The effects of spatial separations between water sound and traffic noise sources on soundscape assessment

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

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

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

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

*Send correspondence to: Bhan Lam (blam002@e.ntu.edu.sg)

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

School of Electrical & Electronic Engineering,

Nanyang Technological University

Singapore 639798, Phone: +65 67906901

Abstract

Many studies have investigated the effects of water sound on soundscape with an assumption that target noise coincides with the masker (co-location), while no attention has been paid to spatial separations between target noise and water sound sources. This study aims to explore the effects of spatial separations between target noise and water sound on perceived loudness of target noise (PLN) and overall soundscape quality (OSQ) through laboratory experiments. Traffic noise (target) and a water sound (masker) were recorded as acoustic stimuli and a spherical panoramic video recording of a water fountain was also used as visual stimuli. The audio-visual stimuli were reproduced through a virtual reality head-mounted display and a multichannel ambisonic loudspeaker setup. The traffic noise and water sound were played simultaneously at various azimuthal separations and were combined with a panoramic recording of a water fountain as visual stimulus. Participants assessed the audio-visual stimuli in terms of PLN and OSQ. The effect of the spatial separation between the traffic noise and water sound was significant in both PLN and OSQ. Specifically, the PLN increase at 135° separation was equivalent to an estimated target noise level increment of ~1-2 dB. Similarly, the OSQ decrease at 135° and 180° separation was equivalent to an estimated target noise level increase of ~2-5 dB. Since the typical field of view of users in space is less than 135°, the results suggest that placing water features within a user's field of view could achieve better soundscape.

Keywords: Soundscape; Spatial audio; Virtual reality; Ambisonics; Spatial release from masking; Traffic noise

Declarations of interest: none

1. Introduction

Water features are important elements not only in landscape design but also in the soundscape design of a public space [1–3]. In addition to visual aesthetic values of water features in landscape design, water sounds from the water features have been employed as soundscape design elements to improve noisy environments [4,5]. Over the last decade, both perceptual and acoustic aspects of water sounds on soundscape have been explored through in-situ [6,7] or lab-based experiments [8–15]. Past studies have mainly studied two perceptual aspects of adding water sounds: reducing perceived loudness of a target noise and enhancing the overall acoustic comfort of the environment. Although water sounds are technically ineffective at energetically masking low-frequency traffic noise [16], water sounds have been shown to partially reduce the perceived loudness of target traffic noise [9,13,17,18]. Moreover, there is clear evidence that introducing water sounds can potentially improve the overall soundscape quality [10,12,14,19,20]. Not all types of water sounds, however, can guarantee the enhancement of soundscape quality, due to individual variations in spectral and temporal features [10,11].

In this context, the effects of acoustical characteristics (e.g., sound level, spectral, and temporal features) of water sounds on subjective preference have been investigated to suggest desirable acoustic design factors of water sounds. In previous studies, the effect of water sound level was usually examined by varying the signal-to-noise ratio (SNR) between a water sound and a target noise. It has been shown that water sounds with similar or 3 dB lower sound levels were preferred when combined with a traffic noise [10,12]. In terms of spectral characteristics, water sounds with more low-frequency contentwere evaluated as less pleasant [10,11]. It has also been shown that water sounds with high temporal variability are preferred to steady-state water sounds [10,11,19].

However, the aforementioned studies on the influence of water sounds inherently assume that the target noise source and water sounds were collocated or emitting from the same axial direction in space, which is difficult to realize in actual soundscape applications. This would take into account the physical constraints in terms of sound source placement in a functional three-dimensional space.

The effect of non-collocation of the target and masker is well-established in the field of speech intelligibility research. In general, speech intelligibility improves with increasing spatial separation between the masker and target speech [21–23]. This phenomenon is known as spatial release from masking (SRM), which is usually quantified by taking the difference between the speech reception thresholds of the collocated and separated conditions [21,24]. It has been reported that for speech separation, SRM can occur up to 12 dB in adults depending on the separating conditions [25–27]. In other words, the same speech intelligibility levels can be achieved even when the target levels were 12 dB lower in the spatially separated case (compared to the collocated case).

Inferring from the SRM phenomenon in speech intelligibility, it is plausible that spatial separation between a target noise and a water sound may affect soundscape perception and hence assessment. To the best of our knowledge, no study has attempted to explore the effect of the spatial separation between the target noise and the water feature on soundscape assessment, which is imperative to better predict the influence of water sounds on soundscape when designing water features in urban public spaces.

Therefore, this study examines the efficacy of introducing a water fountain sound at different spatial orientations in the azimuthal plane and to quantify SRM in soundscape assessment through a laboratory experiment. Two widely employed soundscape descriptors (perceived loudness of the target noise and soundscape quality) are evaluated in this study [28]. Specifically, two research questions are addressed: Does spatial separation between the water sound and the target noise affect (1) the judged perceived loudness of the target noise or (2) the subjective evaluation of overall soundscape quality? The results of the experiments associated with the first and second research questions are analysed and discussed in Sections 3 and 4, respectively.

2. Method

2.1 Stimuli

An audio-visual recording of a water fountain was conducted 4 m away from a fountain in the Nanyang Technological University (NTU) campus in clear weather. The fountain was composed of jets and a basin. A spherical panoramic camera (Garmin VIRB 360 Action Camera, USA) and 4-channel first-order ambisonic (FOA) microphone (Sennheiser AMBEO VR 3D Microphone, Germany) were used to capture the omnidirectional audio-visual recordings of the water fountain. The audio-visual capturing system was mounted on a tripod at a height of 1.5 m from the ground, as shown in Figure 1. The video of the fountain was recorded in 4K 30-FPS resolution with a bit-rate of 80 Mbps and post-processed for white balance and exposure compensations (Adobe Premiere Pro CC 2019). The FOA recordings were converted to the B-format AmbiX with the AMBEO A-B converter. The audio-visual recordings of the water fountain were conducted at two situations when the fountain was turned on and off. The 10-s A-weighted equivalent sound pressure level (SPL) when the fountain was turned on ($L_{\mathrm{Aeq,10s}}$) at a distance of 4 m was 67.4 dB.

Due to its pervasiveness, road traffic was selected as the target urban noise [29]. Road traffic sound was recorded at a distance of 40 m from an expressway (2 × 4 lanes) using the same FOA microphone. For the laboratory experiment, 10-s audio samples of the water fountain and traffic sounds were excerpted from the audio recordings. As the visual stimuli, two 10-s spherical video samples of the water fountain were excerpted from the video recordings, one when the fountain was turned on and one when the fountain was turned off.

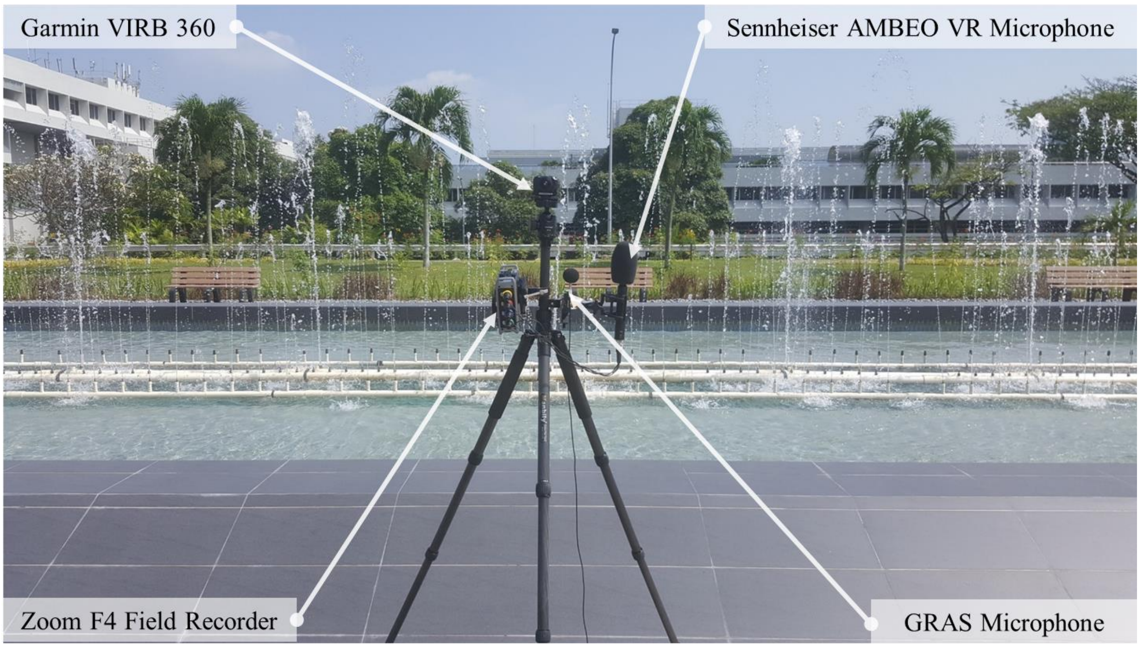


Figure 1. Measurement setup for the omnidirectional audio-visual recordings of the water fountain.

The sound pressure level (SPL) of the acoustic stimuli is plotted as a function of 1/3 octave bands from 63 Hz to 8 kHz in Figure 2(a). The road traffic sound exhibits a relatively constant SPL across all frequencies, whereas the water sound exhibits a roll-off in SPL below 315 Hz. In terms of temporal variability of the stimuli, the traffic noise possesses a relatively lower variance in SPL than the water sound as displayed in Figure 2(b). The SPL ranges of the target noise and the fountain sound over time were 4.9 dB and 9.1 dB, respectively.

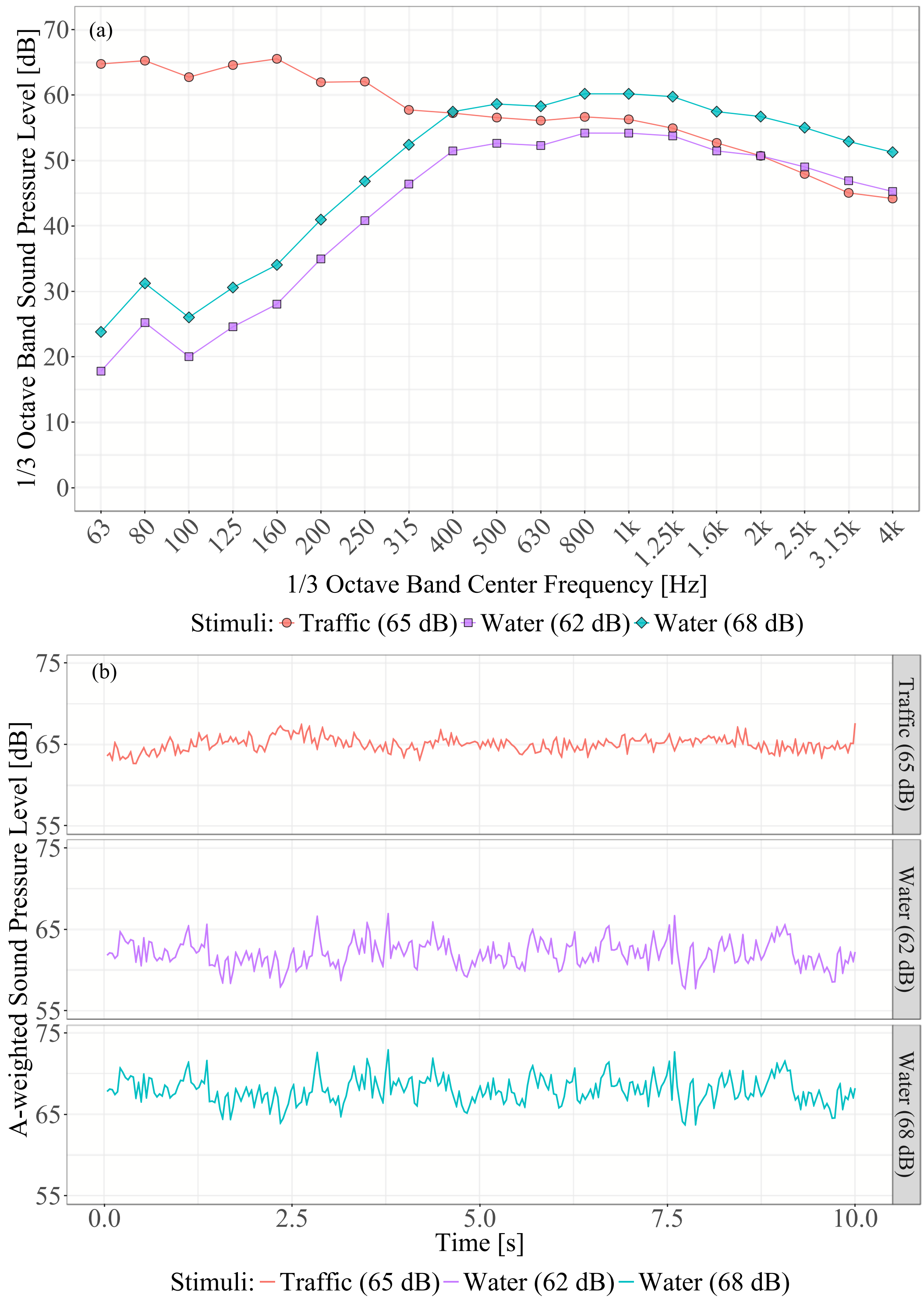


Figure 2. Acoustical characteristics of the 65 dB traffic (—), 62 dB water (—), and 68 dB water (—) stimuli in: (a) 1/3 octave band spectra, and (b) A-weighted sound pressure level as a function of time.

2.2 Subjective evaluation

Perceived loudness of noise (PLN) and overall soundscape quality (OSQ) were assessed for each stimulus. "Perceived loudness of noise" was defined as the subjectively judged auditory loudness of the target noise in this study. The PLN for the road traffic sound was assessed using a fixed-number magnitude estimation method. The target traffic noise of 65 dB was presented to the participants as a reference, and the PLN for the reference was assigned the fixed numerical value of "100".

The participants were instructed to focus on the target traffic noise when evaluating the combined sounds consisting of the target traffic noise and water sound. Subsequently, participants were requested to assign a number to each presented stimulus describing the perceived loudness of the traffic noise relative to the reference (i.e., target traffic noise alone at 65 dB). For instance, if a participant feels that the PLN of the target traffic sound in the following stimulus is three times as loud compared to the reference, the participant would assign it a value of "300". On the other hand, if it is half as loud, the participant would assign a value of "50". Additionally, the overall soundscape quality (OSQ), defined as overall impression of soundscape of both the target noise and water sound, was evaluated for each stimulus using an 11-point scale (0: extremely bad and 10: extremely good).

2.3 Experimental design and settings of VR reproduction

The experiments consisted of two sessions: (I) traffic noise alone and (II) the combined target traffic and water sounds. In session I, the 10-s traffic noise clips were calibrated to five levels from 55 to 75 dB ($L_{\mathrm{Aeq,10s}}$) in 5 dB steps, which represents a range of traffic noise levels in most urban environments [12,30,31]. It was assumed that the target traffic road was always oriented in the frontal direction of the participants, so the traffic noises were fixed at 0° azimuth. Each of the traffic noise alone acoustic stimuli were combined with the omnidirectional visual stimulus with the fountain turned off. In total, five audio-visual stimuli were created for session I, and participants were asked to assess the PLN and OSQ of the five traffic noise alone stimuli.

In session II, a within-subject repeated-measure factorial design was employed to examine the effects of two independent variables, SNR and azimuth separation, between the target traffic noise and water sound on PLN and OSQ. The target traffic noise was set at 65 dB. The SNR between the water (signal) and the target traffic (noise) sounds had two levels (−3 dB and +3 dB) since those SNRs for water sounds were previously found to be effective [10,12]. In other words, the water fountain sounds were set at 62 dB or 68 dB. Regarding the spatial separation between the traffic and water sounds, the traffic sound was always projected from the 0° azimuth position and the water fountain sound was presented at five absolute azimuths: 0° (collocated), 45°, 90°, 135°, or 180°. The azimuth separation interval was set to 45° by inference from a previous study by Marrone et al., where the benefit of spatial separation between the target and maskers for most normal hearing listeners was significant at 45° [32].

For audio-visual congruency, the viewpoint of the fountain in the omnidirectional video in each stimulus was rotated to the same azimuths (location of traffic noise was fixed in front at 0° azimuth) as depicted in Figure 3. Since the target traffic and water fountain sound were asymmetrically separated, two sets of audio-visual stimuli were created to include both the left- (e.g., −45°) or right-hand (e.g., 45°) side separations. For the audio-visual stimuli in set 1, the fountain sound and viewpoint were either positioned at 0°, +45°, −90°, +135°, or 180° azimuth. In set 2, the fountain sound and viewpoint were either position at 0°, −45°, +90°, −135°, or 180° azimuth. In each set, a total of 10 audio-visual stimuli (2 SNRs × 5 azimuth angles) were created and the participants were instructed to evaluate the PLN and the OSQ of all the audio-visual stimuli.

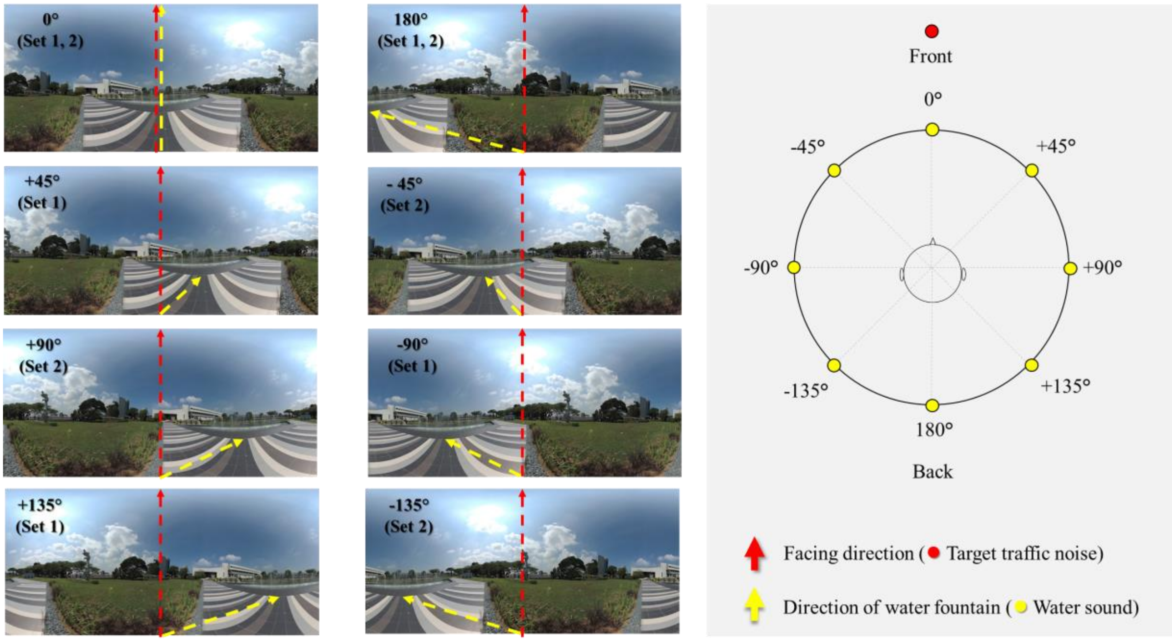


Figure 3. Schematic illustration of the experimental design: (Left) equirectangular panoramic stills from the spherical videos (Right) azimuth separations between the target traffic noise and water sounds

The audio-visual stimuli were presented through a twelve-channel loudspeaker system consisting of double hexagonal arrays and a virtual reality (VR) head-mounted display (HMD) (Pimax 4K VR, China), as shown in Figure 4. According to previous studies [33,34], a multichannel loudspeaker system can reproduce more realistic spatial acoustic cues than a headphone in terms of realism and immersion in soundscape. The B-format FOA tracks were decoded to the FOA-3D hexagonal array using the Ambisonic Toolkit (ATK) plugin for the Reaper DAW [35]. All loudspeakers were placed 1-m away from the center position of the hexagon where the participant was seated. In accordance with the experimental design, the A-weighted equivalent SPL of the 10-s acoustic stimuli were calibrated in an anechoic chamber using a head and torso simulator (Brüel & Kjær 4128-C, Denmark). The loudspeakers (Genelec 8320A Smart Active Monitor, Finland) were calibrated to a flat frequency response with the Genelec Loudspeaker Manager 2.0 (GLM) software at the sitting position (center of the double hexagonal speaker array).

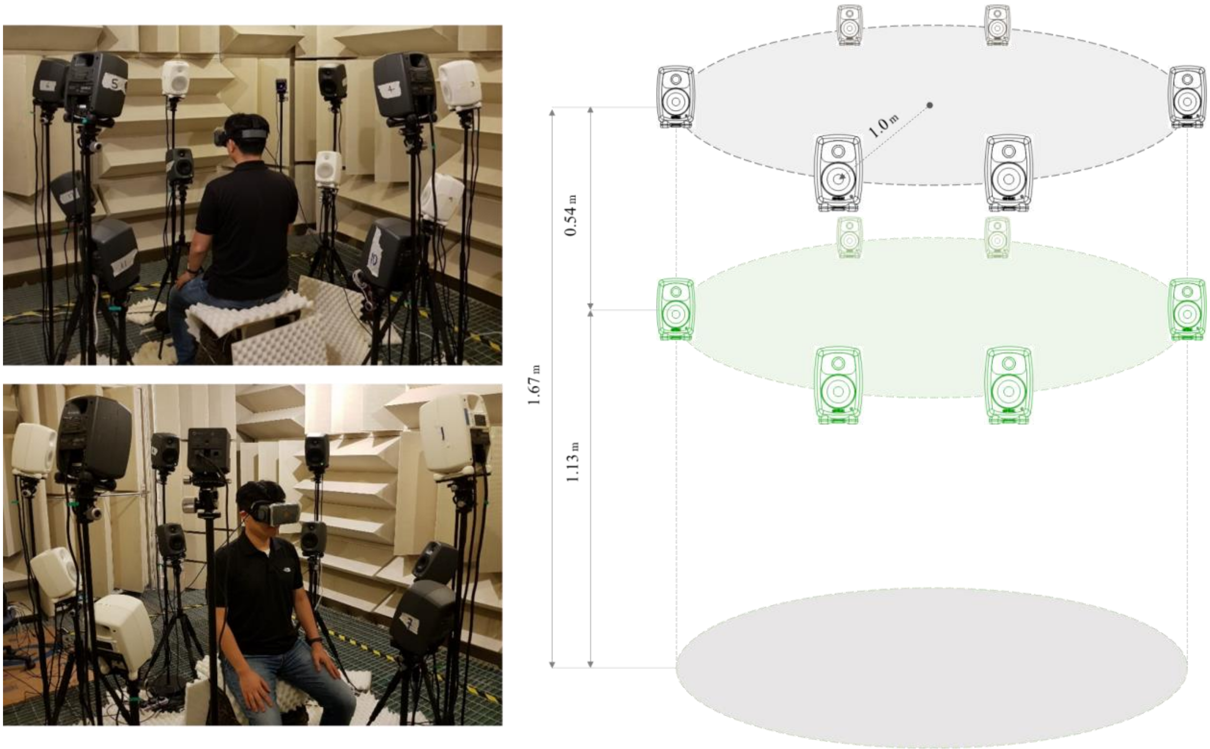


Figure 4. VR experiment setting: (Left) photographs of the audio-visual reproduction system in an anechoic chamber and (Right) the loudspeaker configuration

2.4 Participants

To achieve 80% statistical power, the required minimum sample size for the within-subject design was calculated based on a statistical power analysis using G*Power 3.1 [36]. The power analysis showed that at least 22 participants were required to detect a medium effect of: $f =$

0.3, $\alpha = 0.05$, and $(1 - \beta) = 0.80$. Hence, a total of 23 participants (13 males and 10 females) were recruited for this study, which was slightly more than the required minimum sample size. The age distribution of the participants ranged from 20 to 60 years (Mean = 32.4, SD = 12.4). Most participants were ethnically Chinese (21 Chinese and 2 Malays). Hearing tests were conducted using an audiometer (Interacoustics AD629, Denmark) before the experiment and the results showed that all the participants had normal hearing for all the tested frequencies (0.125, 0.5, 1, 2, 3, 4, 6, and 8 kHz). Among the participants, 11 participants evaluated audio-visual stimuli in set 1 and 12 participants evaluated set 2.

In accordance with ethical procedures, the study was approved by the institutional review board of NTU (IRB-2017-07-025). The study was approved by the local research ethics committee and informed written consent was obtained from all the participants after carefully instructing to them the purpose and procedures used for the experiments.

2.5 Procedure

The audio-visual stimuli were presented to the participants in random order through a VR HMD and loudspeaker reproduction system. After experiencing each audio-visual stimulus, the participants took off their VR HMD and evaluated the PLN and OSQ using a questionnaire. The participants were allowed to replay each stimulus as many times as required. Session I took approximately 10 min, and session II lasted approximately 30 min. There was a break time of approximately 15 min to relieve the boredom and fatigue of the participants [37].

2.6 Data analyses

Two-way repeated-measure (RM) analysis of variance (ANOVA) was conducted to examine the within-subjects effects of the SNR and azimuthal spatial separation, and interaction between SNR and spatial separation on PLN and OSQ. Normality assumptions of the residuals of dependent variables (i.e., PLN and OSQ) for each level of independent variables (SNR and azimuth separation) were examined with Shapiro–Wilk's test. Even though some datasets violated the normality assumption, we conducted RM ANOVA because it has been revealed that ANOVA yields robust and valid results against violation of the normality assumption [38,39]. Partial eta-squared (η_p^2) values were also reported as a measure of effect size. All statistical analyses were performed using the statistical software package, IBM SPSS (version 25.0, IBM, USA).

3. Results

3.1 Perceived loudness of traffic noise alone

Based on the subjective responses from session I, the mean magnitude estimation values of PLN for the traffic noise-alone stimuli were plotted as a function of the A-weighted equivalent SPL in Figure 5. The PLN of traffic noise increased linearly as the A-weighted SPL ($L_{\mathrm{Aeq,10s}}$) of the traffic noise increased, as shown in Figure 5. The mean magnitude estimation values of PLN for the traffic noises from 55 dB to 75 dB varied from 56.52 (SD = 8.84) to 156.52 (SD = 13.69), respectively. The prediction model for PLN of traffic noise-alone was obtained from a simple linear regression analysis as shown in Eq. (1) where $L_{\mathrm{Aeq,10s}}$ is the A-weighted SPL of traffic noise. The coefficient of determination (R^2) for the model was 0.92 ($p < 0.01$).

$$PLN_{traffic} = 4.92\, L_{Aeq,10\mathrm{s}} - 216.43 \qquad (1)$$

Based on the experimental design, the reference traffic noise at 65 dB was chosen for use as the baseline to examine the effects of water sound on PLN in session II. Hence, the regression model for PLN can be utilized to quantify the effects of water sound in terms of SNR and spatial separation.

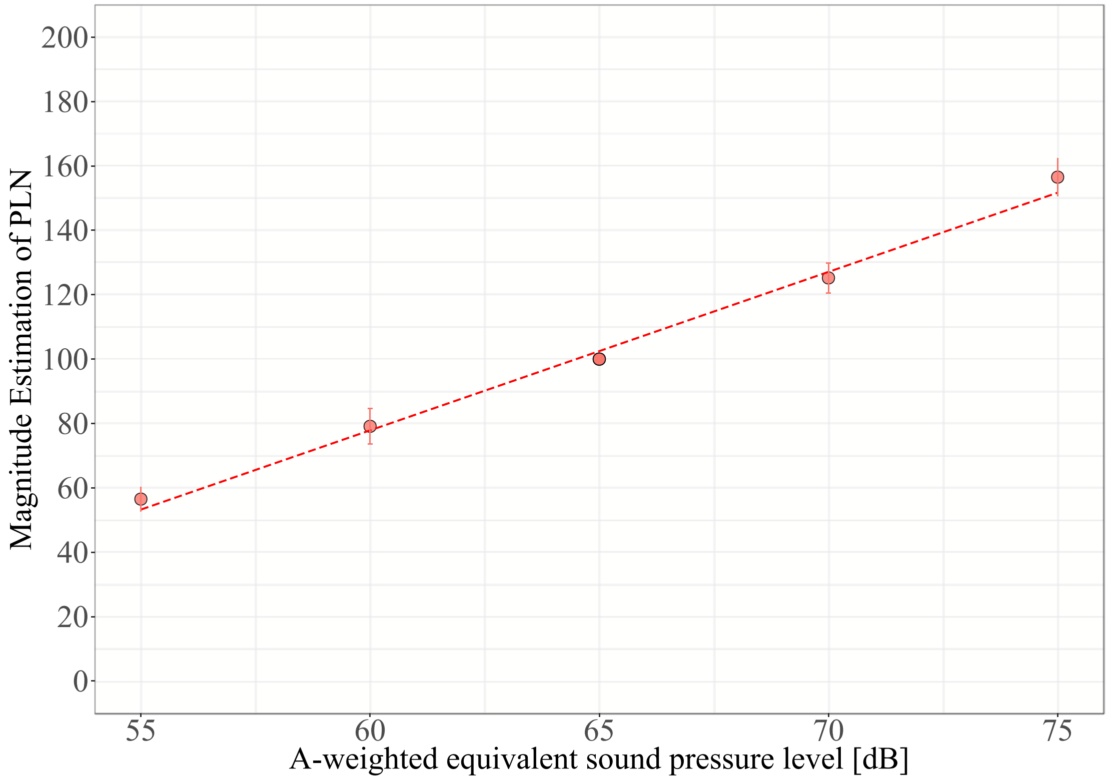


Figure 5. Mean perceived loudness of noise (PLN) scores (●) as a function of A-weighted sound pressure level. The linear regression function is fitted to the data points (- -) and the error bars indicate 95% confidence intervals.

3.2 Perceived loudness of noise in the combined traffic and water sound cases

Mean values of magnitude estimation across all participants for each acoustic stimulus were calculated. The mean values of the combined sounds (target traffic noise + water sound) cases at both SNRs (±3 dB) were plotted as a function of azimuth separation, as shown in Figure 6.

To examine the masking effect of the water sound, pairwise comparisons using Bonferroni correction between PLN of the target traffic and the combined sounds were conducted. The results showed that adding water sound at both SNRs (±3 dB) significantly reduced the PLN of the target traffic sound across the five azimuth angles ($p < 0.01$). This indicates that water sounds at both SNRs (±3 dB) could reduce the perceived loudness of the target traffic sound despite azimuthal separation.

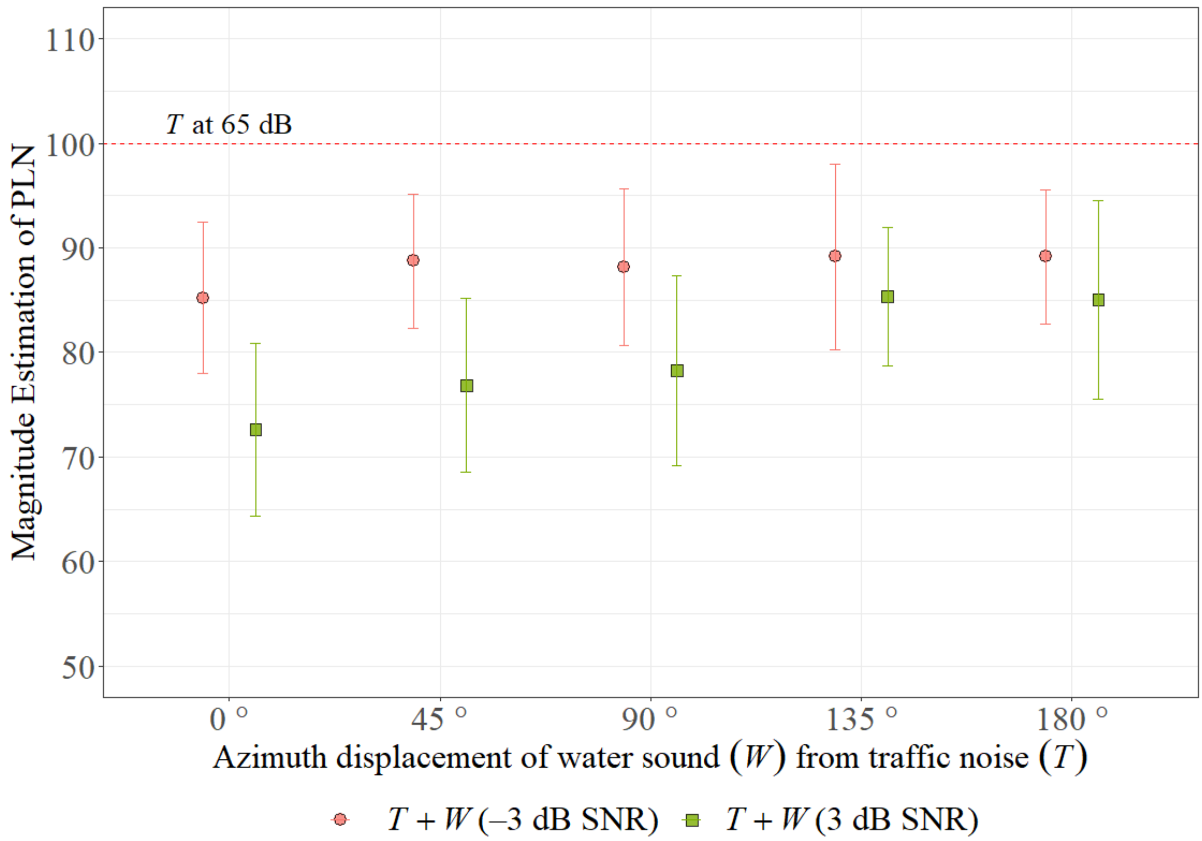


Figure 6. Mean perceived loudness of noise (PLN) as a function of azimuth separations between the target noise and water sound at SNRs of −3 dB ( ) and +3 dB ( ). The variables *T* and *W* designate the target traffic noise and water sounds, respectively; '+' denotes the combination of stimuli. The error bars indicate 95% confidence intervals. The PLN of the target noise at 65 dB from session I, which was fixed at "100", is plotted for reference (--).

Next, a two-way RM ANOVA was conducted to investigate the main effects of the azimuth separation, the SNR, and their4 interactions (azimuth separation × SNR) on the

PLN. The results of the F-tests are summarized in Table 1. The results showed that the main effects of SNR [$F(1,22) = 19.56, p < 0.001, \eta_p^2 = 0.47$] and spatial separation [$F(2.75, 60.61) = 3.25, p = 0.03, \eta_p^2 = 0.13$] were statistically significant. The SNR of +3 dB (Mean = 88.08, SD = 16.76) exhibited a greater reduction in PLN than the SNR of −3 dB (Mean = 78.62, SD = 19.73). This is consistent with previous studies that higher SNRs of water sound to traffic noise yielded greater benefits regarding the reduction of PLN [9,18].

The mean PLN values seemed to increase as the azimuth separation between the target noise and the water sound became larger. Post-hoc tests for PLN in azimuth separation showed that there were no statistical differences in PLN among different azimuth angles except between 0° and 135°. A statistically significant difference in PLN was found between the collocated condition at 0° (Mean = 78.91, SD = 18.85) and 135° spatially separated condition (Mean = 87.24, SD = 18.01) at 0.05 significance level. No significant interaction between SNR and azimuth separation was found [$F(2.29, 50.36) = 1.41, p = 0.25, \eta_p^2 = 0.06$].

Table 1. Summary of RM ANOVA results for the perceived loudness of noise (PLN)

Factors	df_1	df_2	F	p	η_p^2
SNR	1.00	22.00	19.56	< 0.001	0.47
Azimuth [a]	2.75	60.61	3.25	0.03	0.13
SNR * Azimuth [a]	2.29	50.36	1.41	0.25	0.06

[a] Assumption of sphericity was violated and Greenhouse–Geisser correction was applied.

To quantify the effects of the water sound on reducing the PLN of the target traffic noise, an inference method was adopted from a previous study [9]. Using the magnitude estimation scores of PLN in session II, an equivalent target traffic noise level, $SPL_{est,PLN}$, can be estimated from Eq. (1) and expressed as

$$SPL_{est,PLN}(\theta, SNR) = \frac{PLN_{traffic}(\theta, SNR) + 216.43}{4.92}, \quad (2)$$

where $PLN_{traffic}(\theta, SNR)$ refers to the mean magnitude estimation score from session II at the respective azimuthal separation, θ, and SNR of the stimulus in session II.

Hence, the equivalent SPL reduction effect in terms of PLN of the combined stimuli in session II, $SPL_{reduct,PLN}$, is determined by taking the difference between the reference level (65 dB) and the estimated level using Eq. (2) to give

$$SPL_{reduct,PLN}(\theta, SNR) = 65 - SPL_{est,PLN}(\theta, SNR). \quad (3)$$

For clarity, Eq. (3) can be visualized by plotting the mean values of PLN from session II onto Figure 5, as illustrated in Figure 7. For instance, when the SNR was −3 dB, the estimated SPL_{reduct} of the target noise by adding water sound ranged from 2.92 dB (180°) to 3.71 dB (0°), as shown in Figure 7(a). However, a larger reduction effect was attained when the SNR was +3 dB, ranging from 3.69 dB (135°) to 6.28 dB (0°) as shown in Figure 7(b).

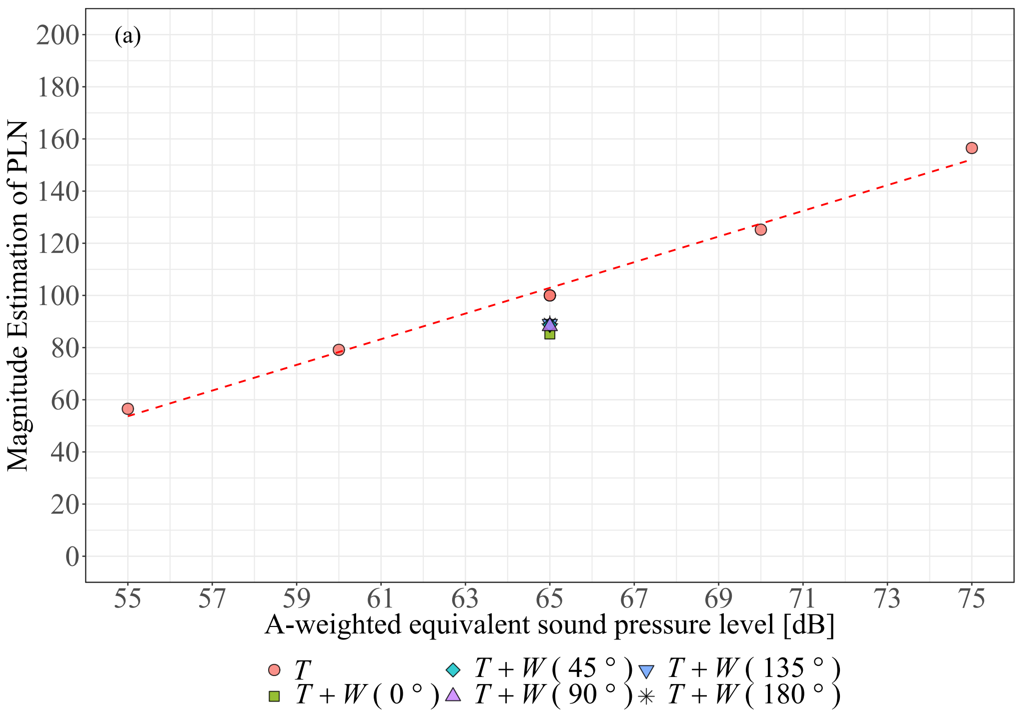
To prevent confusion with spatial release from masking in speech intelligibility studies, the effect of spatial release is quantified in dB for PLN as, SRM_{PLN}. The SRM_{PLN} is defined here as the difference in the estimated equivalent SPL reduction of the target traffic noise between the collocated (i.e., $SPL_{est}(0°, SNR)$), and the non-collocated cases (i.e., $SPL_{est}(\theta, SNR), \theta \neq 0°$), given by

$$SRM_{PLN}(\theta, SNR) = SPL_{reduct,PLN}(0°, SNR) - SPL_{reduct,PLN}(\theta, SNR), \quad (4)$$

where $\theta \neq 0°$. The $SRM_{PLN}(\theta, SNR)$ is computed for all the subjective responses at both SNRs as a function of θ. For clarity, the mean $SRM_{PLN}(\theta, SNR)$ values for all θ and SNR are summarized in Table 2. When the SNR was at −3 dB, the mean SRM_{PLN} were similar across the four azimuths showing $SRM_{PLN}(\theta, -3) \approx 0.7$ dB. However, when the SNR was at +3 dB, the $SRM_{PLN}(\theta, +3)$ at $\theta = 135°$ (Mean = 2.59 dB, SD = 3.11 dB) and $\theta = 180°$ (Mean = 2.53, SD = 4.56 dB) were greater than those at $\theta = 45°$ (Mean = 0.86 dB, SD = 3.90 dB) and $\theta = 90°$ (Mean = 1.15 dB, SD = 4.28 dB).

Table 2. Mean spatial release from masking for perceived loudness of noise (SRM_{PLN}, dB) as a function of azimuth separations between the target noise and water sound at SNRs of −3 dB and +3 dB. The numbers in parentheses represent standard deviations.

SNR (dB)	Azimuth separation				Total
	45°	90°	135°	180°	
−3	0.72	0.60	0.79	0.80	0.73
	(3.00)	(3.52)	(4.17)	(3.02)	(3.41)
+3	0.86	1.15	2.59	2.53	1.78
	(3.90)	(4.28)	(3.11)	(4.46)	(3.98)



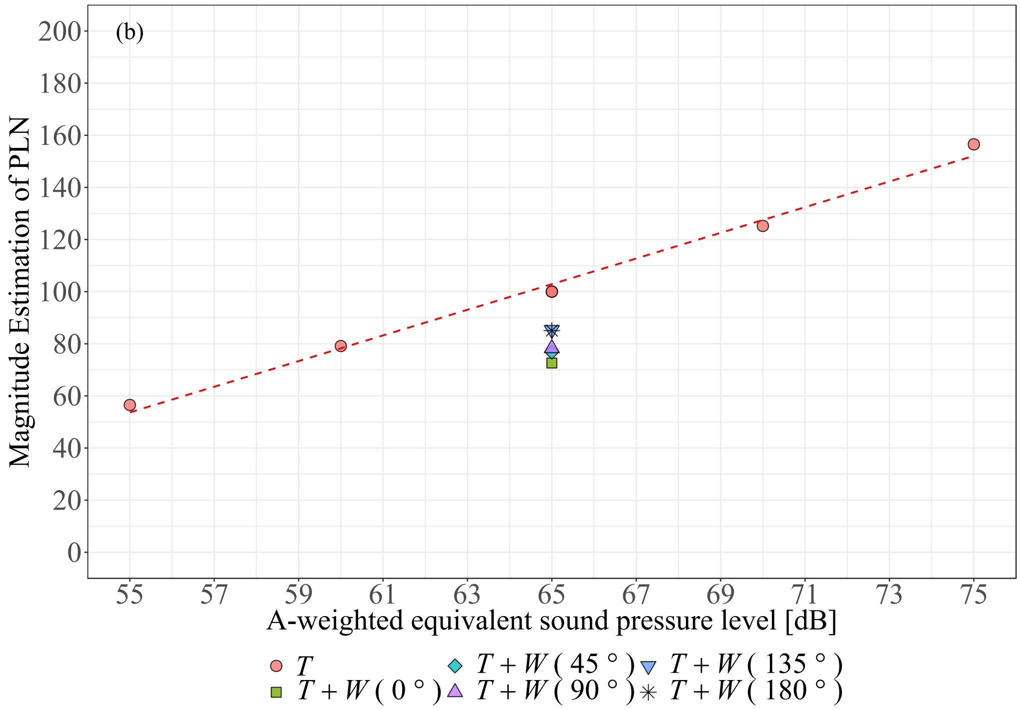


Figure 7. Mean magnitude estimation of perceived loudness of noise (PLN) as a function of A-weighted equivalent sound pressure level of traffic noise. The variables *T* and *W* indicate the target traffic noise and water acoustic stimuli, respectively; '+' denotes the combination of stimuli. The PLN traffic-alone cases (●) and its respective linear regression (- -) from session I is plotted for reference. The PLN of 65 dB target traffic noise combined with water sound (*T*+*W*) is plotted at an (a) SNR of −3 dB and (b) SNR of +3 dB at azimuth separations of 0° (■), 45° (◆), 90° (▲), 135° (▼), and 180° (✳) from session II. The numbers in parentheses denote azimuth separations between the target noise and water sound.

3.3 Overall soundscape quality in traffic noise alone cases

Similar to the PLN in Section 3.1, the mean values of OSQ from session I for the 6 traffic sound stimuli were plotted as a function of the A-weighted equivalent SPL in Figure 8. In contrast to the PLN, the OSQ rating score decreased as the A-weighted SPL of the traffic noises increased, as shown in Figure 8. The mean OSQ scores for the traffic noise cases at 55 dB, 65 dB, and 75 dB were 5.48 (SD = 1.08), 4.04 (SD = 0.98), and 2.61 (SD = 1.20), respectively. The prediction model for OSQ was also drawn from a simple linear regression analysis given by

$$OSQ_{\text{traffic}} = -0.14\, L_{Aeq,10s} + 13.27, \quad (5)$$

where the model explained 47% of the total variance of OSQ ($p < 0.01$).

The OSQ values for the traffic noise at the reference of 65 dB were used as the baseline to examine the effects of water sound on OSQ in the next section. In addition, the regression models for OSQ of the target traffic noise were used in a similar fashion to Section 3.2 to quantify the effects of water sound in terms of SNR and spatial separation.

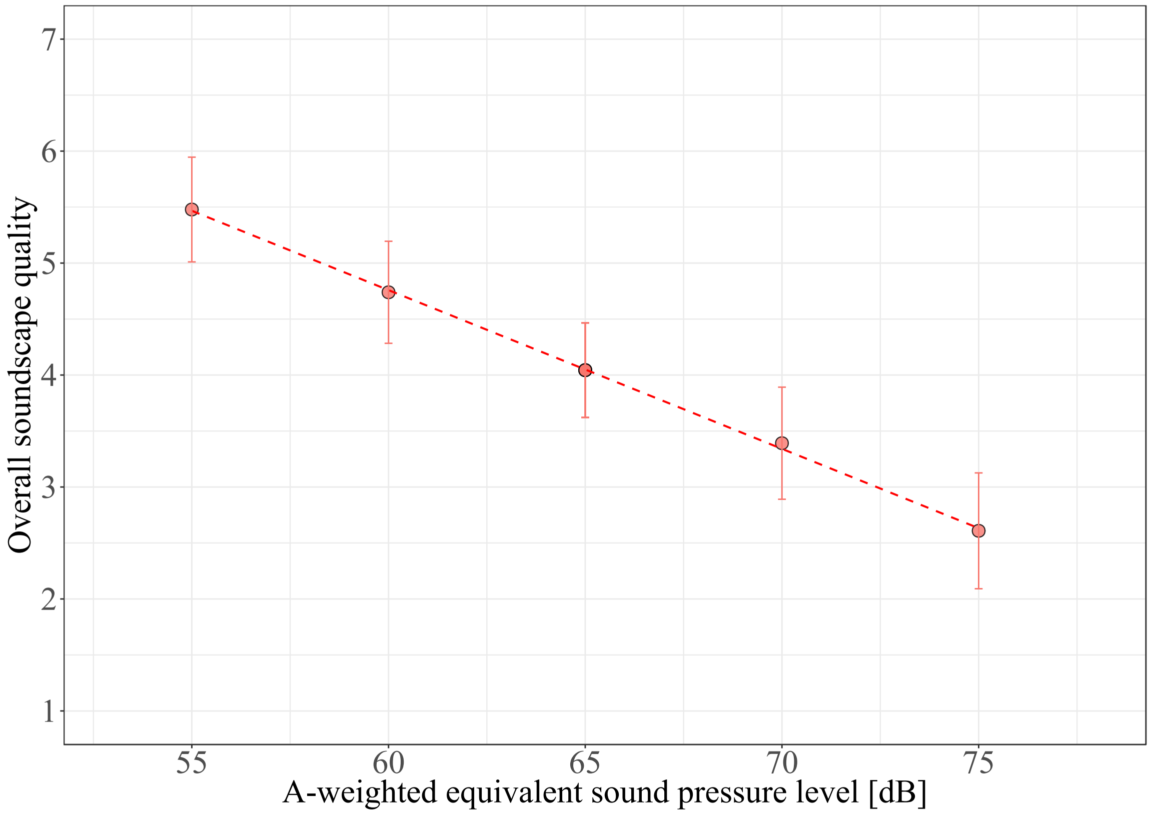


Figure 8. Mean overall soundscape quality (OSQ) score (●) as a function of A-weighted sound pressure level. The linear regression function is fitted to the data points (– –) and the error bars indicate 95% confidence intervals.

3.4 Overall soundscape quality in combined-sound cases

Mean rating scores of OSQ with the two SNRs are plotted as a function of azimuth separations, as shown in Figure 9. Pairwise comparisons for the mean OSQ scores between the target traffic noise-alone and the combined sound cases were conducted to examine the effect of water sound on enhancing the overall soundscape quality. The results showed that the mean OSQ scores for all the combined sounds were significantly higher than that for the target noise ($p < 0.05$). This demonstrates that introducing water sounds to the target noise could significantly increase soundscape quality across all the spatial azimuthal separations.

Regarding the azimuth separation, the mean OSQ scores were the highest when the target traffic noise and water sound were co-located (0°) at +3 dB SNR (Mean = 6.09, SD = 1.44), while the lowest OSQ score was observed when the water sound was located at 135° in azimuth at +3 dB SNR (Mean = 5.26, SD = 1.35).

A two-way RM ANOVA was performed to examine effects of SNR and azimuth separation (main effects and interactions) on the OSQ score. In contrast to the PLN, there were no significant mean differences in OSQ between SNRs of −3 dB (Mean = 5.54, SD = 1.32) and +3 dB (Mean = 5.65, SD = 1.46) $[\mathrm{F}(1.0, 22.0) = 0.82, p = 0.38, \eta_p^2 = 0.04]$, as shown in Table 3. This supports findings from previous studies that a higher SPL of water sound over target noise might not result in higher soundscape quality [10,12]. Meanwhile, the main effect of azimuth separation was statistically significant $[\mathrm{F}(4.0, 88.0) = 6.78, p < 0.001, \eta_p^2 = 0.24]$. The post-hoc tests showed that the mean OSQ scores at 135° (Mean = 5.33, SD = 1.38) and 180° (Mean = 5.41, SD = 1.43) were significantly lower than that at 0° (Mean = 5.89, SD = 1.40) at 0.05 significance level. No significant interaction effects between SNR and azimuth separation were found [$\mathrm{F}(4.0, 88.0) = 0.89, p = 0.47, \eta_p^2 = 0.04$]

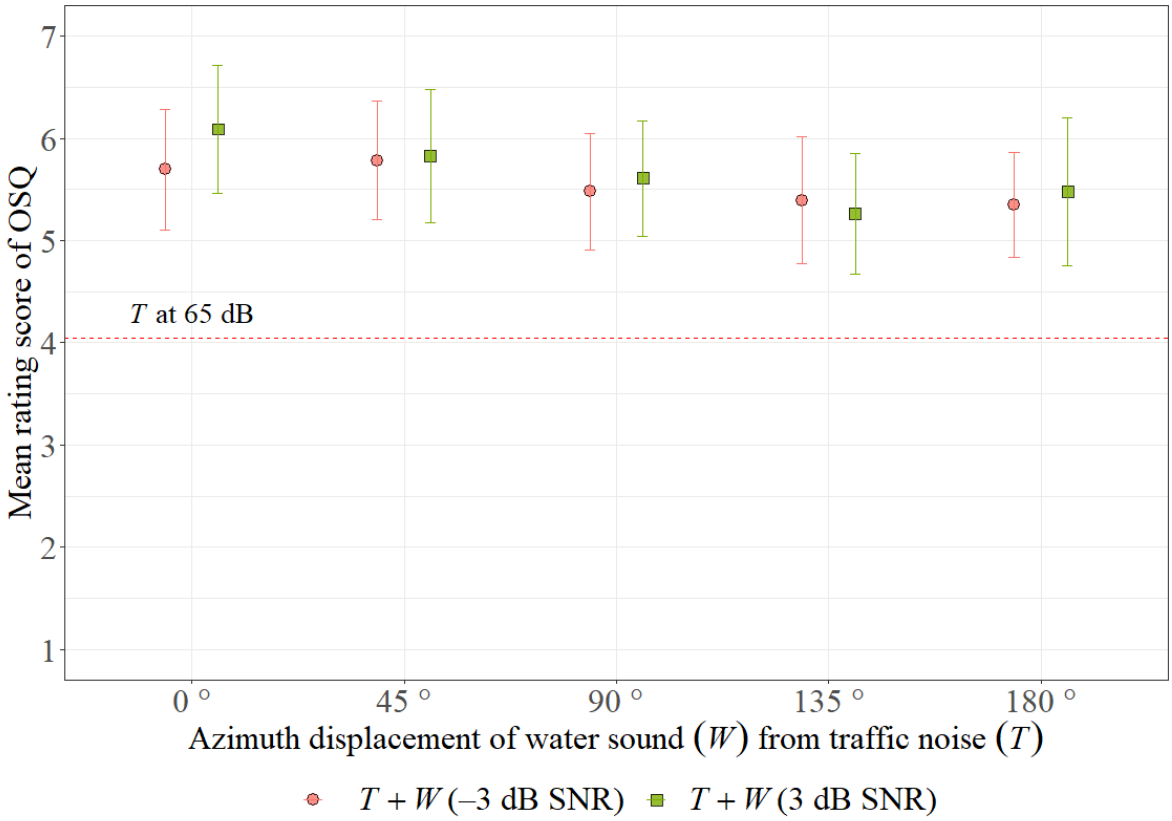


Figure 9. Mean rating scores of overall soundscape quality (OSQ) as a function of azimuth separations between the target noise and water sound at an SNR of −3 dB (●) and +3 dB (■). The variables *T* and *W* designate the target traffic noise and water sounds, respectively; '+' denotes the combination of stimuli. The error bars indicate 95% confidence intervals. The mean OSQ of the target noise at 65 dB (4.04) from session I is plotted for reference (- -).

Table 3. Summary of RM ANOVA results for the overall soundscape quality (OSQ)

Factors	df_1	df_2	F	p	η_p^2
SNR	1.0	22.0	0.82	0.38	0.04
Azimuth	4.0	88.0	6.78	< 0.001	0.24
SNR * Azimuth	4.0	88.0	0.89	0.47	0.04

Based on the inference method adopted in Section 3.2, an equivalent traffic noise level, $SPL_{est,OSQ}$, can be derived from the linear regression in Eq. (5) to give

$$SPL_{est,OSQ}(\theta, SNR) = \frac{13.27 - OSQ_{traffic}(\theta, SNR)}{0.14}, \quad (6)$$

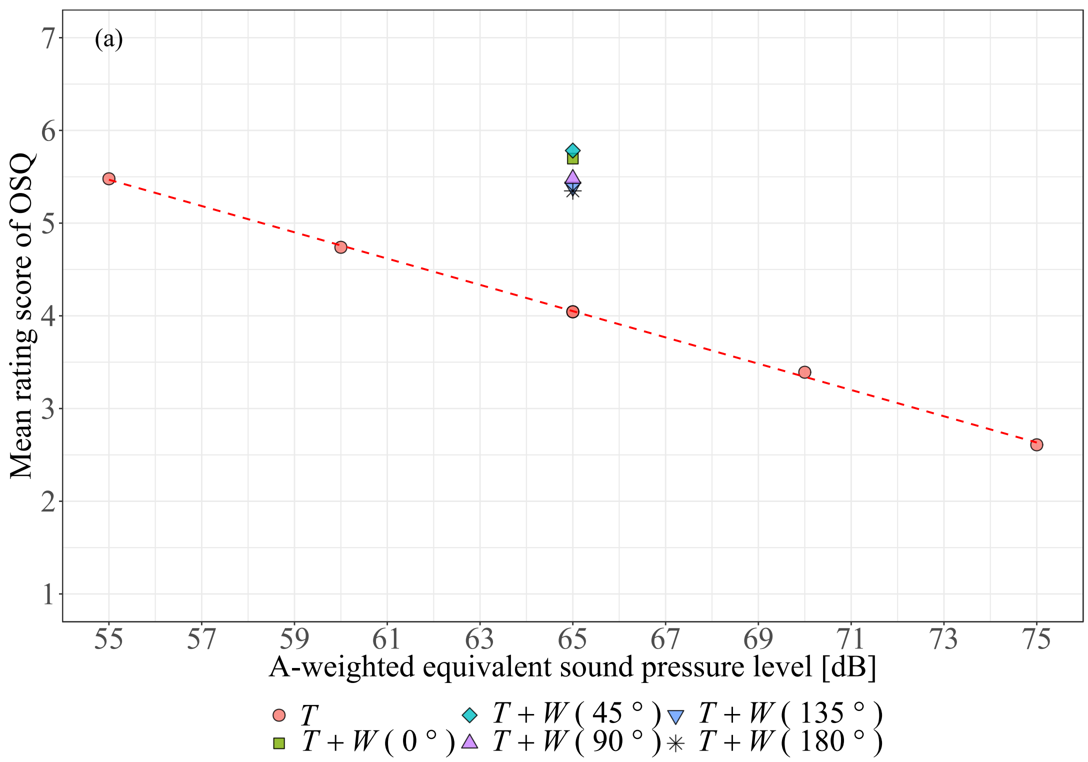
where $OSQ_{traffic}(\theta, SNR)$ refers to the mean OSQ scores at the respective azimuthal separation and SNR of the stimulus in session II.

The equivalent SPL reduction effect of the target noise in terms of the OSQ scores of the combined stimuli in session II, $SPL_{reduct,OSQ}$, is determined by taking the difference between the reference level (65 dB) and the estimated SPL from Eq. (6) to give

$$SPL_{reduct,OSQ}(\theta, SNR) = 65 - SPL_{est,OSQ}. \quad (7)$$

To visualize Eqs. (6) and (7), the mean OSQ scores for both the traffic noise-alone sounds and combined sounds (the target traffic noise at 65 dB + the water sounds) by varying SNR and azimuth separation are plotted as a function of A-weighted SPL, as shown in Figure 10. The red dashed line represents the linear regression line derived from OSQ scores of session I as described by Eq. (5).

As depicted in Figure 10, there were no significant mean differences between SNRs of −3 dB and +3 dB regarding estimated SPL reduction of traffic noise. When the water sound was co-located with the target noise, enhancement in OSQ was equal to a reduction of 14.36 dB at +3 dB SNR. Meanwhile, the minimum SPL reduction was 8.53 dB when the water sound with SNR of +3 dB was separated at 135°. Regarding the OSQ, adding the fountain sound source had the same effect as reducing the A-weighted SPL of the target noise by approximately 11 dB on average, which would be a substantial reduction that might be difficult to achieve using traditional noise control approaches. These results imply that introducing water sounds has more benefits on enhancing the overall acoustic quality of the place than on reducing perceived loudness of the target noise.



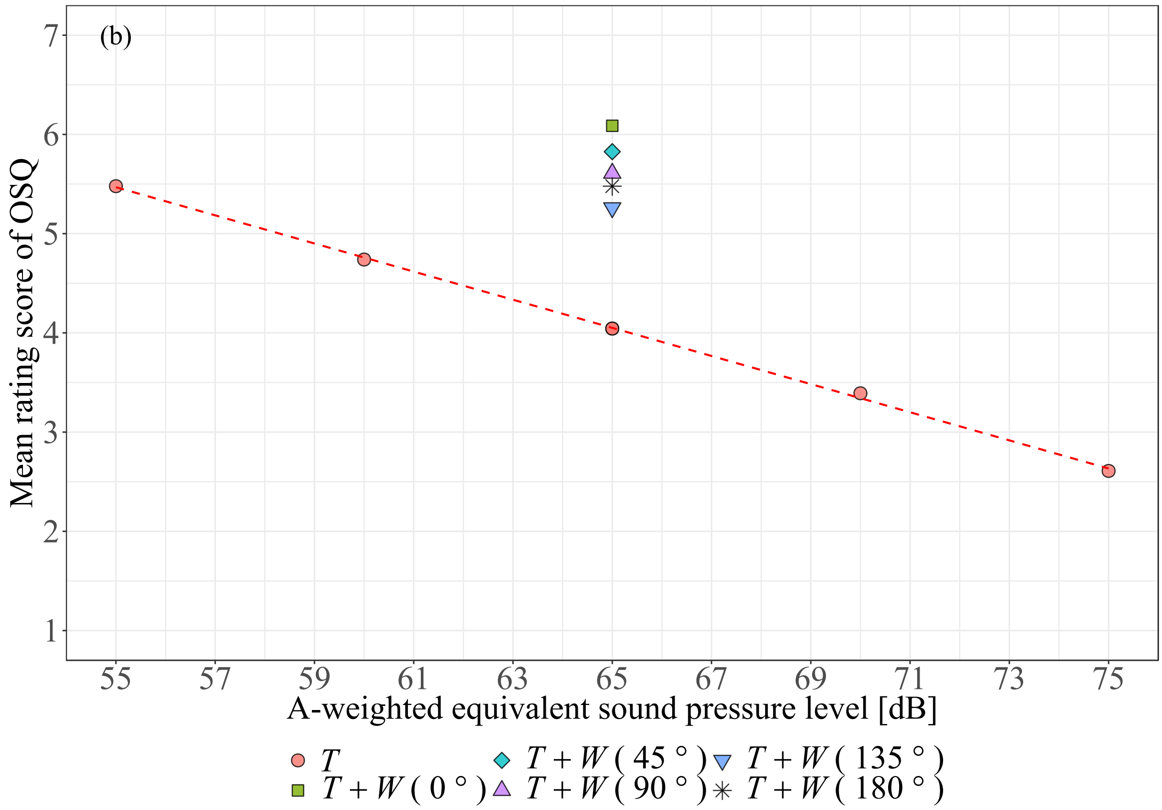


Figure 10. Mean overall soundscape quality (OSQ) as a function of A-weighted equivalent sound pressure level of traffic noise. The variables *T* and *W* indicate the target traffic noise and water acoustic stimuli, respectively; '+' denotes the combination of stimuli. The OSQ traffic-alone cases (●) and its respective linear regression (- -) from session I is plotted for reference. The OSQ of 65 dB target traffic noise combined with water sound (*T*+*W*) is plotted at an SNR of (a) −3 dB and (b) +3 dB at azimuth separations of 0° (■), 45° (◆), 90° (▲), 135° (▼), and 180° (✳) from session II. The numbers in parentheses denote azimuth separations between the target noise and water sound.

The spatial release from masking effect is quantified in terms of the OSQ as a function of azimuth and SNR, $SRM_{OSQ}(\theta, SNR)$. It is defined as the difference in the estimated SPL reduction of the target traffic noise between the collocated (i.e., $SPL_{est,OSQ}(0°, SNR)$) and spatially separated cases (i.e., $SPL_{est,OSQ}(\theta, SNR), \theta \neq 0°$), given by

$$SRM_{OSQ}(\theta, SNR) = SPL_{reduct,OSQ}(0°, SNR) - SPL_{reduct,OSQ}(\theta, SNR), \qquad (8)$$

where $\theta \neq 0°$.

For clarity, the $SRM_{OSQ}(\theta, SNR)$ values are summarized for all θ and SNR in Table 4. Overall, the SRM in terms of OSQ increased as the separation increased. The SRM at +3 dB SNR, $SRM_{OSQ}(\theta, +3)$, was relatively greater than that at −3 dB SNR, $SRM_{OSQ}(\theta, -3)$. When the SNR was −3 dB, the $SRM_{OSQ}(\theta, -3)$ was approximately 2 dB between 90° and 180° ($90° \leq \theta \leq 180°$). Substantial increments in $SRM_{OSQ}(\theta, +3)$ were found at $\theta = 135°$ (Mean = 5.83 dB, SD = 4.14 dB) and $\theta = 180°$ (Mean = 4.29 dB, SD = 5.11 dB) when the SNR was +3 dB.

Table 4. Mean spatial release from masking for overall soundscape quality (SRM_{OSQ}, dB) as a function of azimuth separations between the target noise and water sound at SNRs of −3 dB and +3 dB. The numbers in parentheses represent standard deviations.

SNR (dB)	Azimuth separation, θ(°)				Total
	45°	90°	135°	180°	
−3	-0.61 (4.11)	1.53 (4.00)	2.15 (4.39)	2.45 (3.63)	1.38 (1.92)
+3	1.84 (4.57)	3.37 (3.98)	5.83 (4.14)	4.29 (5.11)	3.83 (2.12)

4. Discussion

In terms of speech recognition, spatial separation between target speech and masker has been shown to improve speech intelligibility. In the design of soundscapes, however, spatial separation produces negative effects on soundscape perception due to spatial unmasking of the target noise. The results of this study revealed that the effects of spatial separation between the target noise and water fountain were significant on both PLN and OSQ evaluations. The quantified SRM effect for PLN and OSQ varied up to ~2.6 dB and ~4.0 dB across the azimuthal separations, respectively. The influence of spatial separation of the water sound from the target noise on PLN and OSQ with their implication on the soundscape design are discussed in Sections 4.1 and 4.2, respectively. Additionally, inherent limitations of this study together with future work are discussed in Section 4.3.

4.1 Effects of spatial separation between the water sound and the target noise on PLN

In terms of the PLN, significant spatial release only occurred when the water sound was separated by $\theta = 135°$. In comparison to studies on speech intelligibility, Marrone et al. [32] revealed that SRM usually occurs from a spatial separation of 15° to the full benefit at 45°. Srinivasan et al. [40] even showed that SRM for normal hearing listeners occurred at smaller spatial separations between target and maskers (2° to 6°). This discrepancy suggests that the judged PLN is less affected by spatial separations between the target and masker than speech intelligibility. This might be owing to a difference in the quantification method of SRM for PLN and speech intelligibility. In this study, the SRM effect in terms of PLN of the target noise was based on the magnitude estimation method – a subjective estimation of loudness. However, the SRM for speech intelligibility is usually quantified in terms of the speech reception threshold (SRT) [21], which could provide more precise differences than the magnitude estimation method.

It is also worth noting that the small effect of SRM in this study may be attributed to the dissimilarity between the target noise and water sound. The road traffic and water fountain stimuli possess different spectra-temporal characteristics. Several studies have proven that target-masker dissimilarity (e.g., different-sex or non-speech masker) results in smaller SRM because different spectra-temporal characteristics of signals can be useful for source segregation [22,41,42].

4.2 Effect of spatial separation between the water sound and the target noise on OSQ

The mean OSQ rating scores were significantly lower for spatially separated conditions at 135° and 180°, as compared to the co-located condition. Interestingly, this result was somewhat different from PLN in that the SRM was not significant when the water sound was displaced by 180°. This could be because sound sources presented from behind produce the same interaural level (ILDs) and time differences (ITDs) [21].

The effect of spatial separation between the target noise and water sound source on OSQ could be explained by visibility of a sound source. The congruency between acoustic and visual environments has a significant influence on soundscape [31,43,44]. When the fountain was located at 135° or 180°, the participants were not able to see the water fountain without head rotation because the VR HMD used in the experiment had a 110° of a binocular field of view, which may decrease OSQ scores than those of collocated condition with the target noise. This corroborates with a previous study [45] that the perceptions of water sounds in terms of their pleasantness and appropriateness are greatly affected by the visibility of water features. Hence, spatial separations from the target noise within a field of view would be effective to enhance acoustic comfort of the place.

4.3 Limitations and future work

This study adopted a laboratory experiment method, which is a widely used soundscape evaluation protocol due to its efficacy in exploring the cause-and-effect relationship between dependent and independent factors under controlled conditions. However, a laboratory experiment in controlled conditions could be limited in reflecting real-life situations, yielding relatively lower ecological validity.

Therefore, acoustic recording and reproduction techniques with sufficient spatial acoustic fidelity are essential to achieve high ecological validity of the results in a laboratory experiment [33,34,46–48]. In this study, B-format FOA audio recordings were reproduced via the FOA-3D hexagonal array speaker system, and it has been previously shown that the multichannel loudspeaker system used in this experiment can reproduce realistic spatial acoustic perceptions regarding directivity and distinctiveness of individual sound sources in a space to a similar extent as in-situ conditions [33]. This demonstrates that the VR acoustic reproduction method used in this study yielded sufficient ecological validity by reproducing sufficient spatial accuracy. In the future, nevertheless, an in-situ study could be conducted to investigate effects of SRM on soundscape assessment in real-life scenarios to cross-validate the results of this study.

It also should be noted that there are several limitations regarding the experimental design. This study deals with limited scenarios in terms of spatial configurations of target noise and water sound to investigate the effects of SRM on soundscape. Specifically, this study assumes that a target traffic noise is fixed at the frontal direction (0° azimuth) of a participant and the directions of the fountain sound are changed. However, the changes in SRM on soundscape might occur in different spatial scenarios between target noise and fountain sounds. Thus, further studies are still necessary to explore the effects of spatial factors of maskers on soundscapes with various spatial configurations. For instance, the cases of target noise and fountain sound can be swapped; the location of the fountain sound could be fixed, and the target noise position is rotated by a given azimuth angle.

In addition, one type of water fountain was used in this study. There are diverse types of water features generating various water sounds with different acoustic characteristics [7,14]. Moreover, it has been found that sociocultural factors (e.g., nationality, ethnicity, or age group) affect appreciation on waterscapes [49,50]. Thus, to build on the findings in this study, the influence of spatial separations between target noise and water sound on soundscape should be examined with a larger set of water features as well as wide ranges of participants with various socio-cultural backgrounds.

Since visual cues play an important role in both SRM [51] and the soundscape [8,15,52], it is necessary to explore audio-visual interactions on SRM in soundscape assessment. Furthermore, this study only focused on directions of masker in a horizontal plane as a spatial acoustic factor of a masker. In addition to the azimuthal separations, future studies could investigate other spatial factors of maskers such as width, elevation, and distances of maskers on soundscapes.

5. Conclusions

The effects of the spatial separations between target traffic noise and water sound on the perceived loudness of noise and overall soundscape quality were examined at five azimuth separations (0°, 45°, 90°, 135°, and 180°) through a VR laboratory experiment. The spatial release effects were quantified by the difference in SPL necessary to induce the corresponding differences in soundscape assessment of the collocated and separated conditions. It was found that the azimuth separations between the target noise and water sound significantly affected both reduction in perceived loudness of the target noise and improvement in overall soundscape quality. In particular, a 135° separation between the water sound and target noise increased the PLN by an amount quantified to be equivalent to a target noise level increase of ~1-2 dB. Similarly, a 135° or 180° separation between the water sound and target noise also decreased the OSQ significantly, by an amount quantified to be equivalent to an increase in target noise level by ~2-5 dB. Since the typical field of view of users in space is less than 135°, the findings suggest that placing water features out of a user's field of view could reduce its effectiveness in soundscape design.

Acknowledgments

This research is supported by the Singapore Ministry of National Development and the National Research Foundation, Prime Minister's Office under the Land and Liveability National Innovation Challenge (L2 NIC) Research Programme (L2 NIC Award No. L2NICCFP2-2015-5). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Singapore Ministry of National Development and National Research Foundation, Prime Minister's Office, Singapore." The first author is currently supported by the 2017 Lee Kuan Yew Postdoctoral Fellowship (Grant No.: 50025000). The authors are also grateful to Nanyang Technological University School of Mechanical & Aerospace Engineering for providing access to their anechoic chamber.

References

[1] J. Nasar, Y.H. Lin, Evaluative responses to five kinds of water features, Landsc. Res. 28 (2003) 441–450. doi:10.1080/0142639032000150167.

[2] M. Tveit, Å. Ode, G. Fry, Key concepts in a framework for analysing visual landscape character, Landsc. Res. 31 (2006) 229–255. doi:10.1080/01426390600783269.

[3] S. Burmil, T.C. Daniel, J.D. Hetherington, Human values and perceptions of water in arid landscapes, Landsc. Urban Plan. 44 (1999) 99–109. doi:10.1016/S0169-2046(99)00007-9.

[4] A.L. Brown, T. Gjestlan, D. Dubois, Chapter 1: Acoustics environments and soundscapes, in: Soundscape Built Environ., 2016: pp. 1–16.

[5] J. Kang, B. Schulte-Fortkamp, Soundscape and the built environment, CRC Press, London, 2015.

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

[7] J. Kang, On the diversity of urban water scape, in: Acoustics, 2012: pp. 3534–3538. doi:10.1093/acprof:oso/9780198240525.003.0002.

[8] J.Y. Hong, J.Y. Jeon, Designing sound and visual components for enhancement of urban soundscapes., J. Acoust. Soc. Am. 134 (2013) 2026–36. doi:10.1121/1.4817924.

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

[10] L. Galbrun, T.T. Ali, Acoustical and perceptual assessment of water sounds and their use over road traffic noise, J. Acoust. Soc. Am. 133 (2013) 227–237.

doi:10.1121/1.4770242.

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

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

[13] B. De Coensel, S. Vanwetswinkel, D. Botteldooren, Effects of natural sounds on the perception of road traffic noise., J. Acoust. Soc. Am. 129 (2011) EL148-L153. doi:10.1121/1.3567073.

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

[15] L. Galbrun, F.M. a Calarco, Audio-visual interaction and perceptual assessment of water features used over road traffic noise., J. Acoust. Soc. Am. 136 (2014) 2609. doi:10.1121/1.4897313.

[16] M. Nilsson, J. Bengtsson, R. Klaeboe, Environmental Methods for Transport Noise Reduction, CRC Press, 2014. doi:10.1201/b17606.

[17] J.Y. Hong, B. Lam, Z.T. Ong, R. Gupta, W.S. Gan, S.H. Chong, J. Feng, Appropriate levels of natural sounds to enhance soundscapes in urban areas, INTER-NOISE NOISE-CON Congr. Conf. Proc. (2017) 49–53.

[18] K. Bolin, M.E. Nilsson, S. Khan, The potential of natural sounds to mask wind turbine noise, Acta Acust. United with Acust. 96 (2010) 131–137. doi:10.3813/AAA.918264.

[19] M. Rådsten-Ekman, Ö. Axelsson, M.E. Nilsson, Effects of sounds from water on

perception of acoustic environments dominated by road-traffic noise, Acta Acust. United with Acust. 99 (2013) 218–225. doi:10.3813/AAA.918605.

[20] M. Lugten, M. Karacaoglu, K. White, J. Kang, K. Steemers, Improving the soundscape quality of urban areas exposed to aircraft noise by adding moving water and vegetation, J. Acoust. Soc. Am. 144 (2018) 2906–2917. doi:10.1121/1.5079310.

[21] R.Y. Litovsky, Spatial release from masking, Acoust. Today. (2012) 18–25. https://acousticstoday.org/wp-content/uploads/2017/09/Article_2of7_from_ATCODK_8_2.pdf.

[22] A.W. Bronkhorst, The cocktail-party problem revisited: early processing and selection of multi-talker speech, Attention, Perception, Psychophys. 77 (2015) 1465–1487. doi:10.3758/s13414-015-0882-9.

[23] A. Ihlefeld, B. Shinn-Cunningham, Spatial release from energetic and informational masking in a selective speech identification task, J. Acoust. Soc. Am. 123 (2008) 4369–4379. doi:10.1121/1.2904826.

[24] K. Allen, S. Carlile, D. Alais, Contributions of talker characteristics and spatial location to auditory streaming, J. Acoust. Soc. Am. 123 (2008) 1562–1570. doi:10.1121/1.2831774.

[25] M.T. Pastore, W.A. Yost, Spatial release from masking with a moving target, Front. Psychol. 8 (2017) 1–8. doi:10.3389/fpsyg.2017.02238.

[26] M.L. Hawley, R.Y. Litovsky, J.F. Culling, The benefit of binaural hearing in a cocktail party: effect of location and type of interferer, J. Acoust. Soc. Am. 115 (2004) 833–43. doi:10.1121/1.1639908.

[27] R.Y. Litovsky, M.J. Goupell, S. Godar, T. Grieco-Calub, G.L. Jones, S.N. Garadat, S. Agrawal, A. Kan, A. Todd, C. Hess, S. Misurelli, Studies on bilateral cochlear

implants at the University of Wisconsin’s binaural hearing and speech laboratory, J. Am. Acad. Audiol. 23 (2012) 476–494. doi:10.3766/jaaa.23.6.9.

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

[29] World Health Organization Regional Office for Europe, Environmental Noise Guidelines for the European Region, The Regional Office for Europe of the World Health Organization, Copenhagen, Denmark, 2018.

[30] J.M.B. Morillas, V.G. Escobar, J.A.M. Sierra, R. Vılchez-Gomez, A categorization method applied to the study of urban road traffic noise, J. Acoust. Soc. Am. 117 (2005) 2844–2852.

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

[32] N. Marrone, C.R. Mason, G. Kidd, Tuning in the spatial dimension: Evidence from a masked speech identification task, J. Acoust. Soc. Am. 124 (2008) 1146–1158. doi:10.1121/1.2945710.

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

[34] J.Y. Hong, J. He, B. Lam, R. Gupta, W.S. Gan, Spatial audio for soundscape design: recording and reproduction, Appl. Sci. 7 (2017) 1–21. doi:10.3390/app7060627.

[35] T. Lossius, J. Anderson, ATK Reaper: The Ambisonic Toolkit as JSFX plugins, in:

Proc. Jt. 40th Int. Comput. Music Conf. 11th Sound Music Comput. Conf., Athens, Greece, 2014: pp. 1338–1345.

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

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

[38] M.J. Blanca, R. Alarcón, J. Arnau, R. Bono, R. Bendayan, Non-normal data: Is ANOVA still a valid option?, Psicothema. 29 (2017) 552–557. doi:10.7334/psicothema2016.383.

[39] E. Schmider, M. Ziegler, E. Danay, L. Beyer, M. Bühner, Is it really robust?: reinvestigating the robustness of ANOVA against violations of the normal distribution assumption, Methodology. 6 (2010) 147–151. doi:10.1027/1614-2241/a000016.

[40] N.K. Srinivasan, K.M. Jakien, F.J. Gallun, Release from masking for small spatial separations: Effects of age and hearing loss, J. Acoust. Soc. Am. 140 (2016) EL73–EL78. doi:10.1121/1.4954386.

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

[42] J.F. Culling, Q. Summerfield, Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay, J. Acoust. Soc. Am. 98 (2005) 785–797. doi:10.1121/1.413571.

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

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

[45] J.Y. Hong, B. Lam, Z.T. Ong, R. Gupta, W.S. Gan, Suitability of natural sounds to enhance soundscape quality in urban residential areas, in: 24th Int. Congr. Sound Vib., London, UK, 2017: pp. 1–6.

[46] C. Guastavino, B.F.G. Katz, D.J. Polack, Jean-Dominique; Levitin, D. Dubois, Ecological validity of soundscape reproduction, Acta Acust. United with Acust. 91 (2005) 333–341.

[47] C. Guastavino, B.F.G. Katz, Perceptual evaluation of multi-dimensional spatial audio reproduction, J. Acoust. Soc. Am. 116 (2004) 1105–1115. doi:10.1121/1.1763973.

[48] E. Hendrickx, P. Stitt, J.C. Messonnier, J.M. Lyzwa, B.F.G. Katz, C. De Boisheraud, Improvement of externalization by listener and Source movement using a “binauralized” microphone array, J. Audio Eng. Soc. 65 (2017) 589–599. doi:10.17743/jaes.2017.0018.

[49] J.Y. Jeon, J.Y. Hong, C. Lavandier, J. Lafon, Ö. Axelsson, M. Hurtig, A cross-national comparison in assessment of urban park soundscapes in France, Korea, and Sweden through laboratory experiments, Appl. Acoust. 133 (2018) 107–117. doi:10.1016/j.apacoust.2017.12.016.

[50] W. Yang, J. Kang, Soundscape and Sound Preferences in Urban Squares: A Case Study in Sheffield, J. Urban Des. 10 (2005) 61–80. doi:10.1080/13574800500062395.

[51] V. Best, E.J. Ozmeral, B.G. Shinn-Cunningham, Visually-guided attention enhances target identification in a complex auditory scene, J. Assoc. Res. Otolaryngol. 8 (2007) 294–304. doi:10.1007/s10162-007-0073-z.

[52] K. Sun, B. De Coensel, G.M. Echevarria Sanchez, T. Van Renterghem, D.

Botteldooren, Effect of interaction between attention focusing capability and visual factors on road traffic noise annoyance, Appl. Acoust. 134 (2018) 16–24. doi:10.1016/j.apacoust.2018.01.001.