DOI: 10.61782/fa.2023.0732

EFFECTS OF VISUAL STIMULUS ON SOUNDSCAPE PERCEPTION: A PILOT STUDY USING IMMERSIVE VIRTUAL REALITY (IVR)

Xinyan Xu[1] **Tin Oberman**[1] **Andrew Mitchell**[1]
Shiqi Huang[1] **Francesco Aletta** [1] **Kang, Jian**[1]

[1] Institute for Environmental Design and Engineering, The Bartlett, University College London, the United Kingdom

ABSTRACT

Due to its rapid development, immersive technologies are being implemented in a growing number of industries. Immersive virtual reality also plays an important role in the field of the built environment, due to its application in visualising imagined spaces, but also recording and reproducing the real-world environments, in laboratory conditions. Since laboratory experiments are one of the essential soundscape research methods where visual context can play an important role, a higher level of immersion is considered crucial to improve the ecological validity and subsequently also the accuracy of the results. Although the impact of visual immersion during soundscape assessments in a laboratory is important, its direct effects on the perceptual models are still underinvestigated. This study proposed a visually immersive approach to soundscape research, which allowed for scene-switching and control, and questionnaire deployment inside a virtual reality environment. In this study, video material from the International Soundscape Database was used. The results have shown a correlation between perceived pleasantness and perceived eventfulness in the absence of the auditory stimulus. In addition, the ratio of road to water in an urban landscape was discovered to be a significant visual factor that influences the perceived soundscape.

Keywords: *Soundscape, Virtual Reality, Visual-Only, Laboratory Experiment.*

1. INTRODUCTION

In contemporary society, the fulfilment of material needs has largely been achieved due to economic and social advancements. Consequently, people's attention has shifted towards the convenience, health, and safety of their living environment on a global scale [1]. However, the increase in urban noise resulting from industrialization and transportation has become a cause for concern [2, 3]. Schafer, a Canadian composer, is often credited with popularising the concept of soundscape, which has since found considerable use in architecture due to the interdisciplinary nature of the field [4]. The relevance of soundscape in architecture is comparable to that of landscape, with the exception that soundscape focuses on the audible aspect, while landscape emphasizes the visual [3]. The International Organization for Standardization (ISO), has proposed an international standard for soundscape research, leading to a broad international consensus on the definition of 'soundscape' [5], which is a sound environment that is perceived or understood by humans in its context. Based on years of experimental research, ISO/TS 12913-2 and ISO/TS 12913-3 [6, 7] have also proposed standardized approaches to the collection of physical data, experimental methods, and emotional evaluation indicators for soundscapes[1].

In recent years, soundscapes have been studied extensively, with the sound source and the surrounding environment being recognized as the two significant physical components of a soundscape [8]. Studies have focused on various acoustic and environmental components, with research conducted both in the field and in the laboratory [1, 9-12]. Experimental findings have been instrumental in designing new soundscapes, enhancing existing sound environments, and formulating new policies [1, 13, 14]. The perception of the environment is frequently multidimensional, and the human senses of hearing, sight, and smell play a vital role in determining participants' evaluations of a soundscape [15, 16]. Participants' evaluations are influenced by the multisensory interaction of the environment, and integrated perceptual awareness of the soundscape helps in better understanding and enhancing future soundscape designs [17]. The audio-visual interaction is a significant factor that affects participants' affective evaluations of environmental sounds [18]. Abdalrahman and Galbrun [19] found that the addition of

[*]Corresponding author: zczlxxu@ucl.ac.uk.

visual stimuli of a water scene via monitor screens enhanced perceived pleasantness at a factor of 2.5.

Two commonly used methods of studying human emotion perception in experimental soundscapes are on-site experiments and laboratory experiments [20, 21]. Field experiments, such as questionnaires, interviews, and soundwalks, are typically employed for data collection in on-site experiments [6]. In contrast, laboratory experiments employ environmental reproduction techniques that eliminate temporal and spatial restrictions, allowing for data collection in a manner similar to that of on-site studies [22]. In addition, outdoor experiments can present a variety of limitations, including requirements for more manpower and dependency on participants' cooperation, as well as the inability to control external factors[10, 17, 21]. Unlike on-site experiments, laboratory experiments can be used to test multiple scenarios on a single participant and for a predetermined period of time [23, 24], providing greater control over exposure conditions and participant sampling. Furthermore, laboratory studies can block the audio to isolate visual stimuli, which cannot be achieved in the field [25]. When reproducing an environment, the results will only be more generalizable if ecological validity is ensured to the fullest extent possible [21, 26, 27]. Although there are many excellent examples of laboratory tests on the reproduction of sound [12, 28], the visual aspect of soundscapes is still an area that requires further research. Visual conditions have been promoted from initial two-dimensional image reproduction to Immersive Virtual Reality (IVR) experiences [18, 29-33]. However, research into the reproduction of realistic scenes, fully immersive experiences, and switching between multiple scenes is still lacking.

The objective of this paper is to investigate the impact of visual stimuli on soundscape perception using an IVR approach in the laboratory. The methods section will provide a detailed description of the approach, and the data analysis will be based on the results obtained from the experiment.

2. METHODOLOGY

The study was carried out in a laboratory setting by reproducing a total of 36 situations across six cities in China, Italy, Spain, and the United Kingdom. The proposed approach involved multi-scene switching in a virtual reality environment complemented with an embedded questionnaire tool.

2.1 Participants

The experiment was conducted with a sample of 32 participants, comprising 16 men and 16 women, recruited from the staff and students of University College London, aged 18 or above. The age distribution of the participants was as follows: 18-28 years (81.25%), 28-38 years (15.6%), and 38-48 years (3.1%). The sample consisted of participants from various ethnic backgrounds, including Caucasian, Asian, and Mixed/Multiple.

Upon arrival at the acoustic laboratory, participants were directed to take a seat in the centre of the laboratory and don a head-mounted display (HMD) while grasping the controller. A training step was implemented to familiarize participants with the virtual reality (VR) environment and controller usage prior to viewing and assessing the 36 videos via the questionnaire deployed in the VR simulations.

2.2 VR environment design

2.2.1 Video footage processing

The video content used in the study was available via the existing International Soundscape Database (ISD), the multimedia part of which was not yet publicly accessible [34]. The selection and processing of video stimuli followed a four-step procedure:

i. The video files were checked to ensure they were usable and in the correct format, with the original field-collected questionnaire data compiled. Here the videos' names were kept the same as the SessionIDs in the ISD.

ii. Ratio of sky, building, road, green and water were selected as contextual factors and were identified using still frames taken from videos and then used to categorise the scenes, as shown in Table 1. When identified, dynamic objects such as people and vehicles were excluded from the categorization. Each site's contextual factors were analyzed using cluster analysis to classify each video into one of four clusters., as in Table 1.

iii. After video selection, the raw footage captured by a 360-degree camera was transformed to equirectangular videos to be applied to spheres in Unity to enable viewing in a VR environment.

iv. To minimize the impact of potential confounding factors on the experimental results, the videos were randomized to prevent repeated sequences and avoid the influence of participants' lack of skill in operation during the early parts of the experiment and physical exhaustion at the end stages of the trial.

Table 1. Scenes selection based on the cluster results

Cluster Group	Description	Videos
1	Broad city streets and squares	CarloV1&CarloV2, PlazaBibRambla1, MonumentoGaribaldi1, SanMarco1, Pingshanstreet1&2, DadongSquare3, OlympicSquare3, ZhongshanPark5, ZhongshanSquare5, SidalinPark1&2, CamdenTown1, CamdenTown2, CamdenTown3, CamdenTown4, TateModern1, TateModern3, TorringtonSq4, PancrasLock1, EustonTap3&2&1, TorringtonSq3&2, StPaulsRow1, PancrasLock2
2	Narrow walkway bordered by buildings and vegetation	MarchmontGarden1, MarchmontGarden2, MarchmontGarden3
3	Semi-open urban park	PingshanPark1&2, StPaulsCross1, RussellSq1, Regents Park Japan_2,
4	Enclosed urban park	LianhuashanParkForest_1, LianhuashanParkForest_2, LianhuashanParkEntrance1, LianhuashanParkEntrance2, RussellSq2, Regents Park Japan_1, RegentsParkFields2

2.2.2 VR environment creation

The experimental setup was implemented using the Unity game engine, which provided a flexible and accessible platform for constructing 3D models and scenes. 36 video scenes were designed to assess using the questionnaire tool, 2 training scenes at the beginning, 38 transition scenes before each video, and 1 scene to collect basic demographic data.

Regarding the interaction design in this experiment, two main scenes were featured, the video scenes with the questionnaire and the transition scene to control for participant's orientation. The video sequence provided pause, play, and questionnaire page functions. The interface in the transitionary scenes for correcting participant's orientation included four panels displaying the back, front, left and right sides relative to the VR space, and a button to play the following video when a participant faces the front view. The interactions were executed using joysticks.

2.3 Data collection

To maintain data compatibility with the previous field experiment, the data collection model for this study was designed to closely resemble that of the ISD. The questionnaire format utilized method A from ISO/TS 12913-2:2018 (ISO, 2018). However, as this experiment focused solely on the visual environment, only part 2 of method A, related to affective quality and the circumplex model, was included in the questionnaire which contained eight sliders to indicate the extent of perception.

While the on-site experiment employed a 5-point Likert scale for data collection, this study opted to use the Visual Analogue Scale (VAS), which recognizes only two poles with equally spaced data points between them. In field research, paper questionnaires are commonly utilized, and the five-point scale, with responses numbered 1 to 5, is simpler to record and process. Conversely, laboratory experiments employ electronic questionnaires, and the VAS facilitates a more detailed delineation of degrees, preventing participants from feeling limited or influenced by marker values. This is considered to lead to higher data accuracy.

The embodied questionnaire was constructed inside the virtual environment. Each affective attribute was represented by a single slider, with the starting position in the middle and 'Strongly Disagree' and 'Strongly Agree' as the left and right endpoints, respectively. The slider consisted of 100 equally spaced points, and participants could use the left and right joystick buttons to adjust the handle's position, indicating how they felt about each attribute. The questionnaire would appear inside the VR environment as in Figure 1 after the participant has experienced the new environment and activated the questionnaire button. The fact that a questionnaire page is visible inside the scene helped participants remain immersed in the environment being assessed. Each participant would offer feedback on 32 scenes, thereby generating data that can be subsequently compared with the on-site results previously collected.

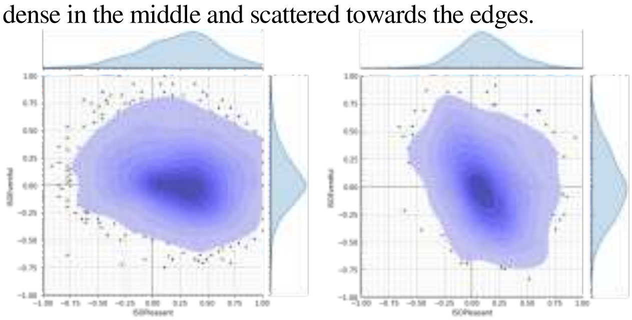
Figure 1. The Questionnaire Page inside the VR environment

3. RESULTS

During the course of one week, data was collected in adherence to established protocols. A total of 1152 responses were received for the 36 surveyed locations, with all responses reflecting genuine perceptions by 32 participants. The International Soundscape Database comprises 3419 valid on-site evaluations for the same locations featured in this experiment.

3.1 Variation in soundscape perception

Following the ISO/TS 12913-3, the raw data obtained from the eight scales were used to calculate ISOPleasant and ISOEventful values. This allowed for the participants' responses to be plotted in the orthogonal system. In this study, the Mitchell et al. soundscape data processing model was used to determine and visualize the values of ISOPleasant and ISOEventful [35]. Results from the same 36 SessionIDs, as collected on-site and in lab, are presented in Figures 3a and 3b, respectively.
The findings demonstrate that there are similarities in the overall trends between on-site and laboratory experiments. Specifically, both ISOPleasant and ISOEventful exhibit a normal distribution in both data sets, as demonstrated in Figure 2. The density plots reveal a similar tendency to be dense in the middle and scattered towards the edges.



(a) On-site experiment (b) Laboratory experiment

Figure 2 The comparison of on-site and laboratory using scatter and density plots

However, there are significant differences between the on-site and laboratory experiments. Figure 3 indicates that the field experiment's overall distribution is more dispersed than that of the laboratory experiment. This could be attributed to the fact that in a multidimensional environment, more components influence emotions, whereas in a visual-only laboratory experiment, evaluations tend to be relatively conservative. On the other hand, there is a large difference in the number of participants, meaning only 32 people assessed two different scenes on-site. The duration of exposure to the scene is also much shorter in laboratory. However, the span of the vertical dimension (ISOEventful) in the laboratory experiment is greater than that of the horizontal dimension (ISOPleasant), while in the field experiment, the opposite is true. In laboratory studies, the sources for evaluating pleasantness are limited, but the degree of serenity may be enhanced by eliminating certain noise components. However, the vividness and attraction of a scene composed solely of pictures are significantly diminished. Participants' experimental experience reveals that the characteristics that could induce negative feelings towards the soundscape were extremely limited in a purely visual setting. As a result, the likelihood of developing a negative perception of the soundscape would be lower in a visual environment than in a combined audio-visual setting, since the source of the noise or unpleasant events is not visible.
In earlier investigations of experiments on audio-visual interaction, there wasn't a result indicate the association between ISOPleasant and ISOEventful. Using the data from both the current laboratory experiment and the previous on-site experiment, correlations between ISOPleasant and ISOEventful were compared. As both of the results follow a normal distribution, Pearson correlation coefficient was calculated. The results are presented in Table 2.

Table 2. Correlation between ISOPleasant and ISOEventful

	On-site	Laboratory
Pearson Correlation	-0.154	-0.313
Sig. (2-tailed)	0.000	0.000

The results presented in Table 2 demonstrate that the significance levels for both the field experiment and the laboratory experiment were lower than 0.05 ($p<0.05$), demonstrating that there was a statistically significant disparity between the two sets of findings. However, the association between pleasantness and eventfulness was only a weak correlation, which is consistent with the results of previous experiments. The correlation coefficient for the laboratory experiment is twice as large as the correlation coefficient for the field experiment. This demonstrates that there is a greater correlation between pleasure and eventfulness with purely visual stimuli than there is in the field trial interaction.
To clarify, the regression analysis found a negative relationship between ISOPleasant and ISOEventful in the laboratory experiment when auditory stimuli were removed. This means that as the level of ISOPleasant increases, the level of ISOEventful decreases, according to the equation ISOEventful = -0.258ISOPleasant + 0.121. It's important to note that this analysis only applies to the laboratory experiment with visual stimuli only, and the relationship between ISOPleasant and ISOEventful may differ in other

contexts, such as the field experiment or experiments with both auditory and visual stimuli presented simultaneously.

Table 3. Regression between ISOPleasant and ISOEventful

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
0.313	0.098	0.097	0.233	1.389

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	6.781	1	6.781	124.599	<0.001
Residual	62.588	1150	0.054		
Total	69.370	1151			

Coefficients

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	0.121	0.007		17.628	<0.001
	-0.258	0.023	-0.313	-11.162	<0.001

3.2 Associations between static objective scene features and perception

Having established a correlation between ISOPleasant and ISOEventful under purely visual conditions, a K-means cluster analysis was conducted on the ISOPleasant and ISOEventful data obtained from the laboratory experiment conducted. According to the elbow plots (Figure 3) derived from the experimental data, clustering can be performed using three groups. The purpose of this analysis was to further investigate the objective scene features that were used for the original clusters. To achieve this, the mean value of each video was used to represent the perception of the place for the clustering computation. Based on the elbow plots derived from the experimental data, it was established that clustering could be performed using three groups. The distribution of videos among the three clustering groups is depicted in Figure 4.

Table 4 provides a detailed breakdown of the three clusters resulting from the K-means cluster analysis using ISOPleasant and ISOEventful data collected. A Kruskal-Wallis H test was then conducted on this clustering result using five objective scene features. Tables 4 and 5 show the statistical results of this test, revealing a significant difference in road ratio between the different clusters, with a $\chi^2(2)$ value of 11.841 and p-value of 0.003. In addition, there was a statistically significant difference in water ratio between different clusters, with a $\chi^2(2)$ value of 6.977 and p-value of 0.031.

The above results suggest that the road ratio and water ratio are crucial objective scene features when evaluating soundscape perception. Specifically, a higher road ratio is associated with more eventful and bothersome affectivity, whereas a higher water ratio is linked to a more pleasant and serene experience. These findings have important implications for the design and management of urban environments, as they highlight the significance of incorporating natural elements such as water into urban landscapes to enhance the quality of soundscape perception.

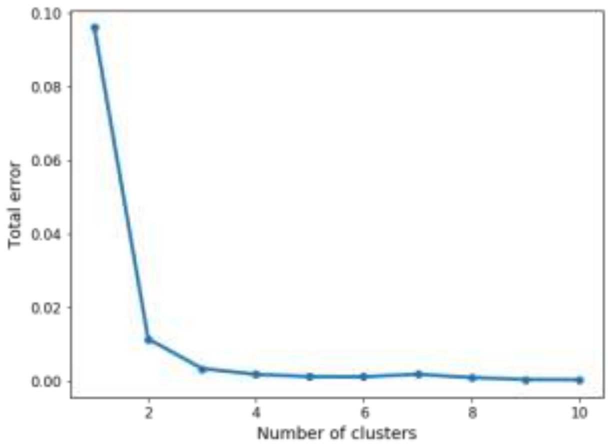


Figure 3. The elbow clustering result

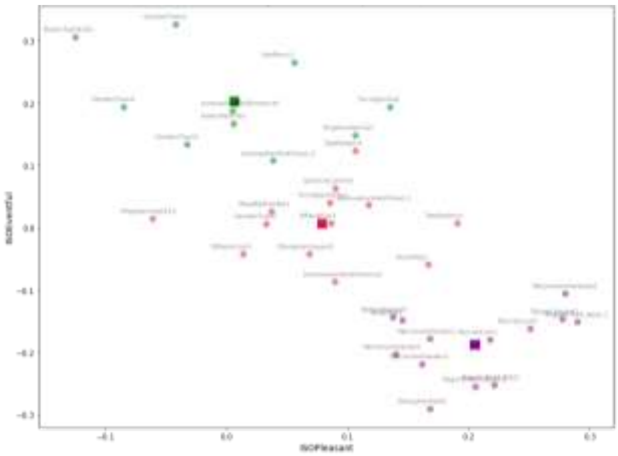


Figure 4. Clustering results based on the laboratory experiment

Table 3. New clusters based on ISOPleasant and ISOEventful

Cluster Group	Videos

1	CamdenTown1, CarloV1&CarloV2, LianhuashanParkEntrance2, LianhuashanParkForest_1, Pingshanstreet1&2, PlazaBibRambla1, RussellSq1, StPaulsCross1, StPaulsRow1, TateModern1, TateModern3, TorringtonSq3&2, ZhongshanSquare5
2	DadongSquare3, MarchmontGarden1, MarchmontGarden2, MarchmontGarden3, MonumentoGaribaldi1 , OlympicSquare3, PancrasLock1, PancrasLock2, Regents Park Japan_1, Regents Park Japan_2, RegentsParkFields2, RussellSq2, ZhongshanPark5
3	CamdenTown2, CamdenTown3, CamdenTown4, EustonTap3&2&1, LianhuashanParkEntrance1, LianhuashanParkForest_2, PingshanPark1&2, SanMarco1, SidalinPark1&2, TorringtonSq4

Table 4. The Ranks of Kruskal-Wallis H test

	Clusters	N	Mean Rank
Sky (%)	1	13	17.77
	2	13	18.69
	3	10	19.20
Building (%)	1	13	19.92
	2	13	16.31
	3	10	19.50
Road (%)	1	13	20.15
	2	13	11.08
	3	10	26.00
Green (%)	1	13	17.69
	2	13	22.00
	3	10	15.00
water (%)	1	13	17.69
	2	13	22.77
	3	10	14.00

Table 5. The Ranks of Kruskal-Wallis H test

	Sky Ratio	Building Ratio	Road Ratio	Green Ratio	Water Ratio
Kruskal-Wallis H	0.111	0.897	11.841	2.616	6.977
df	2	2	2	2	2
Asymp.Sig	0.946	0.639	0.003	0.270	0.031

4. DISCUSSION

4.1 Advantages of soundscape experiments using IVR

This study presents a laboratory-based experiment that successfully reproduces a realistic visual environment, ensuring the participants' sustained immersion throughout the experiment, including scene transitions and questionnaire answering. Prior research has proved the ecological significance of virtual reality (VR) environments. The proposed methodology can be implemented in subsequent experiments using questionnaire tools in VR with 3 degrees of freedom. This replication technique enables a faithful restoration of the visual experiences of an enviornment. The fidelity of the experimental results, including their authenticity, validity, and reliability, directly depends on the degree of real-world environment replication achieved in the study.

The laboratory experiment proved to be a highly effective means of data collection, owing to its convenience and capacity. The experimental design involved a single participant evaluating 36 scenarios, in contrast to the field experiment, where a participant would provide feedback for only one scenario. Consequently, the laboratory experiment could collect 1,152 responses in just one week, while field experiments would require a much longer time frame or manpower for the same number of scenes. Moreover, laboratory experiments offer a more stable and secure environment for data collection.

The selection of participants for laboratory experiments offers greater control than for field experiments. In field research, participants are typically approached randomly while on-the-go and asked to complete questionnaires. People in the same scene may share preferences and habits that can potentially bias the experiment results. Meanwhile, daily commute stress and crowding may influence participants' impressions. In contrast, laboratory experiments allow for a more tailored selection of participants, with considerations for factors such as age distribution, cognitive level, ethnicity, and others based on the specific requirements of the investigation. Further to that, when the same participant evaluates multiple scenes, more consisted context, semantic comprehension, and perception levels are provided. Comparisons, adjustments, and balances can also be made across distinct scenes, allowing for a clearer determination of the origin of specific stimulus and making macro-level comparisons more relevant.

4.2 Limitations

The replication of natural environments within a laboratory setting is a promising avenue of research, but there are limitations to the realism and ecological validity of laboratory reproduction techniques. Despite advancements in technology and past research demonstrating the accuracy of these methods, laboratory experiments can only approximate a sense of immersion to a certain extent. Technological constraints, such as video quality, and the inability to simulate other senses beyond sight and hearing, hinder the accuracy of laboratory experiments. Additionally, participants in this experiment cannot physically interact with their surroundings, which can weaken the feeling of presence and reality. Thus, while laboratory experiments can provide valuable insights, they should be interpreted within the context of their limitations.

In laboratory experiments, it is imperative to ensure that the participants are comfortable in the virtual world. In contrast to field experiments, where participants are already integrated into the real world, forming their first opinions, individuals in a virtual world can struggle to adjust to the surroundings or experience discomfort, such as vertigo, due to strobe or frame skipping, which can adversely affect the experiment's results. Each participant's ability to adapt to the virtual environment without experiencing discomfort is nearly uncontrollable, thus necessitating careful consideration of the timing of the experiments and the order of the videos. Additionally, the equipment itself can cause discomfort, as evidenced by the Oculus Quest2 utilized in this experiment, which includes only three steps for interpupillary distance adjustment, but may not always achieve full focus, resulting in biased results.

5. CONCLUSION

In this study, we have explored how the Method A questionnaire of the ISO/TS 129313 series can be used to investigate the impact of visual context on soundscape perception. Despite some prior efforts to examine the influence of visual stimuli, an optimal method for investigating this relationship has yet to be proposed. The findings hold promise for future investigations.

This pilot approach to soundscape research employs a visually immersive virtual reality (IVR) environment. It provides a means of investigating soundscape that allows for greater control and flexibility while maintaining ecological validity. It was based on presenting visual stimuli only, without any sounds.

Even without the auditory stimuli, meaningful responses were observed using the Method A ISO/TS 12913-2 questionnaire tool. The results reveal a notable correlation between ISOPleasant and ISOEventful. We further investigated this relationship by clustering the ISOPleasant and ISOEventful scores together with urban types based on visual features extracted from the videos, revealing that the presence of road and water in the visual environment significantly influenced soundscape perception. The implications of these findings highlight the potential of considering visual stimuli for future research in soundscape perception.

6. REFERENCES

[1]. Aletta, F., T. Van Renterghem, and D. Botteldooren, *Influence of personal factors on sound perception and overall experience in urban green areas. A case study of a cycling path highly exposed to road traffic noise.* International journal of environmental research and public health, 2018. **15**(6): p. 1118.

[2]. Hong, J.Y. and J.Y. Jeon, *Relationship between spatiotemporal variability of soundscape and urban morphology in a multifunctional urban area: A case study in Seoul, Korea.* Building and Environment, 2017. **126**: p. 382-395.

[3]. Miller, N., *Understanding soundscapes.* Buildings, 2013. **3**(4): p. 728-738.

[4]. Schafer, R.M., *The new soundscape*. 1969: BMI Canada Limited Don Mills.

[5]. ISO, *ISO 12913-1:2014, "Acoustics—Soundscape—Part 1: Definition and conceptual* (International Organization for Standardization, Geneva, Switzer), 2014.

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

[7]. ISO, *ISO/TS 12913-3:2019, "Acoustics—Soundscape—Part 3: Data analysis".* (International Organization for Standardization, Geneva, Switzer), 2019.

[8]. Raimbault, M., C. Lavandier, and M. Bérengier, *Ambient sound assessment of urban environments: field studies in two French cities.* Applied Acoustics, 2003. **64**(12): p. 1241-1256.

[9]. Guastavino, C. and B.F. Katz, *Perceptual evaluation of multi-dimensional spatial audio reproduction.* The Journal of the Acoustical Society of America, 2004. **116**(2): p. 1105-1115.

[10]. Hong, J.Y. and J.Y. Jeon, *Influence of urban contexts on soundscape perceptions: A structural equation modeling approach.* Landscape and Urban Planning, 2015. **141**: p. 78-87.

[11]. Oberman, T., et al., *Using virtual soundwalk approach for assessing sound art soundscape interventions in public spaces.* Applied Sciences, 2020. **10**(6): p. 2102.

[12]. Stienen, J. and M. Vorländer. *Auralization of urban environments—Concepts towards new applications*. in *Proceedings of Euronoise*. 2015.

[13]. Jiang, L., et al., *A demonstrator tool of web-based virtual reality for participatory evaluation of*

urban sound environment. Landscape and Urban Planning, 2018. **170**: p. 276-282.

[14]. Kawai, K., et al., *Personal evaluation structure of environmental sounds: experiments of subjective evaluation using subjects' own terms.* Journal of sound and vibration, 2004. **277**(3): p. 523-533.

[15]. Annerstedt, M., et al., *Inducing physiological stress recovery with sounds of nature in a virtual reality forest—Results from a pilot study.* Physiology & behavior, 2013. **118**: p. 240-250.

[16]. Feng, D. and K.L. O'Halloran, *Representing emotive meaning in visual images: A social semiotic approach.* Journal of Pragmatics, 2012. **44**(14): p. 2067-2084.

[17]. Jo, H.I. and J.Y. Jeon, *Perception of urban soundscape and landscape using different visual environment reproduction methods in virtual reality.* Applied Acoustics, 2022. **186**: p. 108498.

[18]. Memoli, G., G. Hamilton-Fletcher, and S. Mitchell, *Soundscape assessment of aircraft height and size.* Frontiers in psychology, 2018. **9**: p. 2492.

[19]. Abdalrahman, Z. and L. Galbrun, *Audio-visual preferences, perception, and use of water features in open-plan offices.* The Journal of the Acoustical Society of America, 2020. **147**(3): p. 1661-1672.

[20]. Bishop, I.D. and B. Rohrmann, *Subjective responses to simulated and real environments: a comparison.* Landscape and urban planning, 2003. **65**(4): p. 261-277.

[21]. Rossetti, T. and R. Hurtubia, *An assessment of the ecological validity of immersive videos in stated preference surveys.* Journal of choice modelling, 2020. **34**: p. 100198.

[22]. Jo, H.I. and J.Y. Jeon, *Effect of the appropriateness of sound environment on urban soundscape assessment.* Building and environment, 2020. **179**: p. 106975.

[23]. Jiang, L. and J. Kang, *Effect of traffic noise on perceived visual impact of motorway traffic.* Landscape and Urban Planning, 2016. **150**: p. 50-59.

[24]. Lugten, M., et al., *Improving the soundscape quality of urban areas exposed to aircraft noise by adding moving water and vegetation.* The Journal of the Acoustical Society of America, 2018. **144**(5): p. 2906-2917.

[25]. Hong, J.Y. and J.Y. Jeon, *The effects of audio–visual factors on perceptions of environmental noise barrier performance.* Landscape and Urban Planning, 2014. **125**: p. 28-37.

[26]. Sanchez, G.M.E., et al., *Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space.* Landscape and Urban Planning, 2017. **167**: p. 98-107.

[27]. Xu, C., et al. *Ecological validity of immersive virtual reality (IVR) techniques for the perception of urban sound environments*. in *Acoustics*. 2020. MDPI.

[28]. Basturk, S., L. Maffei, and M. Masullo. *Soundscape approach for a holistic urban design*. in *Proceedings of the AESOP 26th Annual Congress, Ankara, Turkey*. 2012.

[29]. Jeon, J.Y. and H.I. Jo, *Effects of audio-visual interactions on soundscape and landscape perception and their influence on satisfaction with the urban environment.* Building and Environment, 2020. **169**: p. 106544.

[30]. Hsieh, C.-H., et al., *The effect of water sound level in virtual reality: A study of restorative benefits in young adults through immersive natural environments.* Journal of Environmental Psychology, 2023: p. 102012.

[31]. Yilmazer, S., P. Davies, and C. Yilmazer. *A virtual reality tool to aid in soundscapes in the built environment (SiBE) through machine learning*. in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*. 2023. Institute of Noise Control Engineering.

[32]. Jo, H.I. and J.Y. Jeon, *Quantification of visual attention using eye-tracking technology for soundscape assessment through physiological response.* 2023.

[33]. Lu, Y., et al., *The interactive effects of traffic sound and window views on indoor soundscape perceptions in the residential area.* The Journal of the Acoustical Society of America, 2023. **153**(2): p. 972-989.

[34]. Mitchell, A., et al., *The International Soundscape Database: An integrated multimedia database of urban soundscape surveys -- questionnaires with acoustical and contextual information [Data Set].* 2022.

[35]. Mitchell, A., F. Aletta, and J. Kang, *How to analyse and represent quantitative soundscape data.* JASA Express Letters, 2022. **2**(3): p. 037201.