EFFECT OF ENVIRONMENTAL CONTEXTS PERTAINING TO DIFFERENT SOUND SOURCES ON THE MOOD STATES

Boyan ZHANG[1] and Jian KANG[2,*]

[1]School of Architecture, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom; Tel: +44 (0)75 0340 8430; Email: bzhang27@sheffield.ac.uk

[2]Institute for Environmental Design and Engineering, The Bartlett, University College London, London WC1H 0NN, United Kingdom; Tel: +44 (0)20 3108 7338 (internal: 57338); Email: j.kang@ucl.ac.uk

*Corresponding author

Date Received: 26th July 2021
Date Accepted: 17th October 2021
Available online: 21st October 2021

DOI: https://doi.org/10.1016/j.buildenv.2021.108456

Abstract

For the improvement of emotional experience in buildings and environments, this study investigated the effect of different non-urban contexts on urban sounds in terms of change in mood state. The current model used tom evaluate emotional response in soundscape (Principle Component Model) makes limited consideration of the formation process of emotional responses and lacks ability to evaluate distinctive moods or emotions. This study adopted the emotional process as a methodology and enabled the analysis of distinctive moods and emotions presented in the subjective evaluation of soundscapes. Three sound types (street music, traffic sound, and fountain sound) and three environmental contexts (courtyard and monk chanting scenarios, specific to a Buddhist temple, and alpine meadows) were combined to create 12 groups of visual and audio stimuli. Multiple audio playbacks were performed for 300 participants. The results indicated that the context significantly influenced the mood scores in combination with the urban sounds. However, the effects varied across different sound types, contexts, and mood categories. Certain common trends were observed: The negative mood scores were reduced when the traffic and fountain sounds were combined with the contexts, while the combination of the considered contexts with street music did not affect the mood states. The results of this study provide several applications in building and environmental design. For example, introducing human activities can significantly improve the mood states (specifically, reducing anger and depression) of users of a space.

Keywords:

Visual contexts; audio contexts; mood state; mood category; mood dimensions; soundscape

1. Introduction

The examination of mood and emotions in the context of soundscape has been mostly centred around the concept of core affect and a two-dimensional model of it [1] in which the two dimensions correspond to pleasure–displeasure (valence) and activation–deactivation (arousal). In this model, core affect is a human neurophysiological state that reflects the raw feeling at the foundation of every mood and emotion. In Russell's study of core affect, he stated that a combination of different intensities of pleasure and arousal could be sufficiently qualified as an emotional episode [1]. However, Russell also clarified that this combination cannot account for every prototypical emotional episode, and any mood or emotional outcome defined using this model should be further analysed through a categorical perspective [1] (distinctive emotions). This concept of core affect was later developed into a principal component model to analyse perceived soundscape qualities [2] which was widely used to evaluate soundscapes. However, the mood states and emotional outcomes were not categorically analysed further in these studies.

Individuals' moods/emotions considerably influence the interpretation of soundscape quality; in turn, environmental contexts are crucial to the forming of emotional responses, and must be examined to understand the mood/emotional process. The existing studies of contexts in the soundscape field have been focused on the relationship of people (the receivers of the sound), activities, and places under the scope of space and time. A change in context can affect both the auditory sensation and its interpretation by the person, thereby influencing their response to the acoustic environment [3]. It can also affect mood/emotion outcomes and alter an individual's response to the acoustic environment. This phenomenon was demonstrated in Schechter and Singer's [4] experiment, in which three groups of participants were injected with adrenaline: one group was provided with accurate information on the effect of the injection, one group was provided with wrong information, and the remaining group was not provided with any information. Participants without any information appeared to attribute their physiological arousal to anger, because they experienced an act of anger during the experiment session. Robinson [5] added that people often label their mood/emotion depending on its development in real time. This labelling can be the result of an initial non-cognitive appraisal of the external or internal change (of the context) or simply a self-induced mood/emotion suggested by environmental information. Both studies indicate that contexts can change the outcome of a mood/emotion.

Due to the importance of environmental contexts in emotional response and the lack of studies categorising emotional responses (distinctive emotions and mood states) in the soundscape field, this study aimed to examine six mood states (anger, confusion, depression, tension, fatigue, and vigour) in relation to the contexts in a soundscape and thus to provide better emotional experience to inform building and environmental design. The contexts primarily differed in terms of location, visual and audio stimuli. By combining several different spatial contexts with urban sounds, the relationship of the mood states and urban sounds across these contexts was identified. Although the demographic aspect (such as the gender difference) of the context was considered, it was not the key focus. The objective was to address the following research questions:

1) Do the contexts affect the interactions between sound types and mood states? What is the overall relation among the sound types, contexts, and mood states?

2) How do the mood states differ among the contexts in terms of the sound types?

3) What is the difference between the audio and visual contexts in terms of their effect on the mood states?

2. Methods and methodology

2.1 Site selection spatial contexts and audio stimuli

To explore the variation in mood states under different environmental contexts combined with urban sounds, three contexts were recorded and used as variables. These contexts consisted of visual cues (static images) and audio recordings. The criteria for site selection were based on these two components (visual and audio environments) and the remoteness of the sites from the urban setting. Specifically, the audio, visual, and functional characteristics of the selected space were expected to have environmental features that were significantly different from those of the urban public setting. In particular, the space was required to have no or limited visuals of urban layout or building types and have no or limited visuals (buildings, traffic etc.) and audio such as those of traffic and machinery; moreover, conventional circulation or gathering spaces were not considered. The environment types of the three contexts selected based on these criteria corresponded to the indoor region and courtyard of a Buddhist temple and an Alpine meadow in a rural setting. A rural setting was selected because among the three contexts, the rural context is most remote from urban environment both physically and contextually (lack of human activity and artificial elements). All the contexts were recorded during summer 2019, under sufficient daylight conditions. The details of these contexts are as follows:

1) The Da Zhao Temple is located in the south-west of the city of Hohhot, in China. A square with greenery surrounds the north, east, and south sides of the temple. A 2-lane main road and commercial structures are present on the west side of the temple. Beyond the square, a 4-lane main road is present in the east, with commercial structures on the north and south sides. Despite the urban environment, the temple still feels disjoint from the city, owing to the wall enclosures and greenery surrounding the temple, as shown in Fig. 1. The temple requires a

ticket to enter and serves as an attraction site, which is a common occurrence for religious spaces in China. Consequently, the temple is not commonly used as a gathering space. The temple itself is a building complex consisting of several temple halls connected through courtyards. In this study, the main courtyard was considered. This courtyard consists of stone flooring surrounded by the temple halls; these halls have a brick-and-wood structure with red walls and orange roof tiles and are structurally considerably different from the urban building typology. Fig. 2 shows that certain statues and a large incense burner are present in the middle of the courtyard, indicating the special functionality (religious purposes) of the space. A picture was taken as a visual cue for the courtyard; the content of the picture consists of human activity (less than 5% of the picture), the front porch of the temple hall (about 25% of the picture), and a panoramic view of the courtyard (about 70% of the picture). There is no natural element in the picture (see Fig. 2). The temple has an extremely low footfall, usually consisting of a few monks and tourists. There is no line of sight of vehicles or the city. The audio of the courtyard consists of birdsong (dominant), human voices, and faint and distant traffic sounds.

Figure 1. Satellite view of the Da Zhao Temple

2) The main temple hall of Da Zhao Temple was selected as the indoor space. The hall is dimly lit through skylights. The central area is surrounded by prayer wheels and is usually where the monks sit, recite chants, and study. An online picture was used as a visual cue for the temple hall; the picture consists of an indoor practice area for monks and a skylight with artificial decoration, and is also devoid of the nature element and human activity (see Fig. 2). The audio of the hall consists mainly of monks chanting and sounds of prayer wheels.

3) The Alpine meadow is a rural area located to the north of Hohhot, about 8 km from the city in the Yin mountain range. A picture was taken as the visual cue for the alpine meadow; the picture consist of a foot path (an artificial element consisting of less than 10% of the picture), a view of the continuing mountain range and sky (each of the two natural elements consisting of just less than 50% of the picture) (see Fig. 2). The audio mainly consists of the sounds of insects and the wind.

In addition, an image of the square where the Sheffield war memorial is situated (Fig. 2) was used as a visual stimulus for the three controlled sound clip groups (traffic sound, fountain sound, and street music).

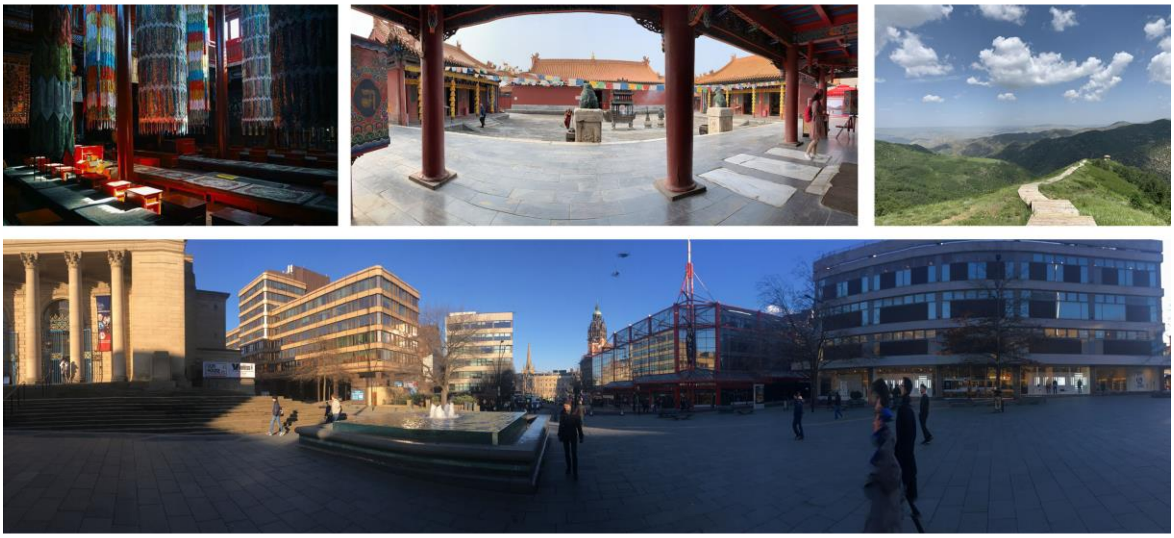

Figure 2. Top left: Main temple hall interior; top middle: Courtyard of the Da Zhao temple; top right: Alpine meadow view; bottom: Panoramic view of the square of the Sheffield war memorial

Twelve audio clips were used as audio stimuli; they were a combination of the three sound types (street music, traffic sound, and fountain sound) and the environmental context sounds (Da Zhao temple courtyard, monk chanting, and Alpine meadow). The three single sound types were recorded by an Edirol R44 binaural recorder in the Peace Garden and Fargate Street, in Sheffield city centre, UK, and the three environmental context sounds were recorded by a Head acoustics Squadriga II binaural recorder in the Da Zhao temple in the city of Hohhot, China, and the mountain range located to the north of the city. The duration of each clip was approximately 10 min. During the experimental session, a single audio clip was played and looped at 65 dBA in A-weighted equivalent continuous sound level (LAeq). The description of the 12 audio clips, along with the assigned notation, is as follows:

1) Street music (M);

2) street music–temple courtyard (MTc);

3)street music–monk chanting; (MC);

 https://doi.org/10.1016/j.buildenv.2021.108456

4) street music–Alpine meadow (MA);

5) traffic sound (T);

6) traffic sound–temple courtyard (TTc);

7) traffic sound–monk chanting (TC);

8) traffic sound–Alpine meadow (TA);

9) fountain sound (F);

10) fountain sound–temple courtyard (FTc);

11) fountain sound–monk chanting (FC);

12) and fountain sound–Alpine meadow (FA). Spectrum of the 12 audio clips is shown in figure 3.

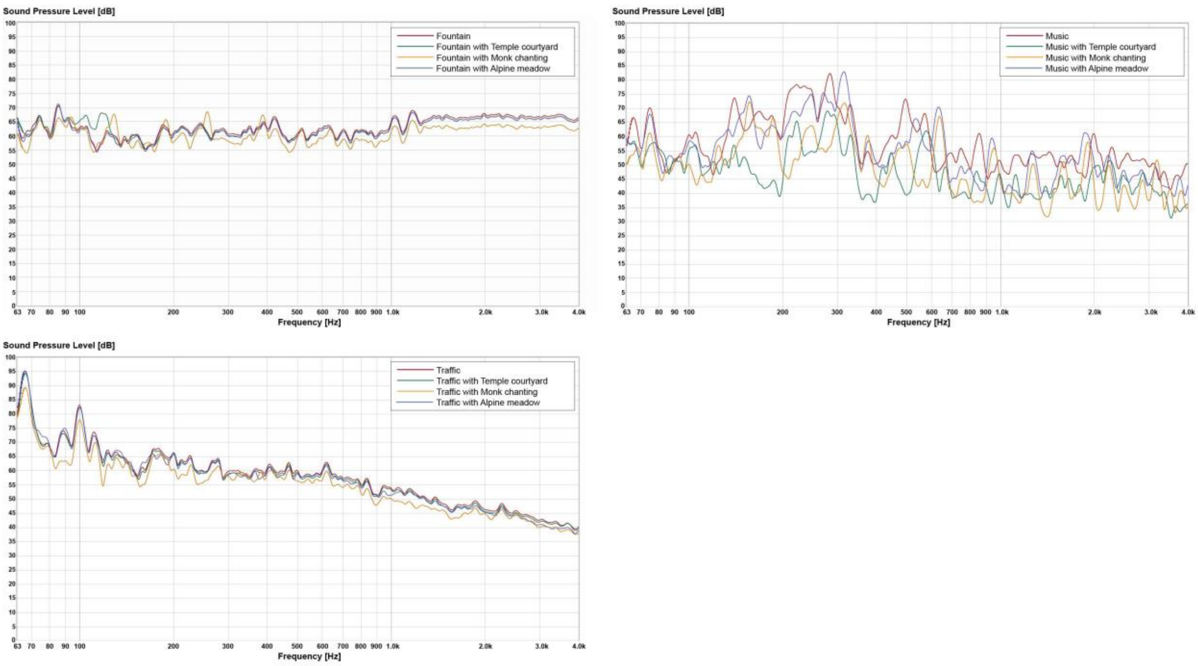

Figure 3. Spectra from 12 audio clips: top left—the fountain sound and the combination of three contexts, top right—the street music and the combination of three contexts, and bottom left—the traffic and the combination of three contexts.

2.2 Implicit process of mood and emotion

The mood and emotion development process, as summarised by Robinson [5], consists of the initial triggering by the affective appraisal of a situation, followed by a physiological change in the body systems. Subsequently, cycles of cognitive appraisal and re-appraisal of the situation occur, in a subconscious and rapid manner [5]. Without conscious labelling, the process can result in a long-lasting mood state, which can affect an individual's subjective evaluation of an environment. Soundscape research, at its foundation, is a study of the interaction between the people and the sound environment. The results of most of the laboratory studies rely on the participants' conscious evaluation of the stimuli, as in the case of commonly used methodologies such as audio replays and sound walks. However, the soundscape may affect people implicitly on a psychophysiological level. Brown et al. [6] termed this phenomenon as the 'experimenter effect'. Specifically, the participants' conscious attention to the soundscape does not necessarily affect the outcome (direct outcomes), and the results (enabled outcomes) may be a result of the interaction of the soundscape with other environmental aspects. In terms of the mood states, psychophysiological or mood changes likely occur owing to a non-conscious reaction towards the environment or stimuli, and such changes later affect the perception of the soundscape through a subjective interpretation.

The current study was designed to account for the subconscious nature of the mood states. The implicit process was realised by limiting the information provided to the participants before and during the experiment session. The participants were told that the experiment was simply a built-environment-related study, and they would be asked to spend some relaxed time in a meeting room without any other specific or dangerous activities. Later, the participants answered a questionnaire, the open questions which were designed to extract information regarding if consciously evaluations were made towards the stimuli. Subsequently, the effectiveness of the method was examined through a data analysis in which the difference in the attention toward different sound stimuli was compared. This method facilitated the generation of the most natural reaction of the participants toward the stimuli and alleviate the experimenter effect.

2.3 Questionnaire and experiment design

Mood states, in comparison to emotions, are less intense, long-lasting [7,8] and lack an object. Only when a person tries to interpret the mood state and attach it to an object (context) does this state become an emotion [1]. A similar concept was described by Robinson [5] as the development process of a mood or emotion. Specifically, this process consists of the triggering, monitoring, and labelling stages. Nevertheless, the mood may be often undefined at the monitoring stage, and thus may not be able to be recorded or analysed. The use of a questionnaire can provide cues pertaining to the attention and help the participants evaluate and define (label) their mood states.

A long form of the profile of mood states (POMS) questionnaire [9] was used as a tool to measure the mood state. This questionnaire is commonly used in medical studies to monitor the patients' mood states over a long duration and has been proved effective for a variety of demographic groups [10–12]. The questionnaire consists of 65 items, 55 of which correspond to one of six mood categories: anger (9 items), confusion (7 items), depression (15 items), fatigue (7 items), tension (9 items), and vigour (8 items). The remaining 10 items correspond to emotion-related terms that do not correspond to specific moods and are designed to examine the attention. The original questionnaire employs a 5-point Likert scale (1: not at all, 5: extremely). The tendency to select the middle scale has the potential to reduce the small effect observed in the data, resulting in potential bias; hence, a 6-point Likert-type scale was used in this study to prevent the participants' tendency to select the middle scale [13]. In addition, the following four open questions were added at the end of the questionnaire to examine the participants' attention and activity during the experiment:

1) During the experiment session, did anything specific attract your attention?
2) If so, how long did you focus your attention on it?

3) During the experiment session, was there anything specific on your mind? If so, please summarise it in a few words (e.g. work-related).

4) Please describe your activities during the relaxation period.

The experiment was performed in a meeting room (see Fig. 4) in the architecture department of the Inner Mongolia University of Technology in China. The meeting room can fit approximately 15 individuals without crowding. For the 300 participants, a total of 24 sessions were conducted for 12 different stimuli (audio clip) groups (12–13 participants per session). To simulate the outdoor sound propagation condition, sound-absorbing foams were placed at the corners and edges of the room walls to achieve a reverberation time of 0.3 s–0.5 s for a frequency of 500 Hz–1000 Hz. As the meeting room is in a quiet corner of the building and has no surrounding traffic there is little to no ambient sound. The ambient sound of the meeting room when occupied consists of a cooling fan sound from the projector and human sounds and the voices of the participants. Tables and chairs were provided for the participants, and tea and coffee were served. The independent variables in this study were the 12 audio clips and their visual cues, noticing of the acoustic stimuli, and gender difference among the participants. The dependent variables were the six mood states (i.e. anger, confusion, depression fatigue tension and vigour).

Figure 4. Panoramic view of meeting room with diagram of experiment setup

A pair of speakers and a laptop were set up at one end of the room for the audio replay. Only one audio clip was played and looped during each session. To reduce the potential bias caused by the difference in the sound levels, all the audio replays were played at 65 dBA, and the sound level was measured using a 01 dB sound level metre before each session. A single static photo was projected on the screen (1.5 × 2 m) in the meeting room to provide a visual cue for each audio clip. A total of four pictures were used, three of which were captured onsite. The remaining picture was downloaded from the internet, as the temple forbids indoor photography. Although all pictures were projected in full screen view, the scale and resolution of the longer pictures (the courtyard of the Da Zhao temple and the panoramic view of the war memorial square) are smaller compared to the other ones. The combinations of the images (Fig. 2) and audio clips were as follows:

1) Photo of the square of the Sheffield war memorial, as a visual cue for sound clips M, T and F;
2) photo of the Da Zhao temple courtyard, as a visual cue for sound clips MTc, TTc and FTc;
3) photo of the Da Zhao temple interior, as a visual cue for sound clips MC, TC and FC;
4) photo of the Alpine meadows, as a visual cue for sound clips MA, TA and FA.

Before the start of each experiment session, the audio clip was played and its related picture was projected before the participants arrived. As the participants arrived, they were asked to sit around the table and relax for 30 min; tea and coffee were served. There were no strict rules on the specific activities that the participants could perform during this period, as long as the activity did not significantly reduce the visual and audio perception of the participants (such as wearing headphones). The purpose of this period was to allow the participants to simulate a relaxation period in a public space and sufficiently absorb the designed stimuli. After 30 min, each of the participants were asked to fill in a POMS questionnaire, to observe their mood states and the activity performed during the experiments.

2.4 Participants

Power calculations were performed to determine the sufficient number of participants. The A-prior power calculation for multi-variate analysis of variance (MANOVA) indicated that to achieve a power of 95% for a small effect size [f = 0.10] with three variables (sound clips, attention to the sound, and gender difference) and six groups (six mood categories), 102 participants were required. In another A-prior power calculation for the analysis of variance (ANOVA), it was noted that 288 participants would be required to achieve a power of 80% for the medium effect size [f = 0.25] with 12 groups (12 sound clips). In other words, 24 participants were required per sound clip group. It was impractical to achieve a small effect size for the ANOVA, which would require 1692 participants. However, a small effect is unlikely to be neglected due to the ability of the MANOVA to detect any small effect and avoid type II errors. Moreover, the results of the MANOVA were considered when interpreting the results of the ANOVA. To compensate for any missing or invalid data, 300 participants were recruited instead of 288.

The ethics approval (160237180) was provided by the University of Sheffield. All the participants were recruited from the Inner Mongolia University of Technology. The recruitment information was announced during lectures. Participants were required to have normal perceptibility (i.e. normal hearing and sight). Any mild disability that the participants were unaware of did not impact the experiment outcome, as the person could still evaluate their surroundings sufficiently. A total of 300 undergraduate students volunteered: 133 male, 179 female, and 8 unspecified (the gender information is empty in the questionnaire, and they were excluded from the data). The ages of the participants ranged from 17 to 25. To secure a large sample size, the undergraduate student (teenager and young adults) demographic was considered as the population of interest; therefore demographic difference must be taken into account when generalising the results.

2.5 Data analysis

The data were explored through MANOVA, discriminant function analysis, and subsequent univariate tests. The initial MANOVA was conducted to identify if a difference in the mood state occurred among the independent variables, namely, the difference in the audio clips (contexts and sound types), gender, and influence of noticing the stimuli. MANOVA was used to prevent the inflation of type I error in the subsequent univariate tests; and a three-way MANOVA was conducted due to the multiple dependent variables (i.e. the six mood states) and independent variables, to explore the interaction between independent variables. As the sound types and contexts that formed the 12 sound clips cannot be analysed independently, discriminant function analysis was conducted to group the dependent (mood states) and independent (sound types and contexts) variables to clarify the trends of sound types and contexts among the groups. The discriminant function analysis also provided supporting data analysis for the univariate tests and the post-hoc test. The relationship between sound types and contexts was further examined by the univariate tests that followed the MANOVA. The subsequent post-hoc tests have made one-to-one comparisons between each sound type and their context combinations. Fig. 5 shows the complete data analysis process.

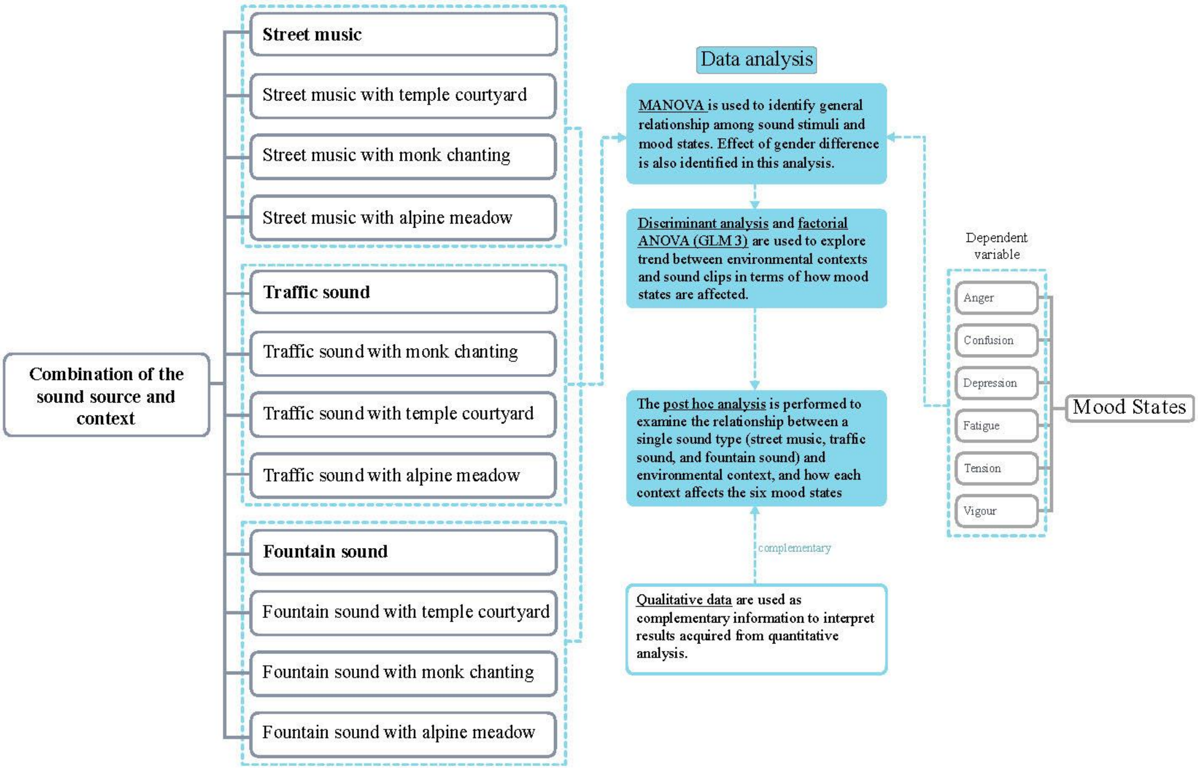

Figure 5. Data analysis process

3. Results

3.1 Overall difference and trends in the contexts and mood states

To identify if any effect existed among the contexts, a MANOVA (Pillai's trace) was performed. The results indicated that a general difference existed in the mood scores for different sound clip groups [$V = 0.43$, $F(66, 1476) = 1.74$, $p < .001$]. This result indicated that the difference among the sound types and their combination with the contexts significantly affected the mood states. The subsequent univariate test confirmed that all six moods were affected by the difference in the sound clips: anger [$F(11, 246) = 3.08$, $p \leqslant .001$, $\omega 2 = 0.036$], confusion [$F(11, 246) = 2.96$, $p \leqslant .001$, $\omega 2 = 0.035$], depression [$F(11, 246) = 1.93$, $p < .05$, $\omega2 = 0.016$], fatigue [$F(11, 246) = 3.27$, $p < .001$, $\omega2 = 0.039$], tension [$F(11, 246) = 3.57$, $p < .001$, $\omega2 = 0.045$], and vigour [$F(11, 246) = 2.36$, $p < .01$, $\omega2 = 0.024$]. Additionally, the participants' noticing the sound [$N = 134$] and the associated interaction with the difference in the sound clips did not significantly affect the mood states [$p < .05$]. This result indicated that without verbal context of the experiment and the sound sources, the participants' attention towards the sound stimuli did not affect their mood states, thereby demonstrating that the implementation of the implicit process was successful.

To examine the influence of the different environmental contexts on the mood states, a discriminant function analysis was conducted to group the sound clips and mood states. The results identified six functions, and the first two functions explained most of the variance. Specifically, function 1 explained most of the variance [49.1%, canonical R2 = 0.22], followed by the function 2, which explained 25.7% of the variance [canonical R2 = 0.13]. Together, the six discriminant functions could significantly differentiate the sound types and their combinations with the environmental contexts [$\Lambda = 0.60$, $X2\ (66) = 149.17$, $p < .001$]. Even when the function 1 was removed, the remaining group could reasonably differentiate the sound clips [$\Lambda = 0.76$, $X2\ (50) = 78.93$, $p < .01$]. However, if the function 2 was removed as well, the sound clips could not be differentiated by the remaining functions [$\Lambda = 087$, X2 (36) = 39.98, $p > .05$]. The correlation between the sound clips and mood indicated that anger [r = 0.84], confusion [r = 0.74], depression [r = 0.61], fatigue [r = 0.71], and tension [r = 0.85] loaded highly on the discriminant function 1, whereas vigour loaded highly on the function 4 [r = 0.73].

Fig. 6, which shows the discriminant function plot, along with Table 1 indicates that the 1st function differentiates the sound clips T, F, TTc and FA from the remaining clips (indicated by the lack of '-' sign in the first row in Table 1 and their separation and occupation on the right side of the graph in Fig. 6). The 2nd function differentiates M, F, MA, and MC, from the remaining clips (indicated by the lack of '-' sign in the second row in Table 1 and their separation and occupation at the top of the graph in Fig. 6). F and its context combinations (FTc, FA, and FC) can be differentiated consistently by the first two functions, while T and its context combinations (TTc, TA, and TC) can be differentiated only by the first function, and M cannot be clearly differentiated by either of the two functions. The results of the function discriminant analysis suggest that the context differences considerably influence the six moods. This difference is particularly notable for T and its context combinations and for F and its context combinations, and does not occur for M and its context combinations.

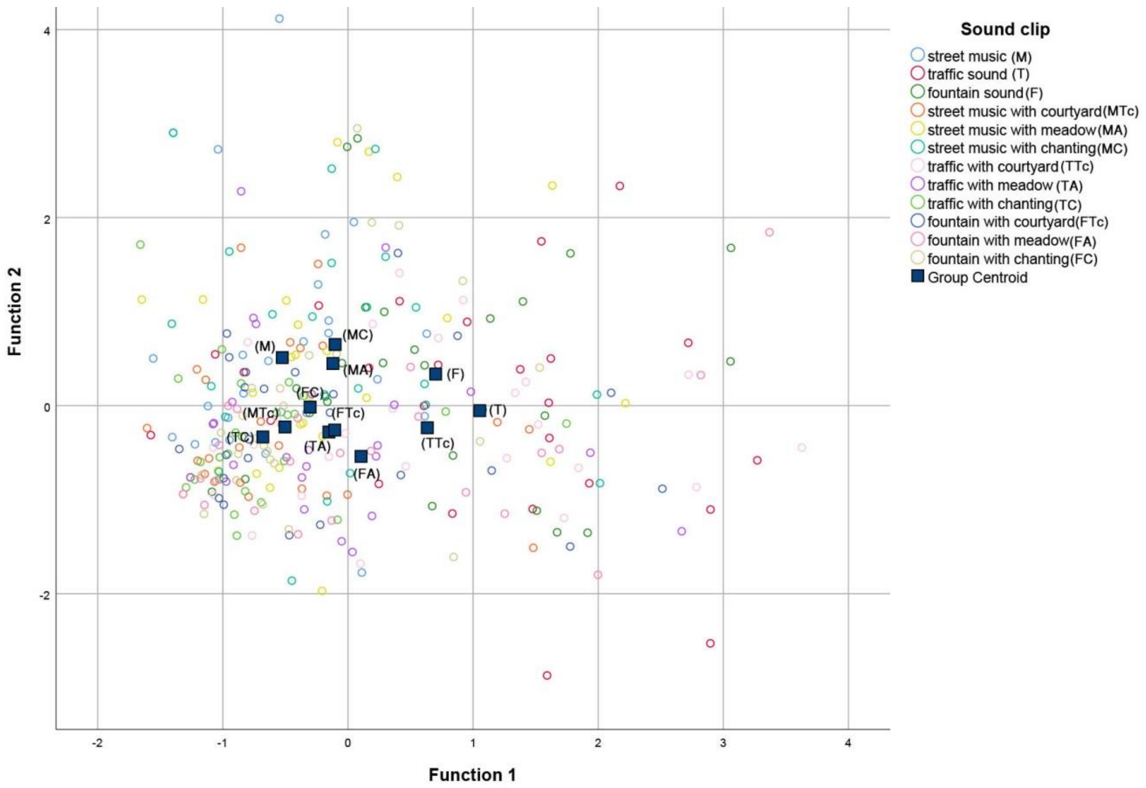

Figure 6. Canonical discriminant functions: Distribution of the 12 sound clips in relation to the first two discriminant functions (major).

Table 1. Unstandardised canonical discriminant functions evaluated at the group means

		M	T	F	MTc	MC	MA	TTc	TC	TA	FTc	FC	FA
Discriminant Function	1	-0.525	1.054	0.701	-0.503	-0.103	-0.120	0.634	-0.680	-0.152	-0.106	-0.303	0.105
	2	0.511	-0.054	0.336	-0.227	0.650	0.448	-0.236	-0.333	-0.280	-0.259	-0.016	-0.540

M: Street music
MTc: Street music with temple courtyard
MC: Street music with monk chanting
MA: Street music with Alpine meadow

F: Fountain sound
FTc: Fountain sound with temple courtyard
FC: Fountain sound with monk chanting
FA: Fountain sound with Alpine meadow

T: Traffic sound
TTc: Traffic sound with temple courtyard
TC: Traffic sound with monk chanting
TA: Traffic sound with Alpine meadow

3.2 Effect of contexts on the different sound type groups

To examine the individual context and its effect on the mood states, a post hoc analysis was conducted after the univariate test (factorial ANOVA). Although the analysis considered every possible two-and-two comparison of the mood states among the sound clip groups, the focus in this study was to compare each single sound type (T, F, M) and their combinations with the contexts. The post hoc analysis result indicated that the different contexts significantly influenced the different mood states depending on the sound type.

a) T: The temple courtyard context did not exhibit a significant effect on any of the moods [$p > .05$], which means that neither the birdsong nor the visual stimulus of the courtyard made a significant difference in how the traffic sound influenced the mood states. The temple chanting context exhibited a significant effect on most of the moods [$p < .05$] (indicated by the '*' signs in the T v TC column in Table 2) except vigour [$p > .05$]. The Alpine meadow context exhibited a significant effect only on anger [$p < .05$] and fatigue [$p < .05$] (indicated by the '*' signs in the T v TA column in Table 2). The mean mood score reduced when the contexts were added.

b) F: The temple courtyard context exhibited a significant effect on confusion and fatigue [$p < .05$] (indicated by the '*' signs in the F v. FTc column in Table 2). The temple chanting context exhibited a significant effect on anger and depression [$p < .05$] (indicated by the '*' signs in the F v. FC column in Table 2). The Alpine meadow context significantly affected only confusion [$p < .05$]. Similar to the case of the traffic sound, the mean mood score was lowered when the contexts were added to the fountain sound.

M: The street music did not exhibit any difference between itself and its context combinations [$p > .05$]. The results indicated that the contexts did not have a significant effect on how the street music influenced the mood states. Usually, street music is aimed to attract the attention of passengers, which means that in most cases, the street music is consciously assessed by the people in the context of urban public spaces. Consequently, the music used in this study was a dominant sound. This dominant nature of street music likely overshadowed any effect that the contexts may have had. Furthermore, relaxing music is known to lower the arousal level in the presence of stress [14], and the music used in this study had a slow tempo, which likely had a calming effect. This effect was observed in the Bonferroni post hoc comparison of the tension score corresponding to M [0.28 ± 6.81 min] and T [12.48 ± 10.06 min, $p < .001$] or F [10.16 ± 11.23 min, $p < .05$]. In both the cases, the tension score decreased when street music was used as the audio stimuli. The positive effect of the context in terms of the reduction in the negative moods scores was observed in the other post hoc comparisons (i.e. traffic and fountain sound groups), which demonstrates that the contexts affect the mood states, even for different sound sources. However, the positive effect of the contexts did not significantly improve the music's ability to reduce the stress, as indicated by the absence of '*' signs in the M v. MTc, M v. MC and M v. MA columns in Table 2.

Table 2. Post hoc analysis among sound clips for the mean score of the mood states

		M v MTc	M v MC	M v MA	T v TTc	T v TC	T v TA	F v FTc	F v FC	F v FA
Games -Howell	Anger	-1.360	-4.200	-2.360	4.320	14.000***	11.760**	8.280	11.120*	6.680
Games -Howell	Fatigue	4.520	-2.440	-1.640	3.840	11.320***	8.400*	8.280*	7.120	7.400
Games -Howell	Tension	-0.280	-3.760	-1.940	2.600	11.920***	8.600	5.600	8.280	5.680
Bonferroni	Confusion	0.120	-1.360	-2.880	1.600	8.360***	5.160	7.000**	5.720	6.440*
Bonferroni	Depression	0.600	-4.320	-0.960	3.480	11.200*	9.920	9.600	11.800*	10.960
Bonferroni	Vigour	2.960	-4.160	-2.280	0.360	8.360	3.320	1.640	4.080	5.880

M: Street music
MTc: Street music with temple courtyard
MC: Street music with monk chanting
MA: Street music with Alpine meadow

T: Traffic sound
TTc: Traffic sound with temple courtyard
TC: Traffic sound with monk chanting
TA: Traffic sound with Alpine meadow

F: Fountain sound
FTc: Fountain sound with temple courtyard
FC: Fountain sound with monk chanting
FA: Fountain sound with Alpine meadow

* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

3.3 Effect of visual and audio stimuli in the contexts

Visually, the temple courtyard is distinct from the urban contexts. The layout is characteristic of a traditional Chinese courtyard (Fig. 1), which means that the space is self-contained and isolated both visually and acoustically. It has been reported that an enclosed space in a religious setting (similar to the temple context used in this study) can enhance the tranquillity in a soundscape [15]. In terms of the principal component analysis model (PCM), tranquillity is located in the uneventful and pleasant category, opposite to the tension mood. However, in this study, the tension score did not exhibit any significant difference when the temple courtyard context was combined with any of the three sound types. This result suggests that the tranquillity provided by the enclosed religious space is limited and depends on the sound types presented in a sound environment. Acoustically, the birdsong of the courtyard context did not affect the mood scores when combined with traffic. This result is expected as the audio playback is at 65 dBA, in which the traffic sound is relatively loud. In Hao et al.'s study [16], the masking effect of the birdsong in terms of the pleasantness was significant only when the traffic sound level was low (<52.5 dBA). Despite the low effectiveness of the birdsong, its perceived pleasantness effect could be observed in F and FTc, for which the fatigue and confusion [$p < .05$] mood scores were significantly different.

In contrast, the effect of the temple chanting context on the mood state was more notable. Visually, the context exists in an indoor environment. The room is dimly illuminated by the surrounding skylights. The colour scheme primarily consists of crimson-coloured furniture and interior walls and columns with multi-coloured religious flags hanging from the roof (Fig. 2). The overall space exhibits a sacred setting, in contrast to the relaxing courtyard. There is a clear shift of functionality from the temple courtyard to the indoor region, which likely contributed to the perceived pleasantness, as shown in the comparison between groups T and

TC. All the negative mood scores were significantly affected. The audio aspects of the context consisted of monks chanting, the occasional sound of the rotating prayer wheels, and quiet conversations of tourists. The acoustic elements were focused on human sounds in contrast to the natural sounds of the courtyard. In a previous case study [15], religious chanting in Buddhist temples was hypothesised to elicit pleasantness in visitors. This finding is in agreement with that of the present study. In the comparison between the F and FC groups, a significant difference was noted only for depression and anger. Along with the result of the T and TC groups, is can be suggested that the effect of the temple chanting context is limited when combined with sounds that already elicit a pleasant evaluation, such as water sounds [17]. Furthermore, human sounds have been reported to have a significant influence on the perceived eventfulness [18]. Moreover, the eventfulness was expected to only have a limited effect on the moods in this study, as most mood categories can only be differentiated in terms of pleasantness and unpleasantness (other than confusion) and not in terms of the eventfulness (i.e. confusion, depression, and vigour; for further discussion, see 4.2).

The Alpine meadow context consists of a natural visual and audio setting. The natural context exhibited a limited effect on anger and fatigue (unpleasant moods) when combined with T, and a limited effect on the confusion when combined with F. The results suggest that the natural context affects the pleasantness scale of the moods (PCM). The results are in agreement with those of previous studies, in which visuals of greenery were reported to improve the environmental preference [19], and the audio of natural sounds such as birdsong and water sounds likely improved the perceived pleasantness of the sound environment [16,20]. However, the effect of the natural context on the perceived pleasantness cannot be widely applied. Specifically, in a study on how the interaction with urban greenery could affect the anxiety of elderly people, it was found that 'awareness of the natural experience' is a determining factor of the outcomes [21]. Furthermore, it has been reported that although the visual stimuli of greenery can induce perceived tranquillity, the effect is small in terms of the anxiety even though the two moods are on the opposite sides of the same axis in the PCM (uneventful–pleasant and eventful–unpleasant for tranquillity and anxiety, respectively) [22]. These findings indicated that there exist more dimensions in the soundscape study of the mood states than simply pleasantness and eventfulness. The PCM also suggests the familiarity of sound as the third dimension for soundscape evaluation; however, the extent of the influence of the familiarity on the mood states is unclear. The results of this study indicate that the pleasantness induced by the natural environment features only in a few mood states, and more dimensions likely contribute toward the differentiation of the sound sources.

3.3 Gender differences

A mood difference was observed in the interaction of the sound clips and gender [$V = 0.32$, $F(60, 1476) = 1.38$, $p < .05$], although this effect only occurred in the case of depression [$F(10, 253) = 1.94$, $p < .05$, $\omega2 = 0.017$] and fatigue [$F(10, 246) = 2.82$, $p < .01$, $\omega2 = 0.031$]. On its own, the gender difference did not significantly affect mood scores [$p > .05$]. Specifically, the gender difference exhibited a significant small to medium effect on the mood states. Despite one study that has shown that gender difference does not affect evaluating subjective acoustic comfort [23], others have reported findings that are in agreement with the result of this paper. In a study of urban squares, Yang and Kang [24] noted that females appeared to be more aroused by mood-related sounds. In a comparison study between urban and rural areas, Pheasant et al. [25] observed that the male rating of tranquillity was higher than that of females in two natural (rural) environment. Another study supported this trend from a physiological basis, in which the heart rate was noted to reduce when people were presented with pleasant sounds compared to the case of unpleasant sounds, and this trend was more significant in males than in females [26]. All these studies reported that the degree of perceived pleasantness of sounds was different for different genders; however, there exists a disagreement in terms of which gender was affected (aroused) the most. It is hypothesised that this difference is largely because of the different contexts and sound sources considered in the various studies.

4. Discussion

4.1 Appropriateness of the contexts

Aletta et al. [27] suggested that the appropriateness of sound should be considered as an additional dimension to complement the PCM. As the current study was designed to explore the context beyond the urban setting, none of the three sound types (urban sounds) was appropriate for any of the contexts, although certain nuances were present within each sound source. 1) Street music is normally limited to an urban public space; however, as a form of music, it is difficult to argue that such music is inappropriate for any environment, as the music is often used as an artificial background sound. For instance, one participant, in the street music group, responded to an open question as follows: 'the well-combined music and imagery helps me to relax from my busy schedule'. Another participant stated that he 'feels like I was there'. This feeling of appropriateness might partially help explain the lack of significant mood changes in the case of street music and the related context groups. 2) A water feature is often present in a Buddhist temple; however, such features usually produce a calming stream sound instead of a high flow rate fountain sound (as in this study). Moreover, in the Alpine meadow context, the visual cue did not include any water feature, which likely contributed to the inappropriateness of the fountain sound. This aspect can help explain the significant change in the confusion score in the FTc and FA groups compared to that of the fountain sound. 3) Finally, the traffic sound was the only sound type unfit for all the contexts, as it had the most significant mood changes among the Alpine meadow and temple chanting context groups. In conclusion, the sound appropriateness likely affected the mood changes; however, the examination of this dimension is beyond the scope of this study and requires further exploration.

4.2. PCM and mood categories

The PCM is a widely used model to evaluate the perceived soundscape quality. The model was developed by Axelsson et al. [2], with reference to Russell's [1] core affect model as a part of the theoretical foundation. The two models include similar components, namely, pleasantness–unpleasantness and eventful–uneventful (or activation– deactivation in the core affect). In the core affect case, the two components are used to explain the perceived moods and emotions from an environment. This ability to evaluate the mood states is also applied in the PCM. However, as noted in this study, this consideration may not be sufficient for certain specific categories. Russell argued that the traditional approach for the categorisation of moods and emotions was unclear and led to confusion in terms of the boundaries and definitions, and he thus developed the core affect model [1]. However, the author specified that the core affect alone is not sufficient to define all the mood states, and a clear mood categorisation is necessary to study the specific moods. To better assess the perceived sound quality, the PCM recommend the use of a third component, familiarity [2]. However, the relevance of the familiarity component in relation to mood states is unclear. It must be noted that the two components (pleasantness and activation/ eventfulness) in the core affect model can be used to directly assess the mood states; however, in the PCM, the two components are used to assess the mood-related aspect of sound. Although this difference does not affect the current discussion, it must be taken into account.

4.3 Requirement of more dimensions to study the mood states

The mood states used in this study can be grouped into unpleasant groups (anger, confusion, depression, tension, and fatigue) and pleasant groups (vigour) based on the desirability and benefit/drawback of such mood states. However, in terms of the eventful/activation, only anger, tension, and fatigue can be vaguely distinguished from the group. Table 3 indicates how the six mood states can be differentiated in terms of the pleasantness and activation based on the associated definition or description in the literature.

	Pleasantness/ Unpleasantness	Activation/ Deactivation	Definition or description of the mood states
Anger	Unpleasant	Activation	A process provoked by triggers such as a perceived threat or unfair treatment. The mood can spread and intensify in a cognitive loop. [28]
Confusion	Unpleasant	-	A mental state in which the reaction to environmental stimuli is inappropriate because the subject is bewildered perplexed, or disoriented. [29]
Depression (mood)	Unpleasant	-	Sadness or unhappiness, usually persistent. This may be a normal reaction to unpleasant events or the environment or may be the result of a genuine depressive illness. [30,31]
Tension (stress)	Unpleasant	Activation	Psychological stress is a particular relationship between the person and environment, which is appraised by the person as taxing or exceeding his or her resources and endangering his or her well-being. [32]
Fatigue (mentally)	Unpleasant	Activation	Boksem and Tops indicated that the feeling of fatigue may result from the subconscious analyses of the cost and benefits to expend or conserve energy. [33]
Vigour	Pleasant	-	In contrast to mental fatigue, vigour is the lack of cognitive activity, resulting in an energetic mental state.

Table 3. Correlation of the six mood states with the pleasantness and activation (eventful) dimensions, based on the definitions presented in the literature.

The two components of the core affect are somewhat interchangeable with the two dimensions of the mood/emotions, that is, valence and arousal [1]. The valence represents whether a mood/emotion is positive or negative, and the arousal represents the intensity. Based on the definition and descriptions of the six studied mood states, it is obvious that while it is fairly easy to identify whether the mood is positive or negative (valence), it is difficult to specify where the mood should be located in terms of the intensity. For example, moods such as vigour and depression can be activating (e.g. energetic) or deactivating (e.g. calm). Moreover, although moods such as anger, tension and fatigue appear on the activation side, they can also be extremely intense (e.g. furiousness and extreme anxiety) or mild. Furthermore, moods such as confusion may be undefinable by the activation dimension. Considering these differences, dimensions in addition to pleasantness and activation/eventfulness must be examined. A reasonable starting point may be familiarity/appropriateness, as discussed previously. This component is relatively less utilised in soundscape studies owing to its small effect; however, this component likely plays a crucial role in the interpretation of the perceived soundscape quality [27]. For mood-related soundscape studies, these additional dimensions may vary across mood/emotion categories.

4.4 Application

Despite the need to expand the varieties of the studied mood categories, sound sources, and types of context, this study may provide guidance for building and environmental design. The results indicate that the introduction of activities with dominant human sounds can significantly reduce negative emotions. The design of an enclosed space that obstructs the visual of the sound sources may help reduce negative emotions, although the effect seems to be limited to the sound types and mood categories. The same limited effect can also be observed by introducing greenery. This finding suggests that the designs of both enclosed environments and greenery can help improve the mood states, although large quantities of both features may have a diminishing effect. Music, however, exerts a dominant effect on stabilising the mood states, despite any contexts. Finally, changing the environmental context does not appear to induce any additional mental energy (vigour) in the visitors.

4.5 Limitation and future studies

This paper only explored a limited number of mood states in relation to certain acoustic environments. Despite the extensive past soundscape studies in the varieties of sound types, environmental contexts, and subjective evaluation of the acoustic environments, nearly none of these studies have examined acoustic environments based on their effect on distinctive emotions or mood states. This lack of focus on distinctive emotions means past findings in soundscape studies cannot be readily adapted into the framework presented by this paper. This limitation also narrows the applicability of current results to environments that are similar to the ones presented in this paper. To reduce such limitations, future studies should include distinctive emotions and mood states as a measurement of the subjective quality of soundscapes.

Furthermore, as the result of this study show, not all mood states were significantly affected under all acoustical environments. These results indicate that there is less likely to be a fixed set of emotions andmood states that is suitable to evaluate every type of acoustic environment. Hence, future studies need to select distinctive emotions and mood states based on their focused acoustic environments. For the purpose of selecting emotions and mood states, it is useful to look into past soundscape studies. Despite most studies having used PCM as the measurement for subjective quality of soundscapes instead of distinctive emotions and mood states, the results of these studies still indicate types of distinctive emotions and mood states that should be included for the evaluation of certain acoustical environments. Additionally, to measure different groups of distinctive emotions and mood states, suitable questionnaires should also be developed and tested in relation to different acoustic environments.

5. Conclusions

Certain significant trends existed among different contexts in terms of their influence on the mood states. Among the six mood states, all five negative moods (anger, confusion, depression, fatigue, and tension) were significantly affected when the contexts were combined with traffic or fountain sounds. These results indicate that altering the environmental context can be effective in changing the mood states experienced in an acoustic environment. This effect, however, did not occur in the case of street music combinations. In contrast to the negative moods, the only positive mood (vigour) was not affected by any of the contexts. This ineffectiveness indicates that the effect observed in certain contexts does not apply to all acoustic environments.

In terms of the influence of the different sound types combined with contexts on the mood score, notable differences were present across all the studied sound types. The significant changes in the mood scores were alleviated when contexts were added to the original sound sources. All three contexts showed different degrees of reduction effect in mood scores of negative mood states when combined with traffic sound and fountain sound. This indicates that the three contexts have reduction effects on negative mood states that can be experienced in an acoustic environment but that the effect varies depends on the sound type that presented in such environments. In terms of the visual and audio stimuli, the contexts exhibited varied effects on the mood scores; however, in the evaluation of each context, no similarity was found across the contexts.

The gender difference had a significant small (depression) to medium (fatigue) effect on the intensity of the mood change caused by the difference in the contexts and sound sources. These results indicate that gender has very limited effect on experienced mood states when considering the evaluation of acoustic environments. However, this implication is limited to the acoustic environment presented in this study.

Moreover, the pleasantness and eventfulness dimensions cannot fully explain the outcomes of the mood states in soundscape studies. New dimensions need to be developed to further understand the mood states. It is hypothesised that the developed dimensions may be different depending on the contexts, sound types, and mood states.

Acknowledgements

This paper has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 740696, project title: Soundscape Indice [SSID]). The experiment was conducted at the Inner Mongolia University of Technology in China. The authors are very grateful to the university students who helped to coordinate the experiment and students who participated in the experiment.

References

[1] J.A. Russell, Core affect and the psychological construction of emotion, Psychol. Rev. 110 (2003) 145, https://doi.org/10.1037/0033-295X.110.1.145.

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

[3] The British Standards Institution, BS ISO 12913-1:2014 - Acoustics — Soundscape Part 1: Definition and Conceptual Framework, ISO, 2014.

[4] S. Schachter, J. Singer, Cognitive, social, and physiological determinants of emotional state, Psychol. Rev. 69 (1962) 379, https://doi.org/10.1037/h0046234.

[5] J. Robinson, Deeper than Reason: Emotion and its Role in Literature, Music, and Art, Oxford University Press on Demand, 2005, https://doi.org/10.1093/0199263655.001.0001.

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

[7] D. Keltner, K. Oatley, J.M. Jenkins, Understanding emotions (3rd ed.). Hoboken, NJ: Wiley-Blackwell, 2013.

[8] Nico H. Frijda, Moods, emotion episodes, and emotions, Handbook of emotions, 381, 1993. In press, https://psycnet.apa.org/record/1993-98937-021, 403.

[9] P.D. O'Halloran, G.C. Murphy, K.E. Webster, Reliability of the bipolar form of the profile of mood states using an alternative test protocol, Psychol. Rep. 95 (2) (2004) 459–463, https://doi.org/10.2466/pr0.95.2.459-463.

[10] Y.H. Shin, K.B. Colling, Cultural verification and application of the profile of mood states (POMS) with Korean elders, West. J. Nurs. Res. 22 (1) (2000) 68–83, https://

doi.org/10.1177/01939450022044278.

[11] P.C. Terry, A.M. Lane, G.J. Fogarty, Construct validity of the profile of mood states - adolescents for use with adults, Psychol. Sport Exerc. 4 (2) (2003) 125–139, https://doi.org/10.1016/S1469-0292(01)00035-8.

[12] J. Van Heerden, J. Hoogstraten, Response tendency in a questionnaire without questions, Appl. Psychol. Meas. 3 (1979). Retrieved from, http://www.copyright.com/.

[13] M.V. Thoma, R. La Marca, R. Brönnimann, L. Finkel, U. Ehlert, U.M. Nater, The effect of music on the human stress response. PLoS one 8 (2013) e70156. https://doi.org/10.1371/journal.pone.0070156

[14] 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

[15] Y. Hao, J. Kang, H. Wörtche, Assessment of the masking effects of birdsong on the road traffic noise environment, J. Acoust. Soc. Am. 140 (2) (2016) 978–987, https://doi.org/10.1121/1.4960570.

[16] 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, https://doi.org/10.1121/1.3298437.

[17] S. Viollon, C. Lavandier, Multidimensional assessment of the acoustic quality of urban environments. In Conf. Proc. “Internoise”, Nice, France, 4 (2000) 2279–2284.

[18] J.Y. Hong, J.Y. Jeon, Designing sound and visual components for enhancement of urban soundscapes. J. Acoust. Soc. Am. 134 (2013) 2026–2836. https://doi.org/10.1121/1.4817924

[19] 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

[20] 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. https://doi.org/10.1121/1.3298437

[21] G. Watts, A. Khan, R. Pheasant, Influence of soundscape and interior design on anxiety and perceived tranquillity of patients in a healthcare setting. Appl. Acoust. 104 (2016) 135–141. https://doi.org/10.1016/J.APACOUST.2015.11.007

[22] W. Yang, J. Kang, Soundscape and sound preferences in urban squares: a case study in Sheffield. J. Urban Des. 10 (2005) 61–80. https://doi.org/10.1080/13574800500062395

[23] W. Yang, J. Kang, Acoustic comfort evaluation in urban open public spaces, Appl. Acoust. 66 (2005) 211–229, https://doi.org/10.1016/j.apacoust.2004.07.011.

[24] R. Pheasant, K. Horoshenkov, G. Watts, B. Barrett, 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 (2008) 1446 https://doi.org/10.1121/1.2831735

[25] K. Hume, M. Ahtamad, Physiological responses to and subjective estimates of soundscape elements. Appl. Acoust. 74 (2013) 275–281. https://doi.org/10.1016/J.APACOUST.2011.10.009

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

[27] N. Alia-Klein, G. Gan, G. Gilam, J. Bezek, A. Bruno, T.F. Denson, T. Hendler, L. Lowe, V. Mariotti, M.R. Muscatello, S. Palumbo, The feeling of anger: From brain

networks to linguistic expressions. Neurosci Biobehav. Rev. 108 (2020) 480–497. https://doi.org/10.1016/j.neubiorev.2019.12.002

[28] Farlex Partner Medical Dictionary. (n.d.). Confusion. (2012). Retrieved June 26 2021 from https://medical-dictionary.thefreedictionary.com/confusion

[29] Collins Dictionary of Medicine. (n.d.). Depression. (2004, 2005). Retrieved June 26 2021 from https://medical-dictionary.thefreedictionary.com/depression

[30] R.S. Lazarus, S. Folkman, The stress concept in the life sciences. Stress, Appraisal, and Coping, 1984. pp. 1–21.

[31] M.A. Boksem, M. Tops, Mental fatigue: costs and benefits. Brain Res. Rev. 59 (2008) 125–139. https://doi.org/10.1016/j.brainresrev.2008.07.001

[32] K. Jambrošić, M. Horvat, H. Domitrović, Assessment of urban soundscapes with the focus on an architectural installation with musical features. J. Acoust. Soc. Am. 134 (2013) 869–879. https://doi.org/10.1121/1.4807805