Comparison between perceived and felt emotions in the soundscape evaluation of urban open spaces

Yan Zhang,[1,2] Jian Kang,[3,a)] and Biao Yang[1]

[1] *School of Architecture, Harbin Institute of Technology, Shenzhen, Shenzhen, 518000, China*

[2] *School of Architecture and Engineering, Yan'an University, Yan'an, 716000, China*

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

This paper is part of a special issue on Advances in Soundscape: Emerging Trends and Challenges in Research and Practice.

Abstract: In the current study on soundscape, the distinction between felt emotion and perceived emotion in soundscape measurement has not been addressed as much as that in music studies. This research was conducted to investigate perceived and felt emotions associated with soundscape evaluation in urban open spaces, through a laboratory audio-visual experiment using photographs and binaural recordings of 16 urban open locations across Harbin, China. In total, 46 participants were required to assess both the "perceived emotion" and "felt emotion" of the soundscapes using a questionnaire (in Chinese). First, five felt emotions and seven perceived emotions associated with the soundscape were identified, among which the dominant factors were *enjoyment* and *excitement* for felt emotions and *comfortable* and *festive* for perceived emotions. Second, when comparing perceived and felt emotions, the holistic soundscape descriptor "preference" is more suitable for predicting

a)Email: j.kang@ucl.ac.uk

through felt emotion, while the holistic soundscape descriptor "appropriateness" is more suitable for predicting through perceived emotion. Third, preference is a more stringent soundscape descriptor than appropriateness, indicating a higher level of requirement in its definition. Meanwhile, preference is a more emotional soundscape descriptor than appropriateness. It may be inferred that for evaluating soundscapes, the more emotional the descriptor, the greater its stringency.

Keywords: felt emotion; perceived emotion; semantic differential analysis; urban open space

(Received 28 October 2023; revised 27 March 2024; accepted 28 March 2024; published online 25 April 2024)

[Editor: James F. Lynch]

Pages: 2756–2768

DOI: 10.1121/10.0025761

I. INTRODUCTION

Soundscapes are defined by the International Organization for Standardization[1] as "[the] acoustic environment as perceived or experienced and/or understood by a person or people, in context." Thus, "soundscape" refers to not only a perceptual construct but also how people actually experience acoustic environments.[2, 3] The effect of cognitive and psychophysiological aspects of users on the soundscape evaluation of urban open spaces has also been an important consideration.[4–6] Better evaluation of soundscapes needs more explicit soundscape descriptors, terms which are used to describe the perception of any acoustic environment.[7] Based on the concept of soundscape, a soundscape descriptor must provide a measure of at least one aspect of how people perceive, experience or understand the acoustic environment.[8] This leads to a paradigm shift in the research methods from physical to linguistic and psychological.[9–11] The recent development of new soundscape descriptors for more in-depth measurement, evaluation, and design of soundscapes is still in progress.[2]

Based on the original philosophy behind soundscape studies,[12] advancements in the field of soundscapes, including its descriptors, can potentially benefit from the research progress made in the domain of music. Felt and perceived emotions have been widely studied in the domain of music. These two emotions can be distinguished according to the experimental instructions given by the experimenter to the participants before the experiment. An emotion is classified as a felt emotion if the instructions state that the participants should focus on their own emotional experiences or the emotions that the music induces or arouses in them. This notion is in contrast to perceived emotional instructions that emphasize the perception and recognition of emotions represented or expressed by music *per se*.[13, 14] In 2002, Gabrielsson[15] first commented upon the distinction between perceived emotion and felt emotion in music. In relation to this topic, Scherer[16] formulated that it is often unclear whether a valence judgment (pleasant or unpleasant) concerns the appraisal of the nature of the stimulus object or event or rather the feeling induced by it. Similarly, arousal or activation ratings may refer to perceived activation in a situation (or image) or to the proprioceptive feeling of physiological arousal induced by the stimulus event.

Several consequent empirical studies also addressed this distinction.[17–19] For instance, Juslin and Laukka[17] conducted a questionnaire study on listening to music every day, in which the participants were asked to respond to the question: "If you perceive that the music expresses an emotion, do you also feel that emotion?" There are also some works that provide strong evidence for the distinction between the two emotions from a physiological perspective. In 2015, using functional magnetic resonance imaging, Tabei[20] found that the underlying neural basis of perceived and felt emotions during listening to music was not entirely consistent. The author stated that activity in the inferior frontal gyrus increased more during the perceived emotion task than during a passive listening task. In addition, the precuneus showed greater activity during the felt emotion task than during a passive listening task.[20] In view of the aforementioned findings from music research, it is promising to investigate the distinction between felt emotion and perceived emotion in the field of soundscape.

When "emotion" about soundscape was mentioned in past studies, whether it refers to felt emotion or perceived emotion is not clear. Both Axelsson *et al.*[21] and Cain *et al.*[22] investigated the relationship between soundscape and emotion, but neither of them distinguished felt emotion or perceived emotion. The emotion in Axelsson *et al.*[21] refers to perceived emotion in the proposed model of perceived affective quality, while the emotion that Cain *et al.*[22] focused on how a person feels when hearing a soundscape, i.e., felt emotion. Watts and Pheasant[23] attempted to investigate the relationship between perception and feeling of soundscape. They asked participants to perceive how wild and tranquil each soundscape is and to respond to how pleasant and excited the soundscape made them feel. However, the only two perceived attributes, wild and tranquil, were not about emotion; and the felt emotions included in their study are general emotion indices that are not specifically developed for soundscapes.

Affective and emotional processing of environmental stimuli is important to the assessment of soundscape quality.[24] Aletta *et al.*[8] noted that a majority of soundscape descriptors are inclined toward emotional evaluation. Axelsson *et al.*[21] and Axelsson[25] found that the two components (pleasantness and eventfulness) of perceived affective quality seem to summarize most of the relevant information of soundscapes. This means that "measures of perceived affective quality provide the same information as a measure of overall soundscape quality," as stated by Aletta *et al.*[8]

The development of more detailed metrics for measuring emotions has the potential to account for more soundscape information, such as "appropriateness," a holistic soundscape descriptor to assess whether or not a soundscape is appropriate for a place. A soundscape descriptor may either refer to a singular underlying dimension of soundscape (e.g., Pleasantness) or to soundscape holistically (e.g., "soundscape quality").[8] The latter is called holistic soundscape descriptors. The principal component analysis results of various soundscape variables, including suitable activity types, sound source kinds, good-bad evaluation, perceived affective quality, and appropriateness, show that appropriateness is the third principal component independent of pleasantness and eventfulness. Axelsson[25] claimed that "appropriateness" is independent of perceived affective quality and regarded it as a potential third dimension for characterizing soundscapes. Further investigations should be put on "appropriateness" for a better understanding of soundscape from emotional perspectives.

Past studies have addressed the issue of how important "appropriateness" is when evaluating soundscape. In two review papers, Brown *et al.*[26] and Brown[27] argued that it is central to assessing whether a soundscape is appropriate for a place in urban planning and design practices. After it was revealed through an experiment that a soundscape might be appropriate even if it is poor, Axelsson[25] raised a question: "Which information should then have priority: the appropriateness of the soundscape or how good or bad it is?" Therefore, this paper proposes to view descriptors from the perspective of "stringency," that is, the strictness and exaction of a descriptor's requirements.

This study aimed to investigate the relationship between emotion and soundscape evaluation in urban open spaces. Specifically, the research was conducted from three perspectives. The first involved defining the felt emotion and perceived emotion in the soundscape and identifying measurement dimensions for the soundscape in urban open spaces. The second involved quantifying the feasibility of evaluating soundscapes by felt and perceived emotions through existing holistic soundscape descriptors in terms of preference and appropriateness. Third, on this basis, we further compared the stringency of the holistic descriptors in the evaluation of soundscapes and explored the mechanism. We also compared and discussed the emotional dimensions of soundscapes in the Chinese and Western contexts. In this study, 16 audio-visual materials of urban open spaces were replayed in a controlled laboratory setting, and the participants' two emotions and their holistic evaluation of the soundscape were investigated through a questionnaire to explore the effect of emotion on soundscape evaluation.

II. METHOD

A. Audio-visual materials

It took approximately 10 min to answer each questionnaire in this study. Audio-visual materials can be played repeatedly in the laboratory to provide a stable audio-visual environment for the entire questionnaire answering process. Therefore, a laboratory-based experiment was conducted in this study. Overall, ten urban open spaces across Harbin were selected to collect audio-visual materials, including three squares, three parks, and four traffic areas. Figure 1 presents the distribution of the urban open spaces in the city. As the perceived acoustic environment varied greatly in different locations of the squares and parks with large areas, more than one piece of audio-visual material were collected in these squares and parks. Notably, 16 audio-visual materials from 16 different locations were collected in this study. The selected city was a large provincial capital city in northern China. The sites were selected from those well-known urban open spaces to the city residents.[28, 29] In addition, different types of spaces, including squares (3), parks (3), and traffic areas (4), were balanced to keep the variety of context. The binaural recording was carried out when the soundscape was typical in relation to the location. Hence, the audio-visual materials utilized in this experiment represent the typical acoustic environment of urban open spaces.

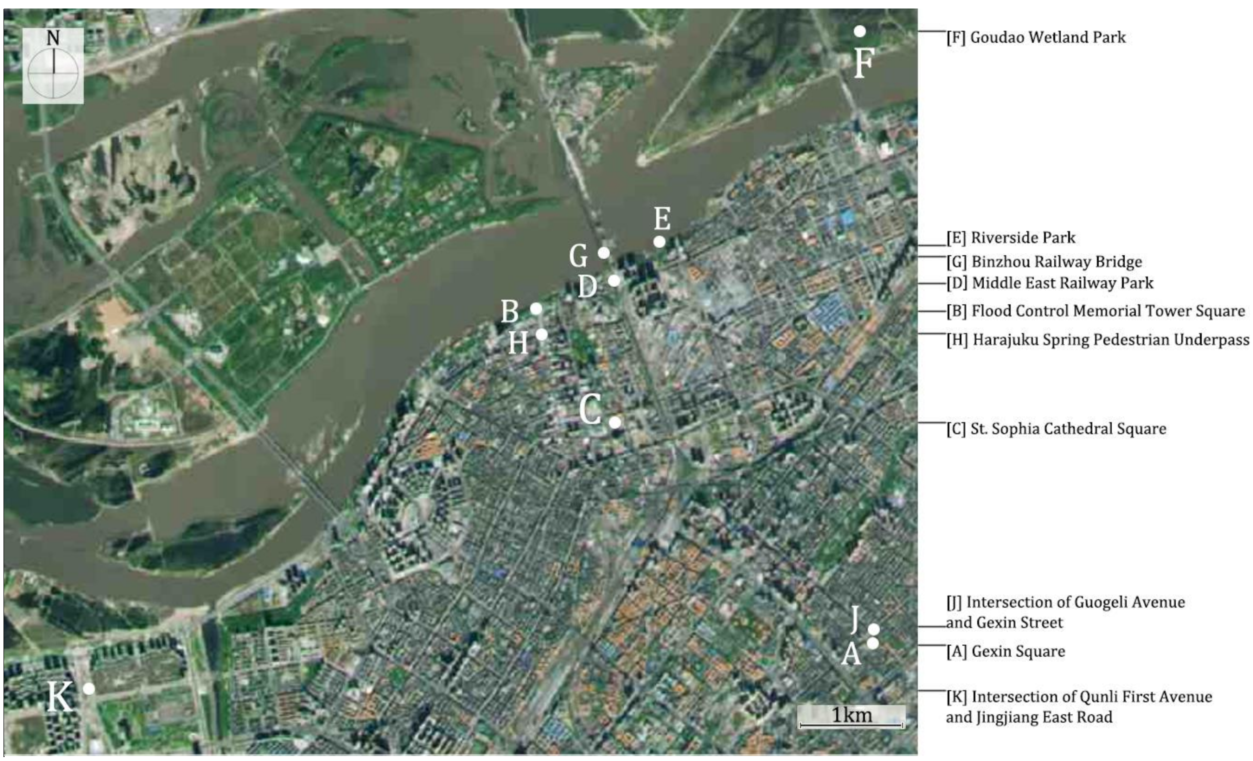

FIG. 1. (Color online) Distribution of the urban open spaces (white dots) in Harbin where the recordings took place and the code assigned to each of the urban open spaces (e.g., B).

To simulate highly complex situations in the field, the materials played in the laboratory included two parts: audio and visual. The materials were captured using audio (SQuadriga II with BHS I) and video (GoPro Fusion) recording, as shown in Fig. 2a. Considering that a majority of physiological indicators exhibit a distinct inflection point at the 60 s mark of the soundscape stimulus,[30, 31] an excerpt lasting for 60 s was cut from a binaural recording at each location. Importantly, to be accepted as a soundscape excerpt, the cut excerpt had to include both a sound event and ambiance. Additionally, to center the participant's attention on the soundtrack, visual information was delivered in a static form (photographs) rather than a dynamic form (videos).[25] A field of view photograph corresponding to the direction where the audio was recorded was cropped from a 360-degree panoramic photograph at each location. Figure 2(b) illustrates the static photograph presented to the participants. Table I shows information on the ten urban open spaces from which the 16 soundscape excerpts came. For each soundscape excerpt, a description of the sound composition is provided.

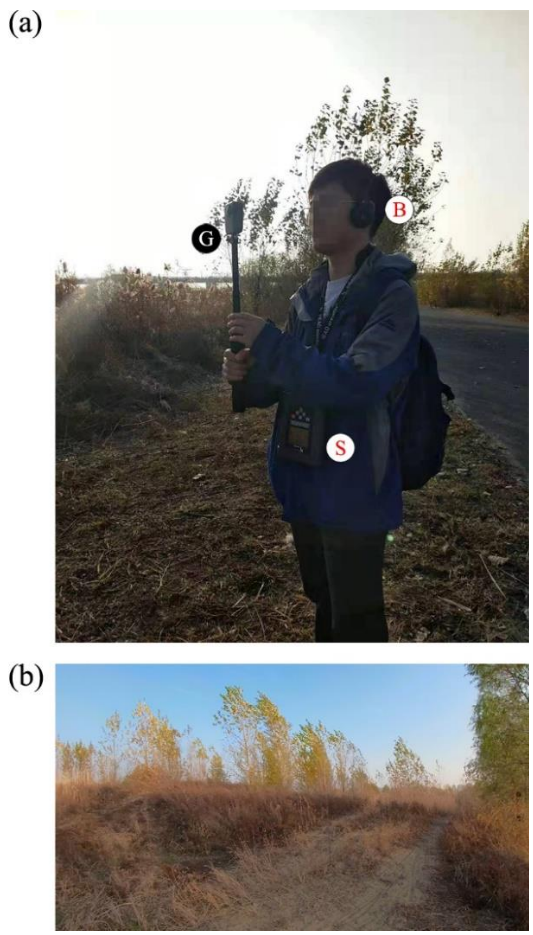

FIG. 2. (Color online) (a) Photograph of the recording setup. S, B, and G indicate, respectively, positions of the SQuadriga II, BHS I, and GoPro Fusion. The height of the BHS I is 1.65 meters. (b) example of the

photo of a soundscape excerpt (No. F2) taken with a spherical panoramic camera (GoPro Fusion) and transformed to a human eye view.

TABLE I. Information for the ten urban open spaces from which the 16 soundscape excerpts came. For each soundscape excerpt, a description of the sound composition is provided.

Category	Code	Name	Coordinates	Number of the Soundscape Excerpt	Time (Beijing Time)	Sound Composition
Square	A	Gexin Square	45.751898°, 126.659766°	A1	October 22, 2018, 1:14 PM	Car traffic, ventilation noise
	B	Flood Control Memorial Tower Square	45.780751°, 126.617225°	B1	November 01, 2018, 8:30 AM	Human voices, hawking, music, loudspeaker advertisements
				B2	November 01, 2018, 8:41 AM	Music, human voices
				B3	November 01, 2018, 8:54 AM	Loudspeaker advertisements, human voices, car traffic
				B4	November 01, 2018, 9:07 AM	Human voices, footsteps
	C	St. Sophia Cathedral Square	45.770382°, 126.627102°	C1	November 03, 2018, 1:43 PM	Human voices, car traffic, vehicle horns, construction works, music
				C2	November 03, 2018,	Human voices,

					1:53 PM	footsteps, bird twittering
Park	D	Middle East Railway Park	45.783078°, 126.626945°	D1	October 21, 2018, 9:52 AM	Technological sound, spraying water
				D2	October 21, 2018, 10:14 AM	Human voices, train traffic, wind
	E	Riverside Park	45.786201°, 126.632657°	E1	October 21, 2018, 12:41 PM	River water hitting the shore, human voices, car traffic
	F	Goudao Wetland Park	45.805212°, 126.658414°	F1	October 21, 2018, 2:53 PM	Bird twittering, car traffic
				F2	October 21, 2018, 3:01 PM	Rustling trees, wind
Traffic area	G	Binzhou Railway Bridge	45.784839°, 126.625721°	G1	October 21, 2018, 9:12 AM	Human voices, bicycle, footsteps, wind
	H	Harajuku Spring Pedestrian Underpass	45.778477°, 126.617984°	H1	October 21, 2018, 11:25 AM	Human voices, footsteps
	J	Intersection of Guogeli Avenue and Gexin Street	45.752057°, 126.660008°	J1	October 22, 2018, 1:24 PM	Car traffic, human voices, loudspeaker advertisements
	K	Intersection of Qunli First Avenue and	45.746582°, 126.560335°	K1	October 22, 2018, 4:45 PM	Technological sound

	Jingjiang East Road

As shown in Table II, a set of acoustic and psychoacoustic measures was calculated for all 16 soundscape excerpts, referring to:

(a) overall loudness: $L_{Aeq, 60s}$ [dB] and N_{10} [sone];

(b) temporal variability: L_{A10}-L_{A90} [dB] and N_{10}-N_{90} [sone]; Fluctuation Strength (FS) [vacil] and Roughness (R) [asper];

(c) spectral content of sound: $L_{Ceq, 60s}$-$L_{Aeq, 60s}$ [dB], and Sharpness (S) [acum].

TABLE II. Category of dominant sounds and acoustic and psychoacoustic measures calculated for all 16 soundscape excerpts.

Category of Dominant Sounds	Number of the Soundscape Excerpt	Acoustic Characterization							
		$L_{Aeq, 60s}$ (dB)	N_{10} (sone)	L_{A10}-L_{A90} (dB)	N_{10}-N_{90} (sone)	$L_{Ceq, 60s}$-$L_{Aeq, 60s}$ (dB)	FS (vacil)	R[a] (asper)	S[b] (acum)
Natural sound	E1	58.50	13.90	5.93	4.29	12.40	0.011	0.298	1.19
	F1	51.60	7.10	5.16	2.52	11.90	0.005	0.288	1.04
	F2	49.30	8.54	5.18	2.68	21.20	0.011	0.180	1.55
Music sound, human sound	B2	68.20	21.60	10.63	8.30	4.50	0.035	0.214	1.22
Human sound, music sound	B1	64.50	17.50	8.48	6.30	5.20	0.027	0.240	1.06
Human sound	H1	76.10	35.40	3.30	6.90	1.50	0.014	0.212	1.33
	B4	71.50	17.80	12.05	7.00	3.30	0.045	0.268	1.32

	C2	61.40	17.20	6.17	5.10	10.60	0.026	0.237	1.34
Human sound, technological sound	G1	58.20	14.50	8.39	6.55	17.60	0.033	0.384	1.11
	D2	65.60	24.50	9.07	10.40	13.10	0.017	0.287	1.21
	B3	64.20	19.10	7.05	5.70	7.50	0.040	0.246	1.21
	C1	61.60	17.70	4.00	3.30	11.00	0.014	0.121	1.44
Technological sound, human sound	J1	68.90	25.70	5.13	6.60	11.30	0.020	0.236	1.31
Technological sound	D1	78.80	53.00	2.47	7.50	5.20	0.011	0.844	1.87
	A1	72.90	31.20	2.19	5.00	10.60	0.011	0.254	1.19
	K1	86.00	91.60	17.31	59.00	4.90	0.042	0.435	1.21

[a]Roughness standard: ECMA-418-2 (1st Edition).

[b]Sharpness method: DIN 45 692.

The 16 soundscape excerpts include a wide range of soundscapes with a large variation in overall sound-pressure level (49.3–86.0 dB $L_{\mathrm{Aeq,\,60\,s}}$) and a great diversity of sounds, including sounds of music, technology, humans, and nature. Furthermore, the analysis of psychoacoustic parameters also contributed to the selection of a rich sample of soundscapes, considering the significance of the differences in psychoacoustic parameters for clustering soundscapes.[32] Sharpness, a sensation of timbre,[33] increases with increasing high frequency content.[34] Roughness is related to the beating phenomenon or relatively quick temporal changes of sound (maximum at 70 Hz).[34] Fluctuation strength, similar to roughness but related to slower changes (maximum at 4 Hz),[34] refers to the sound quality perceived when the individual loudness fluctuations are audible.[35] The sharpness and roughness of the soundscape excerpt No. D1, which features strong technological sounds and spraying water, are the highest. The fluctuation strength of the three excerpts dominated by natural sounds is the lowest. The soundscape excerpt No. B4, characterized by human voices and footsteps of many people, has the highest fluctuation strength.

According to the minimum differences in these metrics which are subjectively perceived: just noticeable differences (JND),[36] the increase in sharpness from the lowest to the highest exceeds 20 times the JNDs of 0.04 acum.[36–38] A just noticeable difference level in both roughness and fluctuation strength is estimated to be 17%.[37, 39] For roughness in this study, the highest is about seven times the lowest. It is worth noting that it is perceptually "no longer rough" at 0.1 asper.[38] The minimum roughness value of the soundscape samples used in this study is 0.121 asper. For fluctuation strength in this study, the highest is nine times the lowest. Therefore, there is sufficient variation in the psychoacoustic parameters in the sampled acoustic environments.

B. Questionnaire design

1. Soundscape emotion measurement

To obtain a tool to measure soundscape emotion, that is, semantic differential scales, a series of "semantic" items need to be achieved first. Accordingly, the following four steps are required.[21, 40]

First, 92 candidate emotional-descriptive terms related to soundscapes were collected from five studies.[4, 21, 22, 26, 40]

Second, 34 university students (11 women, 23 men; mean age: 21.7 years, range: 18–27 years) were asked to assess the suitability of 92 terms, presented in a random order, for the acoustic environment of urban open spaces in a broad sense. The participants were recruited via an experimental information platform at the Harbin Institute of Technology; and the gender bias in the sample is reflective of the gender ratio at the institute, which is significantly skewed towards males. They self-reported having normal hearing and either regular or corrected-to-normal vision. The 92 terms used in the item selection had been translated into Chinese, as all participants were Chinese. Moreover, the final questionnaire of this study was entirely in Chinese. These participants were given a guided soundwalk to experience the acoustic environment in urban open spaces prior to the current experiment. Participants were requested to rate each term according to the following criterion: “According to you, does this adjective describe an affective state with a specific affective ‘color,’ so that, to describe this state, you would choose to use this adjective over another one?” This line of questioning originates from the study on emotions in the field of music, specifically referring to the research conducted by Zentner *et al.*[40] Terms that received “yes” as the answer from >50% of the participants were retained.

Third, according to the dictionaries, the synonyms and antonyms of a word were divided into a group of words. Only the most frequently used word was retained in one group. Accordingly, 63 terms were retained. The frequency here refers to the frequency of being evaluated “yes” mentioned above.

Fourth, a questionnaire was developed for the experiment. Under the topic “How does this sound environment make you feel?” the terms were changed into expressions like “I feel…,” and the 63 terms were reduced to 42 grammatically correct indices. Thus, the questionnaire scales for felt emotions have been set up. Under the topic “For each scale, to what extent do you agree or disagree that the present surrounding sound environment is…,” “xx” (an adjective) was used as an index, thereby reducing the 63 terms to 51. Thus, the questionnaire scales for perceived emotions have also been set up.

In this way, we established two semantic differential scales to measure felt and perceived emotions in soundscapes. Each item was rated on a five-point Likert scale (1 = “strongly disagree,” 2 = “disagree,” 3 = “neither agree, nor disagree,” 4 = “agree,” and 5 = “strongly agree”). All items in each scale (i.e., the 42 items in the felt emotion topic and the 51 items in the perceived emotion topic) were presented in a questionnaire in a prepared random order to avoid scaling order effects.[41] Thus no questionnaire with items was presented in the same order.

Through the questionnaire design, each participant evaluated felt emotion before perceived emotion in the evaluation of each soundscape excerpt. It was considered important that the participants rate felt emotions first (i.e., being a more subjective measure, it may dilute faster than the more objective evaluation of the perceived emotion). Another reason is that the felt emotion may have included physiological responses that decrease as a function of time. Hence, it is imperative to evaluate the emotions felt immediately after listening to the excerpt.[19]

2. Holistic evaluation of soundscape

In this study, we selected two descriptors, namely, preference for the soundscape and appropriateness of the soundscape to the place, to holistically evaluate the soundscapes because these two descriptors have a common ground that "the evaluation of a certain sound environment depends on the context."[25, 26] The former assesses whether an overall soundscape is liked, while the latter evaluates whether a soundscape is appropriate for a place. Preference was measured on a five-point scale (1 = "strongly dislike," 2 = "dislike," 3 = "neither like nor dislike," 4 = "like," and 5 = "strongly like"), and appropriateness was also supplied on a five-point scale (1 = "not at all," 2 = "slightly," 3 = "moderately," 4 = "very," and 5 = "perfectly").

C. Experimental design

For the 48 subsets, each of them consisted of six soundscape excerpts, which were randomly selected from the collection of the 16 excerpts in a counter-balanced manner. For the 48 participants, each of them was randomly assigned one subset of soundscape excerpts (six excerpts). In other words, each one of the 16 soundscape excerpts was evaluated by 18 participants for one time, i.e., 18 evaluations in total to ensure a balance of data: $(6 \times 48 \times 1)/16 = 18$.

The experiment was conducted in an audio-visual laboratory at the Harbin Institute of Technology with background sounds lower than 11 dB(A). The 16 audio files were stored as HEAD Data files and binaurally replayed in headphones (BHS I) from the original recording system

(SQuadriga II) at an authentic sound pressure level.

The distinction between felt emotion and perceived emotion should be clearly explained to the participants, whether in within-subjects or between-subjects design.[15] Therefore, before the commencement of this experiment, the experimenter fully explained the distinction to the participants. The content of the explanation of the meaning of felt and perceived emotions to the participants is the same as the definition of felt and perceived emotions in the second paragraph of Sec. I.

It took approximately 10 min to complete each questionnaire in this study. Each participant scaled six soundscape excerpts. Six alternative versions of the questionnaire were presented to each participant in order to eliminate order effects on the items of the questionnaire. On average, it took approximately 60 min for each participant to finish all six blocks. Each participant received a small monetary compensation for their voluntary participation afterward.

D. Selection of participants

Based on the selection criteria of previous studies,[6] people with normal hearing and regular or corrected-to-normal vision were selected as study participants. The valid participants were 21 female and 25 male university students (19 undergraduates and 27 postgraduates) from the Harbin Institute of Technology, aged 17–34 years [M_{age} = 22.6 years, standard deviation (SD) = 3.9]. They were a completely different set of students from the 34 students who participated in the item selection. Twelve questionnaires were not completed because participants quit midway, and thus they were excluded. Eventually, there were 264 ($6 \times 46-12$) completed questionnaires for further processing. The participants were recruited via adverts at the school. They received a small monetary compensation for their voluntary participation. This study was approved by the Institutional Review Board (IRB Number: HIT-2023027).

E. Data analysis

ArtemiS SUITE 13.5 (a modular software platform; Head Acoustics GmbH) was used to calculate a set of acoustic and psychoacoustic measures. The acoustic and psychoacoustic measures were based on the left or right channels of the binaurally recorded soundscape excerpts with the maximum value of both ears. Sharpness is calculated according to DIN 45692 standard.[42] Roughness is calculated based on ECMA-418-2 (1st Edition) standard.[43] These soundscape excerpts are sorted in Table II by the category of dominant sounds (natural sounds, music sounds, human sounds, and/or technological sounds), and this was assessed by the participants.

IBM SPSS software was used to create a database of results. The expectation-maximization algorithm is suggested as an appropriate method of inputting a small number of missing values in the data.[44] In our study, the percentage of missing values was 0.59%. Accordingly, the reliability and validity of the questionnaire with a total of 103 items were examined.

The reliability of the questionnaire was characterized by the McDonald's omega coefficient, and the Kaiser-Meyer-Olkin (KMO) test was utilized to evaluate its construct validity. The overall reliability of the questionnaire was high (as indicated by McDonald's omega coefficient of 0.933), and its construct validity was also high, corresponding to a KMO criterion of 0.940. The significance level of Bartlett's test of sphericity was 0.000.

The data were analyzed by the following methods. Factor analysis was used to identify factors that characterize the soundscape from the perspectives of felt and perceived emotions by integrating many items in a small number of fundamental dimensions. Then, in order to assess the relationship and difference between the two types of emotions, a correlation analysis was carried out between the felt emotion factors and the perceived emotion factors. Two sets of linear multiple regression analyses using the stepwise method were performed to quantify the prediction accuracy of felt emotion and perceived emotion on the holistic evaluation of the soundscape in terms of preference and appropriateness. A scatterplot was drawn in order to analyze the relationship between the values of appropriateness and preference for each soundscape excerpt, and a simple linear regression model was fitted to the appropriateness-preference scatterplot.

III. RESULTS AND DISCUSSION

A. Semantic differential analysis of felt emotion and perceived emotion in the soundscape

Factor analysis was conducted on the felt and perceived emotions respectively using the data from all 264 valid questionnaires for the 16 soundscape excerpts. Promax rotated principal component analysis has been widely employed to extract oblique factors. As there is no reason to assume that emotions must be organized orthogonally,[40] we ran the same analyses (Kappa = 4) on the 42 items of felt emotion and 51 items of perceived emotion. More information on what input items were used for the two sets of factor analyses is provided in the supplementary material. Consequently, a correlation still exists between these extracted factors after oblique rotation.

1. Felt emotion factors of the soundscape

With a criterion factor of the Kaiser's criterion (eigenvalue > 1)[45] and the visual inspection of the scree plot,[46] five factors of felt emotion in the soundscape of urban open spaces were determined. Before oblique rotation, these factors together accounted for 64% of the total variance, which is acceptable considering the complexity of urban outdoor soundscapes.[9, 47, 48] It also shows that these five factors covered the main aspects of felt emotions associated with the soundscape in urban open spaces.

Specifically, the pattern matrix (refer to the supplementary material) shows that factor 1 is best explained by (I feel...) peaceful, inspired, soothed, fascinated, refreshed, enjoying, and comforted, which we have labeled as *enjoyment*. The term "享受感" (xiǎng shòu gǎn; enjoyment; see the first component in Table III) in the Chinese context tends to lean more towards the concept of "enjoying oneself". Factor 2 is best explained by happy, fiery, thrilled, and excited, and is therefore labeled as *excitement*. Factor 3 is best explained by sad, emptiness, and lonely and is labeled as *desolation*. Factor 4 is best explained by unusual, irritated, scared, and confused and is labeled as *tension*, which has a negative valence as is evident from these main adjectives linked to this dimension; and Factor 5 is labeled *familiarity*. Table III lists the percentage of the total variance covered by each factor and the rotation sums of the squared loadings. The table shows that the first two are the main factors.

TABLE III. Felt emotion factors of the soundscape.

Component		Chinese Phonetic Alphabets	English Translation	Extraction Sums of Squared Loadings		Rotation Sums of Squared Loadings[a]
				Total	% of Variance	Total
1	享受感	xiǎng shòu gǎn	Enjoyment	18.230	43.405	17.981
2	兴奋感	xīng fèn gǎn	Excitement	3.686	8.776	6.719
3	失落感	shī luò gǎn	Desolation	2.561	6.097	4.687
4	紧张感	jǐn zhāng gǎn	Tension	1.482	3.528	4.342
5	熟悉感	shú xī gǎn	Familiarity	1.078	2.567	1.530

[a]When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

Reliability of the scales for felt emotions was analyzed using McDonald's omega ($\omega = 0.896$), and it was acceptable. Meanwhile, KMO = 0.957 > 0.6, which meets the requirements for factor analysis. Factor analysis requires a relatively large sample size. Kang and Zhang[9] indicated that a sample size of 100–150 for factor analysis is generally acceptable for evaluating soundscapes in urban open public spaces. Furthermore, factor analysis requires a sample size to be manyfold to the number of variables. In the field of urban outdoor soundscape assessment, Axelsson *et al.*[21] conducted a factor analysis on 116 attribute scales with a sample size of 500. As 264/42 > 500/116, this analysis satisfied the minimum sample size requirement. These outcomes suggest that the results can be considered reasonable and stable.

2. Perceived emotion factors of the soundscape

Notably, seven factors of perceived emotion in the soundscape of urban open spaces were determined with a criterion factor of the Kaiser's criterion (eigenvalue > 1)[45] and the visual inspection of the scree plot.[46] Before oblique rotation, these factors together accounted for 66% of the total variance, which is acceptable considering the complexity of urban outdoor soundscapes.[9, 47, 48] It also shows that these seven factors covered the main aspects of perceived emotions associated with the soundscape in urban open spaces.

Specifically, the pattern matrix (refer to the supplementary material) shows that factor 1 is best explained by tranquil, soothing, comfortable, disturbing, refreshing, and pleasant, which we have labeled as *comfortable*. Factor 2 is best explained by festive, fiery, exciting, invigorating, and full of life, and is therefore labeled as *festive*. Factor 3 is best explained by empty, lonely, desolate, and gloomy and is labeled as *desolate*. Factors 4–7 are labeled as *familiar*, *various*, *attention-grabbing*, and *nostalgic*, respectively. Table IV lists the percentage of total variance covered by each factor and the rotation sums of the squared loadings. The table shows that the first two are the main factors.

TABLE IV. Perceived emotion factors of the soundscape.

Component	Chinese Phonetic Alphabets	English Translation	Extraction Sums of Squared Loadings		Rotation Sums of Squared Loadings[a]
			Total	% of	Total

					Variance	
1	舒适的	shū shì de	Comfortable	19.141	37.532	18.546
2	喜庆的	xǐ qìng de	Festive	5.788	11.349	9.118
3	荒芜的	huāng wú de	Desolate	3.199	6.272	5.153
4	熟悉的	shú xī de	Familiar	2.025	3.971	3.410
5	多样的	duō yàng de	Various	1.391	2.727	3.692
6	吸引注意的	xī yǐn zhù yì de	Attention-grabbing	1.225	2.401	2.594
7	怀旧系	huái jiù xì	Nostalgic	1.064	2.087	2.389

[a]When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

Reliability of the scales for perceived emotions was analyzed using McDonald's omega (ω = 0.896), and it was acceptable. Meanwhile, KMO = 0.940 > 0.6, which meets the requirements for factor analysis. As mentioned above, because 264/51 > 500/116, this analysis met the minimum sample size requirement. These findings suggest that the results can be considered reasonable and stable.

3. Comparison of the two types of emotions

In one study, both felt emotion factors and perceived emotion factors of soundscapes were obtained to define and compare the two types of emotions. Table V shows the correlation matrix between felt and perceived emotion factors. A certain correspondence between felt and perceived emotions is evident.

TABLE V. Pearson's coefficients of correlation between felt emotion factors and perceived emotion factors. * $p \leqslant 0.05$, ** $p \leqslant 0.01$, and *** $p \leqslant 0.001$.

Felt Emotions / Perceived Emotions	Enjoyment	Excitement	Desolation	Tension	Familiarity
Comfortable	0.929***	0.313***	-0.266***	-0.437***	-0.072
Festive	0.418***	0.777***	-0.339***	-0.123*	0.177**
Desolate	-0.069	-0.466***	0.570***	0.492***	-0.062
Familiar	0.016	0.196**	-0.156*	-0.314***	0.602***
Various	-0.002	0.384***	-0.225***	-0.248***	0.336***
Attention-grabbing	0.123*	0.021	0.029	0.628***	-0.082
Nostalgic	0.141*	0.327***	0.158*	-0.135*	-0.039

4. Discussion

The following two studies are typical, explicit, and complete dimension studies of soundscape emotions:

1. Axelsson *et al.*[21] conducted a laboratory experiment to establish the dimensions of the perceived affective quality of soundscapes. The Swedish participants measured the excerpts of binaural recordings of urban outdoor soundscapes on 116 attribute scales (in Swedish),[49] resulting in two dimensions: pleasantness and eventfulness.
2. Cain *et al.*[22] performed a laboratory experiment to establish the felt emotional dimensions of soundscapes. The British participants measured the excerpts of binaural recordings of typical urban setting soundscapes in five dimensions into which the researchers divided 45 English words, resulting in two dimensions: calmness and vibrancy.

In this study, a laboratory experiment was administered to establish the dimensions of perceived and felt emotions in soundscapes simultaneously. The Chinese participants measured the excerpts of binaural recordings of urban open space soundscapes on both 51 items of perceived emotion and 42 items of felt emotion (in Chinese). The initial candidate term sources included the above two studies, and the resulting dimensions are as described above (refer to Secs. III A 1 and III A 2).

In terms of felt emotion, “calm” is not included in the 42 items in this study. The similar index “I feel peaceful” contributes almost entirely to the first common factor, *enjoyment*. The index “I feel vibrant” contributes roughly equally to the first common factor, *enjoyment*, and the second common factor, *excitement*.

In terms of perceived emotion, the index “pleasant” mostly contributes to the first common factor, *comfortable*. “Eventful” has no vocabulary to correspond to in the Chinese context. In the second step of obtaining semantic items in Sec. II B 1, “多事件的” (duō shì jiàn de; eventful) was included. It was then deleted because the frequency of being evaluated “yes” was only 32%. The term “多事件的” (duō shì jiàn de; eventful) is used to convey the sense of “various events happening”. It does not have the same natural usage or frequency in Chinese as “eventful” does in English.

The circumplex model[50] represents emotions as a mixture of two core dimensions (valence and arousal). Overall, the results of this study and those of Axelsson *et al.*[21] and Cain *et al.*[22] are all within the framework of Russell’s circumplex model,[50] as stated by Fiebig *et al.*[24] in a review paper. From the results of this study, arousal can be felt by Chinese people, that is, Excitement. Meanwhile, the properties of soundscapes that induce feelings of arousal can be summarized as “Eventful” in English, but there is no such equivalent Chinese word. The same difficulty has also been encountered in languages other than English when translating “eventful,” for example, in Spanish there is no similar term either and it seems more accurate to use “lively.”[51] From the results of the current study, “喜庆的” (xǐ qìng de; festive; see the second component in Table IV) appears to be a relatively accurate label in Chinese for the extracted factor of the second dimension of perceived emotion in the soundscape.

In conclusion, in the Chinese context, the core dimensions of soundscape emotion are in agreement with the two dimensions (valence and arousal). However, in the comparison between Chinese and Western contexts, the percentages of total variance in soundscape emotion covered by these two dimensions are different, as shown in Table VI. In particular, the second dimension of emotion (arousal) was relatively weak in the soundscape in the Chinese context.

TABLE VI. Comparison of soundscape emotion studies. This table presents the variance in soundscape emotion explained by core dimensions in both Chinese and Western contexts.

		% of Variance						
		Dimensions						
		1	**2**	3	4	5	6	7
Perceived emotion	Axelsson *et al.*[a,b]	50	18	6				
	Current study	38	11	6	4	3	2	2
Felt emotion	Cain *et al.*[c]	60	20					
	Current study	43	9	6	4	3		

[a]Reference 21.

[b]Apart from the first three components, another eight components satisfied Kaiser's criterion (eigenvalue > 1). Their contribution to the explained variance was marginal at 1%–3%, and they could not be meaningfully interpreted.

[c]Reference 22.

In this study, the accounted overall variance for felt and perceived emotions is only 64% and 66%, respectively. These figures align with some results from previous urban outdoor soundscape studies.[9, 47, 48] The significant variations in the acoustic environment of urban open spaces could potentially account for this level of accounted overall variance.

B. Relationship between emotion and holistic evaluation of the soundscape

To quantify the prediction accuracy of felt emotion and perceived emotion on the holistic evaluation of the soundscape (i.e., preference for the soundscape and appropriateness of the soundscape to the place), the questionnaire data were subjected to six sets of linear multiple regression analyses using the stepwise method. The criteria for selecting independent variables in the stepwise method were set to the default settings of the SPSS system (Probability-of-F-to-enter ≤ 0.050, Probability-of-F-to-remove ≥ 0.100). It should be noted that the constants in the developed regression models were derived based on emotion scales ranging from one to five. The preference and appropriateness scales also ranged from one to five.

1. Felt emotion factors and holistic evaluation of soundscape

Overall, five felt emotion factors of the soundscape were used as independent variables. Factor scores were calculated using the regression method.

In the first regression model, preference for the soundscape was utilized as the dependent variable. The relationship between the dependent and independent variables is shown in Eq. (1),

$$PREF = 0.992*enjoyment - 0.256*tension + 2.932. \tag{1}$$

Preference for the soundscape (*PREF*) was best predicted by felt emotion factor 1, *enjoyment* ($\beta = 0.825$, $t = 30.626$, $p < 0.001$), and felt emotion factor 4, *tension* ($\beta = -0.213$, $t = -7.906$, $p < 0.001$). *Enjoyment* had a positive and significant effect, and *tension* had a negative influence. The negative loading of the later factor in Eq. (1) corroborates a negative valence of the fourth dimension of felt emotion. Together, they explained 82.5% (adjusted R^2) of the variance in the preference for the soundscape ($F = 620.67$, $p < 0.001$). The regression model was statistically significant. The results showed that the accuracy of predicting preference for the soundscape by felt emotion was 82.5%.

In the second regression model, the appropriateness of the soundscape to the place was used as the dependent variable. The relationship between the dependent and independent variables is shown in Eq. (2),

$$APPROPR = 0.359*enjoyment + 0.233*familiarity - 3.485. \quad (2)$$

The appropriateness of the soundscape to the place (*APPROPR*) was best predicted by felt emotion factor 1, *enjoyment* ($\beta = 0.325$, $t = 5.669$, $p < 0.001$), and felt emotion factor 5, *familiarity* ($\beta = 0.210$, $t = 3.670$, $p < 0.001$). Both variables had a positive influence. Together, they explained only 13.8% (adjusted R^2) of the variance in the appropriateness of the soundscape to the place ($F = 21.98$, $p < 0.001$). The regression model was statistically significant. The outcomes showed that the accuracy of predicting the appropriateness of the soundscape to the place by felt emotion was only 13.8%.

2. Perceived emotion factors and holistic evaluation of soundscape

Furthermore, seven perceived emotion factors of the soundscape were employed as independent variables, and factor scores were calculated using the regression method.

In the third regression model, preference for the soundscape was used as the dependent variable. The relationship between the dependent and independent variables is shown in Eq. (3),

$$PREF = 1.036*comfortable + 0.13*nostalgic + 2.932. \quad (3)$$

PREF was best predicted by perceived emotion factor 1, *comfortable* ($\beta = 0.861$, $t = 29.025$, $p < 0.001$), and perceived emotion factor 7, *nostalgic* ($\beta = 0.108$, $t = 3.628$, $p < 0.001$). Both variables had a positive influence, with *comfortable* having a significant influence. Together, they explained 77.1% (adjusted R^2) of the variance in the preference for the soundscape ($F = 443.093$, $p < 0.001$). The regression model was statistically significant. The results showed that the accuracy of predicting preference for the soundscape by perceived emotion was 77.1%, which is lower than that of predicting preference for the soundscape by felt emotion (82.5%).

In the fourth regression model, the appropriateness of the soundscape to the place was used as the dependent variable. The relationship between the dependent and independent variables is shown in Eq. (4),

$$APPROPR = 0.394*comfortable + 0.334*familiar + 0.170*desolate - 3.485. \quad (4)$$

The *APPROPR* was best predicted by perceived emotion factor 1, *comfortable* ($\beta = 0.356$, $t = 6.447$, $p < 0.001$), perceived emotion factor 4, *familiar* ($\beta = 0.302$, $t = 5.318$, $p < 0.001$), and perceived emotion factor 3, *desolate* ($\beta = 0.154$, $t = 2.697$, $p < 0.01$). All variables had a positive influence. Together, they explained 21.5% (adjusted R^2) of the variance in the appropriateness of the soundscape to the place ($F = 25.049$, $p < 0.001$). The regression model was statistically significant, and the findings confirmed that the accuracy of predicting the appropriateness of the soundscape to the place by perceived emotion was 21.5%, which is higher than that of predicting the appropriateness of the soundscape to the place by felt emotion (13.8%).

3. Combination of both felt and perceived emotion factors

To further illustrate, five felt emotion factors and seven perceived emotion factors of the soundscape were combined as independent variables. First, the appropriateness of the soundscape to the place was used as the dependent variable and fitted into a linear multiple regression analysis using the stepwise method. The criteria for selecting independent variables were the same as above. The results showed that the equation of the regression model was exactly the same as Eq. (4), with an adjusted R^2 still at 21.5%.

Next, preference for the soundscape was used as the dependent variable and fitted into a linear multiple regression analysis using the stepwise method. The criteria for selecting independent variables were the same as above. The results showed that the equation of the regression model is as follows:

$$PREF = 0.985*enjoyment - 0.250*tension + 0.063*nostalgic + 2.932. \quad (5)$$

Equation (5) includes perceived emotion factor 7, *nostalgic*, which is not included in Eq. (1).

With the inclusion of the new variable, the unstandardized coefficients of the original variables also changed. The three factors, felt emotion factor 1, *enjoyment* ($\beta = 0.819$, $t = 30.409$, $p < 0.001$), felt emotion factor 4, *tension* ($\beta = -0.208$, $t = -7.711$, $p < 0.001$), and perceived emotion factor 7, *nostalgic* ($\beta = 0.052$, $t = 1.998$, $p < 0.05$), together explained 82.7% (adjusted R^2) of the variance in the preference for the soundscape ($F = 419.854$, $p < 0.001$). Equation (5) explained 0.2% more of the variance than Eq. (1), and the regression model was statistically significant.

4. Discussion

Axelsson[25] found that appropriateness is independent of perceived affective quality. This study attempted to explain the appropriateness of soundscapes, but neither perceived emotion nor felt emotion was found to be effective. However, it can be observed that perceived emotion is slightly stronger than felt emotion in explaining soundscape appropriateness.

As Sec. III B 2 shows, in the predictive model of perceived emotion for appropriateness, there is a significant positive correlation between appropriateness and the perceived emotion factor "familiar." In other words, the more "familiar" the soundscape is perceived, the more "appropriate" it is considered.

The combination of both felt and perceived emotion factors offered negligible improvement in the prediction accuracy of the holistic evaluation of the soundscape. This indicates that the part of the appropriateness of the soundscape to the place that can be explained by felt emotions is entirely contained within what can be explained by perceived emotions. A similar observation was also found for preference, as to appropriateness.

C. Importance of emotion viewed from the stringency of holistic soundscape descriptors

Through an experiment, Axelsson[25] highlighted that a soundscape might be appropriate even if it is poor. Hence, he asked, "Which information should then have priority, the appropriateness of the soundscape or how good or bad it is?"

To find a research approach to answering this question, we drew a scatterplot of 16 soundscape samples. The arithmetic mean of the preference values for each soundscape excerpt was taken as the abscissa. The arithmetic mean of the appropriateness values was used as the ordinate, as shown in Fig. 3.

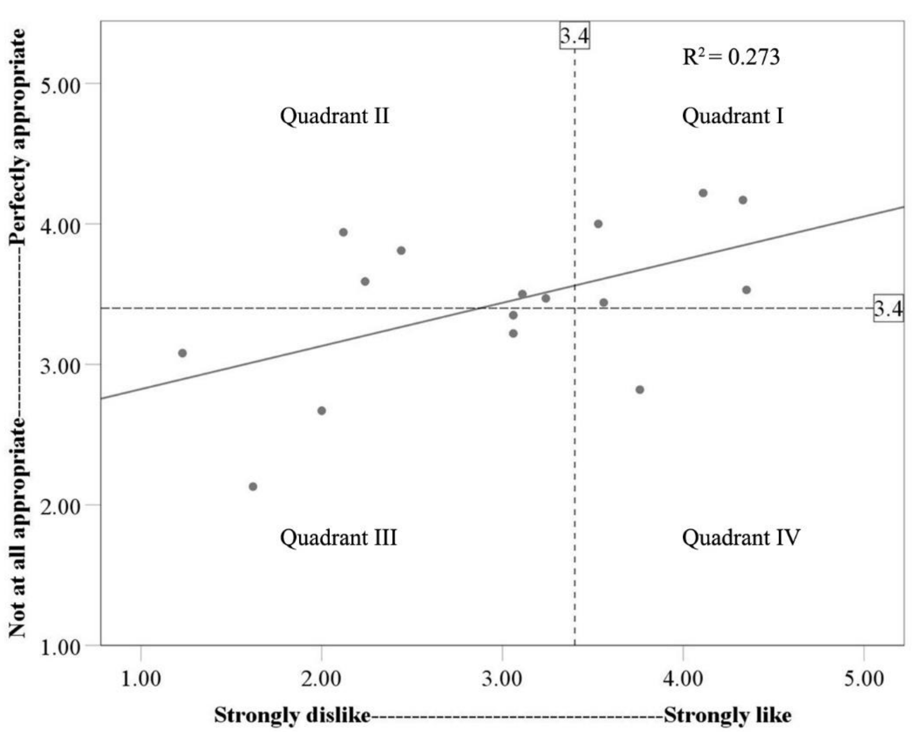

FIG. 3. Scatterplot presenting the extent to which a soundscape is preferred on the x-axis and how appropriate a soundscape is to a place on the y-axis. The data points represent the 16 soundscape excerpts.

A simple linear regression model was fitted to the appropriateness-preference scatterplot. Although the correlation analysis identified a significant positive correlation between the two variables, the determinate coefficient ($R^2 = 0.273$) indicated that the relationship was not close.

The two-dimensional space defined by appropriateness and preference was divided into four quadrants with a criterion of 6/4, as shown in Fig. 3. The first quadrant contained five soundscape samples. The preference and appropriateness values of this kind of soundscape are both relatively high, which means that this kind of soundscape is both preferred and appropriate. The second quadrant contained five soundscape samples. The preference value of this kind of soundscape is relatively low, but the appropriateness value is relatively high, denoting that this kind of soundscape is not preferred but appropriate. The third quadrant contained five soundscape samples. The preference and appropriateness values of this kind of soundscape are both relatively low, signifying that this kind of soundscape is neither preferred nor appropriate. The relationship between appropriateness and preference is illustrated in Fig. 4.

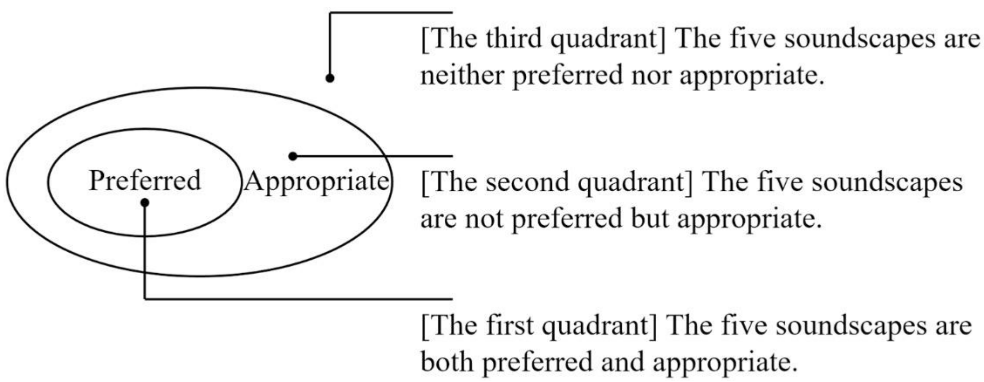

FIG. 4. The preferred soundscape excerpts are a subset of the appropriate soundscape excerpts.

Except for the soundscape sample in the fourth quadrant, all other samples conform to the rule that preferred soundscape excerpts are a subset of appropriate soundscape excerpts. Axelsson[25] drew a scatterplot of 25 soundscape samples, presenting the extent to which a soundscape is good on the x-axis and how appropriate a soundscape is to a place on the y-axis. We still found that the good soundscape excerpts are a subset of the appropriate soundscape excerpts. The soundscape sample in the fourth quadrant is soundscape excerpt No. B2 in Tables I and II. This soundscape excerpt is clearly an outlier, belonging to a square with considerable green spaces dominated by music sounds and human sounds. The sound sources may not be in view, so excerpt No. B2 is exceptionally inappropriate for the place, even though it is preferred.

It can be inferred that if a soundscape is preferred, it will be unconditionally considered to be appropriate. However, an "appropriate" soundscape can be either preferred or not. Preference is a more stringent and rigorous soundscape descriptor than appropriateness.

The prediction accuracies of felt emotion and perceived emotion on preference for the soundscape were 82.5% and 77.1%, respectively. The prediction accuracies of felt emotion and perceived emotion on the appropriateness of the soundscape to the place were 13.8% and 21.5%, respectively. Hence, preference is a more emotional soundscape descriptor than appropriateness. It may be inferred that for evaluating soundscapes, the more emotional the descriptor, the greater its stringency and rigorousness.

IV. CONCLUSIONS

The emotional evaluation mechanism of the soundscape in urban open spaces has been investigated through a laboratory experiment using a questionnaire (in Chinese). From this, the measurement dimensions in Chinese for the soundscape have been proposed. The main conclusions are as follows:

1. Although the soundscape in urban open spaces is rather complicated, it is still possible to identify five felt emotions and seven perceived emotions associated with the soundscape. According to the percentage of total variance covered by each factor, among the five felt emotions (enjoyment, excitement, desolation, tension, and familiarity), enjoyment (43%) and excitement (9%) are the two dominant factors; among the seven perceived emotions (comfortable, festive, desolate, familiar, various, attention-grabbing, and nostalgic), comfortable (38%) and festive (11%) are the two dominant factors. There is a certain correspondence between felt and perceived emotions. In the Chinese context, the core dimensions of soundscape emotion agree with the two dimensions (valence and arousal) proposed by Russell in 1980.[50] However, in the comparison between Chinese and

Western contexts, the percentages of total variance in soundscape emotion covered by these two dimensions are different.

2. When comparing perceived and felt emotions, the holistic soundscape descriptor “preference” is more suitable for predicting through felt emotion, while the holistic soundscape descriptor “appropriateness” is more suitable for predicting through perceived emotion. When explaining the appropriateness of soundscapes, perceived emotion is a slightly stronger factor than felt emotion, but neither is effective.
3. Preference is a more stringent and rigorous soundscape descriptor than appropriateness. Meanwhile, preference is a more emotional soundscape descriptor than appropriateness. It may be inferred that for evaluating soundscapes, the more emotional the descriptor, the greater its stringency and rigorousness.

This study examines soundscape emotions from the perspective of the Chinese context. The emotional-descriptive terms in Chinese are very different from those utilized in Western languages. This study also outlines that the percentages of total variance in soundscape emotion covered by emotional dimensions vary in different cultures. This underscores the necessity for further exploration of the differences between the Chinese and Western contexts. Future research should continue to form China’s overall framework. While the authors have endeavored to summarize the most accurate labels, using only labels to describe dimensions is difficult to convey the full picture.

This study makes a conjecture that “for evaluating soundscapes, the more emotional the descriptor, the greater its stringency and rigorousness.” This was based on the relationship between the two holistic soundscape descriptors, which can be further confirmed by more descriptors in future studies. This notion demonstrates the pivotal role of users’ perceived and felt emotions in the soundscape evaluation process.

SUPPLEMENTARY MATERIAL

See the supplementary material for the pattern matrices of factor analysis on perceived and felt emotions in the soundscape and the laboratory research questionnaire.

ACKNOWLEDGMENTS

The authors would like to express sincere gratitude to every participant in the experiment. This work was supported by the Young Talent Fund of Association for Science and Technology in Shaanxi, China [Grant No. 20220447]; the European Research Council (ERC) Advanced Grant (Grant No. 740696) on "Soundscape Indices" (SSID); the Foundation of the Department of Education of Guangdong Province (HITSZUQP21002); and the National Natural Science Foundation of China (Grant No. 52078161).

AUTHOR DECLARATIONS

Conflict of Interest

There are no conflicts of interest to declare.

Ethics Approval

All study participants provided informed consent, and the study design was approved by the School of Architecture, Harbin Institute of Technology (HIT-2023027).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]ISO, Acoustics — Soundscape — Part 1: *Definition and conceptual framework. In ISO 12913-1* (BSI Standards Publication, Geneva, 2014).

[2]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, and L. Lavia, "Ten questions on the soundscapes of the built environment," Build. Environ. **108**, 284–294 (2016).

[3]J. Kang, "声景：现状及前景" ("Soundscape: Current Progress and Future Development"), 新建筑 (New Architecture) **4**, 4–7 (2014).

[4]W.J. Davies, M.D. Adams, N.S. Bruce, R. Cain, A. Carlyle, P. Cusack, D.A. Hall, K.I. Hume, A. Irwin, P. Jennings, M. Marselle, C.J. Plack, and J. Poxon, "Perception of soundscapes: An interdisciplinary approach," Appl. Acoust. **74**, 224–231 (2013).

[5]M.L. de Ulhôa Carvalho, W.J. Davies and B. Fazenda, "Investigation of Emotional States in Different Urban Soundscapes through Laboratory Reproductions of 3D Audiovisual Samples," in *14th International Postgraduate Research Conference in the Built Environment (IPGRC): Contemporary and Future Directions in the Built Environment*, Salford, UK (December 16–17, 2019), pp. 326–338.

[6]T. Yu, H. Behm, R. Bill, and J. Kang, "Audio-visual perception of new wind parks," Landscape Urban Plan. **165**, 1–10 (2017).

[7]ISO, Acoustics — Soundscape — Part 2: *Data collection and reporting requirements In ISO/TS 12913-2* (International Organization for Standardization, Geneva, 2018).

[8]F. Aletta, J. Kang, and Ö. Axelsson, "Soundscape descriptors and a conceptual framework for developing predictive soundscape models," Landscape Urban Plan. **149**, 65–74 (2016).

[9]J. Kang and M. Zhang, "Semantic differential analysis of the soundscape in urban open public spaces," Build. Environ. **45**, 150–157 (2010).

[10]D. Dubois, C. Guastavino, and M. Raimbault, "A cognitive approach to urban soundscapes: Using

verbal data to access everyday life auditory categories," Acta Acust. United Acust. **92**, 865–874 (2006).

[11]M. Zhang and J. Kang, "Towards the evaluation, description, and creation of soundscapes in urban open spaces," Environment and Planning B Planning and Design. **34**, 68–86 (2007).

[12]R.M. Schafer, *The tuning of the world*, 1st ed. (Knopf, New York, 1977), pp. 13-18.

[13]T. Eerola and J.K. Vuoskoski, "A Review of Music and Emotion Studies: Approaches, Emotion Models, and Stimuli," Music Percept. **30**, 307–340 (2013).

[14]Y. Yu, Y. Jiang, P. Fang, Q. He, and K. Zhang, "音乐诱发情绪测量及其影响因素" ("Measurement for Music-induced Emotions and Its Interacting Factors"), 心理与行为研究 (Studies of Psychology and Behavior) **12**, 695–700 (2014).

[15]A. Gabrielsson, "Emotion perceived and emotion felt: same or different," Music. Sci. **1**, 123–147 (2002).

[16]K.R. Scherer, "What are emotions? And how can they be measured?" Soc. Sci. Inf. **44**, 695–729 (2005).

[17]P.N. Juslin and P. Laukka, "Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening," J. New Music Res. **33**, 217–238 (2004).

[18]P. Evans and E. Schubert, "Relationships between expressed and felt emotions in music," Music. Sci. **12**, 75–100 (2008).

[19]K. Kallinen and N. Ravaja, "Emotion perceived and emotion felt: Same and different," Music. Sci. **2**, 191–214 (2006).

[20]K. Tabei, "Inferior Frontal Gyrus Activation Underlies the Perception of Emotions, While

Precuneus Activation Underlies the Feeling of Emotions during Music Listening," Behav. Neurol. **2015**, 1–6 (2015).

[21]Ö. Axelsson, M.E. Nilsson, and B. Berglund, "A principal components model of soundscape perception," J. Acoust. Soc. Am. **128**, 2836–2846 (2010).

[22]R. Cain, P. Jennings, and J. Poxon, "The development and application of the emotional dimensions of a soundscape," Appl. Acoust. **74**, 232–239 (2013).

[23]G.R. Watts and R.J. Pheasant, "Tranquillity in the Scottish Highlands and Dartmoor National Park – The importance of soundscapes and emotional factors," Appl. Acoust. **89**, 297–305 (2015).

[24]A. Fiebig, P. Jordan, and C.C. Moshona, "Assessments of Acoustic Environments by Emotions - The Application of Emotion Theory in Soundscape," Front. Psychol. **11**, 573041 (2020).

[25]Ö. Axelsson, "How to measure soundscape quality," in *Euronoise 2015 conference*, Maastricht, Netherlands (May 31 – June 3, 2015), pp. 1477–1481.

[26]A. L. Brown, J. Kang, and T. Gjestland, "Towards standardization in soundscape preference assessment," Appl. Acoust. **72**, 387–392 (2011).

[27]A. L. Brown, "A review of progress in soundscapes and an approach to soundscape planning," Int J. Acoust Vib. **17**, 73–81 (2012).

[28]Complete Guide to 22 Unique Tourist Routes in Harbin available at https://wlt.hlj.gov.cn/wlt/c114254/202401/c00_31706577.shtml (Last viewed March 20, 2024).

[29]Enjoy Autumn Scenery in the Ice City! Here are 8 Premium Tour Routes available at https://wlt.hlj.gov.cn/wlt/c114254/202309/c00_31671150.shtml (Last viewed March 20, 2024).

[30]Z. Li and J. Kang, "Sensitivity analysis of changes in human physiological indicators observed in soundscapes," Landscape Urban Plan. **190**, 103593 (2019).

[31]S. Torresin, R. Albatici, F. Aletta, F. Babich, T. Oberman, S. Siboni, and J. Kang, "Indoor soundscape assessment: A principal components model of acoustic perception in residential buildings," Build. Environ. **182**, 107152 (2020).

[32]K. Ooi, Z. Ong, K. N. Watcharasupat, B. Lam, J. Y. Hong, and W. Gan, "ARAUS: A Large-Scale Dataset and Baseline Models of Affective Responses to Augmented Urban Soundscapes," IEEE T. Affect Comput. **15**, 105–120 (2024).

[33]G. von Bismarck, "Sharpness as an attribute of the timbre of steady sounds," Acta Acust. United Acust. **30**, 159-172 (1974).

[34]H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models.* (Heidelberg: Springer Nature, Berlin, 2007), pp. 239-241.

[35]D.A. Hall, A. Irwin, M. Edmondson-Jones, S. Phillips, and J.E.W. Poxon, "An exploratory evaluation of perceptual, psychoacoustic and acoustical properties of urban soundscapes," Appl. Acoust. **74**, 248–254 (2013).

[36]F. Pedrielli, E. Carletti, and C. Casazza, "Just noticeable differences of loudness and sharpness for earth moving machines," J. Acoust. Soc. Am. **123**, 3164 (2008).

[37]G. Fusaro, J. Kang, F. Asdrubali, and W. Chang, "Assessment of acoustic metawindow unit through psychoacoustic analysis and human perception," Appl. Acoust. **196**, 108885 (2022).

[38]B. Lam, K. C. Q. Lim, K. Ooi, Z. Ong, D. Shi, and W. Gan, "Anti-noise window: Subjective

perception of active noise reduction and effect of informational masking," Sustain. Cities Soc. **97**, 104763 (2023).

[39]P. Daniel and R. Weber, "Psychoacoustical Roughness: Implementation of an Optimized Model," Acta Acust. United Acust. **83**, 113–123 (1997).

[40]M. Zentner, D. Grandjean, and K.R. Scherer, "Emotions evoked by the sound of music: characterization, classification, and measurement," Emotion. **8**, 494–521 (2008).

[41]G.A. Gescheider, *Psychophysics: The Fundamentals*. (Lawrence Erlbaum Associates, Hillsdale, 1997), pp. 1–15.

[42]DIN 45692: *Measurement technique for the simulation of the auditory sensation of sharpness* (Deutsches Institut fur Normung E.V. (DIN), Berlin, 2009).

[43]ECMA 418-2: *Psychoacoustic metrics for ITT equipment — Part 2 (models based on human perception) (1st ed.)* (Ecma International, Geneva, 2020).

[44]E. Thulare, R. Ajoodha and A. Jadhav, "An Empirical Analysis and Application of the Expectation-Maximization and Matrix Completion Algorithms for Varying Degrees of Missing Data," in *2021 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa (SAUPEC/RobMech/PRASA)*, Potchefstroom, South Africa (January 27–29, 2021), pp. 1–7.

[45]H.F. Kaiser, "The application of electronic computers to factor analysis," Educ. Psychol. Meas. **20**, 141–151 (1960).

[46]R.B. Cattell, "The scree test for the number of factors," Multivariate Behav, Res. **1**, 245–276 (1966).

[47]B. Yu, J. Kang, and H. Ma, "Development of indicators for the soundscape in urban shopping streets," Acta Acust. United Acust. **102**, 462–473 (2016).

[48]G. Zhu, J. Kang, H. Ma, and C. Wang, "Characterization of soundscape assessment in outdoor public spaces of urban high-rise residential communities," J. Acoust. Soc. Am. **154**, 3660–3671 (2023).

[49]Ö. Axelsson, M.E. Nilsson and B. Berglund, "A Swedish instrument for measuring soundscape quality," in *Euronoise 2009: Action on noise in Europe*, Edinburgh, UK (October 26–28, 2010), pp. 1477–1481.

[50]J.A. Russell, "A circumplex model of affect," J. Pers. Soc. Psychol. **39**, 1161–1178 (1980).

[51]K. Herranz-Pascual, I. Iraurgi, I. Aspuru, I. Garcia-Pérez, J.L. Eguiguren, and Á. Santander, "Development of the Acoustic Comfort Assessment Scale (ACAS-12): Psychometric properties, validity evidence and back-translation between Spanish and English," PLoS One. **18**, e281534 (2023).