Soundscape evaluation: Binaural or monaural?

Chunyang Xu, and Jian Kang

ARTICLES YOU MAY BE INTERESTED IN

Soundscape evaluation: Binaural or monaural?

Chunyang Xu and Jian Kang[a)]

Institute for Environmental Design and Engineering, The Bartlett, University College London,
London WC1H 0NN, United Kingdom

The aim of this study is to explore the performance of binaural and monaural recordings in soundscape evaluation. Twelve sites with different acoustic scenarios were chosen, where binaural and monaural recordings were simultaneously made. Nine soundscape indicators were assessed by residents through a laboratory-based auditory test. The results showed that the two recording methods present good agreement on most soundscape evaluation indicators including overall impression, acoustic comfort, pleasantness, annoyance, eventfulness, and loudness. The two recording methods were found to be correlated with different indicators in a similar way. For most sites, the two recording methods were significantly correlated excluding for directionality. For both recording methods, the A-weighted sound pressure level was found to have a weak impact on soundscape evaluation. Reverberation time significantly affects reverberance through binaural recordings. Overall, for most soundscape indicators, it is feasible to use both recording methods, although when "realism," "reverberance," and "directivity" are involved in evaluation, binaural recordings will render corresponding perception more consistently than the monaural. © 2019 Acoustical Society of America.
https://doi.org/10.1121/1.5102164

[SKT]

I. INTRODUCTION

Soundscape evaluation is an approach that is involved with the reaction of people to sound environments. Various perceptual indicators are used to describe urban soundscapes. Both binaural and monaural recordings are widely applied in the assessment of urban soundscapes. The laboratory-based auditory test through rendering recorded sounds is conventional to model human perception on sound environments.[1–3]

For noise control in urban spaces, monaural microphones are applied in sound level monitoring, and a set of objective single-value parameters can be obtained. Berglund and Nilsson,[4] in 2004, assessed adverse perceived indicators. They utilized binaural and monaural recordings for the listening test, but there was no specific conclusion between two recordings in soundscape evaluation drawn from their research. Berglund and Nilsson,[5] in 2006, used monaural recordings to calculate the acoustic parameters of soundscapes, and these results were correlated with perceived soundscape quality evaluation from the structured soundwalk. The sounds of birds and fountains were applied to improve the overall soundscape quality for freeways, minor, and major roads by Coensel et al.[6] in 2011. They designed and recorded a binaural birdsong sound by playing a monaural birdsong in a reverberation chamber, and mixed this artificial birdsong with soundscapes. Hao et al.,[7] in 2016, conducted an assessment of the masking effect of birdsong for traffic noise, and monaural recordings were made in their research to compare occurrence frequencies of birdsong and the distance to the road. It was stated that owing to the high frequency of birdsong, binaural recordings, including more spatial characteristics, would cause more uncontrolled

variables. For indoor sound environments, monaural sound sources and directivity-applied sound sources have been examined through recent research.[8,9]

There are more studies relevant to soundscape assessments conducted by binaural recordings. Some research focused on soundscape categorization. A principal components model was conducted to identify dimensions of soundscape perception by an auditory test through binaural recordings.[10] Rychtáriková and Vermeir,[11] in 2013, utilized binaural recordings to categorize soundscapes based on an automatic clustering algorithm. Binaural recordings were made by Jeon et al.[12] to assess cross-national urban soundscapes under different cultural backgrounds with the use of principal component analysis and cluster analysis. Other studies emphasized the perceptual qualities accompanied with typical urban sound environments. A head and torso simulator (HATS) was used to investigate the effect of natural sounds on traffic noise by binaural recordings.[6] Genuit and Fiebig,[13] in 2006, recorded soundscapes by the artificial head to explore the use of psychoacoustics in the evaluation of soundscape quality, and they stated that the binaural recordings could reproduce aurally accurate acoustic scenarios. Soundscapes were binaurally recorded by Cain et al.[14] in 2013 to study the emotional dimensions, e.g., calmness and vibrancy. The soundscapes of three urban parks in Rome were assessed through binaural recordings to investigate a place with the higher sound level, and this place still led to a "good" environment.[15] Jambrošić et al.,[16] in 2013, assessed urban soundscapes combining on-site surveys and these two recording methods, and binaural recordings were only used to calculate the sound level differences between two ears. Recently, a drafted international soundscape standard ISO/DIS 12913–2 (Ref. 17) proposed an approach in soundscape measurements based on binaural recordings. Spatial information could be

---

a)Electronic mail: j.kang@ucl.ac.uk

recorded through the mean of calibrated binaural measurement systems, and subjects in the guided interview would be rendered with the same acoustic stimuli.

There is still a debate between binaural and monaural recordings in soundscape evaluation, but the research discussed above showed no systematical conclusions explaining the performance of binaural and monaural recordings with various perceived indicators especially in outdoor sound environments. The subjective auditory test was conducted to compare binaural and monaural recordings in soundscape evaluation. Soundscape indicators, psychoacoustics, and acoustic parameters were assessed in binaural and monaural recordings to determine the proper recording approach according to the auditory test.

The aim of the study is to explore (1) the overall comparison between two recording methods on soundscape evaluation, (2) the relationship among different perceived indicators, (3) the relationship among various sites, (4) the effect of acoustic parameters, and (5) the effect of contextual parameters on soundscape evaluation given by the two recording methods.

II. METHODOLOGIES

A. Site selection

Twelve public sites of representative functions were chosen in Sheffield, United Kingdom, and the views, typical sound events, and functions in these sites are shown in Fig. 1. These sites crossed a wide geographical range in Sheffield (1000 m × 2500 m) from the railway station, university campus, city center, city hall, cathedral to local parks. There is a distinct variation of acoustic performances, space functions, soundscape composition and building installations in these 12 sites chosen in Sheffield. Water features exist in site 1 (Crookes Valley Park), site 2 (Weston Park), site 4 (City Hall), sites 5 and 10 (railway station), and sites 7 and 12 (Peace Gardens). Multiple water features, including the pools, fountains, and water curtains, are abundant in the center of Sheffield, which are mixed with urban spaces of relaxation, recreation, culture, offices, etc. Site 11 (Winter Garden) is an indoor space open all the year round to the public, and it is also considered as a public space visited by citizens with the functions of relaxation and culture. From the perspective of room acoustics, it should have the highest reverberation time (RT) owing to its closed space and glass façades. Thus, it was chosen as a particular sample among other urban sound environments.

B. Audio and visual recordings

At each location, the acoustic environment was recorded by a four-channel digital recorder (Roland R-44, Hamamatsu, Japan) connected with in-ear microphones (DPA 4060, Allerød, Denmark) and an omnidirectional

Site 6- Crookes Valley Road
Cars-passing, birdsong (2)
[Relaxation, traffic]

Site 4- City Hall
Fountain, construction noise, footstep, speech, cars-passing (5)
[Square, commercial, cultural, traffic]

Site 9- Cathedral
Cars-passing, tram-passing, tram ring, church bell, footstep, speech (6)
[Traffic, cultural, commercial]

Site 11- Winter Garden
Speech, footstep, high-heeled shoes step, metal impact sound (4)
[Relaxation, cultural, commercial]

Site 1- Crookes Valley Park
Birdsong, duck, jogger-step (3)
[Relaxation, recreation]

Site 2- Weston Park
Birdsong, duck, footstep (3)
[Relaxation]

Site 10- Railway Station
Water curtain (stream), fountain, speech, footstep, cars-passing (5)
[Transport]

Site 5- Railway Station
Fountain, footstep, car horn (3)
[Transport]

Site 8- Devonshire Green
Cars-passing, birdsong, construction noise (3)
[Relaxation, traffic, recreation]

Site 12- Peace Gardens
Fountain, footstep, speech, cars-passing (4)
[Square, commercial, traffic, office]

Site 7- Peace Gardens
Fountain, bell, speech, footstep, cars-passing (5)
[Square, commercial, traffic, office]

Site 3- Hallam Campus
Cars-passing, footstep, speech, pedestrian crossing beep tone (4)
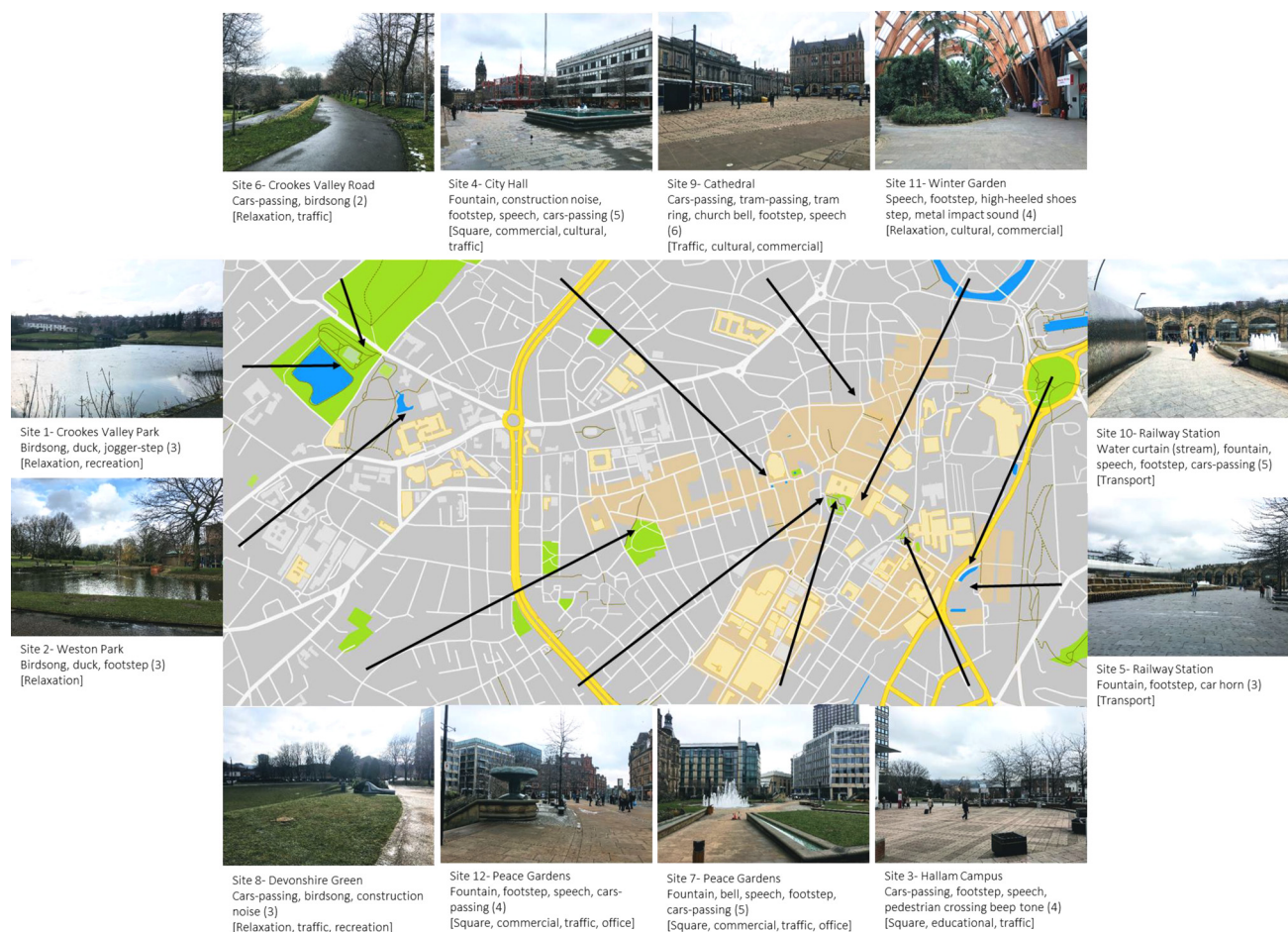[Square, educational, traffic]

FIG. 1. (Color online) Twelve sites chosen in Sheffield with views and the description of sound events and representative functions. (The number of sound events in these spaces is counted in parentheses, and the functions of these public spaces are illustrated in brackets.)

microphone (BSWA MP201, Beijing, China) simultaneously. Recorded sound events were the same for both recordings, and the duration of recordings was 3 min, containing typical sound events in these public spaces. The sampling rate was 48 kHz, and the depth of the recordings was 24 bits. A tripod fixed with the omnidirectional microphone was used to ensure the same height of the operator. Meanwhile, videos at these 12 locations were recorded by a camera (Sony Handycam DCR-DVD115E, Tokyo, Japan) as visual stimuli with the same head orientation of the operator. Residents in Sheffield should be familiar with the recorded views when they saw these videos.

C. Acoustic parameter measurements

Acoustic parameters were also recorded in these sites by a sound level meter (01 dB Solo, Limonest, France), and to describe temporal variability and low frequency contents, L_{A10}–L_{A90} and L_{Ceq}–L_{Aeq} were calculated. The results of acoustic parameters are illustrated in Table I. The numerical order of sites illustrated in this research is determined by L_{Aeq}. The overall range of A-weighted equivalent sound pressure levels for these sites is from 49.9 to 70.8 dB, and L_{A10}–L_{A90} and L_{Ceq}–L_{Aeq} show a large variation presenting a wide range of sound environments chosen. The lowest L_{Aeq} is in site 1 (Crookes Valle Park) near the pool, and the highest is in site 12 (Peace Gardens) near Pinstone Street. The time duration of measurements was kept the same with soundscape recordings. All recordings and measurements were made during weekdays from 10:00 a.m. to 2:00 p.m. to ensure that these public spaces were recorded under their commercial, residential, cultural, relaxation, or transport functions. Moreover, the sound level difference of binaural recordings and loudness were calculated from recorded sounds by ArtemiS SUITE (HEAD acoustics, Herzogenrath, Germany) shown in Table I.

D. RT of sound environments

To investigate the perceived reverberance, the RT of these spaces was calculated by the empirical formula proposed by Kang,[18,19]

$$RT = \frac{0.16V}{-S_0 \ln(1 - \bar{\alpha}) + 4MV}\left(88.6 + 49\alpha_b + 2.7\frac{\sqrt{LW}}{H}\right),$$

(1)

where V is the volume of the space, S_0 is the total surface area (m^2), $\bar{\alpha}$ is the mean absorption coefficient, M is the sound attenuation constant in air, α_b is the average absorption coefficient of boundaries, and L, W, and H are the length, width, and height, respectively, of urban areas. Absorbers mainly include the sky, trees, vegetation, bricks, etc. The RT for 12 sites was estimated as shown in Table II with the description of major absorbers.

E. Evaluation indicators

Perceptual auditory indicators are key components describing how people perceive, experience, or understand a soundscape.[20–22] Previous studies[23–25] selected overall impression and acoustic comfort as major indicators to evaluate soundscapes. Therefore, these two indicators chosen in this study will examine binaural and monaural recording performances in overall soundscape evaluation. Pleasantness, annoyance, eventfulness, loudness, reverberance, and directionality were also addressed by numerous research to assess soundscapes.[10,12,17,24–27] Owing to the fact that laboratory-based studies cannot render all stimuli compared with on-site studies, realism affected by acoustic and non-acoustic factors is also addressed to explore the realism perception difference between binaural and monaural recordings in soundscape evaluation. These non-acoustic factors in sonic environments are also reported by plentiful research.[25,28–30] For convenience, nine soundscape indicators are categorized into four groups: O, overall evaluation indicators (overall impression and acoustic comfort); G, generally perceived indicators (pleasantness, annoyance, eventfulness, and loudness); R, reproduced perceived indicator (realism); and T, technically perceived indicators (reverberance and directionality).

F. Auditory experiment and procedure

The active noise-canceling headphone (Bose QuietComfort 35, Framingham, MA) was used with a

TABLE I. Acoustic parameters (dB), binaural sound level difference Δ (dB), and loudness (sone) in 12 sites.

Site	L_{Aeq}	L_{A10}	L_{A50}	L_{A90}	L_{A10}–L_{A90}	L_{Ceq}–L_{Aeq}	Δ	Loudness
1	49.9	51.0	47.6	46.1	4.9	20.8	2.1	14.0
2	51.6	52.8	51.4	49.9	2.9	15.6	2.6	12.3
3	58.1	60.7	56.9	53.9	6.8	6.3	2.9	19.6
4	59.1	60.7	57.9	55.4	5.3	19.2	3.6	19.9
5	60.3	62.1	59.8	57.3	4.8	14.1	3.0	34.4
6	60.7	65.1	58.2	50.5	14.6	11.9	5.6	24.9
7	62.9	64.6	61.7	60.0	4.6	19.1	3.1	41.7
8	63.2	65.4	61.8	58.0	7.4	15.3	3.1	13.2
9	67.1	68.7	66.6	65.0	3.7	15.3	3.3	19.3
10	67.8	68.0	67.4	67.0	1.0	12.0	1.4	33.3
11	68.5	71.2	67.2	64.8	6.4	12.1	3.2	14.4
12	70.8	71.6	71.0	68.5	3.1	2.7	1.8	31.3

TABLE II. Major absorbers and RTs for 12 sites.

Site	S_0 (m^2)	Major absorbers	RT (s)
1	11 310	Sky, trees, and vegetation	1.98
2	10 850	Sky, trees, vegetation, and bricks	1.64
3	10 460	Sky, trees, ground, and bricks	2.70
4	6500	Sky, ground, and bricks	2.74
5	8200[a]	Sky, ground, and bricks	3.26
6	11 310	Sky, trees, vegetation, and ground	1.92
7	9630	Sky, ground, and bricks	3.37
8	11 310	Sky, trees, vegetation, and ground	1.92
9	11 870	Sky, ground, and bricks	3.38
10	8200	Sky, ground, and bricks	3.26
11	7200	Trees, vegetation, and ground	4.81
12	8900	Sky, ground, and bricks	3.43

[a]Same as site 10 due to the same semi-open space.

headphone amplifier (Lake People PHONE-AMP G109, Konstanz, Germany) connected to a laptop via an audio interface (Roland UA-101, Hamamatsu, Japan). The background noise in the acoustic laboratory was 34.0 dB(A), and the sound level was below 20.0 dB(A) for subjects when the active noise-canceling was switched on. The recorded videos were shared on a monitor by streaming with the laptop. Thus, the participants could evaluate sound environments according to visual and audio stimuli. One of the most obvious differences between binaural and monaural recording methods for laboratory-based soundscape evaluation is that the process of playback is two channel signals for binaural recordings and one channel for monaural recordings. All sound recordings were calibrated through an artificial head (Neumann KU100, Berlin, Germany) before the auditory experiment.

Five-point unipolar continuous category scales were used in evaluation questionnaires suggested by ISO/DIS 12913–2,[17] and the verbal labeling was provided below each scale as shown in the Appendix. Twenty-five subjects aged from 18 to 30 yr, living in Sheffield, gave their subjective evaluation to these sound environments, and they were familiar with these places. The total number of participants in the auditory test was in a considerable and conventional range according to previous soundscape research.[7,25,31] The hearing of subjects was tested before they gave their evaluation, and all subjects had normal hearing with the normal threshold for 125, 1000, and 4000 Hz. In addition, they all received simple acoustic training before the formal evaluation, and they had a basic understanding of acoustic indicators used in the evaluation. The inter-rater reliability among the subjects is 0.896 (Cronbach's α). Meanwhile, the Spearman-Brown coefficient is 0.881 for the split half method. The reliability analyses reflect the high consistency of the subjects' evaluation results. The participants heard 12 pairs of sounds in total. They could directly compare two sounds by different recording methods, one after the other, but the playback sequence of two sounds in each pair was randomized. The auditory procedure was approved by the ethics committee from the university. The consent forms and appraisal forms were obtained from participants.

G. Statistical analyses

In order to assess the correlation between subjective evaluation and acoustic parameters in this study, SPSS Statistics 24 and OriginPro 2017 were utilized to analyze Spearman's rho correlation coefficients, the independent t-test, and linear regression.

III. RESULTS

The main results of soundscape subjective evaluation are presented in five parts: (1) the mean subjective evaluation comparison between binaural and monaural recordings, (2) the comparison among evaluation indicators, (3) the comparison among different sites, (4) the effect of acoustic parameters, and (5) the effect of contextual parameters, including sound events, binaural sound level difference, and RT, on soundscape evaluation by the two recording methods.

A. Overall comparison between two recording methods

Figure 2 shows the correlation coefficients of mean ratings between binaural and monaural recordings and the mean subjective ratings of 9 indicators over 12 sites in Sheffield. Twelve sites in spatial scales are considered as unordered categorical variables, so Spearman's rho is utilized to analyze this rank correlation. For the evaluation of overall impression, acoustic comfort, pleasantness, annoyance, eventfulness, and loudness, binaural and monaural recordings are statistically significant ($p < 0.01$) with the correlation coefficients over 0.5. There are also significant correlations for realism, reverberance, and directionality between these two recording methods with lower coefficient values of 0.362 ($p < 0.01$), 0.496 ($p < 0.01$), and 0.243 ($p < 0.01$), respectively.

According to the results of the independent t-test, the mean binaural subjective ratings of overall impression, acoustic comfort, pleasantness, annoyance, and eventfulness are approximate to the monaural. The t-test results also reveal that the mean rating differences are statistically significant ($p < 0.01$) for loudness, realism, reverberance, and directionality. For these four indicators, the mean binaural ratings are 9%, 19%, 22%, and 39% higher than the monaural. Unsurprisingly, the overall realism and directionality subjective ratings of the binaural recordings are higher than the monaural. The audio information delivered by the monaural recordings is only through one channel less than the binaural.

The standard deviation of binaural and monaural recordings for overall impression, acoustic comfort, pleasantness, annoyance, eventfulness, and loudness is approximate shown as error bars. It reveals that the overall variations of two recording methods to evaluate these perceived indicators are similar, although these subjective fluctuations in some sites present a slight difference.

B. Comparison among evaluation indicators

A comparison between different evaluation indicators was performed by binaural and monaural ratings shown in Table III. Correlations between overall impression, acoustic comfort, pleasantness, annoyance, and loudness are statistically significant ($p < 0.01$) for the two recording methods, and these positive or negative correlation coefficients are approximate with the diagonal contrast. There are also slight differences between two recording methods in the interactions between these five indicators. Notably, the comparison between binaural and monaural recordings for acoustic comfort-annoyance ($r_s = -0.528$ and -0.439) and acoustic comfort-loudness ($r_s = -0.425$ and -0.323) implies that binaural recordings show a more negative tendency than the monaural for these two indicators' interactions. In addition, the correlation coefficients for eventfulness-loudness and reverberance-directionality under the monaural recordings are higher than in the binaural. The significant correlations between directionality and the other indicators in the binaural recordings are less than the monaural.

J. Acoust. Soc. Am. **145** (5), May 2019
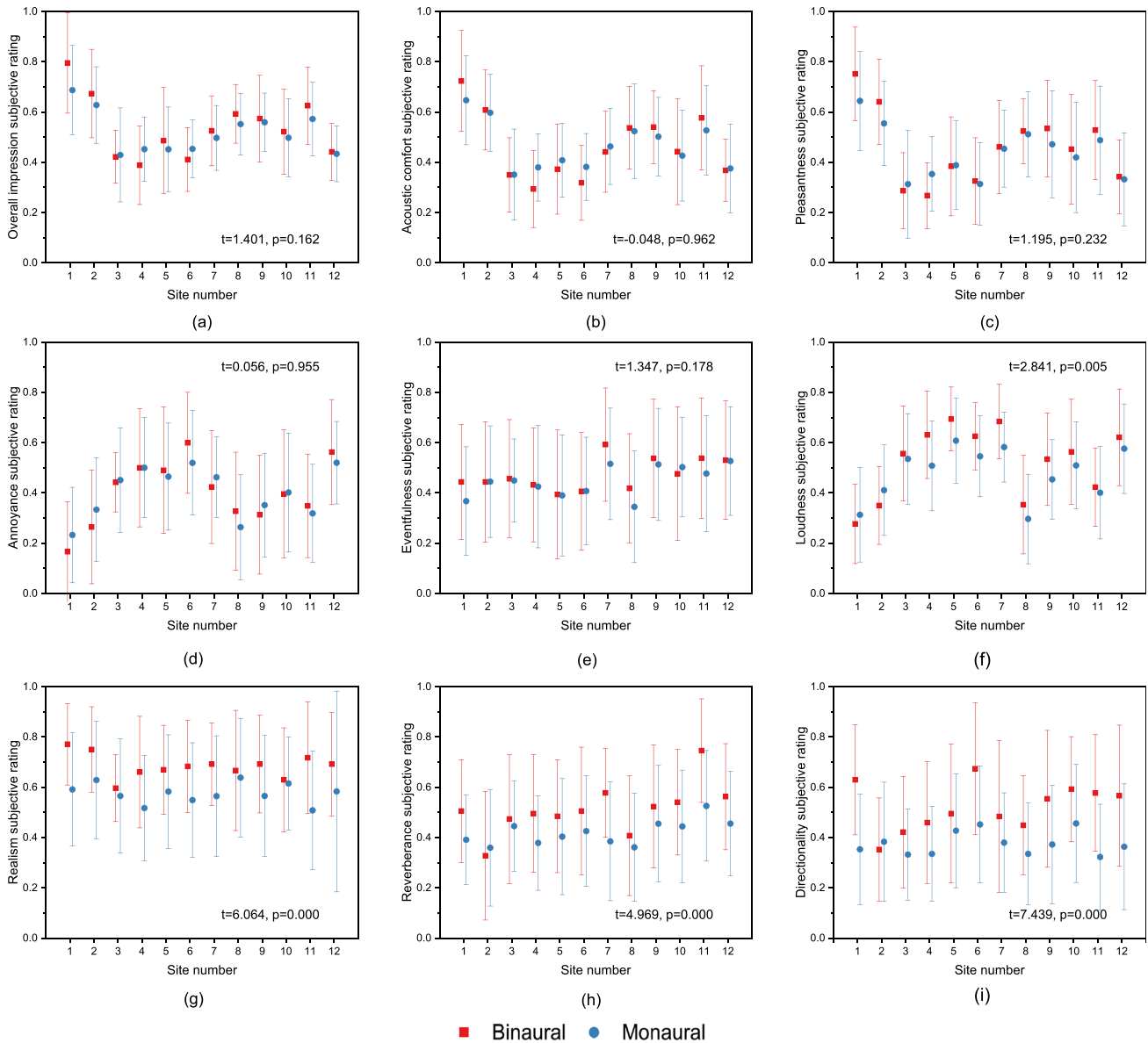
Chunyang Xu and Jian Kang   3211

FIG. 2. (Color online) Comparison between binaural and monaural recordings in 12 sites with 9 perceived indicators. t and p values of the independent t-test and Spearman's rho coefficients are listed (**$p < 0.01$). (a) Overall impression, $r_s = 0.614$**; (b) acoustic comfort, $r_s = 0.569$**; (c) pleasantness, $r_s = 0.670$**; (d) annoyance, $r_s = 0.606$**; (e) eventfulness, $r_s = 0.739$**; (f) loudness, $r_s = 0.712$**; (g) realism, $r_s = 0.362$**; (h) reverberance, $r_s = 0.496$**; (i) directionality, $r_s = 0.243$**.

C. Comparison among different sites

Ratings of each subject were taken into consideration, and Spearman's rho correlation coefficients and the independent t-test for each site are shown in Table IV. Not all sites show correlations with statistical significance between two recording methods for the evaluation of overall impression, acoustic comfort, pleasantness, annoyance, and loudness in spite of their correlation coefficients in the overall comparison over 0.5 ($p < 0.01$). All sites show correlations ($p < 0.01$) for eventfulness, while the correlation coefficients ranged from 0.616 to 0.888. There are six sites whose mean realism differences are significant ($p < 0.05$). These sites with typical acoustic scenarios can be perceived as entirely different through the rendering of two recording methods resulting in mean rating differences for realism.

Eight sites show positive correlations ($p < 0.01$ or $p < 0.05$) between two recording methods on reverberance. The mean binaural reverberance subjective ratings are 22% higher than the monaural, and the highest subjective rating of reverberance occurs in site 11 (Winter Garden) shown in Fig. 2(h). Site 11 was installed with closed glass façades resulting in the longest RT, and indeed, this setup in 12 sites made subjects perceive the reverberance difference from binaural and monaural recordings. The mean reverberance difference on site 11 is also significant according to the result of the independent t-test ($p = 0.001$). There is only one site with statistical significance in the directionality correlation, and seven sites present mean directionality differences with statistical significance ($p < 0.05$). The mean rating difference of directionality is noteworthy, and binaural recordings still dominate directionality in the soundscape evaluation.

TABLE III. Comparison of binaural and monaural ratings between indicators. (The lower left of the matrix is the correlation coefficients between different indicators in binaural recordings, and the upper right is the correlation coefficients between different indicators in monaural recordings.) $*p < 0.05$, $**p < 0.01$ (two-tailed test of statistical significance).

Monaural	Overall impression	Acoustic comfort	Pleasantness	Annoyance	Eventfulness	Loudness	Realism	Reverberance	Directionality
Binaural									
Overall impression		0.724**	0.688**	−0.538**	0.042	−0.323**	0.203**	0.203**	0.119*
Acoustic comfort	0.775**		0.791**	−0.439**	0.063	−0.323**	0.236**	0.194**	0.181**
Pleasantness	0.763**	0.790**		−0.432**	0.112	−0.254**	0.284**	0.248**	0.205**
Annoyance	−0.593**	−0.528**	−0.565**		0.069	0.520**	−0.006	−0.077	0.002
Eventfulness	0.004	0.006	0.040	0.124*		0.486**	−0.240**	0.527**	0.086
Loudness	−0.422**	−0.425**	−0.358**	0.544**	0.148*		0.276**	0.147*	0.260**
Realism	0.170**	0.234**	0.304**	−0.121*	0.247**	0.142*		0.415**	0.414**
Reverberance	0.118*	0.164**	0.128*	0.008	0.310**	0.166**	0.194**		0.505**
Directionality	0.039	0.105	0.120*	0.004	0.141*	0.138*	0.392**	0.321**	

D. Effect of acoustic parameters on subjective evaluation

Table V shows the correlation coefficients of binaural and monaural subjective ratings to acoustic parameters, including nine indicators, sound levels measured by the sound level meter, and loudness calculated by ArtemiS SUITE according to DIN 45631/A1 (HEAD acoustics, Herzogenrath, Germany). The sound pressure level correlated with both recording methods is based on the measurement by a sound level meter.

L_{Aeq} is an essential acoustic parameter to describe a sound environment, and relatively low correlations were found on the relationship between L_{Aeq} and perceived indicators for both recording methods. The agreement tendencies of overall impression, annoyance, loudness, reverberance, and directionality are higher for binaural recordings. The increasing temporal variability of L_{A10}–L_{A90} causes the decline of overall impression and pleasantness with $r_s = -0.114$ ($p < 0.05$) and -0.169 ($p < 0.01$) for binaural recordings. L_{A10}–L_{A90} shows no agreement tendency with reproduced indicator and technically perceived indicators. L_{Ceq}–L_{Aeq} renders a series of significant correlations for two recording methods. The reduced low frequency content also resulted in a decrease in perceived annoyance and loudness. It is interesting to note that the highest correlation

TABLE IV. Spearman's rho correlation coefficients (r_s) and independent t-test (t and p) between two recording methods for each site. $*p < 0.05$, $**p < 0.01$ (two-tailed test of statistical significance).

Site number		1	2	3	4	5	6	7	8	9	10	11	12
Overall impression	r_s	0.585**	0.650**	0.544**	0.288	0.701**	0.516**	0.359	0.545**	0.525**	0.580**	0.571**	0.348
	t	0.227	1.330	−0.923	−1.621	0.355	−0.228	0.857	0.871	1.425	0.542	0.537	2.581
	p	0.825	0.199	0.370	0.119	0.729	0.824	0.405	0.398	0.169	0.598	0.601	0.015
Acoustic comfort	r_s	0.751**	0.494*	0.597**	0.282	0.400*	0.358	0.284	0.316	0.503*	0.398*	0.377	0.384
	t	1.632	0.202	−0.024	−2.117	−0.837	−1.422	−0.458	0.315	0.698	0.288	1.134	−0.180
	p	0.109	0.841	0.981	0.039	0.407	0.161	0.649	0.754	0.489	0.775	0.262	0.859
Pleasantness	r_s	0.675**	0.598**	0.558**	0.092	0.581**	0.509**	0.289	0.429*	0.697**	0.755**	0.435*	0.723**
	t	2.047	1.580	−0.608	−2.167	−0.084	0.251	0.121	0.254	1.070	0.530	0.829	0.210
	p	0.047	0.121	0.546	0.035	0.934	0.803	0.904	0.801	0.290	0.599	0.411	0.835
Annoyance	r_s	0.774**	0.489*	0.413*	0.447*	0.642**	0.597**	0.575**	0.653**	0.610**	0.514**	0.543**	0.344
	t	−1.383	−1.257	−0.140	−0.007	0.402	1.301	−0.589	1.199	−0.676	−0.086	0.536	0.733
	p	0.175	0.215	0.889	0.995	0.689	0.200	0.559	0.236	0.502	0.932	0.595	0.467
Eventfulness	r_s	0.720**	0.684**	0.856**	0.721**	0.836**	0.643**	0.777**	0.888**	0.663**	0.663**	0.711**	0.616**
	t	1.332	−0.019	0.109	0.102	0.062	−0.012	1.087	1.148	0.366	−0.393	0.973	0.057
	p	0.189	0.985	0.913	0.919	0.951	0.990	0.282	0.257	0.716	0.696	0.335	0.955
Loudness	r_s	0.472*	0.868**	0.606**	0.609**	0.293	0.692**	0.615**	0.879**	0.585**	0.637**	0.443*	0.540**
	t	−0.698	−1.237	0.408	2.469	1.774	1.686	2.390	1.084	1.657	1.005	0.543	1.079
	p	0.488	0.222	0.685	0.017	0.082	0.098	0.021	0.284	0.104	0.320	0.590	0.286
Realism	r_s	0.468*	0.518**	0.341	0.367	0.623**	0.325	0.265	0.139	0.649**	0.335	−0.068	0.459*
	t	3.404	2.195	0.459	2.341	1.401	2.312	2.213	0.316	1.954	0.259	3.607	1.862
	p	0.002	0.033	0.648	0.023	0.168	0.025	0.032	0.753	0.057	0.796	0.001	0.069
Reverberance	r_s	0.368	0.649**	0.410*	0.106	0.775**	0.641**	0.451*	0.438*	0.386	0.548**	0.451*	0.349
	t	1.822	−0.577	0.437	1.956	1.197	1.159	2.955	0.775	1.128	1.565	3.733	1.589
	p	0.075	0.567	0.664	0.056	0.237	0.252	0.004	0.442	0.265	0.124	0.001	0.119
Directionality	r_s	0.244	0.196	0.334	0.151	0.226	−0.083	0.388	0.569**	0.239	0.340	0.167	0.242
	t	4.815	−0.496	1.583	2.026	0.944	3.451	1.450	1.516	2.903	2.172	3.529	2.893
	p	0.000	0.622	0.120	0.048	0.350	0.001	0.154	0.136	0.006	0.035	0.001	0.006

TABLE V. Spearman's rho correlation coefficients between acoustic parameters and subjective ratings of two recording methods. $*p < 0.05$, $**p < 0.01$ (two-tailed test of statistical significance).

Indicator		L_{Aeq}	$L_{A10}-L_{A90}$	$L_{Ceq}-L_{Aeq}$	L_{A10}	L_{A50}	L_{A90}	Loudness[a]
Overall impression	Binaural	−0.127*	−0.114*	0.325**	−0.134*	−0.131*	−0.151**	−0.365**
	Monaural	−0.104	−0.080	0.276**	−0.105	−0.115*	−0.129*	−0.319**
Acoustic comfort	Binaural	−0.069	−0.107	0.293**	−0.068	−0.080	−0.098	−0.418**
	Monaural	−0.028	−0.001	−0.031	−0.035	−0.030	−0.028	0.041
Pleasantness	Binaural	−0.105	−0.169**	0.337**	−0.110	−0.108	−0.126*	−0.378**
	Monaural	−0.090	−0.098	0.307**	−0.103	−0.089	−0.096	−0.282**
Annoyance	Binaural	0.166**	0.101	−0.297**	0.171**	0.160**	0.143*	0.337**
	Monaural	0.073	−0.009	−0.198**	0.070	0.076	0.094	0.335**
Eventfulness	Binaural	0.142*	−0.101	−0.004	0.130*	0.135*	0.159**	0.061
	Monaural	0.158**	−0.170**	−0.081	0.147*	0.160**	0.192**	0.119*
Loudness	Binaural	0.146*	−0.045	−0.210**	0.131*	0.173**	0.216**	0.568**
	Monaural	0.076	−0.117*	−0.207**	0.063	0.103	0.145*	0.442**
Realism	Binaural	−0.058	−0.027	0.134*	−0.054	−0.071	−0.081	−0.123*
	Monaural	−0.026	−0.091	−0.012	−0.028	−0.010	−0.008	−0.009
Reverberance	Binaural	0.236**	0.032	−0.089	0.226**	0.213**	0.202**	0.142*
	Monaural	0.133*	−0.006	−0.125*	0.142*	0.125*	0.119*	0.036
Directionality	Binaural	0.127*	0.045	−0.072	0.140*	0.112	0.067	0.091
	Monaural	0.001	−0.070	−0.036	−0.001	0.013	0.001	0.100

[a]Loudness of recorded signals calculated by ArtemiS according to DIN 45631/A1.

coefficients are 0.568 ($p < 0.01$) for the binaural subjective ratings and 0.442 ($p < 0.01$) for the monaural regarding the correlation between perceived loudness and calculated loudness. Moreover, the calculated loudness has higher correlation coefficients for the two recording methods compared with other conventional sound levels. This reveals the importance of psychoacoustic parameters, especially loudness in the soundscape evaluation.

As a result, statistically significant correlations were obtained between acoustic parameters and perceived indicators. Binaural recordings are more sensitively correlated with $L_{Ceq}-L_{Aeq}$ than the monaural on overall impression, acoustic comfort, pleasantness, annoyance, and loudness. The time variability of $L_{A10}-L_{A90}$ does not significantly affect the evaluation of perceived indicators. Only a few acoustic parameters show significant correlations with realism and directionality, and these objective acoustic parameters would not directly affect the perception and evaluation of realism and directionality.

E. Effect of contextual parameters on subjective evaluation

Eventfulness, realism, and reverberance are correlated with contextual parameters, respectively, including sound events, sound level difference, and RT.

Eventfulness is a particularly perceived indicator not closely correlated with overall impression, acoustic comfort, pleasantness, and annoyance. Thus, the relationship between the number of sound events and eventfulness subjective rating is analyzed in Fig. 3(a). The results of linear regression present that there is a positive correlation for both binaural and monaural recording methods with $R^2 = 0.444$ and 0.497, respectively. The overall subjective ratings for binaural and monaural recording methods are approximate, and the proximate correlation coefficients in linear regression between the number of sound events and perceived eventfulness are reasonable. The meaning of "eventful" used in the auditory test will cause vagueness for subjects. Although there is a
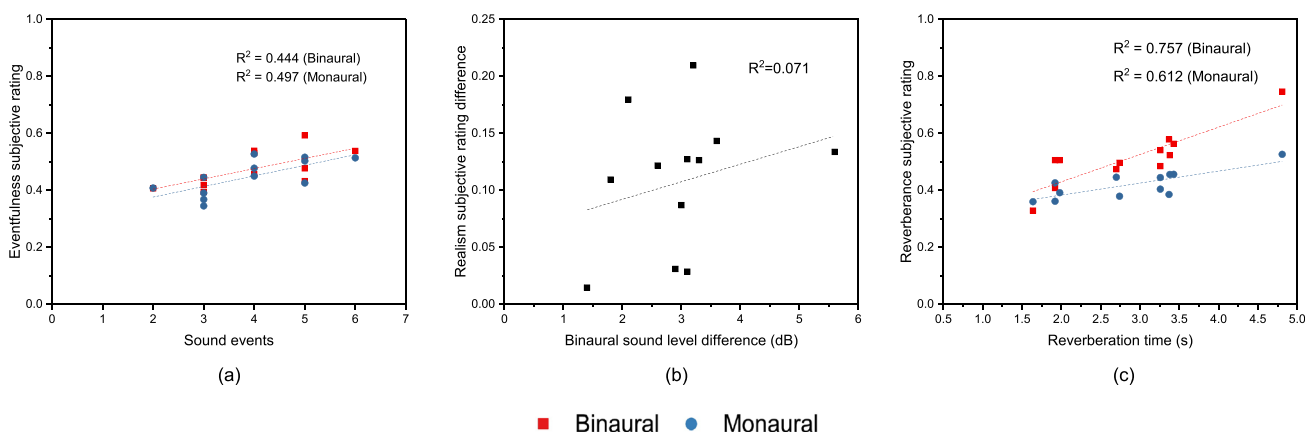

FIG. 3. (Color online) Relationships between (a) the number of sound events and eventfulness subjective rating, (b) the binaural sound level difference and realism subjective rating difference, and (c) RT and reverberance subjective rating.

Chunyang Xu and Jian Kang

positive correlation between the number of sound events and perceived eventfulness, the types of sound events and the sound level in environments are also involved with the evaluation of eventfulness.

Figure 3(b) presents the linear regression between the binaural sound level difference as shown in Table I and the realism subjective rating difference. Indeed, the binaural sound level difference is an impact factor affecting perceived realism, but realism is a reproduced indicator involved with multiple stimuli. Sound environments are time dependent, and the single-value binaural sound level difference cannot fulfill the realism gap between binaural and monaural recording methods. Eventfulness and realism are more complicated in the auditory test with the involvement of the sound composition, sound levels, personal understandings, and other non-acoustic factors.

Reverberance and directionality are categorized as technically perceived indicators in Sec. II E. The relationship between RT in these public spaces and perceived reverberance is shown in Fig. 3(c). The coefficients of determination are 0.757 and 0.612 for binaural and monaural recording methods, respectively. When the RT is longer than 2.5 s, the variation of the two recording methods occurs obviously. Reverberance in these public spaces is dominantly perceived through binaural recordings.

IV. DISCUSSION

A. Overall comparison among different analyses

A summary of binaural and monaural recording methods of different analyses is illustrated in Table VI. The comprehensive performance of the two recording methods on overall impression, acoustic comfort, pleasantness, annoyance, eventfulness, and loudness is similar. Eventfulness is an indicator depending on how subjects understand the meaning of eventful, and it presents a different tendency on the comparison among other usual indicators. It also revealed that eventfulness was previously classified into independent scales compared with pleasantness and annoyance.[12] Thus, these indicators were classified into the high fidelity for both recording methods according to the overall performance crossing different analyses. When perceived indicators in soundscape evaluation include realism, reverberance and directionality, the fidelity of binaural recordings comprehensively performs better than the monaural. Especially, according to the large mean rating difference, monaural recordings are categorized into the zero fidelity for directionality.

B. Directionality and directivity

Directionality in urban sound environments is still a complicated indicator associated with multiple objective and subjective factors. Most sound sources in public spaces are not stationary. The sound levels and positions of these sound sources will vary with time, and these movable sources, such as pedestrians or cars-passing within a close distance could also generate higher perceived directionality. The variations of sound source strength, numbers, frequency contents, direction, distance, and other conditions will all contribute to the evaluation of directionality. The subjects in this study could give their perception for a sound environment depending on their hearing localization abilities and familiarity to these spaces.

As expected, the subjects cannot perceive any directivity by listening to monaural recordings, but a certain number of participants did imagine it as shown in Fig. 2(i). Several factors jointly influence this phenomenon: (1) the subjects are familiar with these public spaces, and they imagine the directionality generated from sound sources; (2) visual stimuli rendered by videos imply the orientation of sound sources in these environments; and (3) the subjects perceived the strong directionality from certain sites during the experiment indeed, and they cannot distinguish the environments with low directivity. For instance, there was no strong directional sound source near the operator in site 2. The subjects would not distinguish these two sounds recorded by two methods under this scenario. The mean rating of directionality for site 2 is approximate, and the mean difference of subjective

TABLE VI. Summary of binaural and monaural recordings over different analyses. B, binaural; M, monaural. *$p < 0.05$, **$p < 0.01$ (two-tailed test of statistical significance).

Group[a]	Indicator	r_s[b]	Rating difference[c]		Site $(n/12)$[d]	Indicator comparison $(h/8)$[e]		Acoustic parameters $(i/7)$[f]		Fidelity[g]	
			t	p		B	M	B	M	B	M
O	Overall impression	0.614**	1.401	0.162	9/12**	5/8** 1/8*	6/8** 1/8*	3/7** 4/7*	2/7** 2/7*	++	++
	Acoustic comfort	0.569**	−0.048	0.962	2/12** 4/12*	6/8**	7/8**	2/7**	0/7	++	++
G	Pleasantness	0.670**	1.195	0.232	8/12** 2/12*	5/8** 2/8*	7/8**	3/7* 1/7*	2/7**	++	++
	Annoyance	0.606**	0.056	0.955	8/12** 3/12*	4/8** 1/8*	4/8**	5/7* 1/7*	2/7**	++	++
	Eventfulness	0.739**	1.347	0.178	12/12**	2/8** 3/8*	3/8**	1/7* 3/7*	4/7** 2/7*	++	++
	Loudness	0.712**	2.841	0.005	9/12** 2/12*	5/8** 3/8*	7/8** 1/8*	4/7** 2/7*	2/7** 2/7*	++	++
R	Realism	0.362**	6.064	0.000	3/12** 2/12*	6/8** 2/8*	7/8**	2/7*	0/7	++	+
T	Reverberance	0.496**	4.969	0.000	4/12** 4/12*	5/8** 2/8*	6/8** 1/8*	4/7** 1/7*	5/7*	++	+
	Directionality	0.243**	7.439	0.000	1/12**	2/8** 3/8*	5/8** 1/8*	2/7*	0/7	++	○

[a]Indicator classification in Sec. II E.
[b]Spearman's rho between two recording methods for 12 sites.
[c]t and p values of the independent t-test between two recording methods.
[d]n is the number of correlation coefficients with statistical significance between two recording methods for 12 sites.
[e]h is the number of correlation coefficients with statistical significance correlated with other eight indicators in the comparison between indicators.
[f]i is the number of correlation coefficients with statistical significance between subjective ratings and seven acoustic parameters.
[g]○, zero fidelity; +, low fidelity; ++, high fidelity.

ratings is also not significant ($p = 0.622$) as shown in Table IV. In addition, the adjectives of "directional" and "reverberant" are also jargon for subjects to some extent, and this may also result in the bias in evaluation results.

Furthermore, moving sound sources like birds in site 1 and cars in site 6 have a significant impact on directionality given by binaural recordings. For site 1, the ducks were close to the pool bank and these sound events occurred near the operator. Moving cars were running parallel to the direction of the operator in site 6. These moving sound sources dominate these two sound environments, and the sound localization of the subjects could detect these sound compositions through binaural hearings. Sound localization from moving sources in urban public spaces is also an essential factor on directionality, and the large subjective difference was found among these acoustic scenarios.

Interaural cross correlation (IACC) is one of the binaural acoustic parameters used to analyze spatial impression and characteristics of sounds, having been applied in various spaces, including concert halls,[32] high-speed train noise,[33] and urban soundscapes.[34] The correlation between the IACC of the early sound field within 80 ms and directionality was examined and no significant correlation was found (binaural: $r_s = -0.048$, $p = 0.407$; monaural: $r_s = -0.060$, $p = 0.302$). The IACC of the late sound field after 80 ms was also not significantly correlated with directionality (binaural: $r_s = 0.017$, $p = 0.768$; monaural: $r_s = -0.030$, $p = 0.599$).

C. Multi-factorial interaction on realism

There is a notable difference between two recording methods for realism generated from the multi-factorial interaction. Realism is influenced by the two recording methods in typical sites, and it is involved with different sound contexts, building installations, and public functions. Site 1 (Crookes Valley Park) and site 11 (Winter Garden) have the largest mean subjective differences in realism between two recording methods, and the t-test in Table IV also shows that the mean differences for these two sites are significant ($p < 0.01$). Site 1 has the lowest sound level and the highest overall impression, acoustic comfort, and pleasantness ratings. Natural sound events, e.g., ducks and other water birds near the pool bank, will attract subjects' attention and increase the overall positive ratings. Meanwhile, owing to these ducks and water birds being within a close distance, subjects will be easily able to distinguish the difference of realism between the two recording methods. They could perceive the environment as more real under the dominant natural sound events with the low background noise in spite of its strong directivity of the sound source by binaural recordings. For site 11 of the highest RT among other sites, binaural recordings will increase the sense of localization and spaciousness resulting from reflected sounds recorded by binaural microphones. The multi-factorial interaction among different indicators reflects the internal connection between objective environments and the subjective evaluation given by the recording methods.

In addition, other physical conditions, including lighting, vibration, temperature, etc., are involved in realism.

These conditions cannot be reproduced accurately in the laboratory, but they did have multiple and significant impacts on realism in soundscapes. Therefore, although binaural recordings performed better than monaural recordings shown in Fig. 2(g), the subjective ratings for realism did not reach the maximum scores.

V. CONCLUSIONS

This study examined the performance of two recording methods in soundscape evaluation. The subjective evaluation and comparative analyses of indicators and sites along with the effects of acoustic and contextual parameters revealed the following:

(1) Binaural and monaural recording methods showed good agreement in mean ratings of overall impression, acoustic comfort, pleasantness, annoyance, eventfulness, and loudness. In contrast with the monaural results, overall binaural subjective ratings were significantly higher in realism, reverberance, and directionality evaluations.
(2) The two recording methods were correlated with different perceived indicators in a similar way.
(3) Most sites showed no correlations in directionality between the two recording methods. It revealed that these two methods performed differently for the evaluation of directionality in most urban spaces.
(4) The A-weighted sound pressure level had a weak impact on soundscape evaluation for both recording methods.
(5) The correlation between eventfulness and the number of sound events was similar with the two recording methods. The difference in realism generated from two recording methods did not significantly depend on the binaural sound level difference. Reverberance was perceived as more consistent with RT through binaural recordings in soundscape evaluation.

Overall, this work suggests that monaural recordings are sufficient to evaluate most soundscape indicators including overall impression, acoustic comfort, pleasantness, annoyance, eventfulness, and loudness. When some special acoustic scenarios (e.g., moving birdsongs or cars passing near the subjects) and built environment (e.g., RT > 2.5 s) occur in soundscapes, the corresponding perception, i.e., directionality and reverberance, would be much better evaluated by binaural recordings.

ACKNOWLEDGMENTS

APPENDIX: EVALUATION QUESTIONNAIRE

Name:
Gender: Male/Female
Age:

Chunyang Xu and Jian Kang

Occupation:

Whether do you live in Sheffield: Yes/No

(1) To what extent is the overall impression of the present surrounding sound environment?

very good—bad—neither good nor bad—bad—very bad

(2) Overall, to what extent represents your feeling of acoustic comfort?

not at all—slightly—moderately—very—extremely

(3) How pleasant is it here?

(Same as the scale of question 2 for questions 3–9.)

(4) How annoying is it here?

(5) How eventful is it here?

(6) How loud is it here?

(7) How real is it here?

(8) How reverberant is it here?

(9) How directional is it here?

[1]C. Guastavino, B. F. Katz, J.-D. Polack, D. J. Levitin, and D. Dubois, "Ecological validity of soundscape reproduction," Acta Acust. Acust. **91**(2), 333–341 (2005).

[2]J. Y. Jeon, P. J. Lee, J. You, and J. Kang, "Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds," J. Acoust. Soc. Am. **127**(3), 1357–1366 (2010).

[3]L. Jiang and J. Kang, "Effect of traffic noise on perceived visual impact of motorway traffic," Landsc. Urban Plan. **150**, 50–59 (2016).

[4]B. Berglund and M. E. Nilsson, "Soundscape perceived in built environments," in *Proceedings of the International Congress of Acoustics*, Kyoto (2004).

[5]B. Berglund and M. E. Nilsson, "On a tool for measuring soundscape quality in urban residential areas," Acta Acust. Acust. **92**(6), 938–944 (2006).

[6]B. D. Coensel, S. Vanwetswinkel, and D. Botteldooren, "Effects of natural sounds on the perception of road traffic noise," J. Acoust. Soc. Am. **129**(4), EL148–EL153 (2011).

[7]Y. Hao, J. Kang, and H. Wörtche, "Assessment of the masking effects of birdsong on the road traffic noise environment," J. Acoust. Soc. Am. **140**(2), 978–987 (2016).

[8]J. Y. Jeon and H. I. Jo, "Three-dimensional virtual reality-based subjective evaluation of road traffic noise heard in urban high-rise residential buildings," Build. Environ. **148**, 468–477 (2019).

[9]H. I. Jo and J. Y. Jeon, "Downstairs resident classification characteristics for upstairs walking vibration noise in an apartment building under virtual reality environment," Build. Environ. **150**, 21–32 (2019).

[10]Ö. Axelsson, M. E. Nilsson, and B. Berglund, "A principal components model of soundscape perception," J. Acoust. Soc. Am. **128**(5), 2836–2846 (2010).

[11]M. Rychtáriková and G. Vermeir, "Soundscape categorization on the basis of objective acoustical parameters," Appl. Acoust. **74**(2), 240–247 (2013).

[12]J. Y. Jeon, J. Y. Hong, C. Lavandier, J. Lafon, Ö. Axelsson, and M. Hurtig, "A cross-national comparison in assessment of urban park soundscapes in France, Korea, and Sweden through laboratory experiments," Appl. Acoust. **133**, 107–117 (2018).

[13]K. Genuit and A. Fiebig, "Psychoacoustics and its benefit for the soundscape approach," Acta Acust. Acust. **92**(6), 952–958 (2006).

[14]R. Cain, P. Jennings, and J. Poxon, "The development and application of the emotional dimensions of a soundscape," Appl. Acoust. **74**(2), 232–239 (2013).

[15]G. Brambilla, V. Gallo, F. Asdrubali, and F. D'Alessandro, "The perceived quality of soundscape in three urban parks in rome," J. Acoust. Soc. Am. **134**(1), 832–839 (2013).

[16]K. Jambrošić, M. Horvat, and H. Domitrović, "Assessment of urban soundscapes with the focus on an architectural installation with musical features," J. Acoust. Soc. Am. **134**(1), 869–879 (2013).

[17]ISO/DIS 12913-2: *Acoustics-Soundscape-Part 2: Data Collection and Reporting Requirements* (International Standard Organization, Geneva, Switzerland, 2017).

[18]J. Kang, "Acoustic comfort in non-acoustic spaces: A review of recent work in Sheffield," in *Proceedings of the Institute of Acoustics* (2003), Vol. 25, pp. 125–132.

[19]J. Kang, *Urban Sound Environment* (CRC Press, London 2006), pp. 133–134.

[20]Y. Smyrnova and J. Kang, "Determination of perceptual auditory attributes for the auralization of urban soundscapes," Noise Control Eng. J. **58**(5), 508–523 (2010).

[21]F. Aletta, Ö. Axelsson, and J. Kang, "Towards acoustic indicators for soundscape design," in *Proc. Forum Acusticum Conf.*, Krakow, Poland (2014).

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

[23]M. Zhang and J. Kang, "Towards the evaluation, description, and creation of soundscapes in urban open spaces," Environ. Plann. B Plann. Des. **34**(1), 68–86 (2007).

[24]J. Kang and M. Zhang, "Semantic differential analysis of the soundscape in urban open public spaces," Build. Environ. **45**(1), 150–157 (2010).

[25]J. Yong Jeon, P. Jik Lee, J. Young Hong, and D. Cabrera, "Non-auditory factors affecting urban soundscape evaluation," J. Acoust. Soc. Am. **130**(6), 3761–3770 (2011).

[26]A. Brown, J. Kang, and T. Gjestland, "Towards standardization in soundscape preference assessment," Appl. Acoust. **72**(6), 387–392 (2011).

[27]C. Lavandier and B. Defréville, "The contribution of sound source characteristics in the assessment of urban soundscapes," Acta Acust. Acust. **92**(6), 912–921 (2006).

[28]S. Viollon, C. Lavandier, and C. Drake, "Influence of visual setting on sound ratings in an urban environment," Appl. Acoust. **63**(5), 493–511 (2002).

[29]J. L. Carles, I. L. Barrio, and J. V. de Lucio, "Sound influence on landscape values," Landsc. Urban Plan. **43**(4), 191–200 (1999).

[30]J. Y. Hong and J. Y. Jeon, "The effects of audio-visual factors on perceptions of environmental noise barrier performance," Landsc. Urban Plan. **125**, 28–37 (2014).

[31]J. Y. Hong and J. Y. Jeon, "Designing sound and visual components for enhancement of urban soundscapes," J. Acoust. Soc. Am. **134**(3), 2026–2036 (2013).

[32]J. S. Bradley and G. A. Soulodre, "The influence of late arriving energy on spatial impression," J. Acoust. Soc. Am. **97**(4), 2263–2271 (1995).

[33]P. J. Lee, J. Y. Hong, and J. Y. Jeon, "Assessment of rural soundscapes with high-speed train noise," Sci. Total Environ. **482–483**, 432–439 (2014).

[34]L. Hermida and I. Pavón, "Spatial aspects in urban soundscapes: Binaural parameters application in the study of soundscapes from Bogotá-Colombia and Brasília-Brazil," Appl. Acoust. **145**, 420–430 (2019).