Effects of the aural and visual experience on psycho-physiological recovery in urban and rural environments

Sang Hee Park[1], Pyoung Jik Lee[2], Timothy Jung[3], and Alasdair Swenson[3]

[1] Department of Living and Built Environment Research, Korea Institute of Civil Engineering and Building Technology, South Korea (sangheepark@kict.re.kr)

[2] Acoustics Research Unit, School of Architecture, School of the Arts, Faculty of Humanities and Social Sciences, University of Liverpool, Liverpool, UK (p.j.lee@liverpool.ac.uk)

[3] Creative AR & VR Hub, Department of Operations, Technology, Events and Hospitality Management, Manchester Metropolitan University, Manchester, UK (t.jung@mmu.ac.uk; a.swenson@mmu.ac.uk)

Corresponding author: Pyoung Jik Lee

- Address: Acoustics Research Unit, School of Architecture, School of the Arts, Faculty of Humanities and Social Sciences, University of Liverpool, Liverpool, UK

- Email: p.j.lee@liverpool.ac.uk

- Telephone: +44 (0)151 794 4944

Abstract

This study examined the effects of aural and visual experience on psycho-physiological recovery in simulated environments. Laboratory experiments were conducted with 32 participants. Ten horror videos were used as stressor clips, while five videos depicted urban or rural settings as recovery clips. The videos were presented via VR (head-mounted display) and on a monitor screen. The effect of the audio presentation was tested by presenting the stimuli with or without sound. Psychological recovery was assessed using a set of questions (e.g., perceived preference). Physiological recovery was measured with five physiological responses (e.g., fEMG and HR) monitored throughout the experiment. It was found that the rural setting led to a better psycho-physiological recovery than the urban setting when the stimuli were presented both in the VR and Screen conditions. In particular, the rural setting with water features evoked the greatest recovery. The rural setting presented with water sound showed significant differences in psychological recovery between Audio-Visual and No-Audio conditions. Compared to the Screen condition, stimuli presented in the VR condition did not have any main effect on the psychological recovery; however, it showed main effects on some of the physiological responses. The audio presentation had significant impacts on all psychological recovery ratings but it showed an impact on only one physiological response, fEMG of the zygomatic muscle.

Keywords: Psycho-physiological recovery; Urban; Rural; Virtual Reality; Audio-visual interaction

1 Introduction

Researchers have been interested in the association between pleasant landscapes and human well-being for a long time [1, 2]. In particular, restoration has been suggested as one of the benefits of natural environments; it is a recovery from attentional fatigue and reflection upon daily or life issues [1]. In general, research on restoration has followed two main theoretical explanations which have their own interpretations of restoration [3]. The first theoretical framework is the Stress Recovery Theory (SRT) proposed by Ulrich [4]. It explains how nature reduces psycho-physiological stress. Ulrich [4] suggested that humans have long evolved in nature and viewing natural environments has been of importance for their adaptation and survival. For example, the view of plants or water has been known to be vital for humans [4, 5]. The fulfillment of human survival is associated with the recovery from fatigue or stress, which also links to the preservation of cognitive resources [6]. Based on this evolutionary perspective, it has been suggested that humans are well adapted to nature and it helps them to recover from stress more quickly and completely than urban or built environments [2]. Another theoretical framework of restoration is the Attention Restoration Theory (ART) through which Kaplan [1] separated attention into two components, directed-attention and involuntary-attention. Directed-attention is a cognitive mechanism which requires cognitive efforts and awareness, while involuntary-attention is an effortless status where the attention is captured by intriguing stimuli. Kaplan [1] also argued that directed-attention needs to be restored by interactions with nature since natural environments are rich in the key characteristics for the restorative experience which consist of the following four components: being away, fascination, extent, and compatibility. The ART proposes that natural environments require less directed-attention, whereas urban contexts involve demands on attentional resources [7].

A number of psycho-physiological studies have been conducted based on the SRT or the ART and supported the benefits of natural environments. For example, Wells [8] investigated the relationship between the well-being of children and natural environment nearby their homes. The study used a longitudinal method involving before and after the relocation of their homes. The results showed a close association between the children's cognitive functioning and the greenness of the environment nearby their houses. Hartig et al. [9] measured psycho-physiological responses in natural and urban environments. The study found a more rapid decrease in blood pressure when trees were seen compared to a viewless setting. The natural setting also evoked greater stress reduction, improved performance on an attentional test, increased positive affect, and reduced anger than the urban setting. More recently, Gidlow et al. [10] measured participants' psycho-physiological restoration responses after walks in natural and residential environments. Average noise levels of the natural environments were lower than the residential environment which involved road traffic noise. They reported that the natural environments evoked greater restoration experiences; in particular, the rate of perceived exertion (RPE) and heart rate were found to be the lowest in the natural environment with the water view. However, measured noise level at each site was not introduced to examine its effect on participants' responses.

Sound is one of many aspects of the environment that influences human health and well-being [11]. In particular, noise has been known to have adverse health effects such as sleep disturbance and cardiovascular disease [12]. Therefore, quietness has been considered as one of the critical features of the environment for human well-being [13]. Practically, there are legislation and suggestions for securing quietness in certain areas by protecting them from noise sources (e.g., EU Directive 2002/49/EC) and for improving tranquillity in open spaces [14]. There are some studies which have demonstrated the restorative effects of sounds. Given that natural sounds are generally preferred to urban sounds [15], researchers have reported a close association between natural sounds and restoration experience of soundscapes. Alvarsson et al. [16] conducted a laboratory study and measured participants' physiological responses when they were exposed to sounds from natural (water and bird sounds) or noisy (road traffic noise) environments after a stressful mental arithmetic task. The study found a fast recovery of skin conductance levels during the natural sound compared to the noisy environment. de Coensel et al. [17] carried out a laboratory study to assess perceived loudness, pleasantness, and eventfulness of road traffic noise combined with either water or bird sounds. The study found that the water sound reduced the perceived loudness of road traffic noise with low temporal variability. In addition, it was found that the bird sound increased the pleasantness and eventfulness. Ratcliffe et al. [18] performed semi-structured interviews to explore the restorative perceptions of bird sounds and suggested bird sounds as the natural sound most commonly associated with stress recovery and attention restoration. Furthermore, Aletta et al. [19] carried out a systematic review on the relationship between health-related effects and perception of soundscape and they reported that positive perceptions of soundscape is likely to correlate with positive health effects. Sun et al. [20] focused on audiovisual aptitude among personal factors. They found that audiovisual aptitude had an impact on the relationship between the visibility of vegetation and the self-reported noise annoyance. It was also reported that audiovisual aptitude may play an important role in the appraisal of living environments. More recently, Erfanian et al. [21] conducted a review on psycho-physiological outcomes of soundscape. According to the review, HR was the most commonly adopted measure in soundscape studies. The review also emphasised that the current level of understanding of psycho-physiological outcomes of sounscape is still inconclusive, thus further investigation is needed to extend the current understanding. Li and Kang [22] investigated an association between physiological parameters and subjective restorative evaluation of soundscape. They measured several physiological data and tested their correlations with participants' responses to the perceived restorativeness. Although several associations between physiological parameters and the subjective evaluation were relatively strong, it was limited to interpret the data accurately.

Given that it has been known that natural soundscapes have restorative effects and that both acoustic and visual features have impacts on restoration experiences, there is a limited number of studies that investigated the interaction between audio-visual features of soundscape. It has been known that visual scene plays an important role in auditory perception [23]. Jeon et al. [24] performed a laboratory study where they presented stimuli in audio-only and audio-visual conditions. The sounds were presented with or without visual images of water features in urban environments. Both sounds and visual images showed significant impacts on perceived preference. It was also found that the contribution of the visual images to improvement in preferences was limited in the urban environment with a high level of road traffic noise. Hong and Jeon [25] conducted a laboratory study to assess the effects of audio-visual components of soundscape. Different views of a street combined with water or vegetation features were presented, while the presented sounds included road traffic noise combined with water and bird sounds. The stimuli were presented in audio-only, visual-only, and audio-visual conditions. The view of greenery enhanced the perception of streetscape, whereas the view of water did not. Concerning the auditory aspect, bird sounds improved the perception of soundscape when it was added to a low level of road traffic noise. Although the previous studies have examined the interaction between the audio and visual cues, there is still a need for further investigation of how audio presentation influences people's physiological as well as psychological responses. This research gap raised the first research question (RQ):

RQ1: Would audio presentation (Audio-Visual vs. No-Audio) have different impacts on psycho-physiological recovery?

Existing experimental research on natural environment's recovery effects has predominantly projected stimuli as still-pictures, 2D videos, or 3D videos on a flat-screen. For example, Ulrich et al. [26] used a wall-mounted flat-screen to presented videos of nature and urban settings, daytime television, and a black screen. They measured participants' blood pressure and pulse rate and found that the natural setting had the greatest restorative effect. Further, Kweon et al. [27] even found that posters on the wall may reduce stress at workplace when they showed natural environments. Since the technology of VR helps users to immerse themselves in the environment and interact with the simulated world in real-time [28], studies on restoration have attempted to apply VR in research. For instance, Valtchanov et al. [29] examined the restorative effects of natural setting in VR. Participants viewed either nature or abstract paintings in a VR setting using a head-mounted display (HMD). The use of the HMD supported 360-degree videos and allowed the participants to be fully immersed in the surroundings. The results showed that the natural setting increased positive affect and decreased stress. However, these studies were not on soundscapes and thereby did not examine how the sounds particularly affected the responses. Annerstedt et al. [30] assessed stress recovery in natural environments and the effect of audio presentation in VR. They measured participants' physiological recovery in two different virtual environments (i.e. with and without exposure to sounds of nature) when stimuli were projected by CAVE System where projectors were directed to the walls and the floor of a room-sized cube. The study found significant links among nature, the sounds of nature, and stress recovery. Payne et al. [31] assessed restoration responses to urban park soundscapes. In a VR laboratory, a few participants were seated and faced a 180-degree curved screen. The videos were projected on the center of the screen. The video was presented either with no sound or one of natural and traffic sounds. Their findings indicated that louder acoustic environment can increase noise annoyance but did not always decrease perceived restoration. In other words, lowering sound levels may or may not have positive effects on individuals' perceptions. Since there had been a need for research on soundscape using fully immersive 360-degree videos, Yu et al. [32] recently adopted HMD to present 360-degree panorama VR videos. They measured psycho-physiological restoration in simulated natural and urban environments, and found that blood pressure and heart rate were not significantly changed across different videos. The study discussed that the results might be due to the lack of senses involved in the simulated environments since VR only stimulates visual and auditory senses. However, it would have been worth comparing the measurements in between VR and traditional non-VR approach of presenting stimuli on a flat-screen. Amongst existing literature on restoration effects of soundscape, there is still a lack of experimental research which uses HMD realising fully immersive VR environment. Moreover, there is also a need for research which compares restoration effects in both VR and traditional non-VR settings such as using a flat-screen. The present study, therefore, aimed to investigate the responses both in VR using HMD ('VR') and in a traditional setting where visual stimuli being presented on a monitor screen ('Screen'). The following is the second research question:

RQ2: Would visual reproduction (VR vs. Screen) have different impacts on psycho-physiological recovery?

The first and second research questions raised another research question: if the audio presentation and the visual reproduction conditions have significant impacts on psycho-physiological recovery, what would be the most effective soundscape enhancing the restoration experience? It has long been known that natural environments enhance restoration [1, 4]. Furthermore, studies have found that sounds of nature closely associate with restoration [16, 17]. In particular, water sound has been reported to improve tranquillity of the soundscape [33], to be the most effective sound masker in urban environments [34] reducing perceived loudness of road traffic noise [17], and to be the most preferred natural sound [35]. Rådsten-Ekman et al. [36] tested the effects of water sounds on the perception of noisy environments. Overall pleasantness increased when road traffic noise was presented with a highly pleasant water sound. Particularly, flowing water sound with a view of less still water has been reported to improve restoration experience [37]. Of recent studies that have adopted VR settings, Masullo and Pascale [38] performed a laboratory study and examined the audio-visual effects of water in an urban environment. For auditory cues, the study presented road traffic noise either combined with water sounds or nothing. When it comes to visual cues, simulated urban environments with different water features were presented. The study found effective masking effects of the water sounds on the traffic noise. Moreover, the study also reported that the view of different water features had notable impacts on participants' subjective responses. Ong et al. [39] conducted a laboratory study where audio-visual environments of a rooftop garden were presented with high traffic noise. They added pleasant sound maskers, either sound of bird or water stream. The study found that the sound of water stream had better masking effect for traffic noise than the bird sound. Although studies have mainly tested water sounds as a type of sound masker, a question remains concerning the most effective soundscape on restoration in VR. Previously, Ulrich et al. [2] presented two 10-minute videos. The first video was a stressor which depicted serious injuries, blood, and mutilation. After presenting the stressor, they presented either urban or natural environments in order to measure stress recovery effects. Given that Ulrich et al. [2] presented the stimuli via a monitor screen with a supplementary speaker for sound presentation, it is worth replicating the experimental design in this study by using the latest technology (i.e. VR). The third research question was as follow:

RQ3: What kind of scene would have the largest impact on psycho-physiological recovery?

Figure 1 illustrates the research questions of this study. The present study aimed to examine the impact of the three conditions (Audio Presentation, Visual Reproduction, and Scene) on psychological and physiological responses.

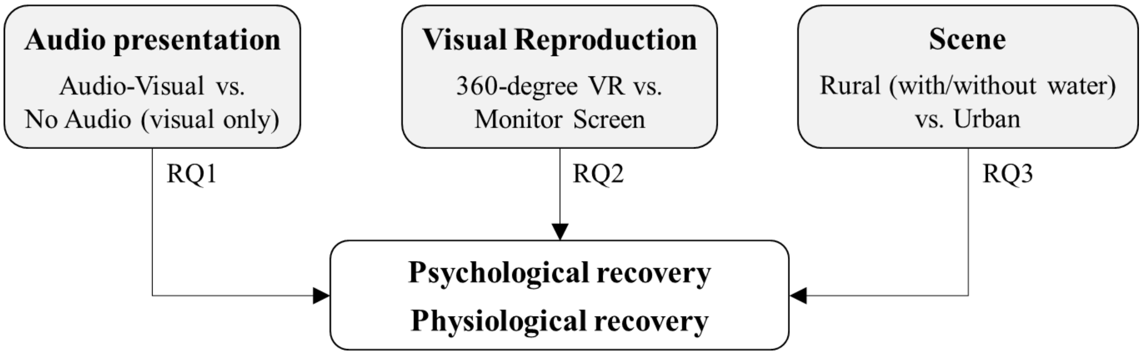

Figure 1. An illustration of the research questions.

The present study also compared the groups split by different demographic characteristics and personal variables (e.g., age, noise sensitivity) to see if any of the variables had moderating effects on the participants' responses. It was expected that these additional analyses might yield a further understanding of psycho-physiological recoveries.

2 Methods

2.1 Participants

Participant recruitment procedure was carried out mainly through an online study advertisement. Those who had an interest in taking part in the study were invited to contact the researcher via email. The researcher answered back to the emails with general information about the study and screened them if they were eligible. Only those who had self-reported normal hearing without any history of hearing, cardiovascular, respiratory, musculoskeletal, and stress/panic-related psychiatric health problems were recruited.

The sample size was estimated based on expected medium effect size, alpha level = 0.5, and statistical power = 0.8 using calculations provided by Cohen [40] and G*Power [41, 42]. The present study was mainly analysed by means of the repeated measures analyses of variance (RM ANOVA) with three groups (Audio Presentation × Visual Reproduction × Scene) and 12 measurements (seven psychological and five physiological measures) when a total of ten stimuli were presented to the participants. The sample size estimation showed that the study needed 27 participants. A total of 32 participants (16 males and 16 females) aged from 20 to 39 (M = 27.8; SD = 4.9) took part in the study. Half of them were in their 20s and the other half were in their 30s. The post-hoc power analysis indicated that the number of participants allowed adequate statistical power to detect the differences in groups.

2.2 Demographic and personal variables

Before the experiment started, each participant responded to a paper questionnaire which measured their personal characteristics. Although the research questions did not include the effects of any personal factors, the present study tested the effects of the personal factors since it might lead to further discussions for a better understanding of the findings or future research insights. It was worth assessing them to see any tendency in each of the groups' responses. Median values were used to divide the groups because the data were not normally distributed and median values are helpful for describing data which is not normally distributed [43]. The questionnaire included items asking age, gender, and three following factors: liking of horror movies, frequency of visits to countrysides, and noise sensitivity. First, the liking of horror movies was measured by a question "How much do you like watching horror movies?" with a 5-point scale (1 = *"Not at all"* and 5 = *"Extremely"*). Later the participants were grouped into low (n = 17) and high (n = 15) horror movie liking groups for the analysis using the median score of 2.0 as a cut-off point. Second, the questionnaire asked how frequently they had visited countryside areas in a year. Seven options were given: once a week, once every two weeks, once a month, once every few months, once a year, less than once a year, or never. The participants who reported they had visited countryside areas once a month or more were grouped as high (n = 18) frequency group and those with less frequency were grouped as low

(n = 14) frequency group. Third, self-reported noise sensitivity was measured. Noise sensitivity is regarded as a stable personality trait that captures attitudes towards a wide range of environmental sounds [44]. In addition, it has been found to have relationships with subjective perception of soundscape [45, 46]. Noise sensitivity was assessed using 21 question items from the scale developed by Weinstein [47] using 6-point scales (1 = *"Not agree at all"* and 6 = *"Extremely agree"*). The median noise sensitivity score (86.0; Min. = 61; Max. = 108) was used as a cut-off point and those with low (n = 15) and high (n = 17) noise sensitivity were split into groups.

2.3 Stimuli

Each stimulus lasted for three minutes. It consisted of a baseline clip, a stressor clip, and a recovery clip, for one minute each. The baseline clip presented a dark grey screen with white noise at 40 dBA ($L_{\text{Aeq,1-minute}}$). The stressor clips showed one of ten horror movie-clips in a randomised order. The horror movies were chosen as stressors because they would increase the participant's stress level [48] and thereby the study might examine the following recovery responses more clearly. All the horror movies were downloaded from online and were 360-degree videos. All the clips were cautiously chosen to avoid violent scenes but they did contain audio and/or visual startling effects. The recovery clips were four rural scenes and one urban scene; all of them were 360-degree videos and recorded on-site or downloaded from online. According to the definition of rural landscape [49], the rural scenes were a mixture of natural and human-managed landscape. The first and second recovery clips (Rural 1a and 1b) depicted rural environments with no water feature, and one was recorded in the Peak District National Park and the other was from online. They were dominated by natural features such as natural landscape but there were also farmhouse with woods around, farmland, sheep, and paths. The third and fourth (Rural 2a and 2b) showed rural environments with visual and auditory water features, one recorded in the Peak District National Park and the other was from online. They were dominated by streams, hills, and trees but small cultivated fields in the distance were also visible. Although existing studies have used different types of water features such as a fountain, waterfalls, and streams., the present study only chose to use water streams based on a previous report that water stream is a positive soundmark that is easily identifiable [50]. The fifth clip presented an urban setting (traffic road) without any natural features and was recorded in an urban area of Manchester. The scene included all the features used by Ulrich [5] such buildings on both sides of the lane, heavy traffic of vehicles, traffic lights, and pedestrians. Compared to the urban scenes in Ulrich's study, the urban scene of the present study showed more heavy vehicles such as buses (eight of 20 vehicles/min) and greater noise level (75 dBA). A 360-degree camera (Samsung Galaxy Gear) was used for the video recordings. All the video file formats were identical. They were all monoscopic 360 videos, 2:1 aspect ratio equirectangular video container at a resolution of 3840 × 2160 with a bit-rate of 80 Mbps. Screen captures of the sample stressors and recovery clips can be found in the Supplement Figure S1.

The sounds were recorded using a B-format Soundfield microphone (ST450 MkII). The recorded sounds were down-mixed to produce the static binaural signals using the KEMAR with small pinna HRTF because the sounds of the videos from online were static binaural signals. All the recovery clips were presented at 55 dBA ($L_{\text{Aeq,1-minute}}$) as it has been reported that perceived tranquillity of urban and rural areas significantly decreased above 55 dB [51]. It was assumed that tranquillity of the places would play a key role in the recovery responses; thus, the sound pressure levels of the recovery clips were set to be 55 dBA to maximise the perceived tranquillity of the sounds. In addition, all the stressor clips were presented at 65 dBA ($L_{\text{Aeq,1-minute}}$) across all the conditions based on a previous finding that human behavioural responses including noise disturbance begin to evoke above 65 dBA [52].

Before the analysis, the sounds were recorded using a head and torso simulator (HATS, Brüel & Kjær Type 4100) to mimic the sounds at participants' ears. The HATS was positioned on a chair in the chamber and sounds were presented using headphones (Sennheiser HD 518). Table 1 shows the temporal and spectral characteristics of the sounds presented to the participants. Rural scenes were less fluctuated than the urban scene with lower L_{10}-L_{90}, while the stressors showed similar L_{10}-L_{90} values except for the stressor #10 with high background noise level. In case of the stressor #10, the screen presented a full view of a snow-covered forest and a man wearing a mask. The man suddenly appeared from nowhere which induced the startling effect. In terms of the video's sounds, a scary and loud background music was constantly being played whereas the man did not make much louder startling sound. Thus, it had similar L_{10}-L_{90} values. In addition, spectograms of the sounds presented for the recovery clips can be found in the Supplement Figure S2.

Table 1. Temporal and spectral characteristics of the stimuli.

		Temporal characteristics					Spectral characteristics						
		L_{eq}	L_{max}	L_{90}	L_{50}	L_{10}	Frequency [Hz]						
							31.5	63	125	250	500	1k	2k
Stressor	1	65.0	81.1	53.7	59.1	67.9	54.0	54.4	59.7	49.2	48.7	52.2	58.0
	2	65.0	76.5	50.4	61.0	68.9	52.8	64.3	64.6	59.5	56.3	54.1	49.5
	3	65.0	77.2	50.5	57.0	69.5	55.5	68.1	64.9	60.5	58.6	53.4	48.3
	4	65.0	80.1	49.6	56.1	68.0	48.7	64.7	63.7	58.3	53.8	55.9	46.6
	5	65.0	77.6	44.8	57.4	69.5	47.2	66.9	64.1	63.1	51.7	51.6	51.5
	6	65.0	79.8	53.2	58.9	68.7	52.7	58.2	62.7	58.8	54.7	53.5	48.2
	7	65.0	78.9	53.9	61.4	67.9	49.4	61.0	63.8	60.9	55.6	54.7	46.1
	8	65.0	75.4	50.9	60.7	69.0	41.9	47.0	55.6	47.2	51.4	62.1	40.4
	9	65.0	75.4	55.7	62.0	68.8	43.3	54.0	63.6	62.3	58.9	52.7	49.0
	10	65.0	78.2	61.8	64.3	66.1	59.0	63.2	58.3	51.0	51.6	43.2	42.0
Recovery	Rural 1a	55.0	59.6	51.0	51.8	55.9	52.6	49.0	51.3	47.6	45.6	45.2	41.9
	Rural 1b	55.0	58.6	49.1	49.5	54.2	48.8	45.1	47.4	35.5	37.3	38.0	39.1
	Rural 2a	55.0	58.2	52.8	54.8	55.7	52.7	45.1	54.0	40.6	44.9	45.8	42.2
	Rural 2b	55.0	60.3	52.2	53.6	56.1	50.0	42.2	48.6	41.0	39.4	45.5	41.8
	Urban	55.0	64.0	49.4	52.0	59.2	56.2	45.6	52.7	50.0	45.7	44.9	41.7

2.4 Experimental design

The stimuli were presented with one of two audio presentation conditions (Audio-Visual or No-Audio) via one of two visual reproduction conditions (VR or Screen). First, the audio presentation conditions consisted of (1) Audio-Visual condition where both sound and vision were presented and (2) No-Audio condition where videos were presented without any sound. Sound stimuli were presented binaurally using headphones (Sennheiser HD 518) without any head tracking system. Second, the visual reproduction conditions consisted of (1) VR condition where 360-degree videos were presented via a VR headset (Oculus Rift) and (2) Screen condition where videos were presented on a monitor screen. Three types of recovery clips (Rural without water, Rural with water, Urban) were used in the VR condition. In the Screen condition, two recovery clips (Rural with water and Urban) were presented. The Rural 1a and 1b (those without water) were only included in the VR condition and not in the Screen condition because it was assumed that rural sounds would be more effective in recovery than urban sounds [53, 54] and the difference between rural sounds might be minor [53]. Furthermore, the duration of the whole experiment including the training session was already about one hour without including the Rural 1a and 1b in the Screen condition. In order to minimise the fatigue due to the long experiment, only the Rural 2a and 2b (those with water) and the Urban were chosen for the Screen condition. Table 2 and Figure 2 illustrate the design of the experiment.

Table 2. The number of stimuli in different conditions of audio presentation and visual reproduction.

Visual and Audio Conditions		Audio Presentation	
		Audio-Visual	No-Audio
Visual Reproduction	VR	3	3
	Screen	2	2

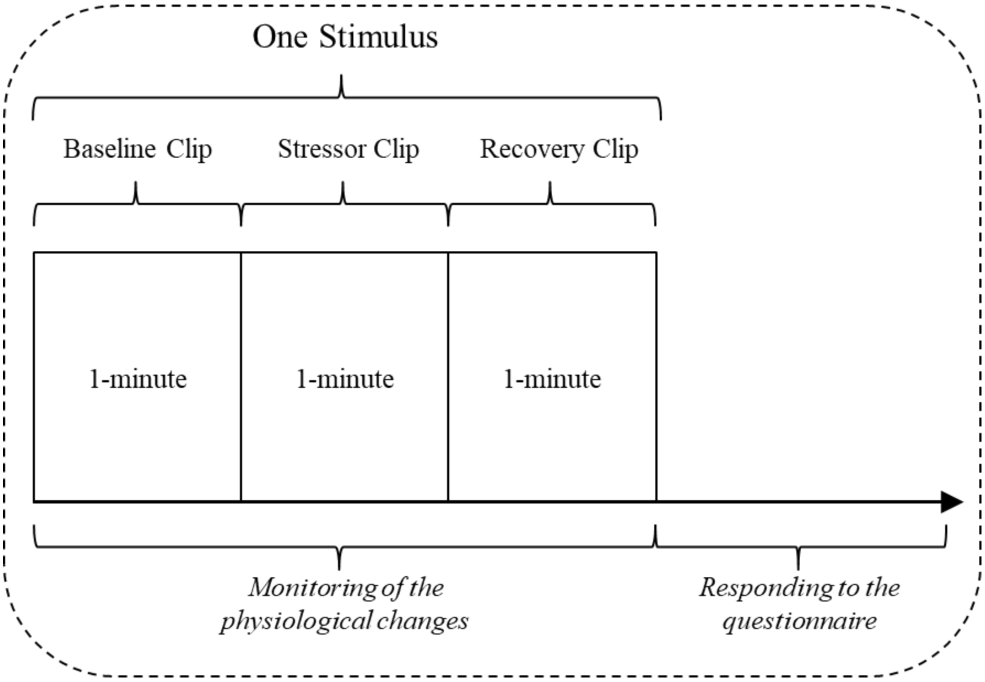

Figure 2. An illustration of the baseline, stressor, and recovery clips consisting of one stimulus.

2.5 Psycho-physiological recovery measurements

At the end of each recovery clip, a set of question items was presented to the participant for measuring psychological recovery. Each question was presented either on the VR headset or on the monitor screen depending on which visual reproduction condition (VR or Screen) the participant was taking part in. The participant responded to the questions verbally in order to minimise their body movement. Since there were a microphone and a speaker connected in and outside the test chamber, the researcher who was outside the chamber could listen to the participant's responses and write down the responses on the response sheet. The participant was asked to repeat the same response twice just in case. Once one question item was answered, the researcher clicked a keyboard button to proceed to the next question item. Responses to the whole questionnaire were collected in this procedure so there was no fixed time length for responding to the questionnaire. Given that the whole experiment lasted for about an hour, it can be guessed that it did not last longer than 3 minutes on average to answer nine questions.

The participant rated tranquillity, pleasantness, and preference for the recovery clips of the stimuli. In addition, six more question items were asked particularly on the sounds. The six items were chosen from the Perceived Restorativeness Soundscape Scale (PRSS) developed by Payne [55]. The number of the original questions in PRSS was nineteen but the present study selected six items which had high factor loadings [55] to reduce the amount of time for the experiment and thereby to minimise the participants' fatigue and loss of concentration. The items measured fascination, being-away-to, being-away-from, and compatibility of the soundscapes. Table 3 presents the question items used for measuring psychological recovery. All question items used 11-point scales (0 = *"least"* and 10 = *"most"*). The six PRSS items were not presented to the participant when the No-Audio stimuli were presented. In other words, the PRSS were only asked when the audio-visual stimuli were presented.

Table 3. The questionnaire used for evaluating psychological recovery. PRSS items were only asked when the audio-visual stimuli were presented.

Please rate the tranquillity of the last scene.		
Please rate the pleasantness of the last scene.		
Please rate your preference for the last scene.		
PRSS:	Fascination:	I find this sonic environment appealing.
		These sounds make me want to linger here.
		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.
	Being- Away-From:	When I hear these sounds I feel free from work, routine and responsibilities.
	Compatibility:	This sonic environment fits with my personal preferences.

Five physiological responses were recorded throughout the experiments: two facial electromyography data (fEMG), heart rate (HR), respiration rate (RR), and electrodermal activity (EDA). All physiological signals were acquired using the Biopac MP150 physiological data acquisition system and AcqKnowledge 4.4 (BIOPAC Systems). Firstly, fEMG signals were measured using five electrodes placed over facial muscles, associated with two different emotion expressions. In particular, positive emotion was expected to activate zygomatic muscle which pulls up the cheek, while negative emotion was expected to activate corrugator supercilii muscle which raises the inner eyebrow [56, 57]. Secondly, a photoplethysmography (PPG) sensor was attached to one finger for measuring HR. The periodicity of the PPG signal represents HR and it has been known to be easy to use, low cost, and convenient measure [58, 59]. Thirdly, a transducer belt was worn around the chest in order to measure RR. Lastly, two electrodes were attached to two adjacent fingers for measuring EDA. Finger has been known to be one of the most responsive locations on the body when it comes to measuring EDA [60] and this study chose to attach two fingers to measure the signal based on previous studies [61, 62]. HR, RR, and EDA were expected to measure arousal responses evoked and recovered by the stressors and the recovery stimuli [63, 64]. The VR headset used in the present study was not wireless so its cable might have touched the fEMG electrodes' cables by any chance; thus, all the electrodes' cables were tightly fixed on the participant's arm using tapes. The participants were also asked to avoid large body movements. In addition, the researcher continuously monitored the participant via a monitor outside the chamber and marked large and unusual body movements that possibly affected the physiological data (e.g., cough or sneeze, yawning, and any body movement which might touch the electrode cables)."

2.6 Procedure

The experiment was carried out in a test chamber where the background noise level was approximately 25 dBA. The dimensions of the chamber were 2.8 m (width) × 1.8 m (length) × 2.4 m (height). The participants took part in the experiment individually. A participant information sheet and a written consent form were provided to each participant upon his/her arrival, and only those who provided consent took part in the experiment. Before obtaining informed consent, the researcher explained the purpose of the study and provided answers to the participant's questions. The participant was assured of complete anonymity and was given pseudonym. All the electrodes were attached to the participant's face, fingers, and chest once the participant gave their consent to take part in the experiment. The participant was then asked to respond to the paper questionnaire which contained question items regarding their age, gender, liking of horror movies, frequency of visits to countrysides, and noise sensitivity. The time for the participant to respond to the questionnaire allowed the gel on each electrode to be fully absorbed into the skin before the experiment commenced. Once the participant completed the questionnaire, he/she was helped to be seated comfortably on a chair inside the test chamber. The researcher then checked whether all electrodes were connected and acquiring physiological data properly using an electrode impedance checker and the data using the data acquisition software. Next, the participant was helped to wear headphones which were used for presenting sounds. The VR headset and a 23-inch monitor screen were used for presenting the videos in the VR and Screen conditions, respectively. The room light was turned off to allow the participant to focus on the videos and avoide visual distraction. The participants sat around 60 cm away from the screen and the brightness of the screen matched the light levels of VR conditions, varying approximately 130 lux to 180 lux. A training session was carried out before the actual experimental session began in order to help the participant to get used to the environment and the measurements, as well as to double-check if all the experiment settings were working properly. There was break time in the middle of the experiment, particularly when the participants needed to switch the Visual Reproduction conditions from the VR condition to the Screen condition, or vice versa. During the break time, the researcher helped the participants to take off/on the HMD used for the VR and turned on/off the screen used for the Screen condition. In addition, the participants were free to ask for additional rest anytime they wanted. There was a camera inside the chamber so that the researcher could monitor the participant on a screen outside the chamber, particularly for the safety reasons.

2.7 Data analysis

Any erroneous data were discarded before the analysis [65]. The respiratory irregularities were used for judging and removing the remnant artifacts in the EDA and HR [66]. Due to the variations of the participants' physiological responses, percentage changes (%) of the physiological responses were calculated [67]. Wilcoxon signed-rank tests were performed in order to test the significance of differences between the responses (e.g., recovery evoked by Rural and Urban). The repeated measures analyses of variance (RM ANOVA) were used to evaluate the main effects of the different conditions on the responses (e.g., VR vs. Screen). Mann–Whitney U tests were conducted to compare the responses between two groups (e.g., responses between age groups). The present study considered *p* values of less than 5% ($p < 0.05$) as statistically significant.

3 Results

3.1 Psychological recovery

Overall, the psychological recovery ratings showed similar patterns. Thus, factor analysis was carried out to test whether all the ratings loaded onto a common factor. All the ratings were found to be loaded onto a common factor (Table 4). In order to present the results in a simple and coherent way, this paper only presents the mean ratings of the perceived preference; the other ratings can be found in the Supplementary Figure S3. As shown in Figure 3, the preference for the Rural (water) was the highest and that for the Urban was the lowest in both VR and Screen conditions. The ratings were significantly different across the scenes in each condition. In the VR condition, Wilcoxon signed-rank tests indicated that the preference score of the Rural (water) was significantly higher than that of the Rural (no water) (Z = -2.81, p < .01) and that of Urban (Z = -4.05, p < .01). In addition, the preference score of the Urban was significantly lower than that of the Rural (no water) (Z = -3.79, p < .01). In the Screen condition, it was found that the preference score of the Rural (water) was significantly higher than that of the Urban (Z = -4.01, p < .01). Further, the score of the Rural (water) in the VR condition was significantly higher than that of the Rural (water) in the Screen condition (Z = -2.11, p < .05).

Table 4. The results of the factor analysis of the self-reported psychological recovery.

Audio presentation	Visual reproduction	Scene	Rating	Factor loading*	Cronbach's Alpha
Audio-Visual	VR	Rural (no water)	Tranquillity	0.929	0.903
			Pleasantness	0.943	
			Preference	0.907	
		Rural (water)	Tranquillity	0.94	0.904
			Pleasantness	0.914	
			Preference	0.823	
		Urban	Tranquillity	0.956	0.955
			Pleasantness	0.896	
			Preference	0.960	
	Screen	Rural (water)	Tranquillity	0.961	0.922
			Pleasantness	0.949	
			Preference	0.888	
		Urban	Tranquillity	0.924	0.941
			Pleasantness	0.964	
			Preference	0.906	
No-Audio	VR	Rural (no water)	Tranquillity	0.929	0.909
			Pleasantness	0.952	
			Preference	0.849	
		Rural (water)	Tranquillity	0.946	0.917
			Pleasantness	0.934	
			Preference	0.876	
		Urban	Tranquillity	0.929	0.936
			Pleasantness	0.940	
			Preference	0.903	
	Screen	Rural (water)	Tranquillity	0.900	0.824
			Pleasantness	0.936	
			Preference	0.725	
		Urban	Tranquillity	0.721	0.83
			Pleasantness	0.971	
			Preference	0.896	

*Extraction method: Principal Component Analysis; Rotation method: Varimax with Kaiser Normalization

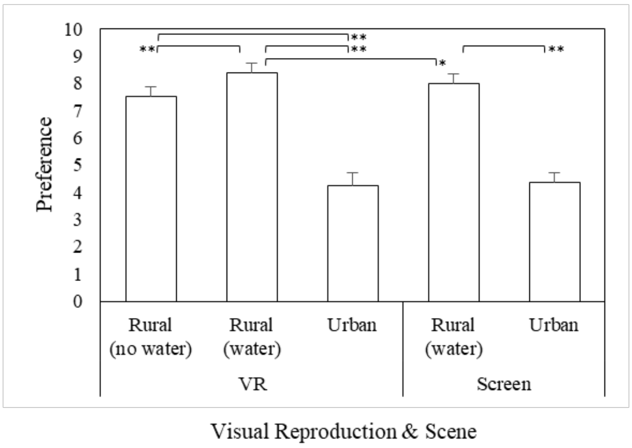

Figure 3. Mean ratings of the perceived preference for the soundscapes in VR and Screen conditions (*$p < 0.05$; **$p < 0.01$). The error bars indicate standard errors.

Figure 4 compares the mean ratings of the perceived preference for the recovery clips when the stimuli were presented as Audio-Visual and No-Audio in the VR and Screen conditions. As shown in Figure 4(a), the mean preference for the Rural (no water) and the Urban did not change much with the different audio presentation conditions in VR. However, the mean preference for Rural (water) was significantly greater when it was presented as Audio-Visual compared to No-Audio. Wilcoxon signed-rank tests indicated that the preference score of the Rural (water) in the Audio-Visual condition was significantly higher than that in the No-Audio condition ($Z = -3.64$, $p < .01$). Figure 4(b) also presents similar tendencies in the Screen condition. It was found that the preference score of the Rural (water) in the Audio-Visual condition was significantly higher than that in the No-Audio condition ($Z = -2.47$, $p < .05$). The results of the other ratings compared between the different conditions of audio presentation can be found in the Supplementary Figure S4.

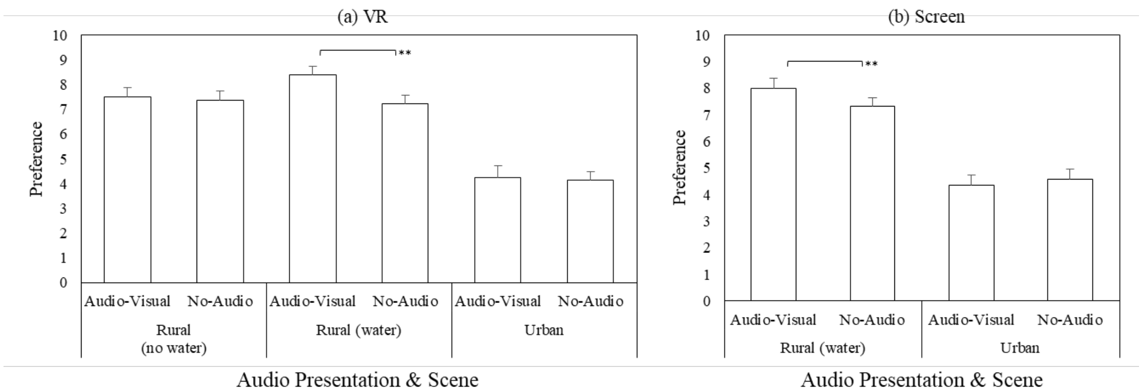

Figure 4. Mean ratings of the preference for the recovery clips in different conditions of audio presentation and visual reproduction (*$p < 0.05$; **$p < 0.01$). The error bars indicate standard errors.

The results showed that Rural (water) evoked greater psychological recovery in the VR condition. The study assessed the main effects of the different experimental conditions by testing RM ANOVA of 2 auditory conditions (Audio-Visual vs. No-Audio) × 2 visual conditions (VR vs. Screen) × 2 scenes (Rural (water) vs. Urban). The results of the main and interaction effects on each of the psychological recovery ratings are presented in Table 5. The scene had significant main effects on all the ratings with large effects sizes. For example, it had a main effect on tranquillity [F (1, 31) = 93.32, $p < 0.0001$] with a large effect size ($\eta p^2 = 0.75$). However, the visual reproduction did not have any significant main effect on the psychological recovery, while the auditory presentation showed significant main effects and large effect sizes on pleasantness [F (1, 31) = 5.63, $p = 0.024$, $\eta p^2 = 0.15$] and preference [F (1, 31) = 6.88, $p = 0.013$, $\eta p^2 = 0.18$]. The interaction between the visual reproduction and the scene had significant impacts on tranquillity and pleasantness. The interaction between the audio presentation and the scene also showed significant effects on all the measured ratings with large effect sizes. However, the interaction between the audio presentation and the visual reproduction and the interaction between all the three conditions (i.e. audio, visual, scene) did not have significant impacts on the ratings.

Table 5. The results of the RM ANOVA showing the main and interaction effects on the psychological recovery.

Source	Psychological recovery	*df*	*F*	*p*	η_p^2
Auditory presentation	Tranquillity	1	4.048	0.053	0.116
(Audio-Visual vs. No-Audio)	Pleasantness	1	5.626	0.024	0.154
	Preference	1	6.880	0.013	0.182
Error		31			
Visual reproduction	Tranquillity	1	0.372	0.546	0.012
(VR vs. Screen)	Pleasantness	1	2.250	0.144	0.068
	Preference	1	0.128	0.722	0.004
	PRSS: Fascination	1	0.035	0.854	0.001
	PRSS: Being-away-to	1	0.061	0.806	0.002
	PRSS: Being-away-from	1	0.168	0.685	0.005
	PRSS: Compatibility	1	0.309	0.582	0.010
Error		31			
Scene	Tranquillity	1	93.322	0.000	0.751
(Rural with water vs. Urban)	Pleasantness	1	64.832	0.000	0.677
	Preference	1	42.792	0.000	0.580
	PRSS: Fascination	1	57.746	0.000	0.651
	PRSS: Being-away-to	1	43.947	0.000	0.586
	PRSS: Being-away-from	1	81.256	0.000	0.724
	PRSS: Compatibility	1	63.835	0.000	0.673
Error		31			
Audio x Visual	Tranquillity	1	0.023	0.882	0.001
	Pleasantness	1	0.075	0.785	0.002
	Preference	1	2.940	0.096	0.087
Error		31			
Audio x Scene	Tranquillity	1	7.476	0.010	0.194
	Pleasantness	1	12.027	0.002	0.280
	Preference	1	6.830	0.014	0.181
Error		31			
Visual x Scene	Tranquillity	1	4.159	0.050	0.118
	Pleasantness	1	7.528	0.010	0.195
	Preference	1	2.417	0.130	0.072
	PRSS: Fascination	1	0.123	0.728	0.004
	PRSS: Being-away-to	1	0.069	0.794	0.002
	PRSS: Being-away-from	1	0.048	0.828	0.002
	PRSS: Compatibility	1	0.320	0.575	0.010
Error		31			
Audio x Visual x Scene	Tranquillity	1	0.003	0.956	0.000
	Pleasantness	1	0.014	0.905	0.000
	Preference	1	0.109	0.744	0.004
Error		31			

Most of the psychological ratings showed no significant difference between the groups of age (20s vs. 30s), gender (male vs. female), the degree of horror movie liking (low vs. high), the frequency of countryside visits (low vs. high), and noise sensitivity (low vs. high). However, some measurements showed significant differences between the groups. For example, Mann-Whitney test indicated that the tranquillity rated to Rural (water) in the Audio-Visual and VR conditions was higher for those who had frequently visited countryside (*Median* = 10) than those who had not (*Median* = 8) ($U = 45.0$, $p = 0.002$).

3.2 Physiological recovery

Similar to the results of the psychological ratings, there were common tendencies in the patterns of the five physiological measurements. Figure 5 shows three physiological responses (two fEMGs and HR) measured in the Audio-Visual and VR conditions. The whole physiological responses measured in all the experimental conditions can be found in the Supplementary Figure S5. First, fEMG of the corrugator supercilli muscle increased during the presentation of the stressor clips, indicating negative emotion [56], and decreased during the presentation of the recovery clips. Among the three recovery clips, the Rural (water) and the Urban showed the quickest and slowest recoveries, respectively. However, there was no significant difference between the responses. Second, fEMG of the zygomatic muscle increased during the presentation of the stressor clips. Although the increasing fEMG response of this muscle is generally associated with positive emotions [56], this result is mainly because many participants moved their mouths when they were frightened by the startling visual or acoustic cues in the stressors. During the recovery, this muscle increased much further with the Rural (no water) and the Rural (water) by evoking positive emotions. In particular, the Rural (water) evoked significantly bigger increase than the Rural (no water). On the other hand, presentation of the Urban dropped the fEMG of the zygomatic muscle even below the stress response. This indicates that the audio-visual presentation of the rural setting enhanced positive emotions and those in the built/urban environment decreased positive emotions. Third, the stressors increased the HR demonstrating the arousal responses [63], and the response decelerated during the recovery. In general, the Rural (water) led to the quickest physiological recovery in all of the experimental conditions.

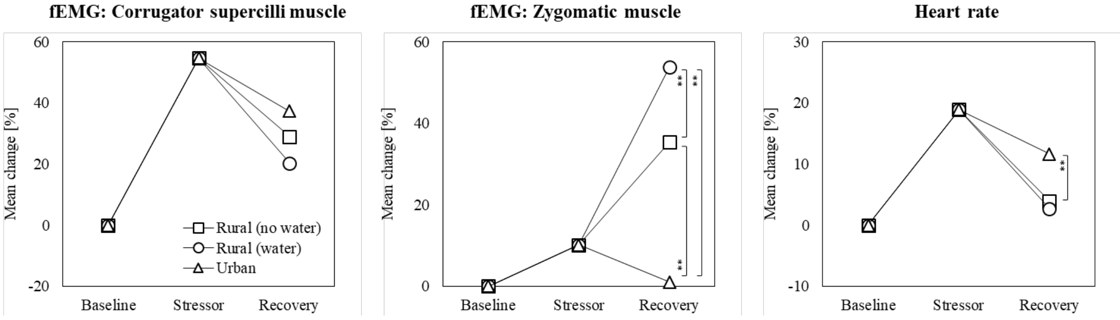

Figure 5. Two fEMGs (corrugator supercilli muscle and zygomatic muscle) and HR changes between baseline, stressor, and recovery clips measured in the Audio-Visual condition in VR (* $p < 0.05$; ** $p < 0.01$).

The main effects of the experimental conditions on the physiological responses were tested with RM ANOVA of 2 auditory conditions (Audio-Visual vs. No-Audio) × 2 visual conditions (VR vs. Screen) × 2 scenes (Rural (water) vs. Urban). The results of the main and interaction effects on each of the physiological responses are presented in Table 6. The audio presentation showed significant impacts on fEMG of zygomatic muscle [F (1, 31) = 53.41, $p < 0.01$, ηp^2 = 0.63] and RR [F (1, 31) = 4.55, p = 0.041, ηp^2 = 0.13]. It was observed that the visual reproduction had a main effect on fEMG of zygomatic muscle [F (1, 31) = 26.52, $p < 0.01$, ηp^2 = 0.46]. It was also found that the scene had significant main effects on fEMG of corrugator supercilli muscle [F (1, 31) = 17.94, $p < 0.01$, ηp^2 = 0.37], HR [F (1, 31) = 6.77, p = 0.014, ηp^2 = 0.18], and RR [F (1, 31) = 18.39, $p < 0.01$, ηp^2 = 0.37]. Moreover, fEMG of zygomatic muscle was significantly affected by all the interaction conditions. The interaction between the audio presentation and the scene also showed a significant impact on HR.

Table 6. The results of the RM ANOVA showing the main and interaction effects on the physiological recovery.

Source	Physiological recovery	*df*	*F*	*p*	η_p^2
Auditory presentation	fEMG-CS (corrugator supercilli)	1	2.519	0.123	0.075
(Audio-Visual	fEMG-Zygo (zygomatic)	1	53.411	0.000	0.633
vs. No-Audio)	HR	1	1.322	0.259	0.041

	RR	1	4.550	0.041	0.128
	EDA	1	1.252	0.272	0.039
Error		31			
Visual reproduction (VR vs. Screen)	fEMG-CS	1	1.628	0.212	0.050
	fEMG-Zygo	1	26.524	0.000	0.461
	HR	1	2.749	0.107	0.081
	RR	1	0.422	0.521	0.013
	EDA	1	0.907	0.348	0.028
Error		31			
Scene (Rural with water vs. Urban)	fEMG-CS	1	17.939	0.000	0.367
	fEMG-Zygo	1	3.313	0.078	0.097
	HR	1	6.773	0.014	0.179
	RR	1	18.389	0.000	0.372
	EDA	1	1.159	0.290	0.036
Error		31			
Audio x Visual	fEMG-CS	1	2.423	0.130	0.072
	fEMG-Zygo	1	41.864	0.000	0.575
	HR	1	1.040	0.316	0.032
	RR	1	0.198	0.659	0.006
	EDA	1	1.075	0.308	0.034
Error		31			
Audio x Scene	fEMG-CS	1	1.001	0.325	0.031
	fEMG-Zygo	1	31.548	0.000	0.504
	HR	1	4.422	0.044	0.125
	RR	1	0.583	0.451	0.018
	EDA	1	0.883	0.355	0.028
Error		31			
Visual x Scene	fEMG-CS	1	0.328	0.571	0.010
	fEMG-Zygo	1	4.336	0.046	0.123
	HR	1	0.028	0.868	0.001
	RR	1	0.137	0.714	0.004
	EDA	1	0.966	0.333	0.030
Error		31			
Audio x Visual x Scene	fEMG-CS	1	2.717	0.109	0.081
	fEMG-Zygo	1	7.367	0.011	0.192
	HR	1	0.245	0.624	0.008
	RR	1	0.003	0.960	0.000
	EDA	1	1.139	0.294	0.035
Error		31			

Similar to the psychological recovery, most of the physiological responses showed no significant difference between the groups of age (20s vs. 30s), gender (male vs. female), the degree of horror movie liking (low vs. high), the frequency of countryside visits (low vs. high), and noise sensitivity (low vs. high). Only some showed significant differences between the groups. For example, Mann-Whitney test indicated that the RR response to the Urban in the Audio-Visual and VR conditions was higher for those who had frequently visited countryside (*Median* = 6.5) than those who had not (*Median* = -2.0) (U = 61.0, p = 0.013).

Figure 6 shows the general tendencies of the physiological responses changed from the stressor clips to the recovery clips across the different experimental conditions. In general, the presentation of the Rural (water) evoked strong recovery than others, whereas Urban had weak effects on the recovery. Moreover, most physiological recovery were greater in the Audio-Visual and VR conditions. Specifically, the fEMG changes in the corrugator supercilli muscle were the largest when the rural scenes were presented in the Audio-Visual and VR conditions. The fEMG changes in the corrugator supercilli muscle were greater in the VR than the Screen condition. The fEMG of the zygomatic muscle also showed the greatest increase when the Rural (water) was presented in the Audio-Visual and VR conditions. Rural (no water) in the VR condition and Rural (water) in the Screen condition also showed strong recoveries when the sounds were presented. On the other hand, the Urban decreased the fEMG of the zygomatic muscle in the VR condition regardless of the audio presentation. HR showed the largest decelerations in the VR condition when the rural scenes were presented with the sounds. The smallest change in HR was found when the Urban was presented as No-Audio in the Screen condition. The changes in the RR and the EDA can be seen in Supplementary Figure S6.

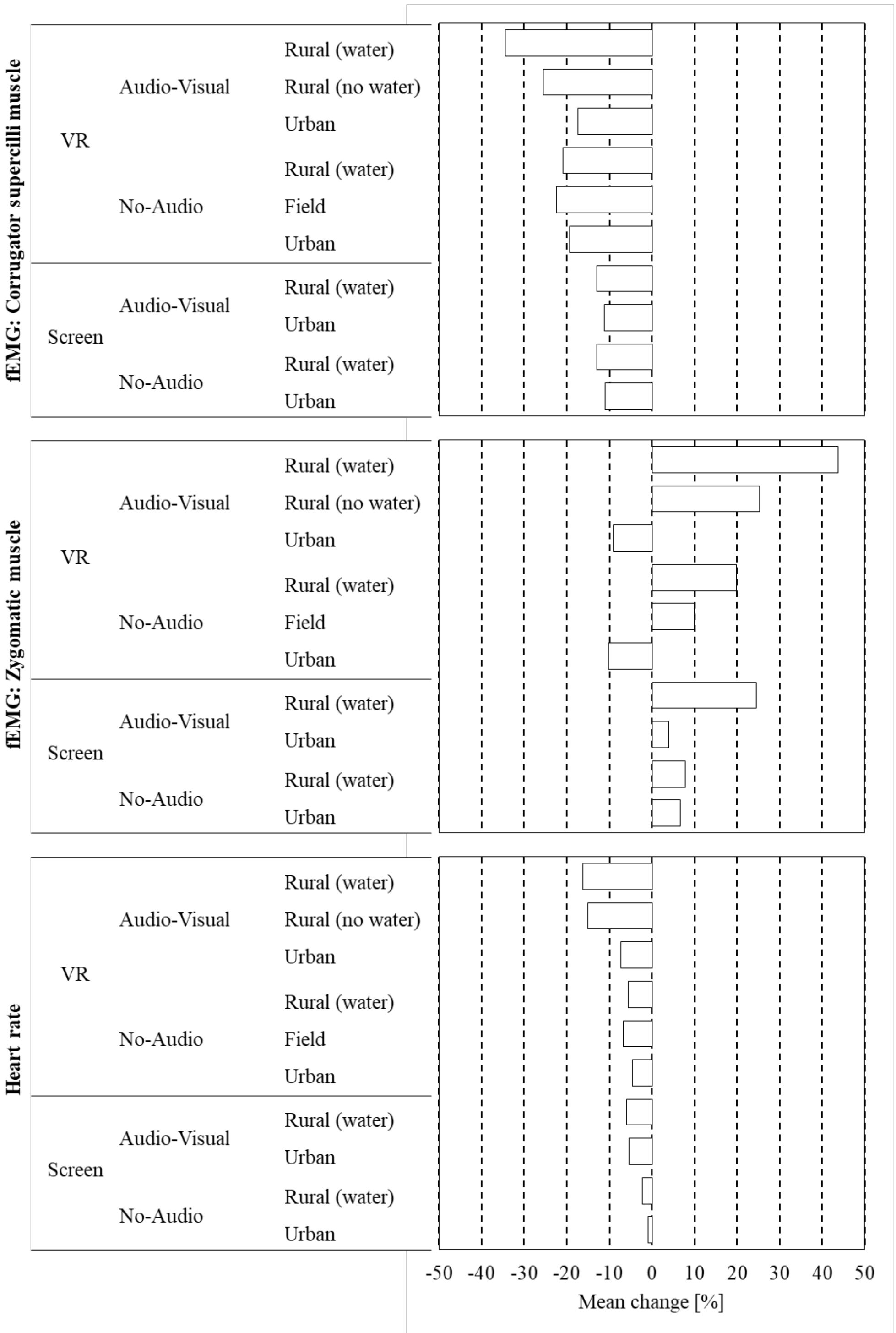

Figure 6. Two fEMGs (corrugator supercilli muscle and zygomatic muscle) and HR changes from the stressor clips to the recovery clips.

4 Discussion

The results of this study supported the previous findings of theoretical frameworks such as SRT and ART [1, 2]. Compared to the urban environment, the rural environments evoked greater psycho-physiological recovery. This result can be explained as a better stress recovery by adopting the approach of the SRT [2] or as a less directed-attention through the lens of the ART [1]. In other words, the urban environment would have led worse stress recovery and required more directed-attention and thereby the participants might have experienced a significant amount of cognitive effort and directed attention fatigue [68].

It has been known that soundscapes have significant impacts on restoration and recovery [69]. This study examined the effects of the audio presentation on the psycho-physiological responses. In general, the rural settings had a better recovery when they were presented as Audio-Visual. In particular, the psychological recovery rated to the Rural (water) was significantly high when the sound was presented in both VR and Screen conditions. Moreover, the Urban evoked greater psycho-physiological recovery when it was presented as No-Audio compared to the Audio-Visual presentation. This result supports what the ART has suggested on the directed-attention as well as the effects of the audio presentation. Since the urban setting contained various noise sources such as noise from the traffic and pedestrians, the absence of noise might have influenced the participants to take less directed-attention [1, 70]. Hence, unlike the rural setting, the urban setting evoked a better recovery when it was presented as No-Audio. In the present study, the Audio-Visual stimuli led to a greater psycho-physiological recovery than the No-Audio stimuli. This implies that the exposure to soundscape is more likely to evoke a greater psycho-physiological recovery compared to landscape presentation without any sound and potentially mere acoustic environment without any visual cue.

This study also examined the effects of the visual reproduction on the psycho-physiological recovery. The results showed greater physiological recovery in VR compared to the Screen condition. Contrary to the physiological responses, the psychological responses between the VR and Screen conditions were not much different. It can be discussed in the following ways. This needs to be further explained by looking into the differences between the psychological responses and the physiological responses. It is known that physiological responses are useful parameters for assessing emotions, particularly when the perceiver is unaware of them [71]. Therefore, the participants' physiological responses could be changed even though they are not aware of the changes in their emotions while their psychological ratings could be similar to each other even though they were measured in different experimental settings (i.e. VR vs Screen). On the other hand, a completely opposite assumption can be made that psychological responses might be more reliable than physiological measures in assessing recovery. This assumption has been made mainly because subtle emotional feelings may not be determined through physiological measurements due to the influences of physical activities [72]. For instance, in the present study, the participants looked around fairly often during the presentation of the stressor clips in the VR condition, whereas their body movements considerably slowed down during the presentation of the urban or rural scenes. Therefore, it is arguable that the bigger physiological changes in VR compared to the Screen condition could be caused by more frequent, faster, or bigger body movements made by the participants. Previous studies have argued that physiological data should be collected by considering measures to detect and remove motion artifacts [73, 74]. Particularly, the potential impacts of participants' head/body movements on their physiological data would be more critical in VR conditions [75]. However, this study recorded only unusual and large movements to remove potential artifacts, whereas minor movements were not recorded. Thus, it was not possible to assess the effects of the body movements on the physiological responses thoroughly. Future research could further investigate the differences in the physiological recovery between the VR and Screen conditions by evaluating this association. Furthermore, this study used only six of nineteen PRSS items [55] in order to reduce the amount of time for the experiment. It would be of worth for future research to use the full nineteen items of PRSS and examine whether the psychological recovery are different between the VR and Screen conditions.

The study also investigated the recovery effects of each of the scenes. It was found that the rural setting with a water stream had greater stress recovery impacts than the rural with a green open-field in VR. Some psychological and physiological responses showed significant differences between the Rural (water) and the Rural (no water) such as the perceived preference and the fEMG of the zygomatic muscle. These results supported the previous findings that the sound of water was one of the most preferred sounds and it resulted in greater psychological restorations [15, 34]. Contrary to the Rural (water) which evoked significantly a higher psychological recovery when the sound was presented together, the Rural (no water) did not show any significant difference between Audio-Visual and No-Audio conditions. This difference might be related to the characteristics of the sound sources in each stimulus. The major two sound sources heard in the Rural (water) were gentle breeze and water flowing. The two major sources in the Rural (no water) were gentle breeze and some sheep baaing. Unlike the sound sources in the Rural (water), the sound of sheep baaing heard in the Rural (no water) was intermittent as it was occasionally heard. This might have evoked higher directed-attention [1] with the participants and caused weaker recovery. Moreover, an audio-visual congruency might also have had an impact. Westman and Walters [76] earlier suggested that the combination of the audio and visual cues should make sense otherwise it may potentially evoke negative reactions. In the Rural (water), the view of the water stream was congruent with the water sound. However, the sound of sheep baaing was heard without the visual presence of the sheep in the Rural (no water). In a recent review paper, van Renterghem [77] addressed that people generally expect audio-visual congruency so that the source visibility plays a significant role in perceptions of sonic environments. Thus, a lack of congruency between the acoustic and visual cues in the Rural (no water) might have negatively affected the recovery responses. Future research could further test the effects of the visibility of sound sources.

Noise sensitivity has been known to be one of the most significant factors affecting subjective responses to noise including psychological [78, 79] and physiological responses [80, 81]. For instance, those who are sensitive to noise report greater noise annoyance [78] and more frequent sleep disturbance [82]. The present study compared the psychological and physiological recovery responses between the low and high noise sensitivity groups and did not find any difference between the groups. Noise annoyance may have negative impacts on the evaluation of soundscape quality [83]; however, it doesn not mean that one's self-reported noise sensitivity can predict his/her restorative experience in different soundscapes. Noise sensitivity is a moderating factor between noise and annoyance but there has been no evidence that it also moderates the relationship between soundscape quality and psycho-physiological restorations. Previous studies have also suggested other measures affecting the perception of soundscape quality instead of noise sensitivity. For example, van Kamp and Davies [84] introduced environmental sensitivity as a larger construct, representing not only sensitivity to noise but sensitivities to various environmental elements. Thus, it would be valuable to examine whether other measures such as environmental sensitivity moderates the relationships between soundscape quality and psycho-physiological restorations in various conditions. Kliuchko et al. [85] reported that noise sensitivity associates with passive/background listening to music but does not correlate with the amount of active/attentive listening to music. Their findings can be further extended to an understanding of the relationship between noise sensitivity and psychological recovery effect of soundscape. The participants in this study responded to the question items about the perceived recovery toward each soundscape so that they listened to the sounds actively. Since they were more focused on the sound itself, noise sensitivity might not show a significant association with the recovery responses.

There are limitations in this study which need to be discussed. First of all, as Southworth [86] said, for the blind, sound is an important way to obtain information about the world. He also noted that our ears are less effective than our eyes in terms of gathering information and thus, the blind may construct the image of the world simpler than those who see the world by sight. In this sense, additional experiment with the audio-only condition would have given far more interesting insights in terms of stress recovery in the future. Second, the videos used in this study were either recorded by the authors or downloaded from online. All the video file formats were identical. Although the audio recordings were conducted in B-format using a Soundfield microphone (ST450 MkII) and thereby ambisonic reproductions were available, all sounds were reproduced as the static binaural signals because the sounds of the videos downloaded from online were only static binaural signals. It is well known that the head-tracked binaural signals and multi-channel/ambisonic loudspeaker reproductions are more immersive and realistic than the static binaural reproduction [87]. Thus, future research on soundscape may consider adopting other sound reproduction systems. Third, sounds were reproduced via headphones in this study. However, sound reproduced through headphones is likely to be different from that reproduced through loudspeakers. In particular, sound reproduced with headphones is generally localised inside the listener's head [88]. Future VR experiments may adopt multi-channel loudspeaker arrays to reproduce sounds and to confirm the results. Lastly, the video clips presented for measuring recovery response in this study depicted either rural or urban settings. The rural settings barely presented features of built environment while the urban setting barely showed features of nature. Thus, future research would find further insights by systematically addressing various qualitative and quantitative factors of different soundscapes. For instance, the research may use urban parks with different features (e.g., large or small and loud or quiet).

5 Conclusions

This study investigated the effects of the urban and rural settings on the psycho-physiological recovery responses. Each participant was exposed to a series of stimuli which contained a one-minute baseline clip, a one-minute stressor clip, and a one-minute recovery clip each. The stressor clips presented horror movie-clips and the recovery clips showed either rural or urban scenes. The stimuli were presented in (1) Audio-Visual and No-Audio conditions and (2) VR and Screen conditions. The study examined the impacts of these conditions on the psychological and physiological responses. Participants rated their perceived recovery towards each of the recovery clips. Their physiological responses were recorded during the whole experiment. The results showed that the rural setting evoked a better psycho-physiological recovery compared to the urban setting. Among the rural settings, those with water features led to a greater recovery than the rural setting with a green open-field. The VR environment generally had significant main effects on the physiological responses but not much on the psychological responses. On the other hand, the audio presentation mainly had more significant impacts on the psychological responses than the physiological responses. The study also discovered that some personal factors such as noise sensitivity and frequency of visiting the countryside had very limited impacts on the recovery responses. The findings from this study supported the theoretical frameworks on the restoration (i.e. SRT and ART). The study also extended the existing suggestions of the different recovery effects of nature and urban environments and the recovery effects of the water soundscape. Moreover, this study added further evidence on the effectiveness of the VR environment and the audio presentation on either psychological or physiological recovery or on both. The findings from this study would also be useful to the landscape planners and sound designers to enhance the understanding of the role of sounds in urban and rural environments.

Acknowledgements

This research was funded by the Early Career Researcher and Returners Fund of the University of Liverpool. We thank Dr Ralph Pawling at the School of Natural Sciences and Psychology, Liverpool John Moores University for advising fEMG measurements.

References

[1] Kaplan S. The restorative benefits of nature: Toward an integrative framework. J Environ Psychol. 1995;15:169-82.

[2] Ulrich RS, Simons RF, Losito BD, Fiorito E, Miles MA, Zelson M. Stress recovery during exposure to natural and urban environments. J Environ Psychol. 1991;11:201-30.

[3] Joye Y, van den Berg AE. Restorative environments. Environmental Psychology: An introduction. 2018:65-75.

[4] Ulrich RS. Aesthetic and Affective Response to Natural Environment. In: Altman I, Wohlwill JF, editors. Behavior and the Natural Environment. Boston, MA: Springer US; 1983. p. 85-125.

[5] Ulrich RS. Visual landscapes and psychological well- being. Landscape Research. 1979;4:17-23.

[6] Ulrich RS. Biophilia, biophobia, and natural landscapes. In: Kellert SR, Wilson EO, editors. The Biophilia Hypothesis. Washington, D.C.: Island Press; 1993. p. 73-137.

[7] Kaplan R, Kaplan S. Well-being, Reasonableness, and the Natural Environment. Applied Psychology: Health and Well-Being. 2011;3:304-21.

[8] Wells NM. At home with nature: Effects of “greenness” on children’s cognitive functioning. Environ Behav. 2000;32:775-95.

[9] Hartig T, Evans GW, Jamner LD, Davis DS, Gärling T. Tracking restoration in natural and urban field settings. J Environ Psychol. 2003;23:109-23.

[10] Gidlow CJ, Jones MV, Hurst G, Masterson D, Clark-Carter D, Tarvainen MP, et al. Where to put your best foot forward: Psycho-physiological responses to walking in natural and urban environments. J Environ Psychol. 2016;45:22-9.

[11] Veitch JA, Gifford R. Assessing Beliefs about Lighting Effects on Health, Performance, Mood, and Social Behavior. Environ Behav. 1996;28:446-70.

[12] Cameron P, Robertson D, Zaks J. Sound pollution, noise pollution, and health: Community parameters. Journal of Applied Psychology. 1972;56:67.

[13] Öhrström E, Skånberg A, Svensson H, Gidlöf-Gunnarsson A. Effects of road traffic noise and the benefit of access to quietness. J Sound Vib. 2006;295:40-59.

[14] Stevens M. Quietening open spaces towards sustainable soundscapes for the City of London. In: Stevens M, editor. UK: Environmental Protection; 2010.

[15] Yang W, Kang J. Soundscape and sound preferences in urban squares: A case study in Sheffield. Journal of Urban Design. 2005;10:61-80.

[16] Alvarsson JJ, Wiens S, Nilsson ME. Stress recovery during exposure to nature sound and environmental noise. Int J Env Res Pub He. 2010;7:1036-46.

[17] de Coensel B, Vanwetswinkel S, Botteldooren D. Effects of natural sounds on the perception of road traffic noise. J Acoust Soc Am. 2011;129:EL148-EL53.

[18] Ratcliffe E, Gatersleben B, Sowden PT. Bird sounds and their contributions to perceived attention restoration and stress recovery. J Environ Psychol. 2013;36:221-8.

[19] Aletta F, Oberman T, Kang J. Associations between Positive Health-Related Effects and Soundscapes Perceptual Constructs: A Systematic Review. Int J Env Res Pub He. 2018;15:2392.

[20] Sun K, Echevarria Sanchez GM, De Coensel B, Van Renterghem T, Talsma D, Botteldooren D. Personal Audiovisual Aptitude Influences the Interaction Between Landscape and Soundscape Appraisal. Frontiers in Psychology. 2018;9.

[21] Erfanian M, Mitchell AJ, Kang J, Aletta F. The Psychophysiological Implications of Soundscape: A Systematic Review of Empirical Literature and a Research Agenda. Int J Env Res Pub He. 2019;16:3533.

[22] Li Z, Kang J. Sensitivity analysis of changes in human physiological indicators observed in soundscapes. Landscape and Urban Planning. 2019;190:103593.

[23] Pheasant RJ, Fisher MN, Watts GR, Whitaker DJ, Horoshenkov KV. The importance of auditory-visual interaction in the construction of 'tranquil space'. J Environ Psychol. 2010;30:501-9.

[24] Jeon JY, Lee PJ, You J, Kang J. Acoustical characteristics of water sounds for soundscape enhancement in urban open spaces. J Acoust Soc Am. 2012;131:2101-9.

[25] Hong JY, Jeon JY. Designing sound and visual components for enhancement of urban soundscapes. J Acoust Soc Am. 2013;134:2026-36.

[26] Ulrich RS, Simons RF, Miles MA. Effects of environmental simulations and television on blood donor stress. Journal of Architectural and Planning Research. 2003;20:38-47.

[27] Kweon B-S, Ulrich RS, Walker VD, Tassinary LG. Anger and Stress:The Role of Landscape Posters in an Office Setting. Environ Behav. 2008;40:355-81.

[28] Burdea GC, Coiffet P. Virtual reality technology. 2 ed. Hoboken, NJ: John Wiley & Sons, Inc.; 2003.

[29] Valtchanov D, Barton KR, Ellard C. Restorative Effects of Virtual Nature Settings. Cyberpsychology, Behavior, and Social Networking. 2010;13:503-12.

[30] Annerstedt M, Jönsson P, Wallergård M, Johansson G, Karlson B, Grahn P, et al. Inducing physiological stress recovery with sounds of nature in a virtual reality forest — Results from a pilot study. Physiology & Behavior. 2013;118:240-50.

[31] Payne SR, Nordh H, Hassan R. Are urban park soundscapes restorative or annoying? European congress and exposition on noise control engineering, Euronoise. Maastricht, Netherlands2015. p. 823-7.

[32] Yu CP, Lee HY, Luo XY. The effect of virtual reality forest and urban environments on physiological and psychological responses. Urban Forestry & Urban Greening. 2018;35:106-14.

[33] Watts GR, Pheasant RJ, Horoshenkov KV, Ragonesi L. Measurement and subjective assessment of water generated sounds. Acta Acust United Ac. 2009;95:1032-9.

[34] Jeon JY, Lee PJ, You J, Kang J. Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds. J Acoust Soc Am. 2010;127:1357-66.

[35] Tedja YW, Tsaih L. Water soundscape and listening impression. J Acoust Soc Am. 2015;138:1750-.

[36] Rådsten-Ekman M, Axelsson Ö, Nilsson ME. Effects of Sounds from Water on Perception of Acoustic Environments Dominated by Road-Traffic Noise. Acta Acust United Ac. 2013;99:218-25.

[37] Zhao J, Xu W, Ye L. Effects of auditory-visual combinations on perceived restorative potential of urban green space. Appl Acoust. 2018;141:169-77.

[38] Masullo M, Pascale A. Effects of combinations of water sounds and visual elements on the traffic noise mitigation in urban green parks. INTER-NOISE and NOISE-CON Congress and Conference. Hamburg, Germany: Institute of Noise Control Engineering; 2016. p. 3910-5.

[39] Ong ZT, Hong JY, Lam B, Gupta R, Gan WS, Chong SH, et al. Effect of masker orientation on masking efficacy in soundscape applications. INTER-NOISE and NOISE-CON Congress and Conference. Hong Kong, China: Institute of Noise Control Engineering; 2017. p. 4916-22.

[40] Cohen J. Statistical Power Analysis. Current Directions in Psychological Science. 1992;1:98-101.

[41] Faul F, Erdfelder E, Lang AG, Buchner A. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods. 2007;39:175-91.

[42] Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods. 2009;41:1149-60.

[43] Campbell MJ, Gardner MJ. Statistics in Medicine: Calculating confidence intervals for some non-parametric analyses. British Medical Journal (Clinical research ed). 1988;296:1454.

[44] Ellermeier W, Eigenstetter M, Zimmer K. Psychoacoustic correlates of individual noise sensitivity. J Acoust Soc Am. 2001;109:1464-73.

[45] Novak CC, La Lopa J, Novak RE. Effects of sound pressure levels and sensitivity to noise on mood and behavioral intent in a controlled fine dining restaurant environment. Journal of Culinary Science & Technology. 2010;8:191-218.

[46] Jeon JY, Lee PJ, Hong JY, Cabrera D. Non-auditory factors affecting urban soundscape evaluation. J Acoust Soc Am. 2011;130:3761-70.

[47] Weinstein ND. Individual differences in reactions to noise: A longitudinal study in a college dormitory. Journal of Applied Psychology. 1978;63:458-66.

[48] Yin J, Levanon D, Chen JDZ. Inhibitory effects of stress on postprandial gastric myoelectrical activity and vagal tone in healthy subjects. Neurogastroenterology & Motility. 2004;16:737-44.

[49] Forman RTT, Gordon M. Landscape Ecology. New York: John Wiley; 1986.

[50] Galbrun L, Calarco FM. Audio-visual interaction and perceptual assessment of water features used over road traffic noise. J Acoust Soc Am. 2014;136:2609-20.

[51] Pheasant R, Horoshenkov K, Watts G, Barrett B. The acoustic and visual factors influencing the construction of tranquil space in urban and rural environments tranquil spaces-quiet places? J Acoust Soc Am. 2008;123:1446-57.

[52] Lambert J, Simonnet F, Vallet M. Patterns of behaviour in dwellings exposed to road traffic noise. J Sound Vib. 1984;92:159-72.

[53] Krzywicka P, Byrka K. Restorative qualities of and preference for natural and urban soundscapes. Frontiers in Psychology. 2017;8:1705.

[54] Zhang Y, Kang J, Kang J. Effects of Soundscape on the environmental restoration in urban natural environments. Noise Health. 2017;19:65.

[55] Payne SR. The production of a perceived restorativeness soundscape scale. Appl Acoust. 2013;74:255-63.

[56] Tassinary LG, Cacioppo JT, Vanman EJ. The skeletomotor system: Surface electromyography. In: Cacioppo JT, Tassinary LG, Berntson GG, editors. Handbook of Psychophysiology. 3rd ed: Cambridge University Press; 2007. p. 267-99.

[57] Carr EW, Winkielman P, Oveis C. Transforming the mirror: Power fundamentally changes facial responding to emotional expressions. Journal of Experimental Psychology: General. 2014;143:997.

[58] Mashhadi MB, Asadi E, Eskandari M, Kiani S, Marvasti F. Heart rate tracking using wrist-type photoplethysmographic (PPG) signals during physical exercise with simultaneous accelerometry. IEEE Signal Processing Letters. 2015;23:227-31.

[59] Aarts LA, Jeanne V, Cleary JP, Lieber C, Nelson JS, Oetomo SB, et al. Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit—A pilot study. Early Human Development. 2013;89:943-8.

[60] van Dooren M, de Vries JJG, Janssen JH. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. Physiology & Behavior. 2012;106:298-304.

[61] Strauss MD. HandWave: design and manufacture of a wearable wireless skin conductance sensor and housing: Massachusetts Institute of Technology; 2005.

[62] Verastegui-Tena L, Schulte-Holierhoek A, van Trijp H, Piqueras-Fiszman B. Beyond expectations: The responses of the autonomic nervous system to visual food cues. Physiology & Behavior. 2017;179:478-86.

[63] Lang PJ, Bradley MM, Cuthbert BN. Motivated attention: Affect, activation, and action. In: Lang PJ, Simons RF, Balaban MT, editors. Attention and Orienting: Sensory and Motivational Processes: Lawrence Erlbaum Associates; 1997. p. 97-135.

[64] Park SH, Lee PJ. Effects of floor impact noise on psychophysiological responses. Build Environ. 2017;116:173-81.

[65] Jennings JR, Gianaros PJ. Methodology. In: Cacioppo JT, Tassinary LG, Berntson GG, editors. Handbook of Psychophysiology. 3rd ed: Cambridge University Press; 2007. p. 812-33.

[66] Schneider R, Schmidt S, Binder M, Schäfer F, Walach H. Respiration-related artifacts in EDA recordings: Introducing a standardized method to overcome multiple interpretations. Psychological Reports. 2003;93:907-20.

[67] Kaiser L. Adjusting for baseline: Change or percentage change? Statistics in Medicine. 1989;8:1183-90.

[68] Hartig T, Kaiser FG, Bowler PA. Further development of a measure of perceived environmental restorativeness (Working Paper No. 5). Institute for Housing Research, Uppsala University1997.

[69] van Kamp I, Klæboe R, Kruize H, Brown AL, Lercher P. Soundscapes, human restoration and quality of life. INTER-NOISE and NOISE-CON Congress and Conference. Hamburg, Germany: Institute of Noise Control Engineering; 2016. p. 1205-15.

[70] Viollon S, Lavandier C, Drake C. Influence of visual setting on sound ratings in an urban environment. Appl Acoust. 2002;63:493-511.

[71] Öhman A, Soares JJF. Unconscious anxiety: Phobic responses to masked stimuli. Journal of Abnormal Psychology. 1994;103:231-40.

[72] Cacioppo JT, Berntson GG, Larsen JT, Poehlmann KM, Ito TA. The psychophysiology of emotion. In: Lewis R, Haviland-Jones JM, editors. The Handbook of Emotion. 2nd ed. Guilford Press2000. p. 173-91.

[73] Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: A complex problem with many partial solutions. Journal of Magnetic Resonance Imaging. 2015;42:887-901.

[74] Islam MK, Rastegarnia A, Yang Z. Methods for artifact detection and removal from scalp EEG: A review. Neurophysiologie Clinique/Clinical Neurophysiology. 2016;46:287-305.

[75] Otsuka S, Kurosaki K, Ogawa M. Physiological measurements on a gaming virtual reality headset using photoplethysmography: A preliminary attempt at incorporating physiological measurement with gaming. TENCON 2017-2017 IEEE Region 10 Conference: IEEE; 2017. p. 1251-6.

[76] Westman JC, Walters JR. Noise and stress: a comprehensive approach. Environmental Health Perspectives. 1981;41:291.

[77] van Renterghem T. Towards explaining the positive effect of vegetation on the perception of environmental noise. Urban Forestry & Urban Greening. 2018.

[78] Öhrström E, Björkman M, Rylander R. Noise annoyance with regard to neurophysiological sensitivity, subjective noise sensitivity and personality-variables. Psychological Medicine. 1988;18:605-13.

[79] Gille LA, Marquis-Favre C, Weber R. Noise sensitivity and loudness derivative index for urban road traffic noise annoyance computation. J Acoust Soc Am. 2016;140:4307-17.

[80] Park SH, Lee PJ, Jeong JH. Effects of noise sensitivity on psychophysiological responses to building noise. Build Environ. 2018;136:302-11.

[81] Stansfeld SA. Noise, noise sensitivity and psychiatric disorder: Epidemiological and psychophysiological studies. Psychological Medicine Monograph Supplement. 1992;22:1-44.

[82] Marks A, Griefahn B. Associations between noise sensitivity and sleep, subjectively evaluated sleep quality, annoyance, and performance after exposure to nocturnal traffic noise. Noise Health. 2007;9:1-7.

[83] Klæboe R, Kolbenstvedt M, Fyhri A, Solberg S. The impact of an adverse neighbourhood soundscape on road traffic noise annoyance. Acta Acust United Ac. 2005;91:1039-50.

[84] van Kamp I, Davies H. Environmental noise and mental health: Five year review and future directions. International congress on noise as a public health problem, ICBEN. Mashantucket, Foxwoords, CT2008.

[85] Kliuchko M, Heinonen-Guzejev M, Monacis L, Gold BP, Heikkilä KV, Spinosa V, et al. The association of noise sensitivity with music listening, training, and aptitude. Noise Health. 2015;17:350-7.

[86] Southworth M. The Sonic Environment of Cities. Environ Behav. 1969;1:49-70.

[87] Hong JY, Lam B, Ong Z-T, Ooi K, Gan W-S, Kang J, et al. Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications. Build Environ. 2019;149:1-14.

[88] Kawaura Ji, Suzuki Y, Asano F, Sone T. Sound localization in headphone reproduction by simulating transfer functions from the sound source to the external ear. Journal of the Acoustical Society of Japan (E). 1991;12:203-16.