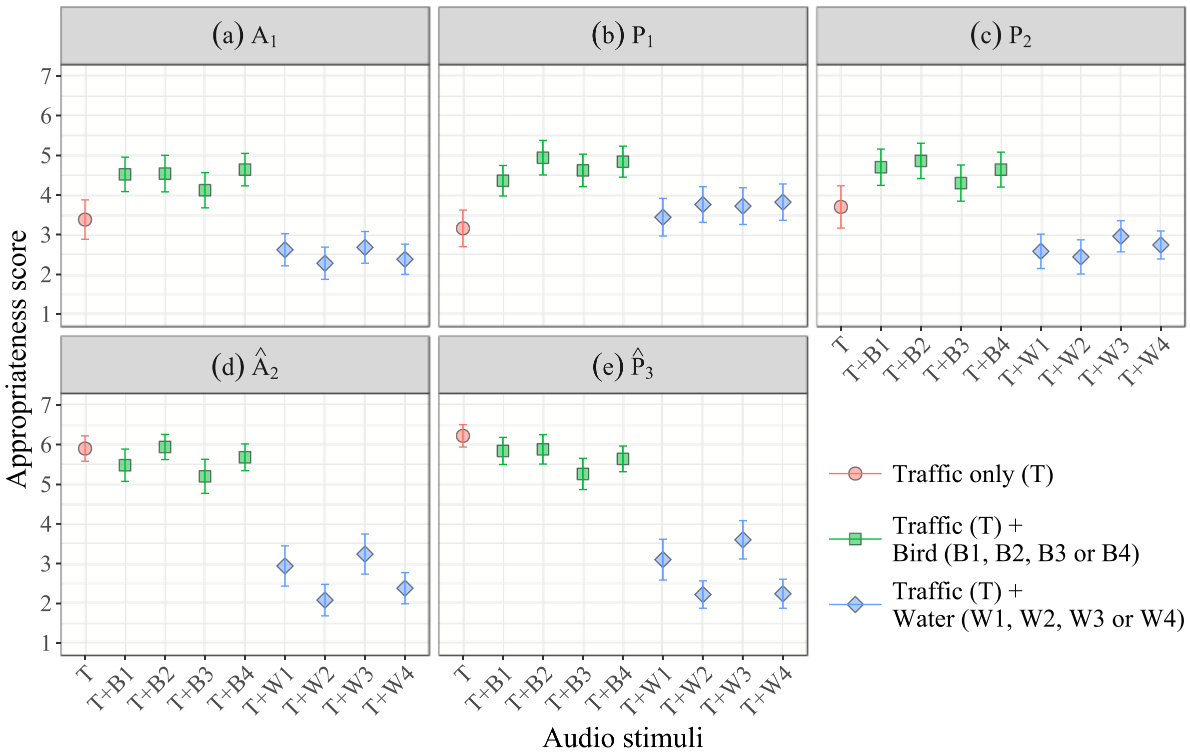

Figure 7. Mean appropriateness scores as a function of audio stimuli across the five locations (denoted by the subplot labels). Error bars indicate 95% confidence intervals.

4. Discussion

Sections 4.1 and 4.2 aim to answer the two main research questions posed in Section 1 in the order that they were posed, and Section 4.3 addresses the implications of the findings of this study and its inherent limitations.

4.1 Effect of expected human activities on perceptions of natural sounds

To explore the relationships between the appropriateness and pleasantness of soundscape and the expected socio-recreational activities in outdoor residential areas, Pearson's correlation coefficients were calculated from the principal component scores of the socio-recreational activities in the visual-only session, and from the appropriateness and pleasantness rating scores in the audio-visual session for birdsongs and water sounds. As summarized in Table 9, the components of socio-recreational activities in urban outdoor residential areas had no or very weak correlations with pleasantness and appropriateness for natural sounds. In particular, the birdsongs and water sounds showed weak correlations with the principal components *social gathering* (r = 0.17, p < 0.01) and *relaxation* (r = 0.18, p < 0.01), respectively. The principal component *outdoor play* was uncorrelated to the appropriateness and pleasantness for both birdsongs and water sounds. This demonstrates that the expected socio-recreational activities in urban residential outdoor areas do not significantly affect pleasantness and appropriateness of natural sounds as maskers.

Table 9. Pearson's correlation coefficients between the appropriateness and pleasantness scores of soundscape, and principal component scores of socio-recreational activities in residential areas. $^{*}p$<0.05, $^{**}p$<0.01

Sound	Soundscape descriptors	Components of socio-recreational activities		
		Relaxation	*Outdoor play*	*Social gathering*
Bird	Appropriateness	-0.07^{*}	-0.01	0.17^{**}
	Pleasantness	-0.02	-0.02	0.05
Water	Appropriateness	0.18^{**}	-0.03	-0.01
	Pleasantness	0.07^{*}	0.03	-0.00

In addition, as presented in Fig. 8, RM ANOVA results showed that there were no significant differences between the active and passive zones in terms of appropriateness and pleasantness scores for each type of sound. This supports the finding that perceptions, at least in terms of appropriateness and pleasantness, of natural sounds are not affected by the functions of spaces in urban residential areas.

These results can be explained by the fact that natural sounds are less psychologically associated with social and recreational activities in a residential place. This is in line with the findings of Hong and Jeon [27] that natural sounds are not directly related to the appropriateness of soundscapes in urban outdoor residential areas. Meanwhile, human-generated sounds such as conversations and sounds of playing children might be more closely associated with socio-recreational activities in outdoor areas because several studies [27,39,51] have found that human-generated sounds play a critical role in constructing an appropriate soundscape, particularly in places for recreation and socializing.

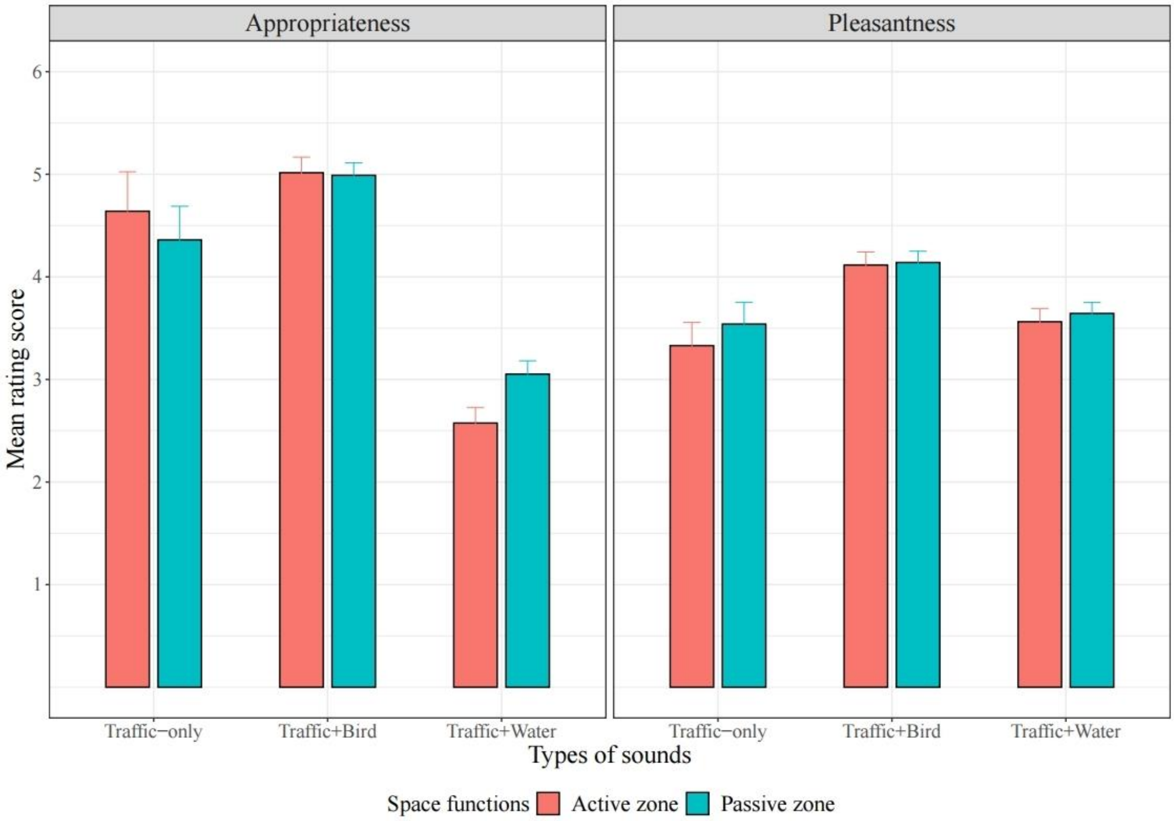

Figure 8. Mean rating scores of appropriateness and pleasantness in terms of types of sounds and functions of outdoor spaces in residential areas. Error bars indicate 95% confidence intervals.

4.2 Effect of sound source visibility on perceptions of sound environments

The results presented here suggest that the congruency between acoustic and visual environment is a critical factor in the perceived pleasantness and appropriateness of soundscape. In terms of pleasantness, at locations where the minor roads were visible (i.e. $\widehat{\mathrm{A}}_2$ and $\widehat{\mathrm{P}}_3$) the traffic noise was judged as less annoying than in those locations where the road traffic was not visible. This result corresponds well with the findings of previous studies. For instance, Watts et al. [43] observed that the judged noise annoyance for the same sound pressure level of traffic noise was higher when the degree of visibility of the traffic source was higher. Aylor and Marks [44] also discovered that the perceived loudness of traffic noise was lower when the traffic was visible, whereas the perceived loudness increased when the traffic was blocked by a noise barrier. The findings of the present study support that people are more sensitive to traffic noise in terms of perceived loudness and annoyance when the source is unseen.

Regarding the appropriateness of soundscape, this study confirms that the appropriateness of traffic noise increased significantly when the traffic source was visible. These results also corroborate the findings of previous studies [32,58] that soundscape could be appropriate to a place although the acoustic quality is poor because the concept of appropriateness depends on the congruency between the acoustic environment and the context of a given place.

For stimuli where traffic noise was combined with birdsongs, there were no significant differences in pleasantness scores between the audio-only and audio-visual sessions (when the sessions were compared as different locations). Additionally, the appropriateness scores for the birdsongs were not significantly affected by the visibility of the sound source. In other words, even though the participants could not see birds in the locations, they evaluated that soundscapes with birdsongs were appropriate in an urban context. These results seem to reinforce the notion that hearing birdsongs in our daily life without the visibility of the birds is

Published article: https://doi.org/10.1016/j.scs.2020.102475

commonplace and hence judged as relatively appropriate. Alternatively, it can be postulated that the appropriateness of birdsongs could be linked to the presence of vegetation such as trees or bushes in a given location. In this study, the participants might have evaluated that the added birdsongs were relevant because all locations in the experiment had visible vegetation (e.g. trees, grass, or bushes) as shown in Table 1 that could indirectly indicate the presence of birds. This is bolstered by a finding by Liu et al. [10], who observed that the perceived loudness of birdsong had a positive correlation with vegetation density. Hao et al. [11] also supported the identification of birdsongs as closely related to urban greenery indicators. Furthermore, Hong and Jeon [35] reported that the combination of images of vegetation and birdsong could have a synergetic effect in improving the soundscape quality in an urban street.

Contrary to the observations regarding birdsongs, both the appropriateness and pleasantness of water sounds were largely determined by the visibility of the water sources. Particularly, the effect of water source visibility on appropriateness was greater than that on pleasantness. This indicates that the visibility of water features is more closely associated with the appropriateness of water sounds. Additionally, it was found that relevant water sounds corresponding to the existing water features in the location could enhance its appropriateness. For instance, only soft-variable water sounds (i.e. W2 and W4) were judged to be significantly more appropriate at location P1 (see Fig. 7); P_1 had a waterway, which is expected to produce soft-variable water sounds. These results regarding water sounds correspond well with the finding of Jeon et al. [13] that the percentage of water features in the visual stimuli showed a positive correlation with the preference of water sounds. The findings of this study demonstrate that as the sound and visual design components in a given location are highly matched, the soundscape designs could be enhanced [34].

4.3 Implications in soundscape design and its limitations

One of the critical findings in this study is that congruency between audio and visual components plays a key contextual role in soundscape design when adding natural sounds, whereas the expected socio-recreational activities in the locations were not. This implies that soundscape design without considering congruency with visual components in the locations might not guarantee the enhancement of pleasantness and appropriateness of the soundscape.

This study also demonstrated that soundscape design approaches by natural sound augmentation should be dependent on the types of natural sounds (e.g., birdsongs or water sounds) due to different interactions between the audio-visual factors regarding types of natural sounds. Soundscape design by natural sound augmentation can be achieved by deployment of real sound sources (e.g., water fountains or trees) or by installing active systems based on loudspeakers reproducing those real sound sources.

In the case of water sounds, using actual water features would be more effective in increasing both the pleasantness and appropriateness of the soundscape due to audio-visual congruency. Presenting water sounds through an invisible loudspeaker system might be less effective owing to the audio-visual incoherence.

Furthermore, planting trees or vegetation could be a valid soundscape design strategy to introduce birdsongs into a real-world setting. However, greenery in urban areas does not always guarantee the presence of birdsongs and unlike water features, the sound levels and types of birdsongs are also beyond the designers' control. Alternatively, introducing birdsongs via a speaker system could a more reliable design strategy if there is surrounding vegetation to justify the appropriateness of the presence of birdsongs.

There remain some inherent limitations in this study. One potential limitation is related to the limited age distribution of the participants in this study. The participants in this study were mainly in their 20s. Although all participants in this study were local residents in Singapore, the expected socio-recreational activities related to the locations and the perceptions of sound environments might be affected by age groups. This might make the findings of this study less generalizable to a broader range of age groups.

The number of locations selected in this study might be another limitation. Although five locations were chosen to represent active and passive activity zones in residential outdoor areas in this study, there is an extremely wide variety of design elements used in active and passive activity zones that may not have been adequately represented in this study. Hence, more diverse locations for passive and active activities could be included in a future study for generalization beyond residential areas.

It should also be noted that the effects of natural sound augmentation were studied using the same traffic noise in the absence of the main activities to control for other audio-visual factors generated by human activity. However, the appropriateness of soundscape could also be affected by active sounds from the main activities in the location [27,32]. Therefore, a future study could investigate the effects of natural sound augmentation in the presence of the main activities.

5. Conclusions

The effects of expected human activities and audio-visual congruency of a location on the perception of a soundscape consisting of traffic noise augmented with natural sounds were investigated in various urban outdoor residential contexts through laboratory experiments. Birdsongs and water sounds were evaluated as soundscape design elements to improve the pleasantness and appropriateness of a traffic soundscape considering the location's context. It was found that three main PCA-derived components of outdoor activities, labelled as *relaxation, outdoor playing*, and *social gathering*, had no significant effect on the perceived pleasantness and appropriateness for the augmented natural sounds.

In contrast, we observed that the degree of congruency between the aural and visual stimuli significantly influenced the judged pleasantness and appropriateness. Overall, when the audio-visual scenes were highly matched, the pleasantness and appropriateness of the soundscape were improved. Interestingly, the perception of water sounds was significantly affected by sound source visibility but not the perception of birdsongs. When water features were not visible, the pleasantness and appropriateness ratings decreased. However, the same ratings for birdsongs appeared to be independent of the visibility of birds, likely due to the presence of vegetation as an indirect visual indicator. The findings of this study suggest that audio-visual coherence is a critical factor in determiningappropriate types of natural sounds for soundscape interventions at a given location.

Acknowledgments

This research is supported by the Singapore Ministry of National Development and the National Research Foundation, Prime Minister's Office under the Land and Liveability National Innovation Challenge (L2 NIC) Research Programme (L2 NIC Award No. L2NICCFP2-2015-5). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Singapore Ministry of National Development and National Research Foundation, Prime Minister's Office, Singapore." The first author is currently supported by the 2017 Lee Kuan Yew Postdoctoral Fellowship (Grant No.: 50025000).

References

[1] M. Yuan, C. Yin, Y. Sun, W. Chen, Examining the associations between urban built environment and noise pollution in high-density high-rise urban areas: A case study in

Wuhan, China, Sustain. Cities Soc. 50 (2019) 101678. https://doi.org/10.1016/j.scs.2019.101678.

[2] A. Karvonen, Toward the Healthy City: People, Places, and the Politics of Urban Planning, Int. J. Urban Reg. Res. 35 (2011) 469–470. https://doi.org/10.1111/j.1468-2427.2011.01040_3.x.

[3] J. Kang, From dBA to soundscape indices: Managing our sound environment, Front. Eng. Manag. 4 (2017) 184–194. https://doi.org/10.15302/J-FEM-2017026.

[4] Eurostat, Sustainable development in the European Union - Monitoring report on progress towards the SDGs in an EU context, 4th ed., Publications Office of the European Union, Luxembourg, 2020. https://doi.org/10.2785/555257.

[5] J. Kang, F. Aletta, T.T. Gjestland, L.A. Brown, D. Botteldooren, B. Schulte-Fortkamp, P. Lercher, I. van Kamp, K. Genuit, A. Fiebig, J.L. Bento Coelho, L. Maffei, L. Lavia, Ten questions on the soundscapes of the built environment, Build. Environ. 108 (2016) 284–294. https://doi.org/10.1016/j.buildenv.2016.08.011.

[6] J.Y. Hong, Z.-T. Ong, B. Lam, K. Ooi, W.-S. Gan, J. Kang, J. Feng, S.-T. Tan, Effects of adding natural sounds to urban noises on the perceived loudness of noise and soundscape quality, Sci. Total Environ. 711 (2020) 134571. https://doi.org/10.1016/j.scitotenv.2019.134571.

[7] D.S. Brungart, Informational and energetic masking effects in the perception of multiple simultaneous talkers, J. Acoust. Soc. Am. 109 (2001) 1101–1109. https://doi.org/10.1121/1.1408946.

[8] Y. Hao, J. Kang, H. Wörtche, Assessment of the masking effects of birdsong on the road traffic noise environment, J. Acoust. Soc. Am. 140 (2016) 978–987. https://doi.org/10.1121/1.4960570.

[9] B. De Coensel, S. Vanwetswinkel, D. Botteldooren, Effects of natural sounds on the

perception of road traffic noise., J. Acoust. Soc. Am. 129 (2011) EL148-L153. https://doi.org/10.1121/1.3567073.

[10] J. Liu, J. Kang, H. Behm, Birdsong As an Element of the Urban Sound Environment: A Case Study Concerning the Area of Warnemunde in Germany, Acta Acust. United With Acust. 100 (2014) 458–466. https://doi.org/10.3813/AAA.918726.

[11] Y. Hao, J. Kang, J.D. Krijnders, Integrated effects of urban morphology on birdsong loudness and visibility of green areas, Landsc. Urban Plan. 137 (2015) 149–162. https://doi.org/10.1016/j.landurbplan.2015.01.006.

[12] Z. Ong, B. Lam, J. Hong, K. Ooi, W. Gan, Selecting birdsongs for auditory masking: a clustering approach based on psychoacoustic praramters, in: Proc. 25th Int. Congr. Sound Vib. ICSV25, Hiroshima, Japan, 2018: pp. 1–8.

[13] J.Y. Jeon, P.J. Lee, J. You, J. Kang, Acoustical characteristics of water sounds for soundscape enhancement in urban open spaces, J. Acoust. Soc. Am. 131 (2012) 2101–2109. https://doi.org/10.1121/1.3681938.

[14] M. Rådsten Ekman, P. Lundén, M.E. Nilsson, Similarity and pleasantness assessments of water-fountain sounds recorded in urban public spaces, J. Acoust. Soc. Am. 138 (2015) 3043–3052. https://doi.org/10.1121/1.4934956.

[15] Ö. Axelsson, M.E. Nilsson, B. Hellström, P. Lundén, A field experiment on the impact of sounds from a jet-and-basin fountain on soundscape quality in an urban park, Landsc. Urban Plan. 123 (2014) 46–60. https://doi.org/10.1016/j.landurbplan.2013.12.005.

[16] M.E. Nilsson, J. Alvarsson, M. Rådsten-Ekman, K. Bolin, Auditory masking of wanted and unwanted sounds in a city park, Noise Control Eng. J. 58 (2010) 524. https://doi.org/10.3397/1.3484182.

[17] J. You, P.J. Lee, J.Y. Jeon, Evaluating water sounds to improve the soundscape of urban areas affected by traffic noise, Noise Control Eng. J. 58 (2010) 477–483.

[18] J.Y. Jeon, P.J. Lee, J. You, J. Kang, Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds, J. Acoust. Soc. Am. 127 (2010) 1357–1366.

[19] J.Y. Hong, B. Lam, Z.-T. Ong, K. Ooi, W.-S. Gan, J. Kang, S. Yeong, I. Lee, S.-T. Tan, The effects of spatial separations between water sound and traffic noise sources on soundscape assessment, Build. Environ. 167 (2020) 106423. https://doi.org/10.1016/j.buildenv.2019.106423.

[20] L. Galbrun, T.T. Ali, Acoustical and perceptual assessment of water sounds and their use over road traffic noise, J. Acoust. Soc. Am. 133 (2013) 227–237. https://doi.org/10.1121/1.4770242.

[21] M. Rådsten-Ekman, Ö. Axelsson, M.E. Nilsson, Effects of Sounds from Water on Perception of Acoustic Environments Dominated by Road-Traffic Noise, Acta Acust. United with Acust. 99 (2013) 218–225. https://doi.org/10.3813/AAA.918605.

[22] International Organization for Standardization, ISO 12913-1: Acoustics — Soundscape — Part 1 : Definition and conceptual framework, International Organization for Standardization, Geneva, Switzerland, 2014.

[23] X. Zhao, S. Zhang, Q. Meng, J. Kang, Influence of contextual factors on soundscape in urban open spaces, Appl. Sci. 8 (2018) 2524. https://doi.org/10.3390/app8122524.

[24] K. Herranz-Pascual, I. Aspuru, I. Garcia, Proposed conceptual model of environmental experience as framework to study the soundscape, in: INTER-NOISE NOISE-CON Congr. Conf. Proceedings, InterNoise10, Lisbon, Portugal, 2010: pp. 2691–3542. https://www.ingentaconnect.com/openurl?genre=article&issn=0736-2935&volume=2010&issue=8&spage=2904&epage=2912&aulast=Herranz-Pascual.

[25] K. Herranz-Pascual, I. García, I. Diez, A. Santander, I. Aspuru, Analysis of Field Data to Describe the Effect of Context (Acoustic and Non-Acoustic Factors) on Urban

Soundscapes, Appl. Sci. 7 (2017) 173. https://doi.org/10.3390/app7020173.

[26] J. Kang, B. Schulte-Fortkamp, Soundscape and the Built Environment, CRC Press, 2016.

[27] J.Y. Hong, J.Y. Jeon, Influence of urban contexts on soundscape perceptions: A structural equation modeling approach, Landsc. Urban Plan. 141 (2015) 78–87. https://doi.org/10.1016/j.landurbplan.2015.05.004.

[28] A. Ståhle, Sociotope mapping : Exploring public open space and its multiple use values in urban and landscape planning practice, Nord. J. Archit. Res. 19:4, s. 5 (2006) 59–71. https://doi.org/10.1086/431492.

[29] I. Babelon, A. Ståhle, B. Balfors, Toward Cyborg PPGIS: exploring socio-technical requirements for the use of web-based PPGIS in two municipal planning cases, Stockholm region, Sweden, J. Environ. Plan. Manag. 60 (2017) 1366–1390. https://doi.org/10.1080/09640568.2016.1221798.

[30] L. Lavia, M. Easteal, D. Close, H. Witchel, O. Axelsson, M. Ware, M. Dixon, Sounding Brighton: Practical approaches towards better soundscapes, in: INTERNOISE NOISE-CON Congr. Conf. Proc., 2012: pp. 436–444.

[31] L. Estévez-Mauriz, J. Forssén, M.E. Dohmen, Is the sound environment relevant for how people use common spaces?, Build. Acoust. 25 (2018) 307–337. https://doi.org/10.1177/1351010X18790832.

[32] Ö. Axelsson, How to Measure Soundscape Quality, in: Euronoise 2015, 2015: pp. 1477–1481.

[33] L. Lavia, Ö. Axelsson, M. Dixon, Sounding Brighton: Developing an applied soundscape strategy, in: Proceeding AESOP 26th Annu. Congr., 2012: pp. 1–9.

[34] S. Viollon, C. Lavandier, C. Drake, Influence of visual settings on sound ratings in an urban environment, Appl. Acoust. 63 (2002) 493–511.

[35] J.Y. Hong, J.Y. Jeon, Designing sound and visual components for enhancement of urban

soundscapes., J. Acoust. Soc. Am. 134 (2013) 2026–36. https://doi.org/10.1121/1.4817924.

[36] J.L. Carles, I.L. Barrio, J.V. De Lucio, Sound influence on landscape values, Landsc. Urban Plan. 43 (1999) 191–200. https://doi.org/10.1016/S0169-2046(98)00112-1.

[37] X. Hong, G. Wang, J. Liu, S. Lan, Cognitive persistence of soundscape in urban parks, Sustain. Cities Soc. 51 (2019) 101706. https://doi.org/10.1016/j.scs.2019.101706.

[38] H. Li, S.K. Lau, A review of audio-visual interaction on soundscape assessment in urban built environments, Appl. Acoust. 166 (2020) 107372. https://doi.org/10.1016/j.apacoust.2020.107372.

[39] H.I. Jo, J.Y. Jeon, Effect of the appropriateness of sound environment on urban soundscape assessment, Build. Environ. 179 (2020) 106975. https://doi.org/10.1016/j.buildenv.2020.106975.

[40] J.L. Carles, I.L. Barrio, J.V. de Lucio, Sound influence on landscape values, 43 (1999).

[41] L. Maffei, M. Masullo, A. Pascale, G. Ruggiero, V.P. Romero, Immersive virtual reality in community planning: Acoustic and visual congruence of simulated vs real world, Sustain. Cities Soc. 27 (2016) 338–345. https://doi.org/10.1016/j.scs.2016.06.022.

[42] Z. Bangjun, S. Lili, D. Guoqing, The influence of the visibility of the source on the subjective annoyance due to its noise, Appl. Acoust. 64 (2003) 1205–1215. https://doi.org/10.1016/S0003-682X(03)00074-4.

[43] G. Watts, L. Chinn, N. Godfrey, The effects of vegetation on the perception of traffic noise, Appl. Acoust. 56 (1999) 39–56.

[44] D.E. Aylor, L.E. Marks, Perception of noise transmitted through barriers., J. Acoust. Soc. Am. 59 (1976) 397–400.

[45] Singapore Ministry of National Development Housing & Development Board, Housing & Development Board landscape guide, HDB Landsc. Guid. (2013) 132.

http://www10.hdb.gov.sg/eBook/landscapeguide/main.html (accessed November 13, 2018).

[46] Planning Department of the Government of the Hong Kong Special Administration Region, Chapter 4: Recreation, Open Space and Greening by planning department, Hong Kong Plan. Stand. Guidel. (2018) 59. https://www.pland.gov.hk/pland_en/tech_doc/hkpsg/full/pdf/ch4.pdf (accessed November 13, 2018).

[47] A.J. Kellett, M.W. Rofe, Creating Active Communities: How Can Open and Public Spaces in Urban and Suburban Environments Support Active Living?, 2009.

[48] Singapore Department of Statistics, Statistics on resident households are compiled by the Singapore Department of Statistics, (2018). https://www.singstat.gov.sg/find-data/search-by-theme/households/households/latest-data (accessed November 13, 2018).

[49] International Organization for Standardization, ISO 532-1 Acoustics – Method for calculating loudness – Part 1: Zwicker method, International Organization for Standardization, Geneva, Switzerland, 2017.

[50] 45692 DIN, Measurement Technique for the simulation of the auditory sensation of sharpness, 2009.

[51] J.Y. Jeon, J.Y. Hong, Classification of urban park soundscapes through perceptions of the acoustical environments, Landsc. Urban Plan. 141 (2015) 100–111. https://doi.org/10.1016/j.landurbplan.2015.05.005.

[52] J.Y. Jeon, I.H. Hwang, J.Y. Hong, Soundscape evaluation in a Catholic cathedral and Buddhist temple precincts through social surveys and soundwalks, J. Acoust. Soc. Am. 135 (2014) 1863–1874. https://doi.org/10.1121/1.4866239.

[53] V. Puyana-Romero, L.S. Lopez-Segura, L. Maffei, R. Hernández-Molina, M. Masullo, Interactive Soundscapes: 360°-Video Based Immersive Virtual Reality in a Tool for the

Participatory Acoustic Environment Evaluation of Urban Areas, Acta Acust. United with Acust. 103 (2017) 574–588. https://doi.org/10.3813/AAA.919086.

[54] J.Y. Hong, B. Lam, Z.-T. Ong, K. Ooi, W.-S. Gan, J. Kang, J. Feng, S.-T. Tan, Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications, Build. Environ. 149 (2019) 1–14. https://doi.org/10.1016/j.buildenv.2018.12.004.

[55] Two Big Ears Ltd., Facebook 360 Spatial Workstation [Computer software], v3.0, 2017.

[56] F. Faul, E. Erdfelder, A.-G. Lang, A. Buchner, G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences, Behav. Res. Methods. 39 (2007) 175–191. https://doi.org/10.3758/BF03193146.

[57] R. Schatz, S. Egger, K. Masuch, The impact of test duration on user fatigue and reliability of subjective quality ratings, J. Audio Eng. Soc. 60 (2012) 63–73.

[58] F. Aletta, J. Kang, Ö. Axelsson, Soundscape descriptors and a conceptual framework for developing predictive soundscape models, Landsc. Urban Plan. 149 (2016) 65–74. https://doi.org/10.1016/j.landurbplan.2016.02.001.