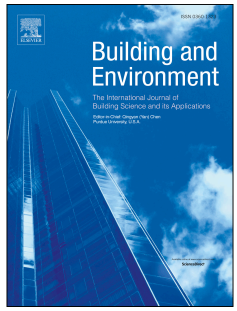
Journal Pre-proof



Effect of sound on visual attention in large railway stations: A case study of st. pancras railway station in London

Chang Liu, Jian Kang, Hui Xie

PII: S0360-1323(20)30549-7

DOI: https://doi.org/10.1016/j.buildenv.2020.107177

Reference: BAE 107177

To appear in: *Building and Environment*

Received Date: 22 March 2020

Revised Date: 24 July 2020

Accepted Date: 30 July 2020

Please cite this article as: Liu C, Kang J, Xie H, Effect of sound on visual attention in large railway stations: A case study of st. pancras railway station in London, *Building and Environment* (2020), doi: https://doi.org/10.1016/j.buildenv.2020.107177.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Ltd.

Journal Pre-proof

Effect of Sound on Visual Attention in Large Railway Stations: A case study of St. Pancras railway station in London

Chang Liu[a,b,**], Jian Kang[b,c], Hui Xie[a,b,**]

[a] *Faculty of Architecture and Urban Planning, Chongqing University, Chongqing, 400045, China*
[b] *Key Laboratory of New Technology for Construction of Cities in Mountain Area, Ministry of Education, Chongqing University, Chongqing 400045, China*
[c] *Institute for Environmental Design and Engineering, The Bartlett, University College London - 14 Upper Woburn Place, WC1H 0NN London, UK*

Abstract

There is a growing body of research into audio-visual interaction in architectural acoustics. The complex physical environment of a large railway station contributes to strong interactions between multi-sensors. This paper investigated how visual attention is shifted when sound level increases; whether the leading effect of sound on visual attention differs under the influence of pleasant, annoying, or information sounds; and whether the leading effect is affected by the correspondence between audio and visual stimuli. The study found that, as the difference between the sound levels of sound signals and background noise increased, the variation in attention rate and evaluation of the audio environment took on a parabola-shaped curve. However, in case of a pleasant sound (music), the inflection points arrived at higher sound level than that in case of an annoying noise (train noise). Pleasant sounds (music) and annoying noise were found to be more 'noticeable' than neutral sounds. In addition, sound that corresponds with visual stimuli had the most significant influence on the leading effect of sound on visual attention.

Keywords: audio-visual interaction, waiting area, large railway stations

1. Introduction

Human senses are not able to work independently [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. Yang and Moon [13] investigated variation in acoustic, thermal, visual, and indoor environmental comfort, under the influence of multisensory interactions. Of these, acoustics has the strongest impact on indoor environmental comfort. Numerous studies have investigated how multisensory interaction in the acoustic perceptions [14, 15, 16, 17]. Ren and Kang investigated the effect of landscape objects, distance to water edge, and the appearence of animals and humans on the evaluation of acoustic comfort [15]. Schaffer et al. found that noise can become less 'annoying' when accompanied by a visualised landscape – and more so when accompanied by the sight of a wind turbine [16]. Many studies have focused on the effect of sound on visual perception [18, 19, 20, 21, 22]. One investigation found that looming structured sounds can enhance visual orientation sensitivity in the hemifield of the sound, even when no information about visual orientation is provided by the auditory stimuli [23]. The performance of visual orientation

*Corresponding author
**Corresponding author
Email addresses: lc87@cqu.edu.cn (Chang Liu), xh@cqu.edu.cn (Hui Xie)

discrimination can be degraded by a burst of white noise through headphones [20]. However, the interaction between audio and visual information in architectural spaces remains unclear.

In architectural spaces, various auditory and visual information interact in the perception of building environments [24, 25]. The interaction phenomenon is especially notable in complex railway stations, which contain multiple-functional spaces and complex sound sources [26]. Existing studies of railway stations focus on ways of evaluating the soundscape with regard to social and behavioural characteristics [27], how the type of space is recognised by the sound information [28], and how local residents' perceptions of loudness and annoyance are affected by the visual characteristics of noise barriers [29]. The functions, spaces, and routines in a large railway station are complicated [30]. Therefore, a clear and straightforward wayfinding system is important for the indoor environment design of such spaces. To enable this, it is necessary to investigate the interaction between audio and visual information. However, the effect of audio-visual interaction in the wayfinding system design of railway stations has seldom been studied. Sound plays a key role in space perception [28, 31, 32]. Few studies have focused on the effect of sound on visual attention [33, 34, 35, 36, 37]. Van der Vurg et al. researched the influence of non-spatial sound on spatial visual searching. They conclude that a visual event is easier to identify with the presence of a synchronized non-spatial sound than with an absence of sound [34]. Coutrot et al. showed that observers' eye positions and movements differ when they watch videos with and without soundtracks [36]. Ye et al. found that the attention of 12% of diners was attracted by videos played in a canteen [37]. However, little research has been done on the influence of auditory characteristics on visual attention.

Sound level is the most significant feature of a sound [14, 16]. For instance, the level of road traffic noise (individual sound) is related to ratings of 'annoyance' in evaluations of urban soundscapes containing combined noise sources [38]. However, little is known about whether the level of combined noise sources affects observers' visual attention in railway stations. The type of sound is the second most influential characteristic. Viollon et al. indicate that, of various kinds of 'natural' noises, the sound of water features such as fountains is the most effective for providing a sense of 'pleasantness' in parks and city squares [39]. However, whether sound preference is related to the leading effect of the sound has not yet been comprehensively investigated. Another characteristic of sound is context [39, 15, 21]. Previous research has showed that co-occurring visual settings influence judgements of sounds [39]. Ren and Kang found that evaluations of acoustic comfort in relation to a natural landscape matched with nature sounds, music, and the sound of a temple bell were more positive than evaluations of an artificial landscape combined with these sounds [15]. However, research on the effect of compatibility between the sound and sound-source object on the leading effect of sound on visual attention is limited.

Little research has investigated the effect of level, type, and context of sound on visual attention in the waiting area of complex railway stations, despite the topic's importance for wayfinding system design in complex environments. In a railway station, the waiting area is the site of the most interactions between people and the environment, as well as being where visitors are likely to spend the longest period of time. The aim of this paper is to investigate the leading effect of sound on visual attention in the waiting areas of complex railway stations, as affected by the level, type, and context of the sound. More specifically, this research contains three research questions: 1) how

Journal Pre-proof

() Waiting area on the ground floor

() Waiting area on the first floor

Figure 1: St. Pancras International Railway Station

visual attention shifts when the sound level increases; 2) whether the leading effect of sound on visual attention differs under the influence of pleasant, annoying, or information sounds; and 3) whether the leading effect is affected by correspondence between a sound and its visual stimuli. The study involves a laboratory experiment, with audio and visual stimuli collected in situ, based on subjective questionnaires and objective in-situ acoustic measurements.

2. Methods

The research strategies included an online questionnaire survey, acoustic measurement, and laboratory experiment. The online questionnaire survey and acoustic measurement were used to investigate the physical environment of a typical large railway station and provide evidence for the laboratory experiment. The laboratory experiment examined the leading effect of sound on visual attention, under the influence of the level, type, and context of sound. The work received generic approval from the ethics committee of The University of Sheffield.

2.1. Case study site

As the research was conducted in a real railway station, the preliminary work included identifying an appropriate location that fulfilled the set criteria. First, the proposed railway station had to be a mainline station, providing at least two million trips per year. The visual and audio environment of the waiting area of the station had to be complex. Finally, it is preferable for the selected railway station to have been constructed (or reconstructed) within the previous 10 years, thereby constituting a ‘modern’ station.

Taking the above requirements into consideration, all the railway stations in the UK were considered. The UK is one of the world’s most developed countries in respect of transportation constructions, and the St. Pancras International Railway Station in London (Figure 1) was ultimately selected for the case study. This site has a central location, it is multi-functional, and it has a complex visual and audio environment. St. Pancras was built in the 1860s, then refurnished to serve as an international station in 2007. It is representative of Victorian architecture, and its planning, interior decoration, and structure are used as reference for the design of large railway stations [40, 41]. St. Pancras contains seven domestic and six international terminal platforms on the first floor, and two through platforms on the ground floor.

Journal Pre-proof

2.2. Questionnaire survey

To explore peoples' traveling habits and subjective assessments of the physical environment of St. Pancras, questionnaires were disseminated through an online survey website, 'Smart Survey', to student volunteers from The University of Sheffield. A total of 192 surveys were completed, and 177 of these were deemed valid. Of the latter, 140 had ever been to St. Pancras and the other 37 had previously travelled through other large railway stations. Four sections of the questionnaire requested the personal information of the respondents, including their age and gender, to investigate whether these characteristics affected the findings. There were also questions about the participants' traveling habits within the UK. The survey then asked for a general evaluation of the physical environment of St. Pancras, including the lighting, sound, temperature, visual aspects, and space organisation. The respondents were then asked about the frequency of 14 sound sources (e.g., voices, announcements, train engines, music, movement of luggage) heard in railway stations ('1' for never, '2' for rarely, '3' for occasionally, '4' for sometimes, '5' for often), and the subjective feelings about these sound sources ('1' for very dissatisfied, '2' for dissatisfied, '3' for neither dissatisfied nor satisfied, '4' for satisfied, '5' for very satisfied) in the waiting area of the station.

2.3. Acoustic measurement

The objective of the acoustic measurement was to explore the physical environment of the St. Pancras waiting area, including the background noise in certain places and the sound level distribution throughout space. The collected data also supported the design of the laboratory experiment.

A 01-dB SOLO Black Edition sound level meter was used to measure the background noise on site. On the ground floor, 85 measurement points were selected, with a further 64 on the first floor. The points were distributed symmetrically across the waiting area to achieve a complete sound level map. Before the measurements were taken, the sound level meter was calibrated and fixed on a tripod (1.2 m height). After collection, the data were transferred and analysed using the modular matching software 01dB Trait.

2.4. Laboratory experiment

The laboratory experiment was the primary component of the study. This sought to clarify the relationship between the leading effect of sound on visual attention and the level, type, and context of the sound stimuli.

2.4.1. Experiment design

A group of 54 students (aged 20-30 years), including 27 females and 27 males, volunteered to participate in the experiment. All reported normal hearing and normal or corrected-to-normal vision, as well as plenty of train-travel experiences.

The primary equipment used in this research was a hi-fi recorder (Edirol R-44) to directly record the audio stimuli, a digital camera to record visual stimuli, and a Tobii TX300 Eye-Tracker to record eye movements. Other tools included a projector and screen to present the visual stimuli; and a headphone (Sennheiser HD590) to listen to the audio stimuli.

Journal Pre-proof

Journal Pre-proof

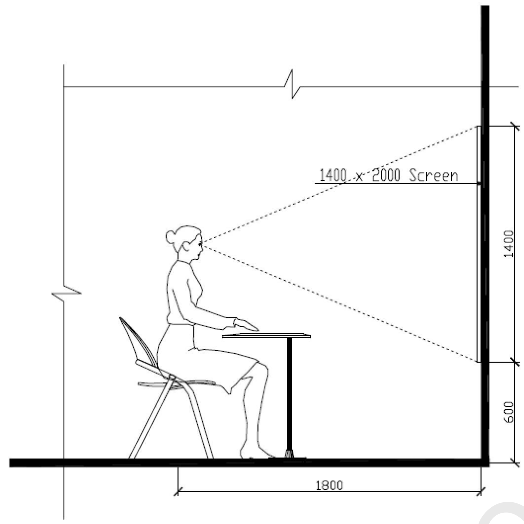


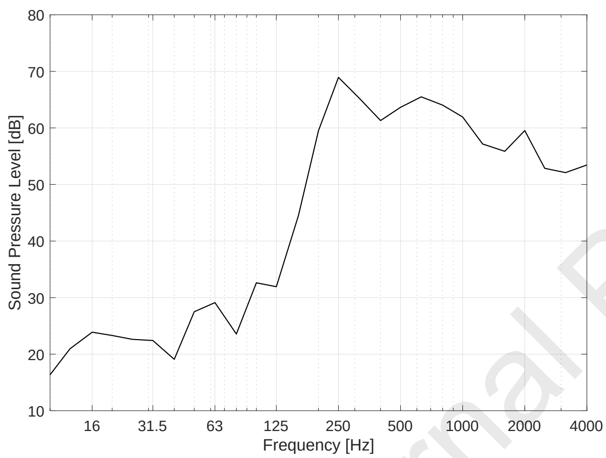
Figure 2: Experiment setup [mm]

Prior to the experiments, pilot tests were untaken with three participants to examine any significant influences of lighting or sudden, large, or bright objects on the video. Each participant undertook their test individually, in a private room. They were informed that they would be watching a series of videos of railway stations, but they had no information about the type of sound or the purpose of the experiment. The participants sat a table in a dimly lit and isolated room, with closed doors and windows. The audio stimuli were presented through headphones to reduce the influence of the environment and to create a feeling of being present at the scene. The videos were presented on a fixed screen (2000 mm x 1400 mm), 1.8 m away from the participant, who was sitting on a fixed chair in the centre of the room (as shown in Figure 2). At the beginning of each test, the participant was asked to check whether the brightness of the screen and the volume of the video were comfortable. They then completed the first part of questionnaire, giving their personal details (gender, age, travelling experience, and visual and hearing conditions). While watching the video, the volunteer was asked to imagine that they were traveling through a railway station. They were informed that some of the videos were silent and some were not. Each video was displayed just once for each participant. After each viewing, the volunteer was asked to complete the second section of the questionnaire, and indicate their subjective evaluation of the perceived loudness, preference of the visual and audio environment shown in the clip. A five-point scale was presented for these questions (for perceived loudness: 1-very low, 2-low, 3-medium, 4-high, 5-very high; for preference of the visual and audio environment: 1-very dissatisfied, 2-dissatisfied, 3-neither dissatisfied nor satisfied, 4-satisfied, 5-very satisfied). In the third section of the questionnaire, the researcher recorded the participants' responses to the following question: 'Please name as many as possible of the items or objects you noticed in video. Please list them in sequence, from the most noticeable to the least noticeable'.

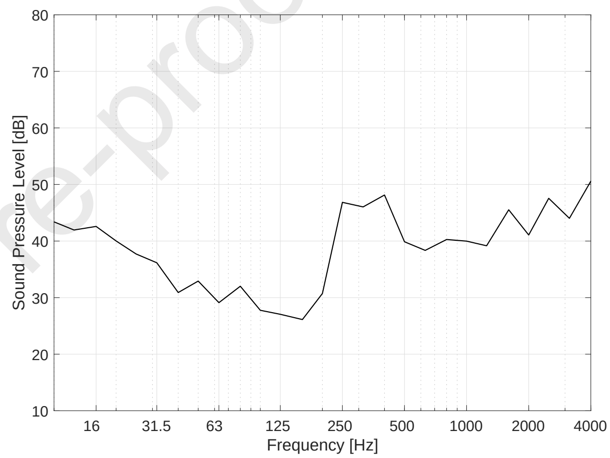
2.4.2. Visual and audio stimuli

A. Audio stimuli

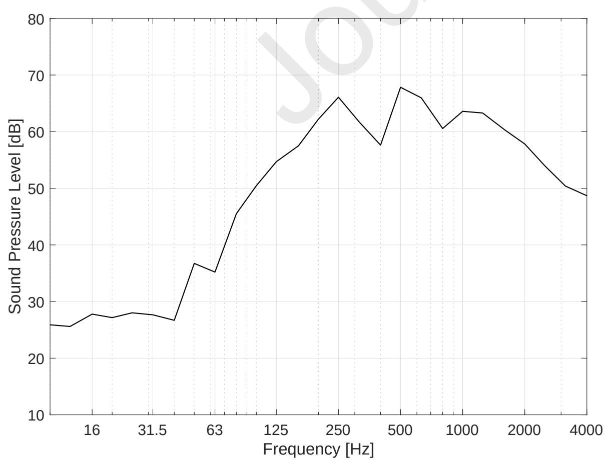
There were four criteria for appropriate sound sources. 1) The most important aspect was that the sound sources occurred in the waiting area of the railway station. According to the online survey, the eight most commonly heard sounds (voices, announcement, background noise, train noise, movement noise of luggage, whistles, restaurant noise, music) were taken into consideration in the selection of audio stimuli. 2) There should be variation in how annoying the sounds are deemed to be – from train noise to music, for example – to highlight the difference in leading effect of the different types. 3) The target sounds should be a mix of sound source objects that would be visible, such as trains, and those that would not, such as station announcements. This was intended to identify whether the leading effect of sound on visual attention has a relationship with the context of the sound. 4) Finally, even though sounds such as voices and the movement noise of luggage are typical in railway stations, they make less sense on architectural design and were not be explored.



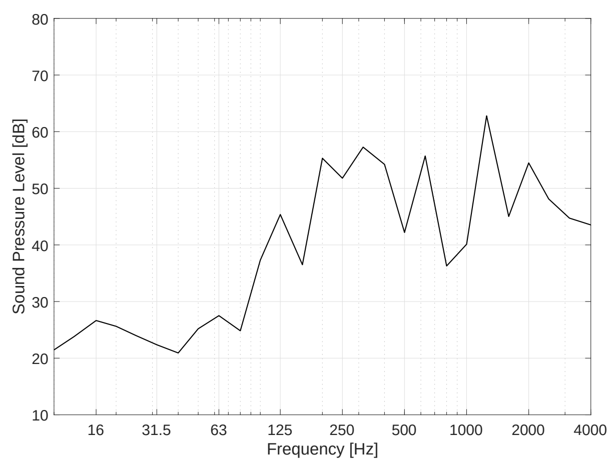
() Announcement.



() Restaurant noise.



() Music.



() Train noise.

Figure 3: Spectrum of recorded sound stimuli.

Announcements, restaurant noise (e.g., dishes, movement of chairs), music (a piano displayed in railway station, which can be played by travelers), and train noise (the arrival and departure of the trains) were selected as the sound

Journal Pre-proof

stimuli. The stereophonic sound samples recorded were of at least two minutes each – with 30 seconds of steady performance in the middle of the recordings used as the edited sound samples. The spectrum of the announcements (information sound), music (positive sound), and restaurant noise and train noise (negative sound) is depicted in Figure 3. Background noise was important for the analyses of the audio stimuli and the effect of sound on visual attention [42]. Therefore, the following three types of edited audio sample were chosen: silent (denoted by 'S'), background noise ('BN'), and the four types of edited sound samples (65 dBA) combined with background noise (60 dBA) in the corresponding scenarios (denoted by 'BN+A65' for announcement, 'BN+R65' for restaurant noise, 'BN+M65' for music, and 'BN+T65' for train noise). The level of the edited sound sample and background noise were decided based on the background noise level of St. Pancras, via physical in-situ acoustic measurement and pilot tests.

To investigate the effect of the sound level on the visual attention leading effect, the edited sound samples were tuned to five different levels (55 dBA, 60 dBA, 65 dBA, 70 dBA, and 75 dBA). Thus, there were another four types of final audio stimuli – labelled 'BN+A75' for background noise combined with 75 dBA announcement, for instance. This meant seven final audio stimuli for each group of sound.

B. Visual stimuli

The four video samples depicted the departure board, a restaurant, the piano, and a train. The video samples (with a resolution of 1440 x 1080 pixels) were each two minutes in length and recorded in the waiting space of St. Pancras. There were five requirements. 1) All the videos were filmed with a fixed focus, in first-person view, to avoid the influence of sight movements and to provide a feeling of being personally on the scene. 2) No obvious large, colourful, bright, or suddenly appearing objects or people were included. 3) People with strange or direction-oriented behaviour were removed from the videos to prevent inappropriate influence. 4) Visual short-term memory proved an important temporary maintenance function when outputting and inputting visual information [43]. Therefore, for each group, three different but similar videos were recorded to provide a combination of audio stimuli. The similar videos depicted the same three key objects. The target object – which corresponded with certain types of sound – took on the same proportions in each video and were located in the same position in all three. 5) To enable a comparison of the different combinations of video and audio stimuli, the videos were recorded in the same lighting environments, thus preventing this factor unduly affecting the results. Unnecessary fragments of the 12 video samples were cut to reduce the clips to 30 seconds each. The 4 x 3 visual stimuli (denoted by 'ad' for the departure board, 'rd' for the restaurant, 'md' for the piano, and 'td' for the train) are presented in Figure 4. Another four video samples (one for each group), with static images (denoted by 'as', 'rs', 'ms', and 'ts') were chosen as the final visual stimuli to compare the effect of the motion status of the visual stimuli on the audio-visual interaction analysis [44].

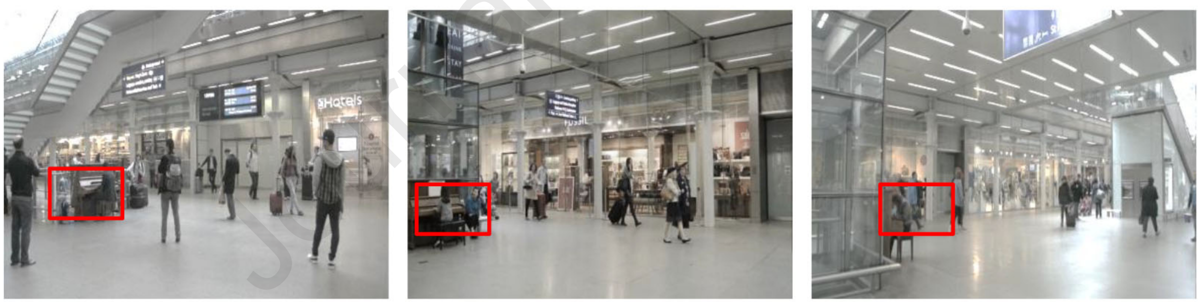
C. Combination of audio and visual stimuli

For each group, four visual stimuli and seven stereophonic audio stimuli were combined (as presented in Table 1). 'GA' represents for Group of Announcement+Departure board, 'GR' represent for Group of Restaurant noise+Restaurant, 'GM' represents for Group of Music+Piano, 'GT' represents for Group of Train noise+Train.

Journal Pre-proof

() Departure board

() Restaurant



() Piano

() Train

Figure 4: Video samples. The red rectangular boxes indicate the target visual objects

Journal Pre-proof

Table 1: Combination of audio and visual stimuli

GA[1]		GR[2]		GM[3]		GT[4]	
Audio	Visual	Audio	Visual	Audio	Visual	Audio	Visual
S[5]	ad[11]	S	rd[12]	S	md[13]	S	td[14]
BN [6]	ad	BN	rd	BN	md	BN	td
BN+A55[7]	ad	BN+R55 [8]	rd	BN+M55 [9]	md	BN+T55[10]	td
BN+A60	ad	BN+R60	rd	BN+M60	md	BN+T60	td
BN+A65	ad	BN+R65	rd	BN+M65	md	BN+T65	td
BN+A70	ad	BN+R70	rd	BN+M70	md	BN+T70	td
BN+A75	ad	BN+R75	rd	BN+M75	md	BN+T75	td
BN+A65	as[15]	BN+R65	rs[16]	BN+M65	ms[17]	BN+T65	ts[18]
BN+M65	ad	-	-	BN+A65	md	-	-

[1,2,3,4] GA represents for Group of Announcement+Departure board, GR represent for Group of Restaurant noise+Restaurant, GM represents for Group of Music+Piano, GT represents for Group of Train noise+Train;

[5,6] S represents for Silent, BN represents for Background Noise;

[7,8,9,10] A55 represents for 55 dBA Announcement, R55 represents for 55 dBA Restaurant Noise, M55 represents for 55 dBA Music, T55 represents for 55 dBA Train Noise;

[11,12,13,14] ad represents for dynamic departure board video, rd represents for dynamic restaurant video, md represents for dynamic piano video, td represents for dynamic train video;

[15,16,17,18] as represents for static departure board video, rs represents for static restaurant video, ms represents for static piano video, ts represents for static train video;

To compare the influence of the context of the sound on the audio-visual interaction, another two videos were also played in the experiments. They were 'BN+A65' combined with 'md', and 'BN+M65' combined with 'ad'. Since the visual stimuli used in eye-tracking system experiments is limited to static images, we used the BN combined with 65 dBA audio stimuli and static visual stimuli. Due to the sample size and experiment time, only 10 of the 12 videos were presented to each participant. The combined videos (each video lasts for around 30 seconds) were played at random during the experiments.

3. Results and Discussions

The background noise and motion status of the visual stimuli could have influenced the evaluation of the leading effect of sound in the analysis [42, 44]. Therefore, this section examines the reliability of the results, based on the analysis and a comparison of objective assessment by the eye-tracking system and subjective assessment by the questionnaire. It then considers the influence of the level, type, and context of sound on its leading effect on visual attenuation.

3.1. Reliability of results

3.1.1. Effect of background

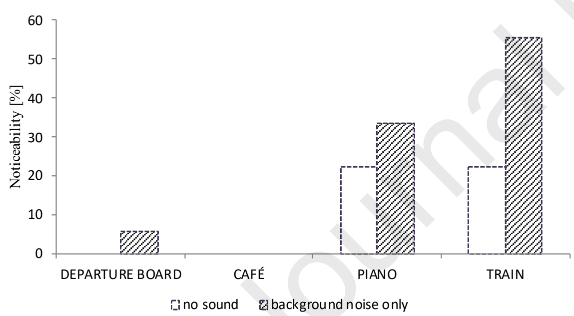


Figure 5: Noticeability on visual objects with and without background noise amongst four groups

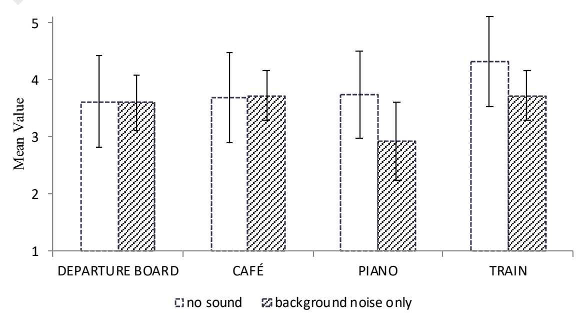


Figure 6: Assessment on visual environment with and without background noise amongst four groups. Error bars depict standard error values.

Table 2: Test statistics

Group	Exact Sig. (2-tailed)	
	Noticeability rank with and without background noise	Assessment on visual environment with and without background noise
GA	0.500	0.816
GR	0.125	0.806
GM	0.727	0.032
GT	0.344	0.083

'Noticeability' here is a single number which refers to the percentage of participants to have selected the relevant object as the 1^{st} noticeable one. Figure 5 shows that, with the presentation of background noise, the noticeability of the departure board, restaurant, piano, and train increased by 5, 0, 11.1 and 33.4 percentage points respectively. In Table 2, 'Noticeability rank is the noticeability ranking scores for departure board, restaurant, piano, and train. The ranking score for each object is '1' for the 1^{st} position, and '0' for the other positions. A Sign Test (Table 2) indicates that noticeability ranking scores with background noise are not statistically significantly higher than that without background noise, $p>0.05$. This is because no information is included in the background noise – so that no directivity can be detected from it. Therefore, the combination of background noise and sound signal cannot have a negative effect on the relationship between the sound and visual noticeability.

'Mean value' represents for the mean value of the participants' preference about the audio environment, visual environment, or perceived loudness. Figure 6 presents a similar mean value in the assessment of the visual environment with and without background noise. In other words, the assessment of the visual environment without background noise is not statistically significantly higher than that with background noise, with p-values (as shown in Table 2) being 0.816, 0.806, and 0.083 for GA, GR, and GT, respectively. For GM, $p=0.032<0.05$, which indicates that the assessment without background noise is statistically significantly higher than that with background noise.

3.1.2. Effect of the motion status of visual stimuli

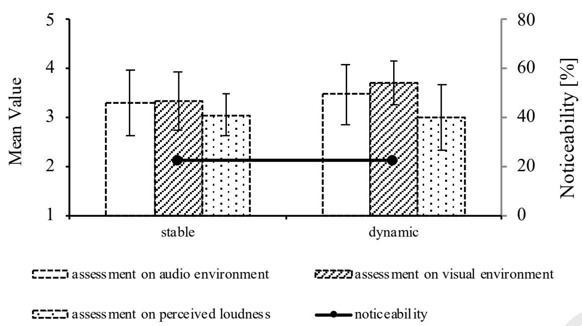
Figure 7 shows no significant differences in the GA, GR, and GM, for noticeability or visual and audio assessment. Therefore, the visual stimuli can be simplified, with stable photos used in place of dynamic videos in future research. However, in the dynamic format, trains were noticed approximately 6.5 times more often than they were in the static format (Figure 7(d)). This is because a train in the scene is moving throughout the period of the experiment, while the other visual objects are static. The movement of the visual object may have an effect on an observer's subjective assessment of involvement and presence [44]. Therefore, it is necessary to use a dynamic video, rather than a static image, for the audio-visual interaction research, especially for moving objects.

3.1.3. Comparison of results from the eye-tracking system and the questionnaire

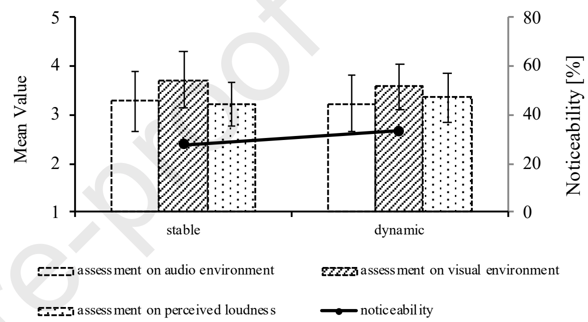
As presented in Figure 7(d) and Figure 8(d), the noticeability on moving objects are affected by the status of the visual stimuli. Since the visual stimuli for the eye-tracking system we used is limited to static images, we only recorded the number of fixations landing in regions of interests achieved by eye-tracking system for static visual stimuli. Figure 8 compares the noticeability of the visual objects, as assessed by the objective eye-tracking system and a subjective questionnaire concerning four audio-visual scenes. The eye-tracker records eye movement and calculates the frequency with which a person's eyes focus on images (Figure 8). The red points suggest more attention on the corresponding points, while green areas indicate less attention. No attention at all is marked with the absence of colour. The upper image in Figure 8 depicts the percentage of participants who selected each object; and the larger the image, the more noticeable the object. The order of the selection is not taken into account.

Figure 8(a) shows that, in this scene, the departure board was one of the most noticeable objects (in addition to the shop, the structure, and people), according to both the eye-tracking results and the subjective questionnaire

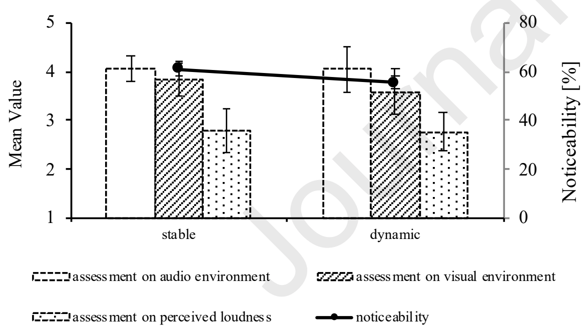
Journal Pre-proof



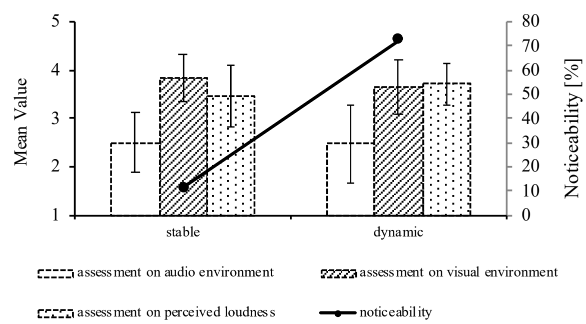
() GA



() GR



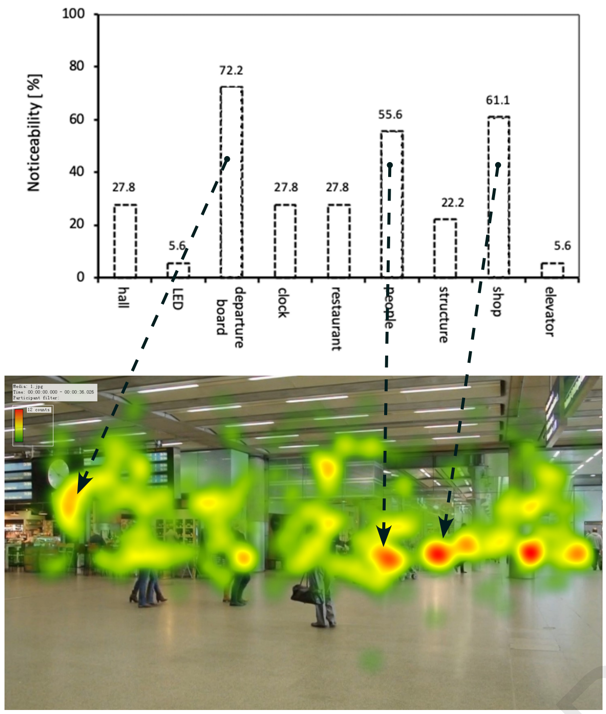
() GM



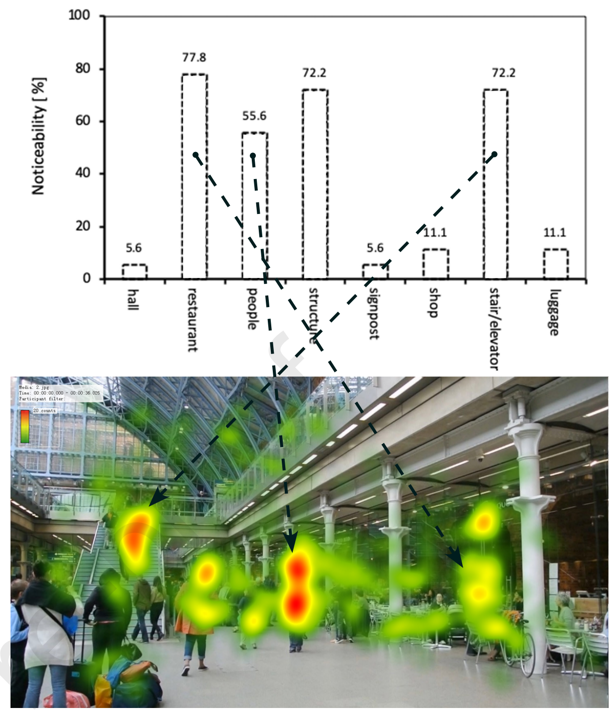
() GT

Figure 7: Assessment on visual/audio environment and percentage of noticeability with dynamic and static visual stimuli. Error bars depict standard error values.

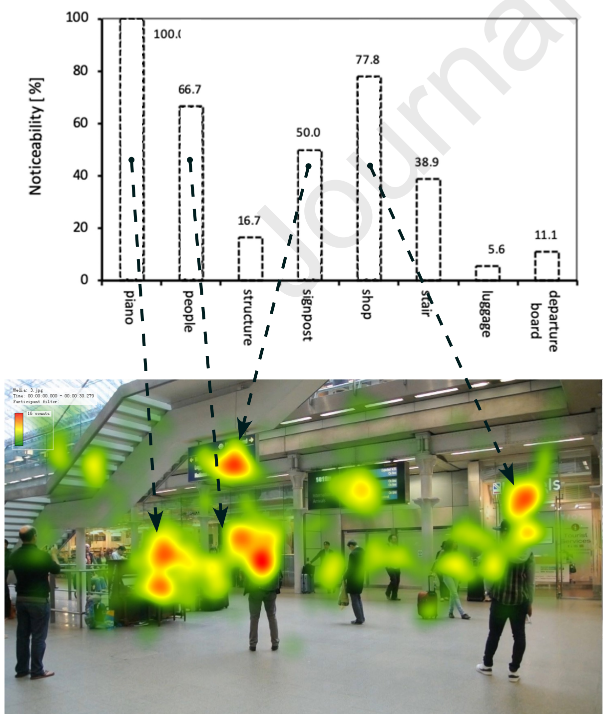
Journal Pre-proof



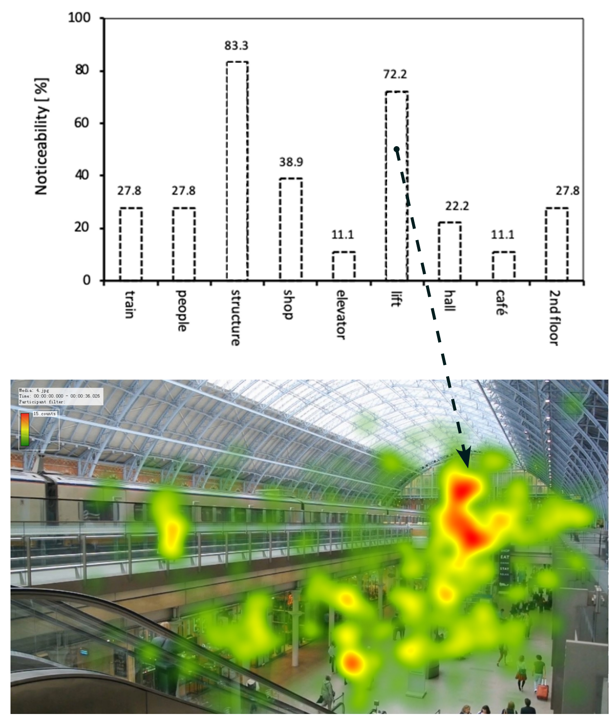
() GA



() GR



() GM



() GT

Figure 8: Comparison of noticeability achieved by subjective questionnaires (bar chart) and eye tracking experiments (color map).

responses. The analysis of the questionnaires groups the 'pillar', 'roof', 'window', and so on, as the 'structure'. Even though, the value of the noticeability for structure is larger than 70% in the scenes of GR and GT, structure was not taken into account because it is not relevant to this study of audio-visual interaction in the design of railway stations. Figure 8(b) suggests that, in this scene, while people and the stairs drew the most attention from participants, the restaurant was also a key object, as indicated by the eye-tracking results. Similarly, the questionnaire results indicate that the restaurant, people, and the stairs attracted the highest levels of visual attention. Both the objective and subjective results presented in Figure 8(c) illustrate that, in this scene, the people and the piano were paid the most attention. In Figure 8(d), the train is not the most noticeable object in both the questionnaire findings and the eye-tracker results, due to the static status of the video stimuli, as discussed in subsection 3.1.2. However, the questionnaire results align well with those of the eye-tracker. Therefore, the subjective assessment is deemed acceptable and reliable for the evaluation of the attendance to the corresponding visual elements.

Figure 8 shows that the departure board, restaurant, piano, and train were not the only visual objects to draw the participants' attention in the corresponding scenes. In Figure 8(a), the departure board had 72.2% noticeability, while people and shops had 55.6% and 61.1%, respectively. Similar situations were identified in all four groups. This finding is meaningful for future investigations of the positive effects of sound on visual attention.

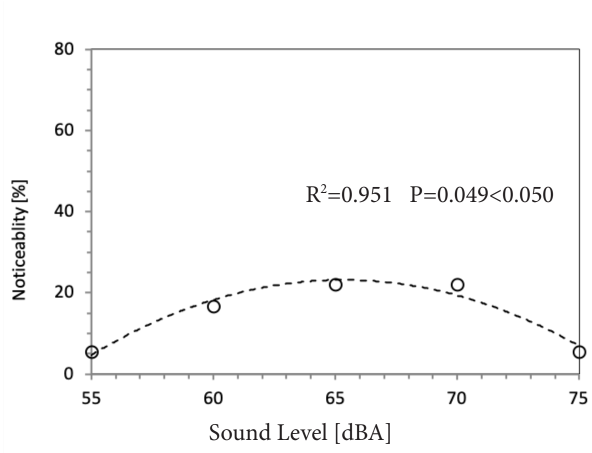
3.2. Effect of the sound level on noticeability and subjective assessment on the physical environment

3.2.1. Effect on noticeability

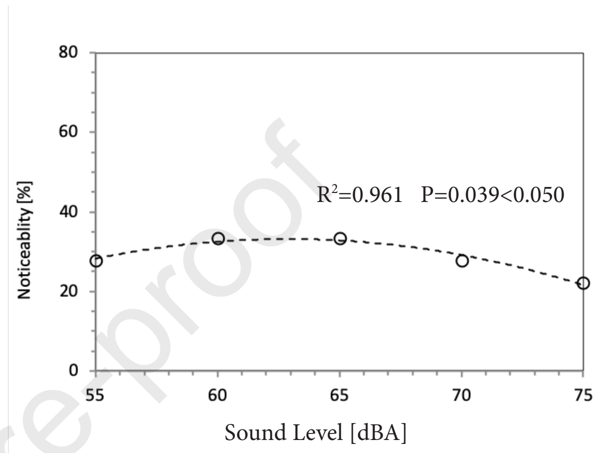
A curve estimation was conducted to compare the effect of the sound level of announcements, restaurant noise, music, and train noise on the noticeability of the departure board, restaurant, piano, and train, respectively. In general, the variation in noticeability takes a parabola shape, as the level of the four types of sound moves from 55 dBA to 75 dBA (Figure 9). An analysis of variance shows that the effects of sound level of four types of sound on noticeability of four objects are significant, $F(2,2)=19.287$, $p=0.049$, $R^2=0.951$ for GA, $F(2,2)=24.814$, $p=0.039$, $R^2=0.961$ for GR, $F(2,2)=47.908$, $p=0.020$, $R^2=0.980$ for GM, and $F(2,2)=23.208$, $p=0.041$, $R^2=0.959$ for GT.

The noticeability of the departure board is 5.6% with a 55 dBA announcement, increases gradually with an increase in sound level, and reaches its highest point at 22.2% when the announcement level is around 65 dBA-70 dBA. Thus, we conclude that it is the announcement that leads participants' visual attention to the departure board. The number falls abruptly to 5.6% when the announcement is raised to 75 dBA, which indicates that the leading effect of announcement is degraded. Similarly, in GR, 27.8% people choose the restaurant as the most noticeable object in the scene when the associated noise was 55 dBA, and this rose to 33.3% when the sound level peaked at 60 dBA-65 dBA. At this point, the curve begins to decline, and it continues its downward trend until 75 dBA, with noticeability of 22.2%. The noticeability of the piano rises from 38.9% at 55 dBA to 77.8% at 70 dBA, as shown in Figure 9(c). There is only a slight decline in noticeability when the level of music climbs from 70 dBA to 75 dBA. The parabolic trend is also evident in Figure 9(d), in relation to GT, where there is a peak of 72.2% at 65 dBA. Although – as discussed in subsection 3.1.2 – the train in the dynamic video attracted 6.5 times more attention than that in the static image, the visual stimuli compared here were all dynamic videos. Therefore, the movement of the visual stimuli had no effect on the findings.

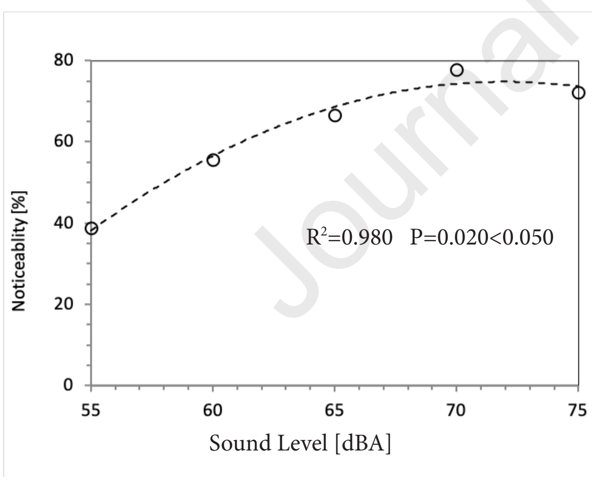
Journal Pre-proof



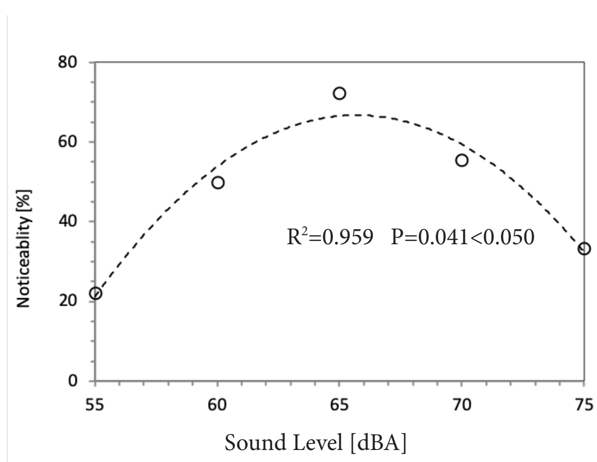
() GA



() GR



() GM



() GT

Figure 9: Noticeability with different type and level of sound. The 2^{nd} order polynomial regression functions are fitted to data points (dashed lines) with coefficient of determination (R^2) and p-value.

Journal Pre-proof

Journal Pre-proof

In conclusion, there is a rise in the leading effect of the four types of sound as the sound level of the signals increases; then, at a certain level, this begins to fall. Previous research into the soundscapes of underground business streets found that, as SPL increases, the acoustic comfort index varies in a parabola shape [45]. Recent paper indicated that the overall effect of nature sound on soundscape perception increases as the signal-to-noise ratio becomes larger, and begins to decrease slightly afterwards [46]. Theses findings are in line with the results in this research. This is understandable, as an increase in the sound level of announcements, restaurant activity, music, and trains enables them to dominate the background noise. These sounds delivered the most important information to the participants, thus their leading effect was the most significant. However, when the sound level exceeded a certain level (the tolerance of the audience), the sound may have become more annoying to the observers [12], preventing the information being received by them.

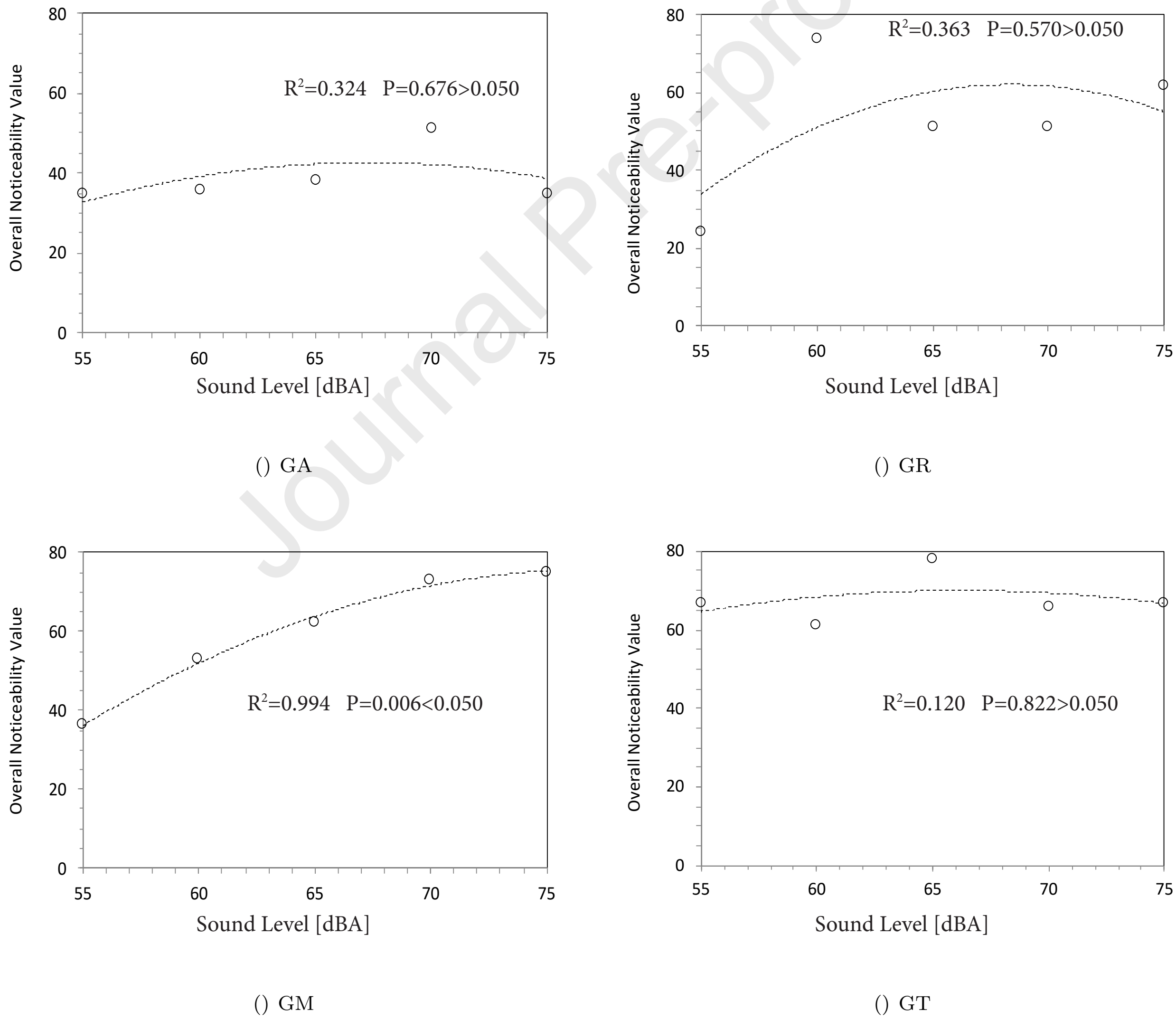


Figure 10: Overall noticeability values with different type and level of sound. The 2^{nd} order polynomial regression functions are fitted to data points (dashed lines) with coefficient of determination (R^2) and p-value.

Journal Pre-proof

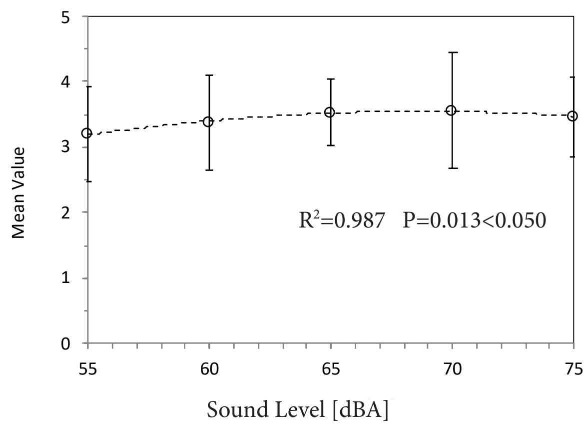
Figure 10 depicts the overall noticeability value of the departure board, restaurant, piano, and train under the influence of sound level. The overall noticeability value is a single number which refers to the sum value of the noticeability ranking scores of departure board, restaurant, piano, and train across all participants. The ranking score of each object is '5' for the 1^{st} position, '4' for the 2^{nd} position, '3' for the 3^{rd} position, '2' for the 4^{th} position, '1' for the 5^{th} position, and '0' for the rest. A curve estimation shows that the effect of sound level of music on the overall noticeability of piano is significant, F(2,2)=158.813, p=0.006, with an R^2 of 0.994. It is noticeable that the overall noticeability of piano varies in a parabola shape as well. The peak point occurs at 75 dBA, which is consistent with that presented in Figure 9(c). However, in terms of the other three groups, no significant relationship can be detected between the overall noticeability and sound level. The noticeability of piano is larger than that of the other three objects as presented in Figure 9. Therefore, the ranking scores for the 2^{nd}- 5^{th} place dominate the overall noticeability values. This is one of the possible explanations for this phenomenon. Moreover, the limited numbers of visual objects in the scenes are likely to be another reason.

3.2.2. Effect on the subjective assessment

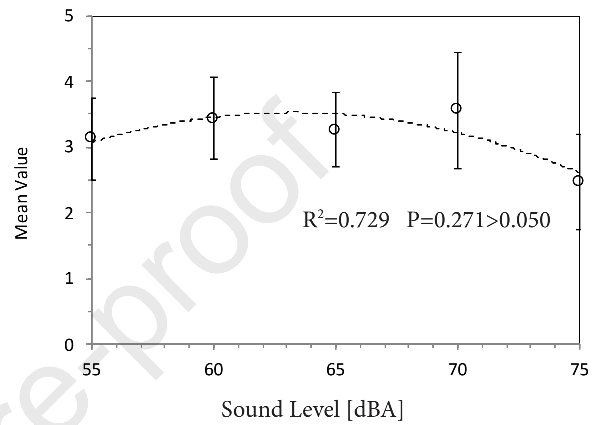
Figure 11 presents the mean value of participants' subjective assessments of the audio environment. There is a similar parabola-shaped variation trend for GA, GR, and GM. A curve estimation shows that the effects of sound level of restaurant noise on the assessment of audio environment is not significant, F(2,2)=2.693, p=0.271, with an R^2 of 0.729. However, the effects of sound level of announcement and music on the assessment of audio environment are significant, F(2,2)=75.157, p=0.013, R^2=0.987 for GA, and F(2,2)=19.613, p=0.049, R^2=0.951 for GM. For GM, the evaluation of the audio environment begins at 3.5 with music of 55 dBA, rising to 4.3 when the music is 70 dBA. The figure then falls, reaching its lowest point at 3.3 when the music is 75 dBA. When the level of the music increases, music (a pleasant sound) becomes a larger proportion of the audio stimuli (BN+music); therefore, the preference on the audio environment is increased. However, for the annoying sound (GR), the peak point moves to the lower level, similar to the results shown in Figure 9. In the last group (GT), as shown in Figure 11 (d), the level of satisfaction with the audio environment shows a downward trend, from neither satisfied nor dissatisfied (3.5) to dissatisfied (2.3). This is related to sound preference. It is also noticeable that only a slight variation (between 3.2 and 3.5) can be detected in the first group (GA). This is likely because the information in the announcement is part of the background noise in the railway station. A curve estimation shows that the effects of sound level of four types of sound on the assessment of visual environment are not significant, $p>0.05$.

Recent papers found a similar parabola-shaped variation of the acoustic evaluation under the influence of sound level. Results in previous paper show that the mean evaluation score of landscape decreases significantly as the traffic noise increases from 40 dBA to 70 dBA [15]. Meng et al. found that the effect of sound level on musical evaluation during communication is positive when sound level is smaller than 50 dBA. However, the scores of musical evaluation sharply decreases when sound level exceeded 50 dBA [8]. The level of the peak points from our results are different from that achieved in previous research, which is due to the differences in the sound source type and background noise level.

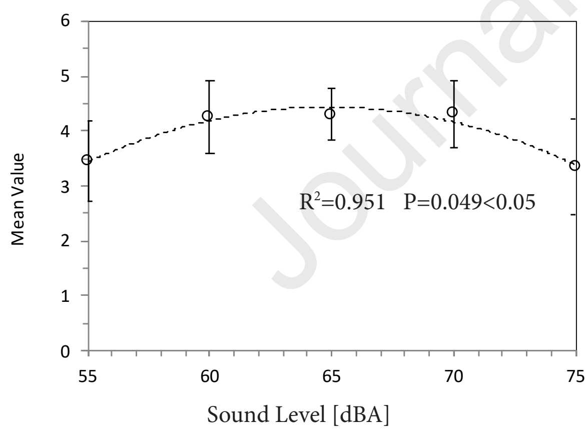
In conclusion, the sounds of announcements, restaurant activity, music, and trains have a leading effect on visual



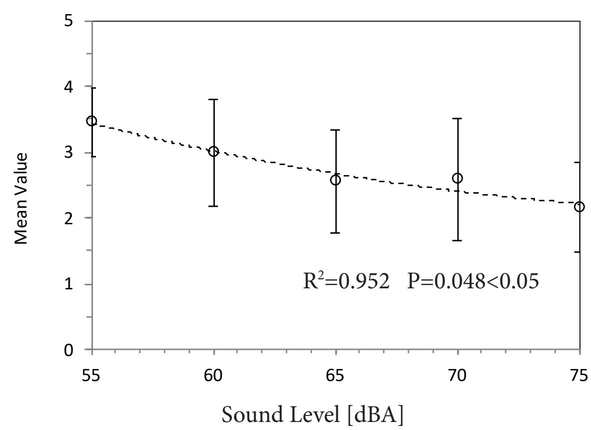
() GA



() GR



() GM



() GT

Figure 11: Assessment on audio environment with different type and level of sound. The 2^{nd} order polynomial regression functions are fitted to data points (dashed lines) with coefficient of determination (R^2) and p-value. Error bars depict standard error values.

Journal Pre-proof

attention to a departure board, a restaurant, a piano, and trains. The variation of this effect and the subjective evaluations of the audio environment reflect a parabolic trend in increases of the sound signals' levels. However, the noticeability begins to decrease at a certain level – namely, 60 dBA for restaurant noise, 65 dBA for train noise, 70 dBA for announcement, and 75 dBA for music. Although there is no downward trend for 75 dBA music, it is clear that the peak point appears here. Furthermore, corresponding with the variation in visual attention, satisfaction with the audio environment increases with the sound level and declines (in varied degrees) at 65 dBA of announcement noise, 70 dBA of restaurant noise, 60 dBA of music, and 55 dBA of train noise.

3.3. Effect of the type of sound on noticeability and the subjective assessments of the physical environment

Figure 9 shows that the peak of the leading effect depends on the type of sound, with 70 dBA for music (a pleasant sound), 65 dBA-70 dBA for an announcement (informative sound), and 60 dBA-65 dBA for restaurant and train noise (negative sounds). Tolerance for pleasant sounds is much higher than for that of relatively annoying sounds, therefore the peak of the parabola moves to the higher level. The noticeability levels of a relatively positive sound (music) and a negative sound (train noise) are higher than that of a neutral sound (announcement).

Similarly, in the subjective assessments, the peak of the mean value depends on the type of sound. For the pleasant sound group (GM), the peak value occurs around 65 dBA-70 dBA. However, for the annoying sound (GR and GT), the peak points move to the lower level. This is because the sound preference is key to the subjective assessment. For the information sound group (GA), the subjective assessment is consistent with the response to the announcement itself (neither satisfied nor dissatisfied).

3.4. Effect of the context of sound on noticeability and the subjective assessment of the physical environment

Although announcements are very common in railway stations, the source object cannot be detected in the same scene. It is revealing to note what happens to visual attention when the sound and context are not matched. This is analysed and discussed in the results for GA and GM.

Figure 12 shows that the noticeability of the piano is higher (56%) when there is a corresponding audio stimuli ('BN+M65'+'md') than when associated with announcement ('BN+A65'+'md'), which is unrelated (40%). Although the departure board is not the source object of announcement, there is potentially a relationship between the information content of these two. Similarly, the noticeability of the departure board is 20% lower when the departure board scene is matched with entirely unrelated music ('BN+M65'+'ad') than when it appears with the announcement ('BN+A65'+'ad'). Furthermore, matched context (piano with music) achieves greater noticeability than partially related context (departure board with an announcement). These findings are consistent with the previous results [39].

Satisfaction with the visual and audio environment is slightly lower with completely unrelated context ('BN+A65'+'md') than with matched context ('BN+M65'+'md'). This is because the preference for music is higher than that for the announcements, which aligns with the questionnaire results discussed earlier. However, the variation in participants' assessments of the visual environment, audio environment, and perceived loudness in the other two cases

Journal Pre-proof

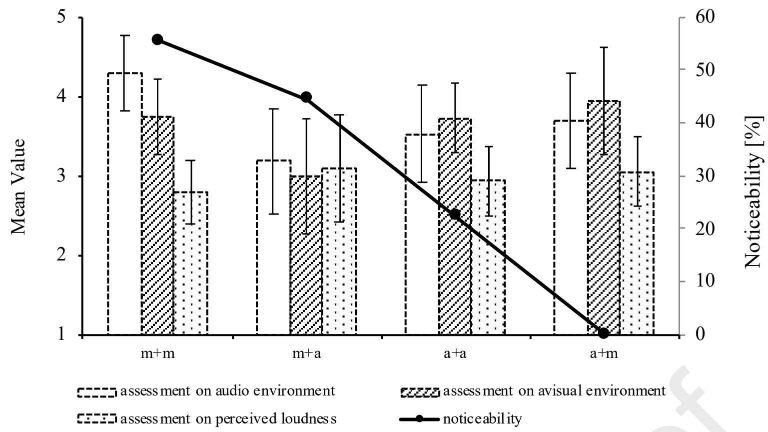


Figure 12: Assessment on visual/audio environment and percentage of noticeability with different context of sound. 'BN+M65'+'md': music(audio)+piano(visual); 'BN+A65'+'md': announcement(audio)+piano(visual); 'BN+A65'+'ad': announcement(audio)+departure board(visual); 'BN+M65'+'ad': music(audio)+departure board(visual). Error bars depict standard error values.

are less pronounced, suggesting no significant difference between the subjective assessments of the physical environment with related and unrelated contexts. This is also likely to be due to the relatively poor correlation between announcements and departure.

Therefore, the leading effect of a corresponding sound is stronger than that of an unmatched sound on a visual object. Satisfaction level is enhanced when music is matched with a corresponding context.

4. Conclusions

This research used three methods to investigate audio-visual interaction in the waiting area of a large railway station: a questionnaire survey, acoustical measurement, and a laboratory experiment. The key findings in relation to the research objectives concern the effect of the level, type, and context of the sound on its leading effect on visual attention.

The noise associated with announcements, restaurant activity, music, and trains was found to have a leading effect on the noticeability of various sound sources. With an increase in sound level, the noticeability of a departure board, restaurant, piano, and train rose, peaking at 22.2%, 33.3%, 77.8%, and 72.2%, respectively, then declining, in a parabola-shaped trend. The subjective evaluations of the audio environment also reflect this parabolic trend, with the variation in sound level corresponding with the change in visual attention.

However, the inflection point of the parabola depends on the type of sound. It appears at 70 dBA for a pleasant sound (music), 65 dBA-70 dBA for information sound (announcement), and 60 dBA-65 dBA for negative sound (restaurant and train noise). In addition, the noticeability of sounds that are relatively positive (music) or negative (train noise) are around 40%-50% higher than that of a neutral sound (announcement). Moreover, compared with annoying sounds, pleasant noise has a positive influence on evaluations of the audio environment, with ratings for such environments around 1.5 higher than those of annoying noise.

Regarding the influence of context, sound which corresponds with its visual stimuli has a more significant influence than unrelated sound on the leading effect of visual attention. In assessments of physical environment, satisfaction with the piano was higher when matched with audio stimuli than it was when paired with the announcement.

The findings of this research could be applied during the acoustic design process and wayfinding system design in large railway stations. To improve the acoustic environment, it is necessary to control the level of announcements, music, restaurant activity, and trains, keeping them to the appropriate levels identified by the parabola described in this research. For example, announcements are functionally important in railway stations, thus their sound levels should be 5-10 dBA higher than that of the background noise, thereby ensuring the noticeability of the departure board and a relatively high assessment of the audio environment (Figures 9 and 11). Similarly, the ideal sound level for music is around 10 dBA above the background noise. However, for negative sound (restaurant activity and train noise), the levels should be kept below those of the 'background'. The wayfinding system can also be improved by applying different types of sound to certain spaces. For instance, positive sound (music) has the most significant leading effect on visual attention and could therefore be employed at crossroads, in transition spaces, or at joint points to guide visitors in the right direction.

The St. Pancras Station is representative of modern large railway stations. Its planning, interior decoration, and structure are used as reference for the design of large railway stations. Moreover, the sound sources and visual objects investigated in St. Pancras are common in modern large railway stations. Therefore, the application of the findings, associated with the same sound sources and visual objects in this research, can be applied to other modern large railway stations. However, this research is only based on one case study in London. Future studies should be conducted in more large railway stations, with different layouts, decorations, and/or functions. Furthermore, the leading effect of sound on visual attention may be affected by other factors, such as the lighting environment or the scale of railway station. Further research is necessary to identify the influence of other factors on the leading effect of sound in complex audio-visual environments.

Acknowledgement

This research was supported by the Key Laboratory of New Technology for Construction of Cities in Mountain Area, Chongqing University (LNTCCMA-20190110), and the Fundamental Research Funds for the Central Universities (2020CDXYJZ008).

[1] L. Huang, Y. Zhu, Q. Ouyang, B. Cao, A study on the effects of thermal, luminous, and acoustic environments on indoor environmental comfort in offices, Building and Environment 49 (1) (2012) 304–309.

[2] T. Yu, H. Behm, R. Bill, J. Kang, Audio-visual perception of new wind parks, Landscape and Urban Planning 165 (May) (2017) 1–10.

Journal Pre-proof

Journal Pre-proof

[3] J. Zhao, W. Xu, L. Ye, Effects of auditory-visual combinations on perceived restorative potential of urban green space, Applied Acoustics 141 (July) (2018) 169–177.

[4] C. Buratti, E. Belloni, F. Merli, P. Ricciardi, A new index combining thermal, acoustic, and visual comfort of moderate environments in temperate climates, Building and Environment 139 (April) (2018) 27–37.

[5] W. Yang, H. J. Moon, Cross-modal effects of illuminance and room temperature on indoor environmental perception, Building and Environment 146 (May) (2018) 280–288.

[6] S. Zhang, X. Zhao, Z. Zeng, X. Qiu, The influence of audio-visual interactions on psychological responses of young people in urban green areas: A case study in two parks in China, International Journal of Environmental Research and Public Health 16 (10).

[7] M. Franěk, L. Režný, D. Šefara, J. Cabal, Effect of birdsongs and traffic noise on pedestrian walking speed during different seasons, PeerJ 2019 (10).

[8] Q. Meng, J. Jiang, F. Liu, X. Xu, Effects of the musical sound environment on communicating emotion, International Journal of Environmental Research and Public Health 17 (7) (2020) 1–19.

[9] G. M. Echevarria Sanchez, T. Van Renterghem, K. Sun, B. De Coensel, D. Botteldooren, Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space, Landscape and Urban Planning 167 (2017) 98–107.

[10] K. Filipan, B. De Coensel, P. Aumond, A. Can, C. Lavandier, D. Botteldooren, Auditory sensory saliency as a better predictor of change than sound amplitude in pleasantness assessment of reproduced urban soundscapes, Building and Environment 148 (September 2018) (2019) 730–741.

[11] J. Y. Hong, B. Lam, Z. T. Ong, K. Ooi, W. S. Gan, J. Kang, S. Yeong, I. Lee, S. T. Tan, The effects of spatial separations between water sound and traffic noise sources on soundscape assessment, Building and Environment 167 (September 2019) (2020) 106423.

[12] A. Haapakangas, V. Hongisto, D. Oliva, Audio-visual interaction in perception of industrial plants – Effects of sound level and the degree of visual masking by vegetation, Applied Acoustics 160 (2020) 107121.

[13] W. Yang, H. J. Moon, Combined effects of acoustic, thermal, and illumination conditions on the comfort of discrete senses and overall indoor environment, Building and Environment 148 (September 2018) (2019) 623–633.

[14] M. Szychowska, R. Eklund, M. E. Nilsson, S. Wiens, Effects of sound pressure level and visual perceptual load on the auditory mismatch negativity, Neuroscience Letters 640 (2017) 37–41.

[15] X. Ren, J. Kang, Effects of the visual landscape factors of an ecological waterscape on acoustic comfort, Applied Acoustics 96 (2015) 171–179.

Journal Pre-proof

[16] B. Schäffer, R. Pieren, U. Wissen Hayek, N. Biver, A. Grêt-Regamey, Influence of visibility of wind farms on noise annoyance – A laboratory experiment with audio-visual simulations, Landscape and Urban Planning 186 (March) (2019) 67–78.

[17] Y. H. Kim, I. H. Hwang, J. Y. Hong, S. C. Lee, Effects of vegetation on soundscape of an urban religious precinct: Case study of Myeong-dong cathedral in Seoul, Building and Environment 155 (March) (2019) 389–398.

[18] B. Cowan, D. Rojas, B. Kapralos, F. Moussa, A. Dubrowski, Effects of sound on visual realism perception and task performance, Visual Computer 31 (9) (2015) 1207–1216.

[19] D. Rojas, B. Cowan, B. Kapralos, K. Colllins, A. Dubrowski, The effect of sound on visual realism perception and task completion time in a cel-shaded serious gaming virtual environment, in: 2015 7th International Workshop on Quality of Multimedia Experience, QoMEX 2015, 2015.

[20] S. Hidaka, M. Ide, Sound can suppress visual perception, Scientific Reports 5:10483.

[21] E. S. Tusch, N. C. Feng, P. J. Holcomb, K. R. Daffner, Task-irrelevant novel sounds have antithetical effects on visual target processing in young and old adults, Frontiers in Aging Neuroscience 9 (OCT) (2017) 1–14.

[22] U. Kaya, H. Kafaligonul, Cortical processes underlying the effects of static sound timing on perceived visual speed, NeuroImage 199 (May) (2019) 194–205.

[23] F. Leo, V. Romei, E. Freeman, E. Ladavas, J. Driver, Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds, Experimental Brain Research 213 (2011) 193–201.

[24] Y. Wani, T. Terashima, Y. Tokunaga, Effect of visual information on subjective impression for sound field in architectural space, in: Proceedings of Meetings on Acoustics, Vol. 19, 2013, p. 40102.

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

[26] Y. Wu, J. Kang, W. Zheng, Acoustic environment research of railway station in China, in: Energy Procedia, Vol. 153, Elsevier, 2018, pp. 353–358.

[27] Q. Meng, P. Ren, W. Liao, Effect of users' social characteristics on soundscape in the waiting halls of railway stations, in: Advanced Materials Research, Vol. 518-523, 2012, pp. 3805–3808.

[28] J. Tardieu, P. Susini, F. Poisson, P. Lazareff, S. McAdams, Perceptual study of soundscapes in train stations, Applied Acoustics 69 (2008) 1224–1239.

[29] L. Maffei, M. Masullo, F. Aletta, M. Di Gabriele, The influence of visual characteristics of barriers on railway noise perception, Science of the Total Environment 445-446 (2013) 41–47.

Journal Pre-proof

Journal Pre-proof

[30] B. Edwards, The Modern Station: New Approaches to Railway Architecture, Taylor & Francis.

[31] J. Tardieu, P. Susini, F. Poisson, P. Lazareff, S. McAdams, Auditory information in the soundscape of a train station, Noise and Vibration Worldwide 40 (8) (2009) 13–19.

[32] J. Dodiya, V. N. Alexandrov, Use of auditory cues for wayfinding assistance in virtual environment, in: Proceedings of the 2008 ACM symposium on Virtual reality software and technology - VRST '08, ACM Press, New York, New York, USA, 2008, p. 171.

[33] S. Onat, K. Libertus, P. König, Integrating audiovisual informatiom for the control of overt attention, Journal of Vision 7 (10) (2007) 1–16.

[34] E. Van der Burg, C. N. Olivers, A. W. Bronkhorst, J. Theeuwes, Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search, Journal of Experimental Psychology: Human Perception and Performance 34 (5) (2008) 1053–1065.

[35] A. Vilaró, A. T. Duchowski, P. Orero, T. Grindinger, S. Tetreault, E. di Giovanni, How sound is the Pear Tree Story? Testing the effect of varying audio stimuli on visual attention distribution, Perspectives: Studies in Translatology 20 (1) (2012) 55–65.

[36] A. Coutrot, N. Guyader, G. Ionescu, A. Caplier, Influence of soundtrack on eye movements during video exploration, Journal of Eye Movement Research 5 (4) (2012) 2.

[37] K. Ye, H. Luo, J. Kang, S. Wu, Indoor sound environments and visual media displays: A case study on canteens, Building and Environment 176 (March) (2020) 106831.

[38] J. Y. Jeon, P. J. Lee, J. You, J. Kang, Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds, The Journal of the Acoustical Society of America 127 (3) (2010) 1357–1366.

[39] S. Viollon, C. Lavandier, C. Drake, Influence of visual setting on sound ratings in an urban environment, Applied Acoustics 63 (5) (2002) 493–511.

[40] R. W. Hayes, Ambivalent and Contradictory: Victorian Architects' Responses to TechnologyLes architectes victoriens face à la technologie : contradictions et ambivalences, Cahiers victoriens et édouardiens (87 Printemps) (2018) 1–17.

[41] E. Riot, A European perspective on the planning of major railway stations: Considering the cases of St Pancras Station and Paris Gare du Nord, Town Planning Review 85 (2) (2014) 191–202.

[42] A. T. Woods, E. Poliakoff, D. M. Lloyd, J. Kuenzel, R. Hodson, H. Gonda, J. Batchelor, G. B. Dijksterhuis, A. Thomas, Effect of background noise on food perception, Food Quality and Preference 22 (1) (2011) 42–47.

[43] S. Grimault, N. Robitaille, C. Grova, J. M. Lina, A. S. Dubarry, P. Jolicoeur, Oscillatory activity in parietal and dorsolateral prefrontal cortex during retention in visual short-term memory: Additive effects of spatial attention and memory load, Human Brain Mapping 30 (10) (2009) 3378–3392.

Journal Pre-proof

[44] W. Ijsselsteijn, H. De Ridder, J. Freeman, S. E. Avons, D. Bouwhuis, Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence, Presence: Teleoperators and Virtual Environments 10 (3) (2001) 298–311.

[45] Q. Meng, Research and prediction on soundscape in underground shopping streets, Ph.D. thesis, Harbin Institute of Technology (2010).

[46] J. Y. Hong, Z. T. Ong, B. Lam, K. Ooi, W. S. Gan, J. Kang, J. Feng, S. T. Tan, Effects of adding natural sounds to urban noises on the perceived loudness of noise and soundscape quality, Science of the Total Environment 711 (2020) 134571.

Journal Pre-proof

Highlights

- The noticeability varies as a “parabola” with the increase of sound loudness.
- The peak noticeability for pleasant sound moves to higher sound loudness.
- The noticeability for pleasant and annoying sound is higher than neutral sound.
- Corresponding sound has more significant leading effect on the visual stimuli.

Journal Pre-proof

Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof