Physiological indicators and subjective restorativeness with audio-visual interactions in urban soundscapes

Zhongzhe LI[a], Meihui BA[a*], Jian KANG[b**]

[a]Pan Tianshou College of Architecture, Art and Design, Ningbo University, Ningbo, China

[b]UCL Institute for Environmental Design and Engineering, University College London (UCL), London, United Kingdom.

[*]Corresponding author: Meihui BA, bameihui@nbu.edu.cn, Pan Tianshou College of Architecture， Art and Design， Ningbo University, Ningbo, China.

[**]Corresponding author: Jian KANG, J.kang@ucl.ac.uk, UCL Institute for Environmental Design and Engineering, The Bartlett, University College London (UCL), London WC1H 0NN, United Kingdom. Tel: +44 (0)20 3108 7338.

Abstract: The present study aimed to identify the trends of changes in physiological indicators and subjective restorativeness in response to audio-visual interactions in the environment. Four scenarios types were presented using four different modalities (video-sound, image-sound, sound-only, and video-only; each modality was evaluated by independent groups of subjects). The physiological responses and subjective restoration of subjects were measured to assess the interactions between the audio-visual modalities. These data were also analysed to determine the physiological and subjective differences between dynamic and static visual presentations. We found that with visual modalities, the heart rate (HR), heart rate variability (HRV) calculated using the standard deviation of the NN intervals (SDNN-HRV), high-frequency band in the HRV power spectrum (HF-HRV), alpha reactivity on electroencephalography, and skin temperature (ST) decreased; however, the beta reactivity on EEG and skin conductance level (SCL) increased. With auditory modalities, the SDNN-HRV, HF-HRV, ST, and respiration depth decreased; however, the respiration rate and SCL increased. Use of static images and sound to reproduce the natural environment evoked more physiological comfort and subjective restorativeness. These findings could provide physiological insights for the theory of the restorative environment.

Keywords: Soundscape, multisensory, physiological indicators, subjective restorativeness

Abbreviations: ANOVA, analysis of variance; ANS, autonomic nervous system; ART, attention restoration theory; ECG, electrocardiography; EEG, electroencephalography; α-EEG, alpha reactivity on EEG; β-EEG, beta reactivity on EEG; HF-HRV, high-frequency band on the HRV power spectrum; HR, heart rate; HRV, heart rate variability; PNS, parasympathetic nervous system; PRSS, Perceived Restorativeness Soundscape Scale; ΔR, amplitude of the R wave; RD, respiration depth; RR, respiration rate; SCL, skin conductance level; SDNN-HRV, HRV calculated using the standard deviation of the NN intervals; SNS, sympathetic nervous system; SPL, sound pressure level; SRT, stress restoration theory; ST, skin temperature; VR, virtual reality

1 Introduction

With increasing urban density and the acceleration of the pace of life, city dwellers' negative emotions and physical stress in the city are also increasing (Laufs et al., 2020). There is growing evidence showing that natural elements such as green spaces or waterscapes in cities play a positive role in the physical and mental health of urban residents (Belmeziti et al., 2018; Hartig et al., 2014; Lu et al., 2021; Wu and Kim, 2021). Environmental quality in urban public open spaces has become a critical factor for improving urban sustainability. Soundscape research evaluates people's perception of the environment and the positive factors in sounds, thereby treating environmental sounds as a resource rather than a waste (noise) (Kang et al., 2016). Therefore, soundscape studies could provide more insights into urban sustainability from the aspect of sound management and planning (Hong et al., 2020).

1.1 Visual and auditory perception in a multisensory environment

A soundscape is defined as an acoustic environment perceived, experienced, or understood by a person or people in context (ISO 2014). However, human perception of the environment is not generated by a single, isolated sense; rather, it is an inherently multisensory experience involving visual, auditory, olfactory, and other sensory stimuli interacting and modifying the overall awareness (Spence et al., 2014; Schreuder et al., 2016; Ba et al., 2020). Therefore, soundscape research does not only focus on sound; instead, it focuses more attention on the 'context' behind the sound, i.e., how people experience sound in a multisensory environment. Multiple sensory organs detect various physical characteristics of the environment simultaneously. Then, they integrate this physical information with multimodal perception to form the bottom-up cognitive process (Kayser and Logothetis, 2007), which is also coordinated by human behaviour and emotions (top-down signals) (Talsma et al., 2010), thus forming a unique and subjective perceptual experience (Choi et al., 2018). Unlike the other senses, both vision and hearing can accurately identify space and orientation. Vision is usually more sensitive to, and more accurately interprets, spatial distribution (Fisher, 1968); furthermore, vision exerts a greater influence on the perception of an object (Godfroy-Cooper et al., 2015) relative to hearing. However, hearing provides more accurate temporal information than vision (Morein-Zamir et al., 2003). The perceptive-affective system is dominated by visual stimuli; a large number of studies have revealed that after evaluating the overall environment, the dominance of visual factors observed in the environment affect the perception of hearing (Yang et al., 2011; Tse et al., 2012). Additionally, the use of vegetation in visual design can mitigate the negative perception of environmental noise (Van Renterghem, 2019). However, some studies have shown that, in some cases, subjective preferences in the environment are dominated by auditory factors (Carles et al., 1992; Liu et al., 2013; Preis et al., 2015), and audition has a larger influence on mood states than visuals (Jiang et al., 2021). The role of auditory preference during landscape assessment is approximately 4.5-times that of visual preference (Gan et al., 2014). Hence, hearing plays a vital, though disproportionate, role in one's overall environmental preference. Many studies on multi-domain approaches have also been carried out in the field of indoor environmental quality, some of which showing that visual and auditory domains interact with each other in influencing human perception and behaviour (Torresin et al., 2018; Schweiker et al., 2020). Therefore, the effects of audio-visual interactions on physiological responses and subjective evaluation warrant investigation.

1.2 Research on audio-visual interaction in the field of soundscape

Most studies of the urban environment are focused on vision; relatively few have focused on hearing. Soundscape research is closely related to 'context,' and visual factors are an important part of the context. Therefore, by studying audio-visual interactions from the perspective of the soundscape, we can study the whole environment in a more auditory manner. Research of audio-visual interactions in the field of soundscapes includes both field and laboratory studies. Field studies consider soundscapes based on real places. Southworth (1969) was the first to use soundwalks in the field; by blocking part of the subjects' eyes or ears, the influences of vision and hearing on each other in the urban acoustic environment were analysed. Thereafter, soundwalks gradually matured as an experimental method (Yong Jeon et al., 2013; Liu et al., 2014). However, due to the limitations of geographical conditions, it is difficult to present the sound environment with different attributes during the same survey; furthermore, this method can downplay other factors in the natural environment, including olfactory and other sensory elements that can affect perception (Ba and Kang, 2019a). Hence, laboratory research involves more widespread use of methods reproducing the sound environment. Through a series of field investigations and laboratory experiments, Anderson et al. (1983) studied the influence of sound on preferences for outdoor environments and found an interaction between sound and vision. People are more sensitive to sound in natural environments and those covered with vegetation. They also compared the differences between field research and laboratory research. Laboratory research has since evolved into a developed research method for audio-visual research. To study audio-visual interactions in the laboratory, researchers now mainly focus on the subjective evaluation of factors that influence the perception of the environment. Although most subjective evaluation studies only consider single evaluation dimensions, such as comfort (Liu and Kang, 2018), pleasantness (Carles et al., 1999; Jiang and Kang, 2017), or annoyance (Chau et al., 2018), others have developed multidimensional adjectives to evaluate the overall environment through semantic differentiation (Jeon et al., 2012；Hong and Jeon, 2013) or one aspect of acoustic quality such as tranquillity (Pheasant et al., 2010) or perceived restorativeness of a soundscape (Payne, 2013).

1.3 Research on restoration theory

The restorative effects of the environment have been given increasing attention by scholars since the attention restoration theory (ART) (Kaplan and Kaplan, 1989) and stress restoration theory (SRT) (Ulrich et al., 1991) were proposed. The hypothesis in both these theories is that, compared with urban environments, natural environments can make it easier for people to recover from physical stress and mental fatigue (Hartig et al., 2003). However, research in this field is focused more on visual stimulation (Berto, 2014). These studies found that green space (Van den Berg, 2017) and blue space (White et al., 2010) have positive effects on restoration and health. However, during many experiments, there was either no sound or no control of the acoustic parameters, such as the sound pressure level (SPL). Additionally, a large number of studies have shown that the soundscape affects our health and quality of life (Shepherd et al., 2013; Von Lindern et al., 2016). Some of the results of these studies revealed that natural sounds have stronger restorative effects (Ratcliffe et al., 2013; Li and Kang, 2019). Some studies have shown that the natural environment can not only relieve stress (from the stress phase to the baseline), but also improve cognitive ability and positive emotions (beyond the baseline, even if there is no stress phase) (Korpela and Ratcliffe, 2021). Perceptions of restorative potential and outcomes without a prior stress intervention are called instorative.

1.4 Physiological research

The SRT suggests that perceiving the particular quality of a place can support recovery from physiological stress. Physiologically, people detect visual and auditory signals through the external environment, which then cause the response of the autonomic nervous system (ANS) and measurable changes in electroencephalography (EEG). Both the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) are part of the ANS. The SNS is activated in response to stress, and the PNS is responsible for maintaining balance (homeostasis) after stress. The existing research of the restorative environment has included many physiological indicators, such as the heart rate (HR) (Ulrich, 1981), heart rate variability (HRV) (Brown et al., 2013), blood pressure (Lee et al., 2009), skin conductance level (SCL) (Yin et al., 2018), and EEG results (Yang et al., 2011; Roe et al., 2013). Therefore, the studies have been gradually proven to be reliable for exploring environmental quality through physiological measurements.

Although previous studies have analysed the influences of visual and auditory stimuli on each other during subjective evaluations, few have considered the influence of physiological indicators on audio-visual interactions (Erfanian et al., 2019; Torresin et al., 2019). The current physiological research of soundscapes involves a stimulus-locked design and passive listening; these are used to collect data regarding human physiological indicators to study the changes in those indicators according to different types of sound and their restorative effects (Bradley et al., 2001; Alvarsson et al., 2010; Irwin et al., 2011; Hume & Ahtamad, 2013; Medvedev et al., 2015). Previous physiological studies have shown a weak correlation between physiological indicators and subjective restorative factors of soundscapes (Li and Kang, 2019). This shows that people's physiological responses and subjective evaluations may not be synchronised; therefore, the quality of the sound environment should not be assessed using only subjective evaluations. The study of the physiological response of the sound environment can better explore the relationship between the soundscape and both health and well-being. Hence, physiological indicators and subjective perception are two aspects of assessments of soundscapes, and the effects of audio-visual interactions on these two aspects should be compared.

1.5 Dynamic vision and static vision

Additionally, with the development of digital technology, recording and reproduction of the audio-visual environment have become gradually enhanced (Li and Lau, 2020). Because of the limitations of experimental conditions, early research usually attempted to reproduce the environment with static pictures (Anderson et al., 1983; Viollon et al., 2002). However, reproduction of the environment with video or a rendered virtual visual environment has become convenient (Hong and Jeon, 2013; Hong et al. 2019; Park et al., 2020). The differential effects of static images and dynamic videos when presenting the audio-visual environment, especially on physiology, have rarely been compared. Although dynamic vision can obviously present the environment more comprehensively and vividly than static images, in the actual audio-visual interactive landscape design, it is worth studying whether using dynamic vision or static vision to design environments can achieve the same effects. Hence, the optimal visual presentation method for creating a comfortable audio-visual environment also warrants study.

1.6 Research questions

Few researchers have considered the physiological effects of the influence of visual and auditory factors on the environment or the differences between dynamic videos and static images during audio-visual environment presentations. The present study measured physiological indicators and subjective restorativeness after the reproduction of common and representative audio-visual scenarios of urban public open space in the laboratory. We addressed the following main research questions: First, how do physiological indicators and subjective restorativeness change with different urban scenarios? Second, how does visual stimulation affect the physiological responses of hearing and vice versa? Lastly, is there any difference between the effects of visually dynamic and static reproductions of the audio-visual components of an environment on physiological indicators and subjective restorativeness? The answers to these questions could provide physiological references for the theory of ART and SRT and provide guidance for experimental methods, such as recording and reproducing the environment in the laboratory. The results of physiological responses can also help designers to design or modify the audio-visual interactions in the urban environment.

2 Material and methods

This study presented an environment by controlling the visual and auditory stimuli. As shown in Figure 1, the subjects were randomly divided into four groups (30 people each group): video-sound, image-sound, video-only, and sound-only. While presenting the scenarios, the physiological indicators of the participants were measured. The participants in the video-sound and image-sound groups were asked to evaluate the restorativeness of the environment after each of its segments was presented. The effects of vision and hearing on physiological indicators during the presentation of the scenarios were analysed by comparing the differences between the video-only, sound-only, and video-sound groups. The physiological and subjective differences between dynamic vision and static vision were analysed by comparing the differences between the video-sound and image-sound groups.

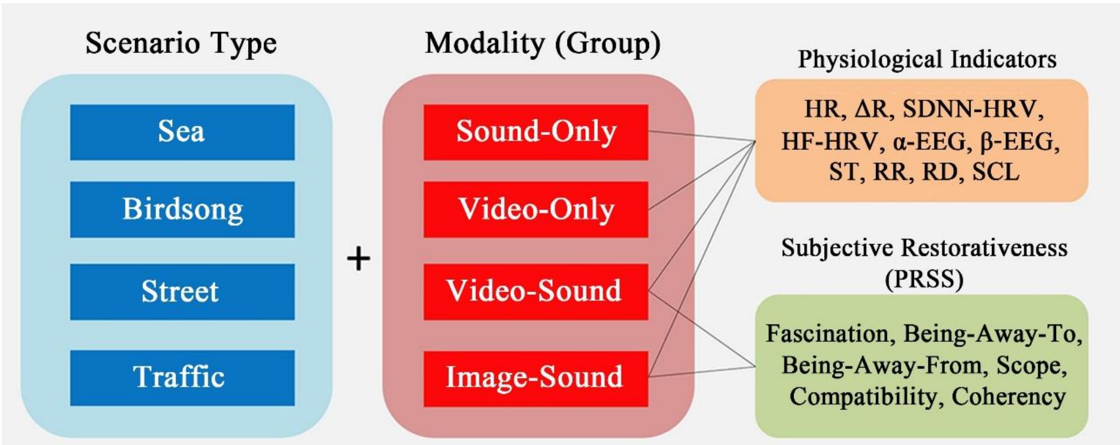


Figure. 1 Diagram of stimuli and modalities

2.1 Participants

For this experiment, we enrolled 66 male and 54 female participants who were unpaid undergraduate and graduate students (average age, 22.72 years; standard deviation [SD], 3.171; range: 18-35). The participants were mainly recruited through two experimental information platforms on campus. A few of them were recruited with the aid of a lecturer involved with undergraduate courses. A lecturer (not one of the authors of this article) publicised this experiment in an undergraduate course, and some students in the course volunteered to participate. The experiment adopted a balanced design. Participants were randomly divided into four groups of 30 participants. All reported that they had normal hearing and did not use any psychotropic drugs. They wore comfortable and loose clothing, did not engage in strenuous exercise during the 2 hours preceding the experiment, and had no obvious sense of fatigue. The participants were instructed not to consume alcohol nor caffeine 12 h before the experiments. We did not control the participants' smoking behaviour, mainly because for smokers, controlling addiction may bring physical discomfort. This study was approved by the Degree Committee of the School of Architecture, Harbin Institute of Technology (this governing body has an ethical review board).

2.2 Stimuli

The stimuli were the same as those used in previous studies (Li and Kang, 2019). We selected common and representative scenarios in urban public open spaces from the four common soundscape categories of biological, geophysical, human, and traffic (Axelsson et al., 2010; Hunter et al., 2010; Medvedev et al., 2015): birdsong (an uninhabited forest with birds at dawn); sea (unattended waves and beaches on a calm, sunny day); street (an outdoor shopping corridor with hurrying pedestrians and hawking vendors); and traffic (an intersection at rush hour on a sunny day). The audio-visual environment was recorded with a combination of video (Sony A7M3) and audio (Head Acoustics HMS IV, binaural recording), both the audio and visual recorded at the same height of 1.5 metres. A representative clip was excerpted from each scenario to populate the stimuli for the experiment. Each scenario was further edited into four types of 1-minute presentations: video-sound, image-sound, video-only, and sound-only. The images used during the experiment are shown in Figure 2. There was no image-only condition during this experiment, mainly because the comparisons between other conditions all had clear practical significance, and the image-only situation is relatively rare in the actual environment. Because of the differences in the recording environments, the background SPL differed. The equivalent SPL for each audio frequency was adjusted to 70 dB (the actual SPLs of traffic and the street and sea scenarios were rather high, so to unify the experimental conditions, the equivalent SPL for each scenario was adjusted) (Ba and Kang, 2019b). The auditory stimuli were reproduced with headphones (Head Acoustics BHS II) that matched the Dummy Head and could realistically play the recording (Hong et al., 2017). The video stimuli were presented on a television screen (Samsung H6400: 166.030 cm × 93.375 cm [75 inches]; resolution, 1920 × 1080 px).

(a) Sea (b) Birdsong

(c) Street (d) Traffic

Figure. 2 Static images used in the experiment

2.3 Physiological measurements

The physiological signals were collected with a BIOPAC MP160 system. The collected physiological indicators included electrocardiography (ECG) results of the positive pole of the left wrist, negative pole of the right wrist, and grounding of the right foot; EEG results of the F7, F8, T3, T4, T5, T6, C3, C4, and Cz leads based on the international 10-20 system; and results of electrodes attached at their proper positions to measure respiration waves (fixed to the chest through a respiration band), skin conductance (connected to the index and middle fingers), and body surface temperature (connected to the back of the hand). With the exception of a special respiration band used for testing the respiration waves and a temperature probe for testing the body surface temperature, all other physiological indicators were determined using standard (Ag/AgCl) electrode sheets connected to the skin. The collected physiological signals were analysed with AcqKnowledge 5.0. The 10 physiological indicators included the following: HR (reciprocal of the peak period value of the R wave), amplitude of the R wave (ΔR), HR variability (HRV) calculated using the SD of the NN intervals (SDNN-HRV), high-frequency band on the HRV power spectrum (HF-HRV), α-EEG (the main frequency of 8–13 Hz filtered after removing the ocular artefact of the EEG), β-EEG (the main frequency of 14–30 Hz filtered after removing the ocular artefact of the EEG), respiration rate (RR) calculated between the consecutive peaks of the respiration wave, respiration depth (RD) calculated by the amplitude of the respiration wave, SCL, and skin temperature (ST).

The aforementioned physiological indicators are all sensitive and related to stress and emotions (Kreibig, 2010). The cardiovascular response, respiration, and skin electrical signals are all controlled by the ANS. Visual and auditory signals affect the SNS and PNS, resulting in changes in related physiological indicators. The increase of SNS activity or the decrease in PNS activity will lead to increased HR, and vice versa. The ΔR reflects the amplitude of the ECG signal, which is related to blood pressure (Rautaharju and Blackburn, 1961) and can be used as an indicator of emotions (Bulagang et al., 2020). Studies have shown that people have lower ΔR when listening to arousing music and higher ΔR when listening to calm music (Dousty et al., 2011). Generally, SDNN-HRV is a time-domain parameter of HRV that can be used as an indicator of stress. Increased stress will lead to decreased SDNN-HRV (Thayer et al., 2012). HF-HRV is a frequency-domain parameter of HRV that is a measure of PNS activity. Increased stress will lead to decreased HF-HRV (Malliani et al., 1991; Kim et al., 2018). The body surface temperature is closely related to the SNS. The activation of the SNS can constrict the peripheral blood vessels, thereby decreasing the body surface temperature (Kistler et al., 1998). The RR and RD are also related to stress. Increased stress can increase the breathing frequency, increase the RR, and decrease the RD (Langewitz et al., 1987; Baekey et al., 2010). The SCL is a physiological indicator that reflects the activity of sweat glands. Because sweat gland activity is also controlled by the SNS, the SCL is used as an index of psychological and physiological arousal and to evaluate stress. Increased stress will increase the skin resistance (Dawson et al., 2016). The α-EEG and β-EEG are two bands that are closely related to emotion and stress. Increased α-EEG usually indicates relaxation (Klimesch, 1999; Fachner et al., 2013), and increased β-EEG is related to positive or anxious thinking and active attention (Fernández et al., 1995; Prinsloo et al., 2013). Therefore, we regard the decrease of HR, β-EEG, RR, ST, and SCL, as well as the increase of α-EEG and RD, as evidence that a certain scenario or modality can make people feel more relaxed and comfortable.

2.4 Subjective ratings

Subjective evaluation is the most common method of performing soundscape research (Aletta et al., 2016). This study used the Perceived Restorativeness Soundscape Scale (PRSS) as a subjective questionnaire to evaluate comprehensively and effectively the restorativeness of the presented soundscapes and obtain references for the physiological indicators (Payne, 2013). The questionnaire was divided into six restorative factors (see Appendix): fascination (the ability of a stimulus to have attention-holding properties); being-away-to (pull factors that shift attention away from the present situation to a different environment); being-away-from (push factors that shift attention away from the present situation to a different environment); compatibility (consistency between environment and individual feelings); coherency (the plausibility of the organisation of and connectedness between the elements in the environment and their structure); and scope. In the study by Purcell et al. (2001), the subscale 'scope' was added as a restorative factor of the environment, and it refers to the scale of the domain in which the perceptual and organisational activities are situated. Each question was measured using a 5-point scale to assess the response to 'how much do you agree with the statement?': not at all (1); a little (2); somewhat (3); a fair bit (4); and completely (5). The questions were edited with E-prime (Version 3.0.3.31, PSYCHOLOGY SOFTWARE TOOLS, INC) and presented to participants on a screen (Stahl, 2006). Because the PRSS evaluates the overall environment, including both visual and auditory factors (i.e., 'The sounds I am hearing seem to fit together quite naturally with this place'), participants are unable to understand these questions if they are in the video-only or sound-only group. Because PRSS is a questionnaire designed for soundscapes, all questions are used to assess the overall environment, so the questionnaire cannot be divided into a visual part and an auditory part. Therefore, only the participants in the video-sound and image-sound groups needed to answer the questionnaire. The overall reliability of the questionnaire was good (Cronbach's alpha=0.895), and the construct validity was very good (Kaiser-Meyer-Olkin test=0.926). The reliability and validity of each restorative factor dimension were approximately 0.8.

2.5 Procedure

The experiment was performed in a closed room (8.3 × 6.4 × 3.5 m, noise floor was 11 dB according to ISO3745 standard) with walls equipped with sound-absorbing materials. Except for the necessary experimental equipment, no objects were placed in the room to avoid distractions (Hermida Cadena et al., 2017; Lindquist et al., 2016). The participants were first asked to sit comfortably at a distance of 1.5 metres away from the screen. Then, the researcher explained the experimental process to the participants and asked them about their physical and mental states. After the participants fully understood the experimental process and completed the consent form, they were connected to the physiological instruments and asked to wear the headphones which can present realistic-sounding binaural recordings. Then, the researcher left the audiometric room and went into the observation room. During the whole process of the experiment (including the stage of relaxation), the participants were asked not to move their bodies as much as possible (to prevent poor contact with the electrodes), not to close their eyes for a long time, and to look ahead in a comfortable way when the scenarios were presented. Participants were asked to imagine themselves in the presented environment and fully experience the scenarios presented. During the first 10 minutes of the experiment, the participants were asked to relax completely. At that time, the baseline values of the physiological indicators were recorded. Baseline values were used to measure the physiological relaxed state of the subjects without any stimulation. There was no process of subjecting participants to stress because this experiment did not focus on the process of the recovery from stress; instead, it focused on the physiological responses to a certain environment. In real environments, not everyone is under stress; therefore, we assume that people enter the scenarios at their baseline state, and that these scenarios affect their physical and mental states. Participants were asked to look at the screen (even in the sound-only group, who only saw a screen displaying the focus symbol '+') and listen to the accompanying audio (with the exception of the visual-only group) on which four scenarios were randomly presented. Each scenario was shown for 1 minute. According to a previous study, all physiological indicators responsive to stimuli appear within 1 minute, there was a greater correlation between the 1-minute physiological response data and subjective restorativeness evaluation results, and the participants had the same physiological responses trend in similar scenarios (Li and Kang, 2019). Therefore, to better observe the physiological indicators, 1 minute should be reasonable. After the presentation, the participants were asked to complete the questionnaire concerning the scenario. Then, the participants were asked to relax for 90 seconds before the next scenario was presented. According to the results of previous experiments, there was no significant difference between the physiological indicators of 90 seconds relaxation after stimuli and those in baseline state. To avoid familiarity, each scenario was presented only once. After the four scenarios were presented, the researcher re-entered the audiometric room, retrieved the headphones and electrodes from the subjects, and ended the experiment. The whole process of the experiment lasted approximately 50 minutes, during which the actual test time was usually less than 30 minutes.

2.6 Data analysis

Because of the great differences between individuals, all physiological signals required normalisation. Each physiological datum was converted to the percentage of its difference from the baseline value using the following formula:

$$\text{Percentage change (\%)} = (\text{raw value} - \text{baseline value}) / \text{baseline value}) \times 100.$$

The baseline values were the benchmark values of the physiological indicators obtained from the subjects during the absence of stimulation. The aforementioned formula was used to normalise the data of all participants by referencing the baseline values, thus rendering the data between the participants comparable.

IBM SPSS 25.0 was used to populate a database with all of the results. Analyses of variance (ANOVAs), including the physiological data collected from the video-only, sound-only, and video-sound groups were performed to analyse the influences of visual and auditory factors on physiological indicators. Variance analyses, including the physiological data and subjective data collected from the video-sound and image-sound groups were performed to evaluate the physiological and subjective differences between dynamic vision and static vision. All the variables were tested for normal distribution (Shapiro-Wilk method). Other physiological indicators and subjective factors except for RD and SCL satisfied the normal distribution. The RD and SCL data were transformed by the square root, and the results of the ANOVA after the transformation were consistent with those of the unconverted data. Because the variance model can tolerate data that are non-normally distributed with only a small effect on the type I error rate, to better understand and discuss the results, the results of the unconverted data were retained for this experiment. The results of the power analysis showed that the samples of all ANOVA models were sufficient (power > 0.8). The threshold for statistical significance was set at $P < 0.05$.

3 Results

3.1 Differences in physiological indicators and subjective restorativeness with different scenarios

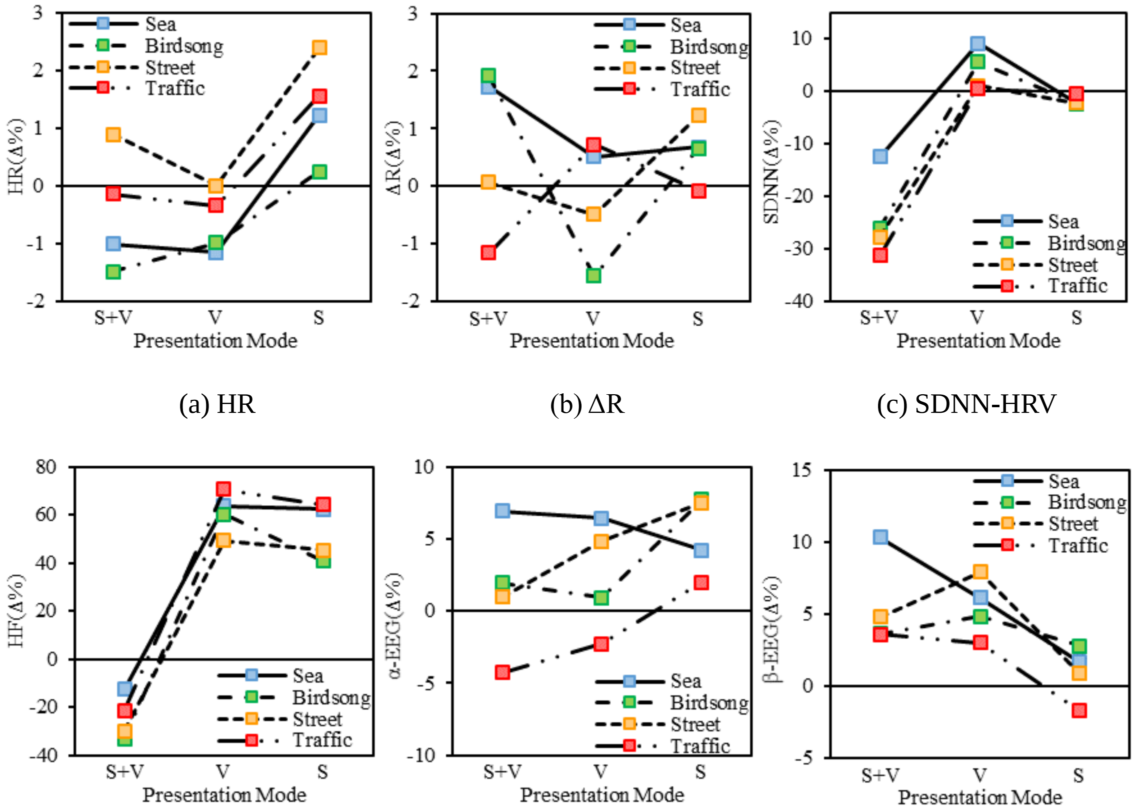
3.1.1 Effects of the scenario type and presentation mode on physiological indicators

The data of the physiological indicators of the participants in the video-sound group were compared with those of the participants in the sound-only and video-only groups. The effects included in the ANOVA model were the presentation mode (sound-video, sound-only, and video-only), scenario type (birdsong, sea, street, and traffic), and their interactions (presentation × scenario). The results are shown in Table 1. The estimated marginal mean values for the cross-terms of the factors in each model were calculated (Figure 3).

Table 1 ANOVAs of the physiological indicators according to presentation mode (video-sound, video-only, and sound-only) and scenario type

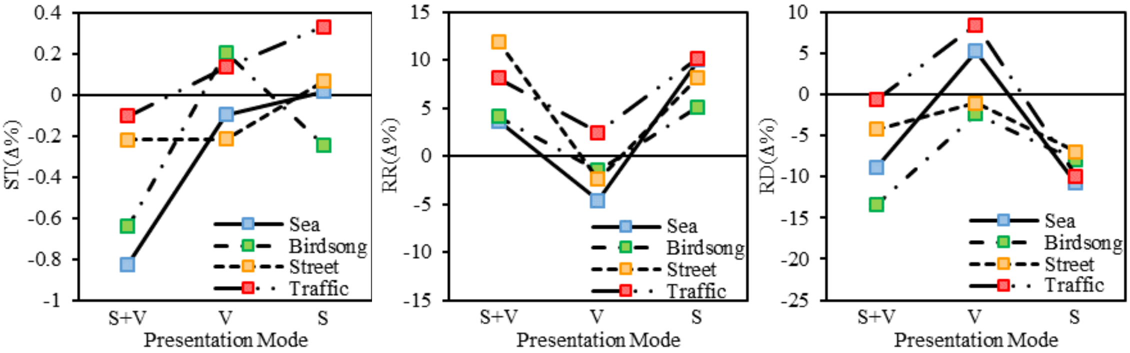
	Effect	Type III Sum of Square	df	Mean Square	F	Sig.
HR	Presentation mode	295.971	2	147.985	12.999	**<.001**
	Scenario type	181.468	3	60.489	5.313	**.001**
	Presentation × Scenario	22.081	6	3.680	0.323	.925
ΔR	Presentation mode	55.369	2	27.684	1.354	.260
	Scenario type	60.274	3	20.091	0.982	.401
	Presentation × Scenario	259.348	6	43.225	2.113	.051
SDNN-HRV	Presentation mode	57235.955	2	28617.977	33.403	**<.001**
	Scenario type	4178.370	3	1392.790	1.626	.183
	Presentation × Scenario	3885.865	6	647.644	0.756	.605
HF-HRV	Presentation mode	564224.408	2	282112.204	66.532	**<.001**
	Scenario type	22938.909	3	7646.303	1.803	.146
	Presentation × Scenario	5175.585	6	862.597	0.203	.976

α-EEG	Presentation mode	1059.111	2	529.556	2.635	.073
	Scenario type	2885.537	3	961.846	4.787	**.003**
	Presentation × Scenario	1263.448	6	210.575	1.048	.394
β-EEG	Presentation mode	1784.344	2	892.172	8.512	**<.001**
	Scenario type	946.448	3	315.483	3.010	**.030**
	Presentation × Scenario	759.152	6	126.525	1.207	.302
ST	Presentation mode	19.253	2	9.627	6.312	**.002**
	Scenario type	9.595	3	3.198	2.097	.100
	Presentation × Scenario	10.273	6	1.712	1.123	.348
RR	Presentation mode	7011.902	2	3505.951	16.413	**<.001**
	Scenario type	1288.745	3	429.582	2.011	.112
	Presentation × Scenario	1471.784	6	245.297	1.148	.334
RD	Presentation mode	9223.919	2	4611.959	6.191	**.002**
	Scenario type	2503.855	3	834.618	1.120	.341
	Presentation × Scenario	3242.874	6	540.479	0.725	.629
SCL	Presentation mode	36837.900	2	18418.950	5.229	**.006**
	Scenario type	6230.104	3	2076.701	0.590	.622
	Presentation × Scenario	41490.966	6	6915.161	1.963	.070

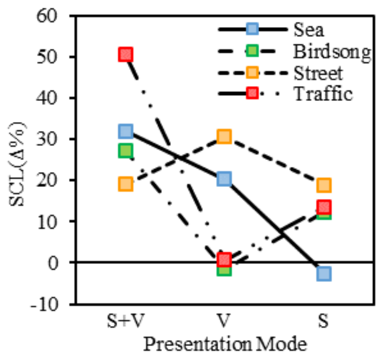


(a) HR (b) ΔR (c) SDNN-HRV

(d) HF-HRV (e) α-EEG (f) β-EEG



(g) ST (h) RR (i) RD



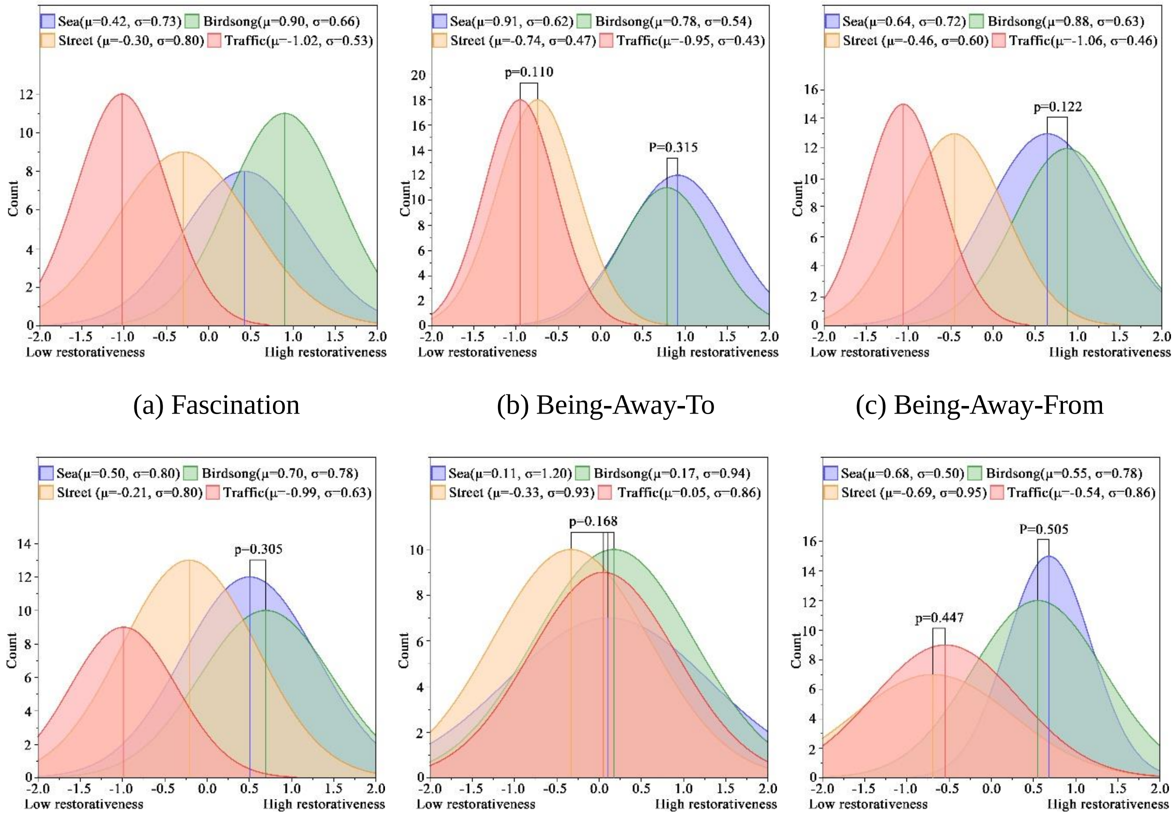
(j) SCL

Figure. 3 Marginal mean values of physiological indicators according to presentation mode (video-sound, sound-only, and video-only) and scenario type

Table 1 shows that the interaction between the presentation mode and scenario type had no significant effect on any of the physiological indicators, demonstrating that the changing trend of the presentation mode was unaffected by the scenario type and vice versa. The presentation mode significantly affected all physiological indicators except for the ΔR and α-EEG. However, the scenario type had a significant effect only on the HR, α-EEG, and β-EEG.

3.1.2 Differences in subjective restorativeness with different scenarios

By analysing the subjective restorativeness distribution of the four scenarios, we could assess the differences in the restorative factors of the four scenarios and obtain a reference for the physiological data. Of the four presentation modes, the video-sound group was the most verisimilar reproduction of the scenarios. The subjective questionnaire completed by the video-sound group was processed using the principal component analysis with the six principal components (as restorative factors) featuring a mean value of 0 and an SD of 1. Figure 4 shows the distribution of each factor according to the four scenarios. The subsets with no significant differences were determined using the Student-Newman-Keuls (S-N-K) test.



(a) Fascination (b) Being-Away-To (c) Being-Away-From

(d) Compatibility (e) Coherency (f) Scope

Figure. 4 Distribution of the restorative factors according to the four scenarios

Figure 4 shows that the scenario types can be easily distinguished by the factor of fascination and that the four scenarios belong to different subsets (Figure 4.a). The being-away-to and scope factors classified scenarios as natural environment (birdsong and sea) and noisy environment (street and traffic) (Figures 4.b and 4.f). The being-away-from and compatibility factors divided scenarios into three categories of natural (birdsong and sea), traffic, and street (Figure 4.c and 4.d). However, the coherency factor failed to distinguish the scenarios with sufficient clarity (Figure 4.e). Furthermore, the SD for coherency of the four scenarios was large, indicating the lack of consistency in the subjective evaluation of this dimension. This may have been because the elements in the scenarios we chose were relatively singular, the visual and auditory elements fully belonged to the environment, and the sound and vision combined naturally. Therefore, each participant's evaluation of the four scenarios in this dimension was very similar. However, because of the different evaluation criteria used by each participant, the overall data were relatively scattered. Of the four scenarios, the birdsong and sea had high restorative potential; furthermore, except for scope, the restorative factors for birdsong were higher than those for sea. The factors of fascination, being-away-from, and compatibility differed significantly between street and traffic. Traffic values were generally lower than street values, indicating that the subjective restorative potential of traffic noise is lower than that of human noise.

3.2 Physiological influence of visual factors on auditory perception

Because the ANOVA model can only test whether one factor has an impact on the dependent variable and cannot specify the changing trend of the dependent variable at different levels, the differences between the video-sound and sound-only groups were compared using Dunnett's post-hoc test to analyse the effects of visual factors on physiological indicators (Table 2). Dunnett's test only compared one group with other groups at one time, thus solving the problem of paired comparisons of multiple control groups with a single control group (Dunnett, 1955). The mean values of the physiological indicators of each group in Table 2 are shown in Figure 3.

Table 2 Differences in physiological indicators between presentation mode

Indicator	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
				Lower Bound	Upper Bound
Differences between the video-sound and sound-only groups[a]					
HR	**1.77***	0.423	<.001	0.832	2.710
ΔR	-0.047	0.572	.995	-1.318	1.224
SDNN-HRV	**22.828***	3.645	<.001	14.728	30.927
HF-HRV	**77.404***	8.110	<.001	59.370	95.439
α-EEG	**4.001***	1.769	.045	0.070	7.932
β-EEG	**-4.669***	1.290	.001	-7.534	-1.803
ST	**0.489***	0.153	.003	0.148	0.829
RR	1.427	1.820	.651	-2.618	5.472
RD	-2.109	3.386	.760	-9.633	5.416
SCL	**-21.684***	7.390	.007	-38.105	-5.262
Differences between the video-sound and video-only groups[b]					
HR	-0.204	0.429	.851	-1.157	0.749
ΔR	-0.867	0.582	.236	-2.158	0.425
SDNN-HRV	**28.680***	3.713	<.001	20.428	36.932
HF-HRV	**85.257***	8.345	<.001	66.699	103.814
α-EEG	1.037	1.795	.789	-2.953	5.027
β-EEG	-0.136	1.306	.992	-3.037	2.765
ST	**0.454***	0.156	.007	0.108	0.800
RR	**-8.461***	1.847	<.001	-12.565	-4.356
RD	**9.394***	3.443	.013	1.743	17.045
SCL	**-19.774***	7.513	.017	-36.469	-3.079

Note: a Mean Difference = video-sound group subtracted from sound-only group, * p<.05

b Mean Difference = video-sound group subtracted from video-only group, * p<.05

As shown in Table 2, there were significant differences in all physiological indicators except ΔR, RR, and RD for the video-sound and sound-only groups. This finding indicates that visual factors had an impact on auditory perception. With the addition of visual stimuli, the HR, SDNN-HRV, HF-HRV, α-EEG, and ST decreased, and the β-EEG and SCL increased. The observation that EEG signals are affected by visual factors is consistent with that of Yang et al. (2011). Compared with sound-only stimulation, visual factors strengthened the participants' emotions, decreased their HR and HRV, and increased their SCL. These indicators were associated with the feeling of stress (Kreibig, 2010), suggesting that the participants were more likely to have stress in a visual-auditory environment than in a pure auditory environment. This may be because the audio-visual environment could create more intense feelings. Although the participants in the sound-only group and the visual-only group were not deprived of hearing or vision, in the video-sound group, the participants were forced to receive more information; these audio-visual stimuli may have forced them to experience more, thus affecting their physiological indicators. The RR and RD were not significantly different between conditions, indicating that the addition of vision did not affect the regulation of respiration.

3.3 Physiological influence of auditory factors on visual perception

As mentioned in Section 3.2, Dunnett's post-hoc test was performed to analyse the impact of auditory factors on visual perception (Table 2). The mean values of the physiological indicators for each group are shown in Figure 3.

As shown in Table 2, the addition of sounds significantly affected most physiological indicators, including SDNN-HRV, HF-HRV, ST, RR, RD, and SCL. Similar to the effects of visual factors on auditory perception, the addition of sound decreased SDNN-HRV, HF-HRV, ST, and RD and increased RR and SCL. In contrast with the effects of vision, auditory factors had no significant effect on EEG; however, auditory factors had a significant effect on respiration. With auditory modalities, the RR increased and the RD decreased (i.e., it made people short of breath).

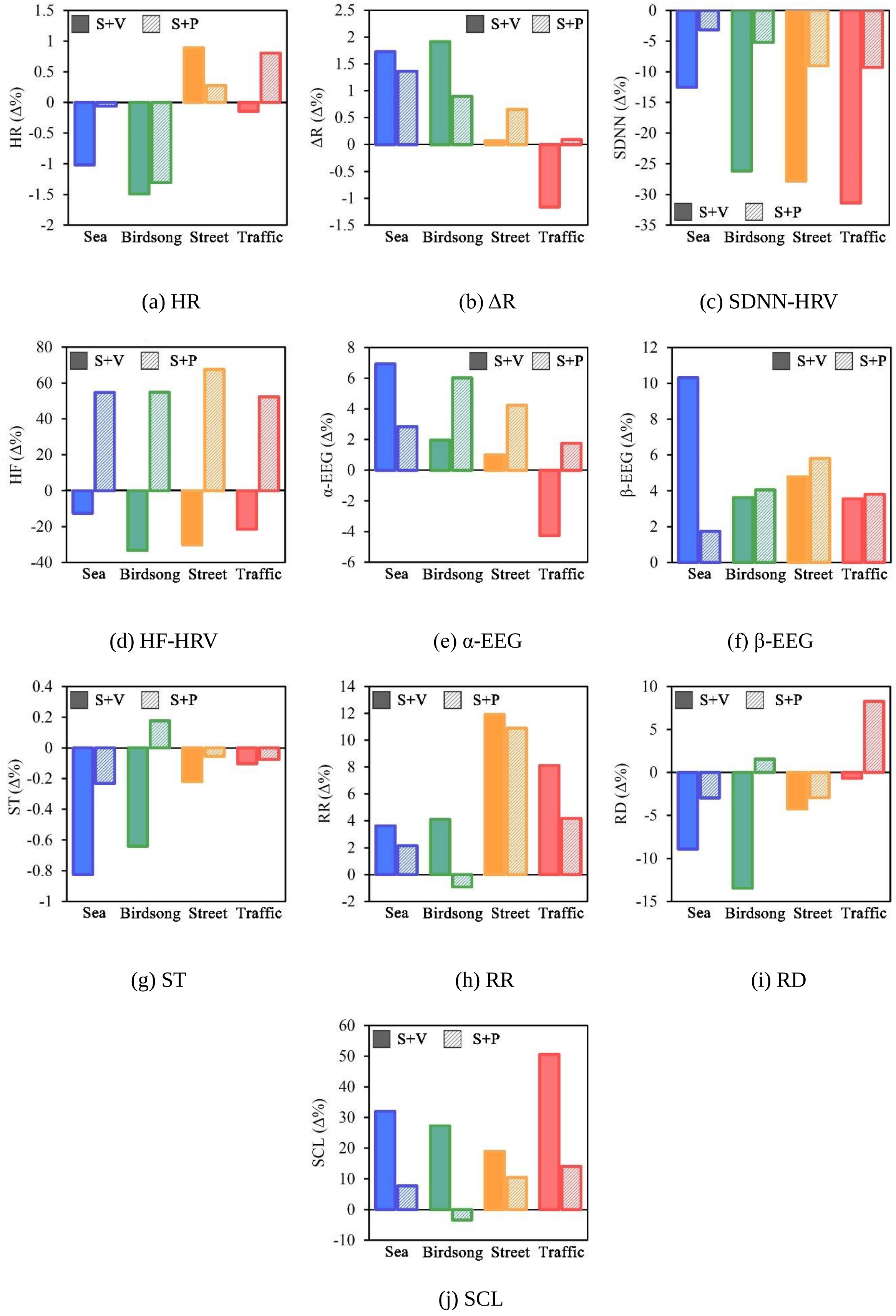
3.4 Differences between dynamic vision and static vision

3.4.1 Differences in physiological indicators

For the ANOVAs of the effects of the presentation mode, scenario type, and their interaction with physiological indicators, two presentation modes (video-sound, image-sound) and the four scenario types (birdsong, sea, street, and traffic) were used as dependent variables, and the 10 physiological indicators were used as independent variables (Table 3). The marginal mean values of the factors were calculated for each model (Figure 5).

Table 3 ANOVAs of physiological indicators and restorative factors by presentation mode (video-sound and image-sound) and scenario type

Source	Effect	Type III Sum of Square	df	Mean Square	F	Sig.
Physiological indicators						
HR	Presentation Mode	8.422	1	8.422	0.570	.451
	Scenario Type	151.380	3	50.460	3.416	**.018**
	Presentation × Scenario	26.194	3	8.731	0.591	.621
ΔR	Presentation Mode	0.773	1	0.773	0.027	.869
	Scenario Type	167.833	3	55.944	1.985	.117
	Presentation × Scenario	45.375	3	15.125	0.537	.658
SDNN-HRV	Presentation Mode	19636.648	1	19636.648	19.493	**<.001**
	Scenario Type	5502.456	3	1834.152	1.821	.144
	Presentation × Scenario	1529.720	3	509.907	0.506	.678
HF-HRV	Presentation Mode	411639.840	1	411639.840	89.381	**<.001**
	Scenario Type	3601.585	3	1200.528	0.261	.854
	Presentation × Scenario	8764.570	3	2921.523	0.634	.594
α-EEG	Presentation Mode	301.656	1	301.656	1.238	.267
	Scenario Type	1247.416	3	415.805	1.706	.167
	Presentation × Scenario	836.686	3	278.895	1.144	.332
β-EEG	Presentation Mode	162.784	1	162.784	1.221	.270
	Scenario Type	217.063	3	72.354	0.543	.654
	Presentation × Scenario	880.502	3	293.501	2.201	.089
ST	Presentation Mode	9.305	1	9.305	4.265	**.040**
	Scenario Type	6.700	3	2.233	1.024	.383
	Presentation × Scenario	6.851	3	2.284	1.047	.373
RR	Presentation Mode	504.620	1	504.620	2.294	.131
	Scenario Type	3546.771	3	1182.257	5.375	**.001**
	Presentation × Scenario	170.940	3	56.980	0.259	.855
RD	Presentation Mode	3821.698	1	3821.698	3.944	**.048**
	Scenario Type	4019.367	3	1339.789	1.383	.249
	Presentation × Scenario	1548.112	3	516.037	0.532	.660
SCL	Presentation Mode	38920.721	1	38920.721	11.277	**.001**
	Scenario Type	15269.266	3	5089.755	1.475	.222
	Presentation × Scenario	6870.223	3	2290.074	0.664	.575
Restorative factors						
Fascination	Presentation Mode	0.243	1	0.243	0.620	.432
	Scenario Type	152.850	3	50.950	130.293	**<.001**
	Presentation × Scenario	4.027	3	1.342	3.433	**.018**
Being-Away-To	Presentation Mode	5.765	1	5.765	20.034	**<.001**
	Scenario Type	172.118	3	57.373	199.379	**<.001**
	Presentation × Scenario	1.155	3	0.385	1.338	.263
Being-Away-From	Presentation Mode	0.054	1	0.054	0.165	.685
	Scenario Type	170.482	3	56.827	173.482	**<.001**
	Presentation × Scenario	0.921	3	0.307	0.938	.423
Compatibility	Presentation Mode	0.353	1	0.353	0.731	.393
	Scenario Type	131.660	3	43.887	90.874	**<.001**
	Presentation × Scenario	2.164	3	0.721	1.494	.217
Coherence	Presentation Mode	0.004	1	0.004	0.004	.948
	Scenario Type	18.121	3	6.040	6.376	**<.001**
	Presentation × Scenario	2.136	3	0.712	0.751	.522
Scope	Presentation Mode	0.085	1	0.085	0.129	.720
	Scenario Type	89.160	3	29.720	45.352	**<.001**
	Presentation × Scenario	1.124	3	0.375	0.572	.634



(a) HR (b) ΔR (c) SDNN-HRV

(d) HF-HRV (e) α-EEG (f) β-EEG

(g) ST (h) RR (i) RD

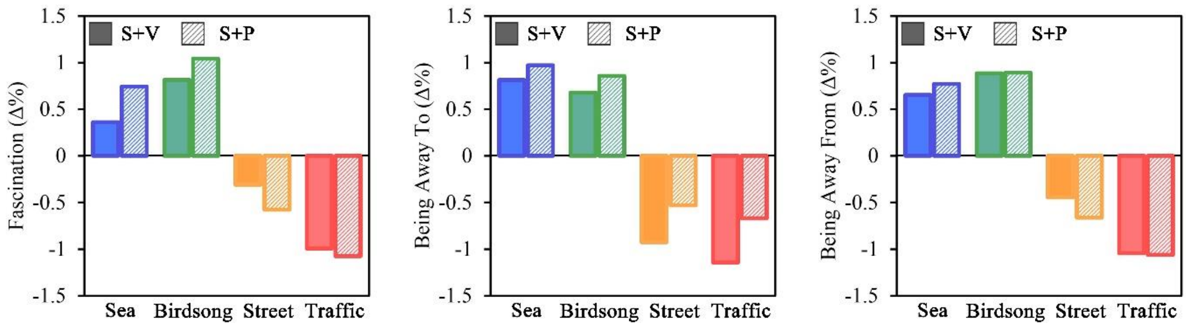
(j) SCL

Figure. 5 Marginal mean values of the physiological indicators under the different presentation modes (video-sound [S+V] and image-sound [S+P]) and scenario types

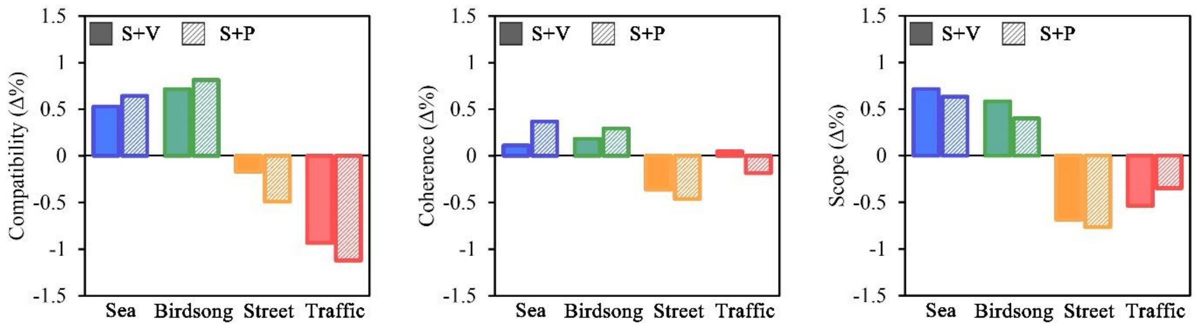
Table 3 shows that the interaction between presentation mode and scenario type was not significant for any of the physiological indicators, indicating that the influence of trends of the scenario type on physiological indicators remains the same regardless of whether the visual stimuli are presented as a video or image. The effects of the presentation mode on SDNN-HRV, HF-HRV, ST, RD, and SCL were significant ($p<0.05$). The marginal means of the SDNN-HRV, HF-HRV, ST, and RD in dynamic vision were lower, and that of the SCL were higher, than those in static vision (Figure 5). The HRV and SCL were both very sensitive indicators, indicating that there were differences in responses to dynamic vision and static vision. Because dynamic vision increased the SCL while decreasing indicators related to vagus nerve activity, such as SDNN-HRV, it might generate more intense stimulation for the subjects (i.e., the subjects were more relaxed when the environment was presented as a static image). Moreover, although there was no statistically significant difference in the RR values for dynamic stimuli and static stimuli, the following trends were revealed: the subjects in the video-sound group had higher RR and lower RD, indicating that the dynamic visual environment might aggravate shortness of breath (Figures 5.h and 5.i). Notably, the differences between α-EEG and β-EEG were nonsignificant, indicating that dynamic vision and static vision have no effect on EEG.

3.4.2 Differences in subjective restorativeness

Similar to the physiological data, a variance analysis was performed for the six restorative factors to assess the effects of the presentation mode, scenario type, and their interaction with subjective restorativeness. The subjective factors were obtained using a principal component analysis of the two groups of data together to compare the differences between the two groups more conveniently (Table 3). The marginal mean values of the factors were calculated for each model (Figure 6).



(a) Fascination (b) Being-Away-To (c) Being-Away-From



(d) Compatibility (e) Coherence (f) Scope

Figure. 6 Marginal mean values of the restorative factors under different presentation modes (video-sound [S+V] and image-sound [S+P]) and scenario types

Table 3 shows that the scenario type had a significant effect on all subjective restorative factors. The presentation mode only affected the being-away-to factor, and the being-away-to factor of dynamic vision was lower than that of static vision (Figure 6.b). The interaction between scenario type and presentation mode had a significant impact on fascination. Regardless of the manner in which the environment was presented, the fascination factors of the birdsong and sea conditions remained positive, whereas those of the street and traffic conditions were negative. Under natural sound conditions (birdsong and sea), the fascination factor was lower in the video-sound group than in the image-sound group. Under the street and traffic conditions, the fascination factor was greater in the video-sound group. In other words, static vision had a greater effect on fascination: static images could increase the fascination of scenarios of high interest and decrease that of scenarios with low appeal. Although the presentation mode had little effect on subjective restorativeness, the subjective evaluation revealed that static vision was slightly more restorative. This is consistent with the results of the physiological data presented in Section 3.4.1, which showed that the natural environment presented as image-sound could physically relax participants to a greater extent and was subjectively more restorative than environments presented as video-sound. The relative lack of visual stimulation provided by static images compared with that provided by video could allow for greater stimulation with sound. When the participants stared at static images, the stimulation of dynamic sounds became more potent, thereby enhancing the factors of fascination and being-away-to in natural scenarios.

4 Discussion

4.1 Differences between physiological indicators and subjective restorativeness

Because the scenarios we chose were representative of specific situations, the differences in subjective restorativeness were significant. Natural sound and noise were polarised by each restorative factor. This was obviously consistent with the results of previous studies of restorative theory that indicated that the natural environment can be more restorative than urban traffic and pedestrian street scenarios (Berto, 2005; Ratcliffe et al., 2013). In contrast, it was difficult to distinguish different scenarios from the physiological data. Many of the changes in physiological data were caused by individual differences. Despite our use of a normalisation method to minimise the differences between subjects, these differences (mainly sensitivity to the environment and the floating range of different scenarios) still exerted an important influence on physiological indicators and were the main cause of error in the ANOVAs. However, the changes in most physiological indicators were relatively stable because the four scenarios lacked any extreme stimuli (there were no sudden shocks or any factors capable of eliciting strong emotional reactions). This shows that people can subjectively feel significant differences in different scenarios, but these differences are difficult to reflect using physiological indicators. However, although the physiological indicators can reflect the change in the presentation mode more sensitively, the difference in the presentation mode was not obvious in the subjective data. A similar conclusion has been reached by a previous study in which changes in physiological indicators more reflected the spectrum of sound (Li et al., 2020). Therefore, physiological indicators may be more sensitive to pure physical stimuli of hearing and vision, whereas subjective evaluations can better reflect the meaning behind the environment.

4.2 Effects of the scenario type on physiological indicators using different presentation modes

Concerning the ANOVA results presented in Section 3.1.1, the differences in HR, α-EEG, and β-EEG in different scenarios were significant. However, only the HR data followed a trend. The HR in response to a natural environment was lower than that in response to noise (Figure 3.a). The overlaps between the line charts generated by the other physiological data were very complex, demonstrating the difficulty classifying the types of scenarios through physiological indicators as well as the significant differences between the physiological responses to a single sensory modality and that of audio-visual interactions. Notably, in the video-sound group, physiological indicators followed the law of natural sound and noise more closely. In other words, in the audio-visual modality, it is more likely to observe that the natural scenarios are more relaxing than the noisy scenarios. In other modalities, the differences between natural scenarios and noise scenarios are relatively small. Furthermore, the physiological indicators obtained during the presentation of natural scenarios to the video-sound group indicated that the participants were more relaxed; conversely, those obtained during the presentation of traffic and street scenarios indicated greater tension. Specifically, natural scenarios decreased the RR and increased α-EEG. Consistent with our previous research (Li and Kang, 2019), these findings suggest that when the environment is presented as an audio-visual interaction, the reaction of the participants is closer to the one evoked by the emulated situation, and their physiological feelings are more consistent with their subjective perceptions. Sound-only and video-only scenarios prompt physiological feelings to change, and subjectively comfortable scenarios are not necessarily comfortable when presented in a single environment. For example, in the video-only group, the physiological data of birdsong were close to those of traffic for many indicators or those of the tensest state of the four scenarios. Similar results were reported by Annerstedt et al. (2013); in a very quiet forest environment, the physiological indicators suggested that people may be more likely to feel nervous and stressed.

According to the ANOVA results presented in Section 3.4.1, the scenario type only affected HR and RR significantly. This greater impact of the presentation mode on physiological indicators may explain this finding. Figure 5 shows that HR and RR were significantly lower in response to natural sounds than in response to street and traffic noise (Figure 5.a and 5.h). Furthermore, the ΔR and SDNN-HRV were lower in response to natural sounds than in response to noise.

In the experiment, the order of the four scenarios appeared randomly, so the order factor is consciously eliminated. The results of previous experiments showed that some physiological indicators which are closely related to the SNS will be affected by the sequence of the scenario. For example, SCL will gradually decrease with the increase of scenario order, which is mainly because the participants have a sense of freshness when perceiving the first scenario, and the responses will tend to be weaker as the order increases gradually. In addition, in the experiment, it was not found that some specific sequential combinations will produce special physiological responses, because there were sufficient intervals between scenarios.

4.3 Differences in the effects of auditory and visual factors on physiological indicators

Our results of the participants' subjective evaluation differ from those of previous studies that found that overall environmental comfort is more affected by auditory stimuli than by visual stimuli (Liu et al., 2013; Gan et al., 2014). It is difficult to determine from our results whether vision or hearing has a greater impact on physiological indicators; vision affected most physiological indicators except ΔR, RR, and RD, indicating that the visual effect does not involve the regulation of respiration. The effects of auditory factors on vision did not involve HR, ΔR, α-EEG, or β-EEG, indicating that the effects of hearing on HR and EEG were nonsignificant. Auditory factors had a greater influence on the indicators related to HRV, whereas visual factors had a slightly greater impact on SCL; however, the differences between these effects were minimal. In other words, the differences between the video-only and sound-only groups were relatively small, and the differences between either group and the video-sound group were relatively larger. This finding suggested that the physiological effects of the sound-only conditions and video-only conditions were similar. Therefore, although vision and hearing affect different physiological changes, their effects were not significantly different.

4.4 Recommendations

The results of this study indicate that physiology and subjectivity are not completely synchronised. Therefore, we suggest that when performing an overall evaluation of the urban soundscape or urban landscape, measurements of physiological indicators can be introduced to supplement the subjective evaluation.

This study verified the instoration theory under the condition of audio-visual environment, which can be reflected by most physiological indicators such as HR; in other words, the natural environment with audio-visual modalities can make people feel more relaxed physiologically than the urban environment. It should be noted that there was no stress phase in this experiment. Therefore, the experiment can only prove that the natural scenarios under audio-visual modalities can physiologically make people have a positive effect beyond the baseline. However, with incomplete vision or hearing, the natural environment may lose its advantage. In some scenarios, using a single sense may cause stress and tension, such as hearing waves without vision or watching the forest without any sound. These phenomena indicate that it is worth considering simultaneous multisensory factors in urban design. In actual urban design, sometimes an inadvertent design could mask part of the visual or sound (for example, some low-frequency noise could mask some sounds that match the landscape), which will lead to the lack of awareness of the overall environment, thereby affecting the physiological response.

Although the differences between dynamic vision and static vision show that it is difficult to subjectively reflect the differences between static scenarios and dynamic scenarios, some physiological indicators can be keenly detected. Static vision can make people feel more physically relaxed. This may because in the static visual modality, people receive relatively fewer visual stimuli so there is less information to cognitively process. On the other hand, people may anticipate what may suddenly appear in dynamic scenarios, so they are more anxious. Therefore, in urban landscape design, the introduction of more static visual elements and control of the more dynamic visual elements may make people feel more physiologically comfortable. Of course, the specific reasons behind this phenomenon are worthy of further verification and exploration.

4.5 Limitations and future research

During this research, a stimulus-locked design was used for the experiment. Participants were asked to take the initiative to experience the presented environment. However, in the real environment, people are more likely to perform other tasks than to fully experience where they are. In that case, vision and hearing may be affected by salience, which is an attention mechanism used to passively or actively filter out elements that stand out from the environment (Hoffman and Singh, 1997; Huang and Elhilali, 2017). Therefore, people will selectively ignore or focus attention on some visual or auditory elements, which is called the ‘cocktail party effect’ in hearing (Hawley et al., 2004). It is worth noting that some research evidence has shown that audio-visual preferences can also affect the perception of the outdoor environment because some people may be visual-dominant, while others may be auditory-dominant (Sun et al., 2018). In our experiment, visual elements and auditory elements were synchronised, and there was no conflict between auditory attention and visual attention. In addition, some research have found that augmenting natural sound in noisy environments has a positive effect on people's subjective perception (Van Renterghem et al., 2020; Hong et al., 2021), and the physiological effect of this phenomenon is also worth studying. Therefore, in follow-up research, the influence of audio-visual interactions on physiological indicators can be studied from many perspectives, such as the coordination of sound and vision, the salience of vision and hearing, and audio-visual preferences.

With the development of technology, an increasing number of audio-visual experiments have presented the environment through virtual reality (VR) technology (Echevarria Sanchez et al., 2017; Puyana-Romero et al., 2017), which includes auditory factors (Hong et al., 2017). The natural environment presented using VR technology has been shown to relieve stress (Valtchanov et al., 2010). In this study, the VR technique was not used because at the time of the experiment, the VR instruments in the laboratory were relatively heavy. In the future, we will consider how to combine physiological measurements with VR instruments.

5 Conclusions

Through the analysis of changes in physiological indicators and subjective restorativeness based on the present audio-visual environment in the laboratory, this study focused on the influence of visual and auditory factors on physiological indicators and the subjective and physiological differences between dynamic vision and static vision during environment presentations. We found the following:

(1) The influence of scenario type was significant only on the physiological indicators of HR, α-EEG, and β-EEG. The subjective instorative abilities of the birdsong and sea conditions were very high (with that of birdsong higher than that of the sea), whereas the subjective instorative abilities of the traffic and street conditions were relatively low (that of traffic was lower than that of the street).

(2) Visual factors in the environment had an impact on auditory physiological responses. With the addition of visual factors, HR, SDNN-HRV, HF-HRV, α-EEG, and ST decreased, and β-EEG and SCL increased.

(3) Most of the physiological indicators were affected by the addition of sound, which decreased SDNN-HRV, HF-HRV, ST, and RD and increased RR and SCL.

(4) There were significant differences between dynamic vision and static vision; namely, the physiological indicators SDNN-HRV, HF-HRV, and ST of the video-sound group were lower than those of the image-sound group. However, the SCL was higher in the video-sound group than in the image-sound group.

Moreover, we found that physiological indicators may be more sensitive to the pure physical stimuli of hearing and vision, whereas subjective evaluations are more often based on empirical cognition and can better reflect the meaning behind the environment. The presentation mode of video-sound prompted more natural changes in physiological indicators. Sound-only and video-only conditions affected changes in physiological indicators, and subjectively comfortable scenarios did not necessarily remain physiologically comfortable when presented as a single perceptual condition. Although vision and hearing elicited different changes in physiological indicators, there were no significant differences between visual and auditory effects on physiology. Compared with environments presented as dynamic videos, those presented as static images could elicit more physiological relaxation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to express our sincere gratitude to every participant in the experiment. The work was supported by the National Natural Science Foundation of China [grant number: 51778169], and also, the European Research Council (ERC) Advanced Grant (no. 55740696) on “Soundscape Indices” (SSID).

Appendix

Table 4 Perceived Restorativeness Soundscape Scale (Payne, 2013)

Fascination
I find this sonic environment appealing.
My attention is drawn to many of the interesting sounds here.
These sounds make me want to linger here.
These sounds make me wonder about things.
I am engrossed by this sonic environment.
Being-Away-To
I hear these sounds when I am doing something different to what I usually do.
This is a different sonic environment to what I usually hear.
I am hearing sounds that I usually hear.
Being- Away-From
This sonic environment is a refuge from unwanted distractions.

When I hear these sounds I feel free from work, routine and responsibilities.
Listening to these sounds gives me a break from my day-to-day listening experience.
Compatibility
These sounds relate to activities I like to do.
This sonic environment fits with my personal preferences
I rapidly get used to hearing this type of sonic environment
Hearing these sounds hinders what I would want to do in this place
Extent (Coherence)
All the sounds I'm hearing belong here (with the place shown).
All the sounds merge to form a coherent sonic environment.
The sounds I am hearing seem to fit together quite naturally with this place.
Extent (Scope)
The sonic environment suggests the size of this place is limitless.

References

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

Alvarsson, J.J., Wiens, S., Nilsson, M.E., 2010. Stress recovery during exposure to nature sound and environmental noise. Int. J. Environ. Res. Public Health. 7(3), 1036–1046. https://doi.org/10.3390/ijerph7031036.

Anderson, L.M., Mulligan, B.E., Goodman, L.S., Regen, H.Z., 1983. Effects of sounds on preferences for outdoor settings. Environ. Behav. https://doi.org/10.1177/0013916583155001

Annerstedt, M., Jönsson, P., Wallergård, M., Johansson, G., Karlson, B., Grahn, P., Hansen, Å.M., Währborg, P., 2013. Inducing physiological stress recovery with sounds of nature in a virtual reality forest - Results from a pilot study. Physiol. Behav. 118, 240–250. https://doi.org/10.1016/j.physbeh.2013.05.023

Axelsson, Ö., Nilsson, M.E., Berglund, B., 2010. A principal components model of soundscape perception. J. Acoust. Soc. Am. 128(5), 2836–2846. https://doi.org/10.1121/1.3493436.

Ba, M., Kang, J., 2019a. A laboratory study of the sound-odour interaction in urban environments. Build. Environ. 147, 314–326. https://doi.org/10.1016/j.buildenv.2018.10.019.

Ba, M., Kang, J., 2019b. Effect of a fragrant tree on the perception of traffic noise. Build. Environ. 156, 147–155. https://doi.org/10.1016/j.buildenv.2019.04.022.

Ba, M., Kang, J., Li, Z., 2020. The effects of sounds and food odour on crowd behaviours in urban public open spaces. Build. Environ. 182, 107104. https://doi.org/10.1016/j.buildenv.2020.107104

Baekey, D.M., Molkov, Y.I., Paton, J.F.R., Rybak, I.A., Dick, T.E., 2010. Effect of baroreceptor stimulation on the respiratory pattern: Insights into respiratory-sympathetic interactions. Respir. Physiol. Neurobiol. https://doi.org/10.1016/j.resp.2010.09.006

Belmeziti, A., Cherqui, F., Kaufmann, B., 2018. Improving the multi-functionality of urban green spaces: Relations between components of green spaces and urban services. Sustain. Cities Soc. 43, 1–10. https://doi.org/10.1016/j.scs.2018.07.014

Bradley, M.M., Codispoti, M., Cuthbert, B.N., Lang, P.J., 2001. Emotion and Motivation I: Defensive and Appetitive Reactions in Picture Processing. Emotion 1(3), 276–298. https://doi.org/10.1037/1528-3542.1.3.276.

Berto, R., 2005. Exposure to restorative environments helps restore attentional capacity. J. Environ. Psychol. 25, 249–259. https://doi.org/10.1016/j.jenvp.2005.07.001

Berto, R., 2014. The role of nature in coping with psycho-physiological stress: A literature review on restorativeness. Behav. Sci. (Basel). 4, 394–409. https://doi.org/10.3390/bs4040394

Brown, D.K., Barton, J.L., Gladwell, V.F., 2013. Viewing nature scenarios positively affects recovery of autonomic function following acute-mental stress. Environ. Sci. Technol. 47, 5562–5569. https://doi.org/10.1021/es305019p

Bulagang, A.F., Weng, N.G., Mountstephens, J., Teo, J., 2020. A review of recent approaches for emotion classification using electrocardiography and electrodermography signals. Informatics Med. Unlocked. https://doi.org/10.1016/j.imu.2020.100363

Carles, J., Bernáldez, F., De Lucio, J., 1992. Audio-visual interactions and soundscape preferences. Landsc. Res. 17, 52–56. https://doi.org/10.1080/01426399208706361

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

Chau, C.K., Leung, T.M., Xu, J.M., Tang, S.K., 2018. Modelling noise annoyance responses to combined sound sources and views of sea, road traffic, and mountain greenery. J. Acoust. Soc. Am. 144, 3503–3513. https://doi.org/10.1121/1.5083833

Choi, I., Lee, J.Y., Lee, S.H., 2018. Bottom-up and top-down modulation of multisensory integration. Curr. Opin. Neurobiol. 52, 115–122. https://doi.org/10.1016/j.conb.2018.05.002

Dawson, M.E., Schell, A.M., Filion, D.L., 2016. The electrodermal system, in: Handbook of Psychophysiology, Fourth Edition. pp. 217–243. https://doi.org/10.1017/9781107415782.010

Dousty, M., Daneshvar, S., Haghjoo, M., 2011. The effects of sedative music, arousal music, and silence on electrocardiography signals. J. Electrocardiol. 44, 396.e1-396.e6. https://doi.org/10.1016/j.jelectrocard.2011.01.005

Dunnett, C.W., 1955. A Multiple Comparison Procedure for Comparing Several Treatments with a Control. J. Am. Stat. Assoc. https://doi.org/10.1080/01621459.1955.10501294

Echevarria Sanchez, G.M., Van Renterghem, T., Sun, K., De Coensel, B., Botteldooren, D., 2017. Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space. Landsc. Urban Plan. 167, 98–107. https://doi.org/10.1016/j.landurbplan.2017.05.018

Erfanian, M., Mitchell, A.J., Kang, J., Aletta, F., 2019. The psychophysiological implications of soundscape: A systematic review of empirical literature and a research agenda. Int. J. Environ. Res. Public Health 16(19), 1–20. https://doi.org/10.3390/ijerph16193533.

Fachner, J., Gold, C., Erkkilä, J., 2013. Music therapy modulates fronto-temporal activity in rest-EEG in depressed clients. Brain Topogr. 26, 338–354. https://doi.org/10.1007/s10548-012-0254-x

Fernández, T., Harmony, T., Rodríguez, M., Bernal, J., Silva, J., Reyes, A., Marosi, E., 1995. EEG activation patterns during the performance of tasks involving different components of mental calculation. Electroencephalogr. Clin. Neurophysiol. 94, 175–182. https://doi.org/10.1016/0013-4694(94)00262-J

Fisher, G.H., 1968. Agreement between the spatial senses. Percept. Mot. Skills 26, 849–850. https://doi.org/10.2466/pms.1968.26.3.849

Gan, Y., Luo, T., Breitung, W., Kang, J., Zhang, T., 2014. Multi-sensory landscape assessment: The contribution of acoustic perception to landscape evaluation. J. Acoust. Soc. Am. 136, 3200–3210. https://doi.org/10.1121/1.4898424

Godfroy-Cooper, M., Sandor, P.M.B., Miller, J.D., Welch, R.B., 2015. The interaction of vision and audition in two-dimensional space. Front. Neurosci. 9. https://doi.org/10.3389/fnins.2015.00311

Hartig, T., Evans, G.W., Jamner, L.D., Davis, D.S., Gärling, T., 2003. Tracking restoration in natural and urban field settings. J. Environ. Psychol. 23, 109–123. https://doi.org/10.1016/S0272-4944(02)00109-3

Hartig, T., Mitchell, R., De Vries, S., Frumkin, H., 2014. Nature and health. Annu. Rev. Public Health 35, 207–228. https://doi.org/10.1146/annurev-publhealth-032013-182443

Hawley, M.L., Litovsky, R.Y., Culling, J.F., 2004. The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. J. Acoust. Soc. Am. 115, 833–843. https://doi.org/10.1121/1.1639908

Hermida Cadena, L.F., Lobo Soares, A.C., Pavón, I., Bento Coelho, J.L., 2017. Assessing soundscape: Comparison between in situ and laboratory methodologies. Noise Mapp. 4(1), 57–66. https://doi.org/10.1515/noise-2017-0004

Hoffman, D.D., Singh, M., 1997. Salience of visual parts. Cognition 63, 29–78. https://doi.org/10.1016/S0010-0277(96)00791-3

Hong, J.Y., He, J., Lam, B., Gupta, R., Gan, W.S., 2017. Spatial audio for soundscape design: Recording and reproduction. Appl. Sci. 7(6). https://doi.org/10.3390/app7060627.

Hong, J.Y., Lam, B., Ong, Z.T., Ooi, K., Gan, W.S., Kang, J., Yeong, S., Lee, I., Tan, S.T., 2020. Effects of contexts in urban residential areas on the pleasantness and appropriateness of natural sounds. Sustain. Cities Soc. 63, 102475. https://doi.org/10.1016/j.scs.2020.102475

Hong, J.Y., Lam, B., Ong, Z.T., Ooi, K., Gan, W.S., Kang, J., Yeong, S., Lee, I., Tan, S.T., 2021. A mixed-reality approach to soundscape assessment of outdoor urban environments augmented with natural sounds. Build. Environ. 194, 107688. https://doi.org/10.1016/j.buildenv.2021.107688

Hong, J.Y., Jeon, J.Y., 2013. Designing sound and visual components for enhancement of urban soundscapes. J. Acoust. Soc. Am. 134, 2026–2036. https://doi.org/10.1121/1.4817924

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

Huang, N., Elhilali, M., 2017. Auditory salience using natural soundscapes. J. Acoust. Soc. Am. 141, 2163–2176. https://doi.org/10.1121/1.4979055

Hume, K., Ahtamad, M., 2013. Physiological responses to and subjective estimates of soundscape elements. Appl. Acoust. 74(2), 275–281. https://doi.org/10.1016/j.apacoust.2011.10.009.

Hunter, M.D., Eickhoff, S.B., Pheasant, R.J., Douglas, M.J., Watts, G.R., Farrow, T.F.D., Hyland, D., Kang, J., Wilkinson, I.D., Horoshenkov, K. V., Woodruff, P.W.R., 2010. The state of tranquility: Subjective perception is shaped by contextual modulation of auditory connectivity. Neuroimage 53(2), 611–618. https://doi.org/10.1016/j.neuroimage.2010.06.053.

Irwin, A., Hall, D.A., Peters, A., Plack, C.J., 2011. Listening to urban soundscapes: Physiological validity of perceptual dimensions. Psychophysiology 48(2), 258–268. https://doi.org/10.1111/j.1469-8986.2010.01051.x.

ISO. ISO 12913-1:2014: Acoustics—Soundscape Part 1: Definition and Conceptual Framework; ISO: Geneva, Switzerland, 2014.

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

Jiang, B., Xu, W., Ji, W., Kim, G., Pryor, M., Sullivan, W. C., 2021. Impacts of nature and built acoustic-visual environments on human's multidimensional mood states: A cross-continent experiment. J Environ Psychol. 77, 101659.https://doi.org/10.1016/j.jenvp.2021.101659

Jiang, L., Kang, J., 2017. Perceived integrated impact of visual intrusion and noise of motorways: Influential factors and impact indicators. Transp. Res. Part D Transp. Environ. 57, 217–223. https://doi.org/10.1016/j.trd.2017.09.027

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

Kaplan, R., Kaplan, S., 1989. The experience of nature: a psychological perspective. Exp. Nat. a Psychol. Perspect. https://doi.org/10.1097/00005053-199111000-00012

Kayser, C., Logothetis, N.K., 2007. Do early sensory cortices integrate cross-modal information? Brain Struct. Funct. https://doi.org/10.1007/s00429-007-0154-0

Kim, H.G., Cheon, E.J., Bai, D.S., Lee, Y.H., Koo, B.H., 2018. Stress and heart rate variability: A meta-analysis and review of the literature. Psychiatry Investig. https://doi.org/10.30773/pi.2017.08.17

Kistler, A., Mariauzouls, C., Von Berlepsch, K., 1998. Fingertip temperature as an indicator for sympathetic responses. Int. J. Psychophysiol. 29, 35–41. https://doi.org/10.1016/S0167-8760(97)00087-1

Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. Brain Res. Rev. https://doi.org/10.1016/S0165-0173(98)00056-3

Korpela, K.M., Ratcliffe, E., 2021. Which is primary: Preference or perceived instoration? J. Environ. Psychol. 75, 101617. https://doi.org/10.1016/j.jenvp.2021.101617

Kreibig, S.D., 2010. Autonomic nervous system activity in emotion: A review. Biol. Psychol. 84, 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Langewitz, W., Ruddel, H., Von Eiff, A.W., 1987. Influence of perceived level of stress upon ambulatory blood pressure, heart rate, and respiratory frequency. J. Clin. Hypertens. 3, 743–748.

Laufs, J., Borrion, H., Bradford, B., 2020. Security and the smart city: A systematic review. Sustain. Cities Soc. 55. https://doi.org/10.1016/j.scs.2020.102023

Lee, J., Park, B.J., Tsunetsugu, Y., Kagawa, T., Miyazaki, Y., 2009. Restorative effects of viewing real forest landscapes, based on a comparison with urban landscapes. Scand. J. For. Res. https://doi.org/10.1080/02827580902903341

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

Li, Z., Kang, J., 2019. Sensitivity analysis of changes in human physiological indicators observed in soundscapes. Landsc. Urban Plan. 190. https://doi.org/10.1016/j.landurbplan.2019.103593.

Li, Z., Kang, J., Ba, M., 2020. Influence of distance from traffic sounds on physiological indicators and subjective evaluation. Transp. Res. Part D Transp. Environ. 87. https://doi.org/10.1016/j.trd.2020.102538

Lindquist, M., Lange, E., Kang, J., 2016. From 3D landscape visualisation to environmental simulation: The contribution of sound to the perception of virtual environments. Landsc. Urban Plan. 148, 216–231. https://doi.org/10.1016/j.landurbplan.2015.12.017.

Liu, F., Kang, J., 2018. Relationship between street scale and subjective assessment of audio-visual environment comfort based on 3D virtual reality and dual-channel acoustic tests. Build. Environ. 129, 35–45. https://doi.org/10.1016/j.buildenv.2017.11.040

Liu, J., Kang, J., Luo, T., Behm, H., 2013. Landscape effects on soundscape experience in city parks. Sci. Total Environ. 454–455, 474–481. https://doi.org/10.1016/j.scitotenv.2013.03.038

Liu, J., Kang, J., Behm, H., Luo, T., 2014. Effects of landscape on soundscape perception: Soundwalks in city parks. Landsc. Urban Plan. 123, 30–40. https://doi.org/10.1016/j.landurbplan.2013.12.003

Lu, Y., Chen, L., Liu, X., Yang, Y., Sullivan, W.C., Xu, W., Webster, C., Jiang, B., 2021. Green spaces mitigate racial disparity of health: A higher ratio of green spaces indicates a lower racial disparity in SARS-CoV-2 infection rates in the USA. Environ. Int. 152. https://doi.org/10.1016/j.envint.2021.106465

Malliani, A., Pagani, M., Lombardi, F., Cerutti, S., 1991. Cardiovascular neural regulation explored in the frequency domain. Circulation 84, 482–492. https://doi.org/10.1161/01.CIR.84.2.482

Medvedev, O., Shepherd, D., Hautus, M.J., 2015. The restorative potential of soundscapes: A physiological investigation. Appl. Acoust. 96, 20–26. https://doi.org/10.1016/j.apacoust.2015.03.004.

Morein-Zamir, S., Soto-Faraco, S., Kingstone, A., 2003. Auditory capture of vision: Examining temporal ventriloquism. Cogn. Brain Res. https://doi.org/10.1016/S0926-6410(03)00089-2

Park, S.H., Lee, P.J., Jung, T., Swenson, A., 2020. Effects of the aural and visual experience on psycho-physiological recovery in urban and rural environments. Appl. Acoust. 169. https://doi.org/10.1016/j.apacoust.2020.107486

Payne, S.R., 2013. The production of a perceived restorativeness soundscape scale. Appl. Acoust. 74, 255–263. https://doi.org/10.1016/j.apacoust.2011.11.005

Pheasant, R.J., Fisher, M.N., Watts, G.R., Whitaker, D.J., Horoshenkov, K. V., 2010. The importance of auditory-visual interaction in the construction of “tranquil space.” J. Environ. Psychol. 30, 501–509. https://doi.org/10.1016/j.jenvp.2010.03.006

Preis, A., Kociński, J., Hafke-Dys, H., Wrzosek, M., 2015. Audio-visual interactions in environment assessment. Sci. Total Environ. 523, 191–200. https://doi.org/10.1016/j.scitotenv.2015.03.128

Prinsloo, G.E., Rauch, H.G.L., Karpul, D., Derman, W.E., 2013. The effect of a single session of short duration heart rate variability biofeedback on EEG: A pilot study. Appl. Psychophysiol. Biofeedback 38, 45–56. https://doi.org/10.1007/s10484-012-9207-0

Purcell, T., Peron, E., Berto, R., 2001. Why do preferences differ between scene types? Environ. Behav. 33, 93–106. https://doi.org/10.1177/00139160121972882

Puyana-Romero, V., Lopez-Segura, L.S., Maffei, L., Hernández-Molina, R., Masullo, M., 2017. Interactive soundscapes: 360°-video based immersive virtual reality in a tool for the participatory acoustic environment evaluation of urban areas, in: Acta Acustica United with Acustica. pp. 574–588. https://doi.org/10.3813/AAA.919086

Ratcliffe, E., Gatersleben, B., Sowden, P.T., 2013. Bird sounds and their contributions to perceived attention restoration and stress recovery. J. Environ. Psychol. 36, 221–228. https://doi.org/10.1016/j.jenvp.2013.08.004

Rautaharju, P.M., Blackburn, H., 1961. Relationship of elevated blood pressure to ECG amplitudes and spatial vectors in otherwise “healthy” subjects. Am. Heart J. 61, 156–160. https://doi.org/10.1016/0002-8703(61)90569-5

Roe, J.J., Aspinall, P.A., Mavros, P., Coyne, R., 2013. Engaging the brain: the impact of natural versus urban scenarios using novel EEG methods in an experimental setting. Environ. Sci. https://doi.org/10.12988/es.2013.3109

Schreuder, E., van Erp, J., Toet, A., Kallen, V.L., 2016. Emotional Responses to Multisensory Environmental Stimuli: A Conceptual Framework and Literature Review. SAGE Open 6. https://doi.org/10.1177/2158244016630591

Schweiker, M., Ampatzi, E., Andargie, M.S., Andersen, R.K., Azar, E., Barthelmes, V.M., Berger, C., Bourikas, L., Carlucci, S., Chinazzo, G., Edappilly, L.P., Favero, M., Gauthier, S., Jamrozik, A., Kane, M., Mahdavi, A., Piselli, C., Pisello, A.L., Roetzel, A., Rysanek, A., Sharma, K., Zhang, S., 2020. Review of multi‐domain approaches to indoor environmental perception and behaviour. Build. Environ. 176, 106804. https://doi.org/10.1016/j.buildenv.2020.106804

Shepherd, D., Welch, D., Dirks, K.N., McBride, D., 2013. Do quiet areas afford greater health-related quality of life than noisy areas? Int. J. Environ. Res. Public Health 10, 1284–1303. https://doi.org/10.3390/ijerph10041284

Southworth, M., 1969. The sonic environment of cities. Environ. Behav. 1, 49–70. https://doi.org/10.1177/001391656900100104

Spence, C., Puccinelli, N.M., Grewal, D., Roggeveen, A.L., 2014. Store atmospherics: A multisensory perspective. Psychol. Mark. 31, 472–488. https://doi.org/10.1002/mar.20709

Stahl, C., 2006. Software for generating psychological experiments. Exp. Psychol. 53, 218–232.

https://doi.org/10.1027/1618-3169.53.3.218

Sun, K., Sanchez, G.M.E., De Coensel, B., Van Renterghem, T., Talsma, D., Botteldooren, D., 2018. Personal audiovisual aptitude influences the interaction between landscape and soundscape appraisal. Front. Psychol. 9. https://doi.org/10.3389/fpsyg.2018.00780

Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, M.G., 2010. The multifaceted interplay between attention and multisensory integration. Trends Cogn. Sci. https://doi.org/10.1016/j.tics.2010.06.008

Thayer, J.F., Åhs, F., Fredrikson, M., Sollers, J.J., Wager, T.D., 2012. A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. Neurosci. Biobehav. Rev. https://doi.org/10.1016/j.neubiorev.2011.11.009

Torresin, S., Pernigotto, G., Cappelletti, F., Gasparella, A., 2018. Combined effects of environmental factors on human perception and objective performance: A review of experimental laboratory works. Indoor Air 28, 525–538. https://doi.org/10.1111/ina.12457

Torresin, S., Albatici, R., Aletta, F., Babich, F., Kang, J., 2019. Assessment methods and factors determining positive indoor soundscapes in residential buildings: A systematic review. Sustain. 11(19), 5290. https://doi.org/10.3390/su11195290.

Tse, M.S., Chau, C.K., Choy, Y.S., Tsui, W.K., Chan, C.N., Tang, S.K., 2012. Perception of urban park soundscape. J. Acoust. Soc. Am. 131, 2762–2771. https://doi.org/10.1121/1.3693644

Ulrich, R.S., 1981. Natural versus urban scenarios: Some Psychophysiological Effects. Environ. Behav. 13, 523–556. https://doi.org/10.1177/0013916581135001

Ulrich, R.S., Simons, R.F., Losito, B.D., Fiorito, E., Miles, M.A., Zelson, M., 1991. Stress recovery during exposure to natural and urban environments. J. Environ. Psychol. 11, 201–230. https://doi.org/10.1016/S0272-4944(05)80184-7

Valtchanov, D., Barton, K.R., Ellard, C., 2010. Restorative effects of virtual nature settings. Cyberpsychology, Behav. Soc. Netw. 13, 503–512. https://doi.org/10.1089/cyber.2009.0308

Van den Berg, A.E., 2017. From green space to green prescriptions: Challenges and opportunities for research and practice. Front. Psychol. https://doi.org/10.3389/fpsyg.2017.00268

Van Renterghem, T., 2019. Towards explaining the positive effect of vegetation on the perception of environmental noise. Urban For. Urban Green. https://doi.org/10.1016/j.ufug.2018.03.007

Van Renterghem, T., Vanhecke, K., Filipan, K., Sun, K., De Pessemier, T., De Coensel, B., Joseph, W., Botteldooren, D., 2020. Interactive soundscape augmentation by natural sounds in a noise polluted urban park. Landsc. Urban Plan.

194, 103705. https://doi.org/10.1016/j.landurbplan.2019.103705

Viollon, S., Lavandier, C., Drake, C., 2002. Influence of visual setting on sound ratings in an urban environment. Appl. Acoust. 63, 493–511. https://doi.org/10.1016/S0003-682X(01)00053-6

Von Lindern, E., Hartig, T., Lercher, P., 2016. Traffic-related exposures, constrained restoration, and health in the residential context. Heal. Place 39, 92–100. https://doi.org/10.1016/j.healthplace.2015.12.003

White, M., Smith, A., Humphryes, K., Pahl, S., Snelling, D., Depledge, M., 2010. Blue space: The importance of water for preference, affect, and restorativeness ratings of natural and built scenarios. J. Environ. Psychol. 30, 482–493. https://doi.org/10.1016/j.jenvp.2010.04.004

Wu, L., Kim, S.K., 2021. Health outcomes of urban green space in China: Evidence from Beijing. Sustain. Cities Soc. 65, 102604. https://doi.org/10.1016/j.scs.2020.102604

Yang, F., Bao, Z.Y., Zhu, Z.J., 2011. An assessment of psychological noise reduction by landscape plants. Int. J. Environ. Res. Public Health 8, 1032–1048. https://doi.org/10.3390/ijerph8041032

Yin, J., Zhu, S., MacNaughton, P., Allen, J.G., Spengler, J.D., 2018. Physiological and cognitive performance of exposure to biophilic indoor environment. Build. Environ. 132, 255–262. https://doi.org/10.1016/j.buildenv.2018.01.006

Yong Jeon, J., Young Hong, J., Jik Lee, P., 2013. Soundwalk approach to identify urban soundscapes individually. J. Acoust. Soc. Am. 134, 803–812. https://doi.org/10.1121/1.4807801