

The influence of audio-visual aptitude on audio-visual interaction in appraisal of the environment

Dick Botteldooren¹, Kang Sun¹, Gemma Maria Echevarria Sanchez¹, Bert De Coensel¹, Timothy Van Renterghem¹

¹ University of Gent, Department of Information Technology, Research Group WAVES, Gent, Belgium

Corresponding author's e-mail address: dick.botteldooren@ugent.be

ABSTRACT

Perception and appraisal of the living environment are multi-sensory processes. People are able only to some extent to isolate one particular sensory input and appraise it separately. Hence, in a more holistic soundscape design, combined stimuli should be included. Of particular interest for appraising the sonic environment, is the role of attention. Attention and gating partly determine whether sound interferes with or promotes instantaneous activities that the listener is engaged in.

As it has become clear that a considerable part of the brain governing auditory and visual perception has a high plasticity, large differences can be expected between persons. Hence we redesigned a classical experiment for an ecologically valid setting to assess one of these personal factors: audio-visual aptitude. Both the ability of a person to distinguish small changes in the sonic environment and its resilience to visual distraction are assessed by the test. Using a noise annoyance experiment with visual context in a mock-up living room, it could be shown that there is an effect of being easily visually distracted, in particular in combination with visibility of natural green elements.

INTRODUCTION

Although most environmental noise research considers sound on itself, it has been well established that contextual factors such as the visual environment influence appraisal of the sonic environment. Indeed, even at the very first milliseconds of sensory processing both modalities interact. At the same time, effects of noise such as annoyance have been related to noticing the intruding sound in a home situation [3] and attention guiding has been put forward as a possible way to improve the soundscape in the urban open public space [13]. Attention itself is multisensory. Congruent sensory stimuli could increase the probability of noticing [9] an intruding sound (at home) or an unwanted sound (in public space). Likewise, in complex sonic scenes specific sound may go unnoticed due to inattentive deafness.

Still, evidence on the interaction of auditory and visual stimuli in the perception of the pleasantness of an urban environment [4], a noise mitigation measure [14], or the annoyance caused at home by a specific sound [5][6][7] remains non-conclusive. In this paper we explore the possibility that this may – at least partly – be caused by inter-individual differences in

audio-visual aptitude. Brain research indeed indicates that plasticity of the brain is much higher than expected. Education and past experience may shape the way the audio-visual environment is perceived even at very basic levels of processing [12]. Medication, diet, and general mental state may alter the way environmental stimuli are gated.

Three carefully designed experiments that run over multiple days were conducted with 72 volunteers to investigate how individual audio-visual aptitude influences the audio-visual interaction in the appraisal of the environment. In this paper we give a brief overview of the experiments and the main results. More details have been and will be published in several journal papers [1][2].

EXPERIMENTAL SETUP

Two experiments related to the perception of the environment were run on four separate days, the third experiment – which is of main interest here – is added on day 4.

Experiment 1: noise annoyance at home

In a mock-up living room (Figure 1) participants are asked to engage in some light activities for 15 minutes during which they hear traffic sound. After 15 minutes a standard IC BEN noise annoyance question with 11-point scale is asked, referring to the past 15 minutes. This experiment is repeated with 4 sound conditions roughly corresponding to 4 different window situations. The following days the same experiment is repeated but while participants are led to believe that they simply evaluate 4 more window glazing each day, what changes in fact is the video that is played in the background to simulate a window (Table 1). With this experimental design, we aim on one hand to go beyond loudness evaluation without evoking a hypothetical situation at home with a short sound fragment, on the other hand we hide the true purpose that is evaluation of the audio-visual interaction. More details on this experiment can be found in [2].



Figure 1: The mock up living room with hidden environmental noise loudspeakers indicated next to the mock-up window

Table 1: Snapshots from the videos played in the mock-up window

	Green elements	No green elements
Sound source visible		



Experiment 2: perception of public space

The second experiment is complementary to the first in two ways. Firstly it considers the public space, more specifically the perceived quality of a bridge crossing a ring road giving access to a park. Secondly, four visual designs – concluding different noise barriers – are evaluated while hiding the fact that our interest is in the noise. For this, each day the participants experience a walk across the bridge in a virtual environment displayed to them using oculus rift. The visual designs are rather different (Figure 2) yet the sound of the ring road stays the same. Participants are asked to rate the pleasantness of the experience without referring to sound. On subsequent days they evaluate more environments but these are visually identical yet the sound changes. More details on this experiment can be found in [1].

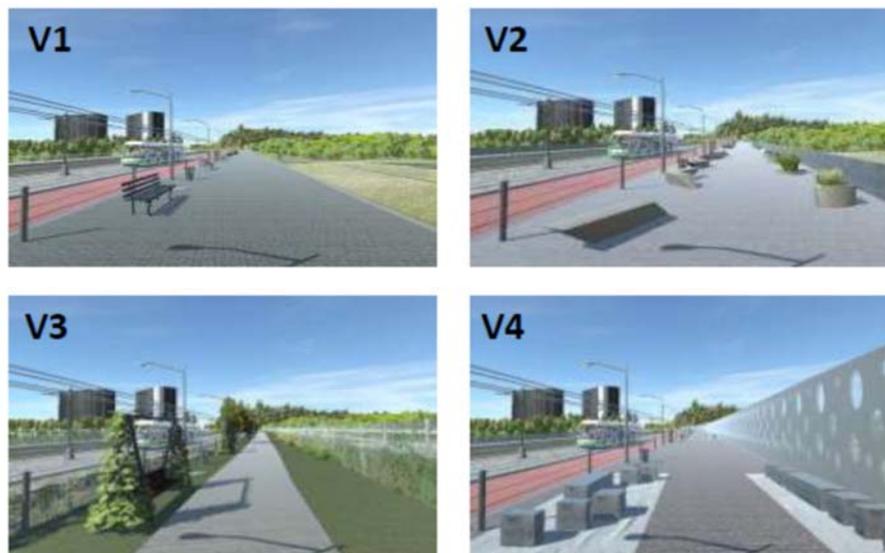


Figure 2: Snap shot of the virtual reality display of the 4 bridge designs; the barrier seen on the right progressively increases in height.

Experiment 3: audio-visual aptitude

The third experiment conducted on the 4th day only, is an extension of a psychological experiment showing audio-visual interaction in attention mechanisms, to ecologically valid, complex environmental situation. During the experiment, participants are asked to point at the sonic environment that differs from the others in a comparison between 3 auditory scenes, each played for 30 seconds or 1 minute. This is repeated for four scenarios. First this is done without visual information, then the experiment is repeated with visual information. However, the change in the visual scene is incongruent with the change in the auditory scene. The outcome of this experiment allows to identify different aptitudes. It sorts out the careful listeners with good auditory memory that are able to detect even the smallest change; it allows to identify the group that does quite well on the auditory task on itself, but gets misled by the visual information; it allows to identify the group that gets completely confused by the combination of incongruent visual and auditory information, that is they think the sound is there when they hear it and/or when they see the source.

RESULTS

Audio-visual aptitude

Because it was first piloted for difficulty, the third experiment allows to differentiate between persons. Figure 3 shows the number of errors made on the purely auditory task versus the number of errors made on the audio-visual task. Errors where the participant points at the video where the source is visible and overall errors are shown. It is clear that adding incongruent visual information results in more errors.

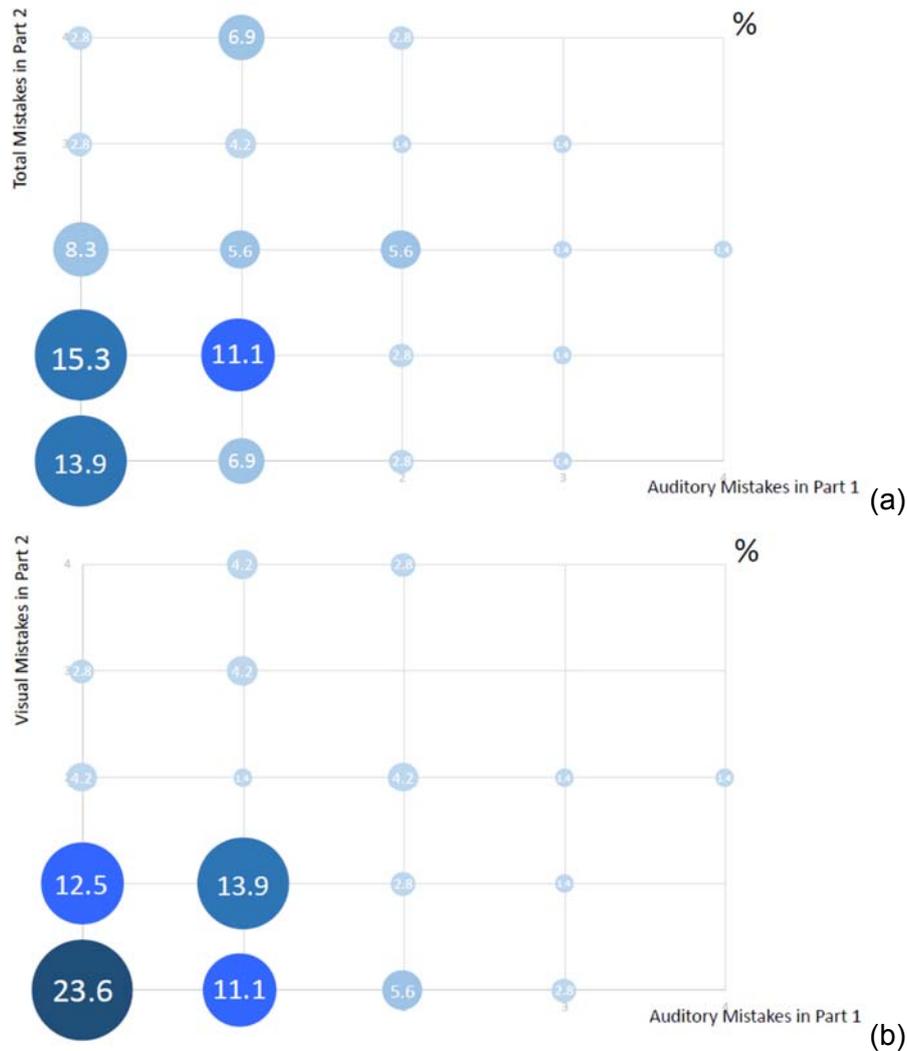


Figure 3: Percentage of the 72 participants that made none to 4 errors on the purely auditory deviant detection (Part 1) versus (a) the number of errors made on the audio-visual task (Part 2); (b) the number of errors by selecting the visual deviant (Part 2)

Detailed analysis, reported elsewhere [2], showed that the group of people that mistakenly point at the video where the source is visible as the deviant rather than the one where the sound of this source is present is the most interesting group for further analysis. This group could be described as a group where vision dominates over audition.

Effect of audio-visual aptitude on audio-visual interaction in annoyance in a home situation

The effect of the view from the window on the reported noise annoyance in the mock-up living room was investigated with noise sensitivity – assessed using Weinstein's noise sensitivity

questionnaire – and audio-visual aptitude as personal factors. Table 2 shows the result of a multiple logistic model with interaction fitted on the data from the first experiment. The expected significant dependence of reported annoyance on indoor sound exposure (labelled SPL but interpreted as an ordinal variable as also the spectrum changes) and noise sensitivity is retrieved as a single factor. In addition, being vision dominated also has a significant effect, yet sound source visibility nor green elements visibility has a single factor effect in this model with interactions. When interactions are removed, sound source visibility becomes a significant factor.

Looking at the interactions, sound source visibility seems to interact with noise sensitivity and visibility of green seems to interact with being vision dominated. The interaction of noise sensitivity and source visibility shows that the effect of source visibility on noise annoyance is different for sound sensitive people and sound insensitive people. For the former source visibility tends to decrease annoyance, for the latter annoyance increases with source visibility. The interaction between being vision dominated and visibility of green shows that the annoyance rating by vision dominated people is more strongly affected by the visibility of green, but green elements seem to increase the annoyance for this group. This could be due to a difference in expectations.

Table 2: Multiple linear regression model for reported noise annoyance with interaction effects and person as a random factor.

Fixed Effects	Target: Annoyance			
	F	df1	df2	Sig.
Source				
Intercept	50.283	15	1.087	.000
Gender	2.438	1	1.087	.119
Education	0.925	1	1.087	.336
Age	2.866	1	1.087	.091
Sensitivity	5.960	1	1.087	.015
Auditory acuity	0.020	1	1.087	.888
Vision dominated	4.129	1	1.087	.042
SPL	236.894	3	1.087	.000
Green	2.254	1	1.087	.134
Sound source	0.352	1	1.087	.553
Sensitivity*Green	1.610	1	1.087	.205
Sensitivity*Sound source	5.941	1	1.087	.015
Vision dominated *Green	4.894	1	1.087	.027
Vision dominated *Sound source	0.098	1	1.087	.754

*Participant is used as random factor

Effect of audio-visual aptitude on perceived quality of the bridge (public space)

The pleasantness of the experience of crossing the bridge is most strongly influenced by the visual design, yet the reduction of ring road traffic noise has a positive influence on the overall appreciation, even if the sound is not explicitly mentioned as a factor. This effect has been explored in detail in [1]. Here we focus on the influence of audio-visual aptitude as a personal factor. Figure 4 shows the pleasantness rating of crossing the bridge for the 4 designs shown in Figure 2 for different subgroups of audio-visual aptitude of participants. Although some interesting differences can be observed, only the difference between the pleasantness rating for V4 by participants in the group that makes no mistakes on the auditory deviant detection task and participants in other groups is statistically significant. While interpreting these results, it should be noted that the matching sound decreases with increasing design number. Design V4 includes a high barrier and therefore has the lowest road traffic noise, yet participants in

general do not like this design as much as the ‘green’ design V3. Participants that are audition dominated (GANoVD) seem to rate the pleasantness for V4 very similar to the designs V2 and V3 that are in general rated as more attractive. The reduction in noise level seems to partly compensate for the less attractive visual design.

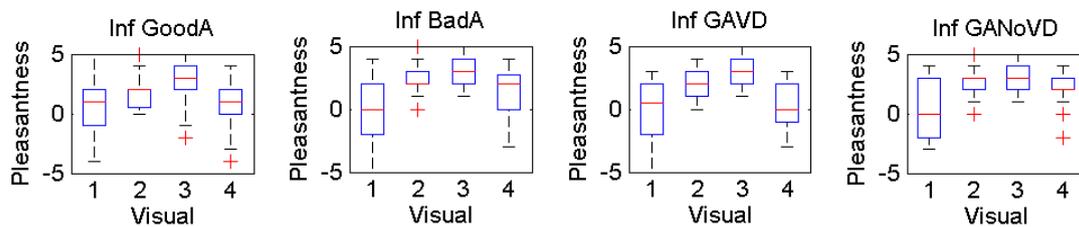


Figure 4: Pleasantness rating of the experience of crossing a bridge over the ring road (in the virtual reality experiment) depending on design (V1 to V4) for different groups of participants: GoodA=those making no mistakes on the auditory deviant detection; BadA=those making at least one mistake in the auditory deviant detection; GAVD=those making no mistake until the incongruent visual is shown; GANoVD=those making no mistake whether or not the visual is present.

DISCUSSION

This work investigates a personal factor that could influence how the interaction between auditory and visual stimuli affects the appraisal of the environment. For this, a deviant detection experiment was designed with on the one hand complex auditory scenes where one sound event is added to generate the deviant, on the other hand the same auditory experiment with incongruent visual information added as a distractor. Because of the complexity of the auditory scene, inattentional deafness or differences in auditory memory capacity could lead to differences between people in the purely auditory task. Note that persons with hearing problems detected via tonal audiometry or reported by the participant were excluded from any analysis.

The difference in effect of the visual distractor may be due to the difference in auditory or visual dominant attention mechanisms even at the first milliseconds of auditory processing [9]. In addition, increasing numbers of errors when visual information is present may also be the result of visually mediated inattentional deafness [10]. This may be affected by individual differences in capacity [11] as observed in dual task (one auditory, one visual) tests.

No matter what the exact underlying mechanism is that describes the differences between people in their audio-visual aptitude, the factor extracted from the experiment seems to explain some of the inter-individual differences in two experiments, one related to noise annoyance at home, one related to the perceived pleasantness in using the public space. In the noise annoyance experiment, being a vision dominated listener is almost as significant as noise sensitivity – a known stable personality threat – but more importantly this personal factor seems to interact with the visibility of green. Age, education level, and gender were also included in the model and hence could be excluded as confounders. Reported noise annoyance by people that show easy distraction in the deviant detection test is more affected by visibility of green. The perceived pleasantness of crossing a bridge leading to a park is influenced by traffic noise and hence including a noise barrier in the design makes sense. However, the highest and thus most effective noise barrier was not appreciated as creating a pleasant environment on average. Persons that are not easily distracted visually or in other words are more apt in detecting auditory differences in complex audio-visual environments seem nevertheless to appreciate the design including the high noise barrier more than their peers.

In summary, if we call the ability to detect deviants in a complex auditory scene, even in the present of a visual distractor, *audio-visual aptitude*, we suspect that:

- High audio-visual aptitude leads to a lower effect of visibility of green from the window on perceived noise annoyance at home
- High audio-visual aptitude would make people less sensitive to visual quality of noise mitigation in the public space.

This personal factor should thus be a subject of future investigation.

Acknowledgements

Gemma Maria Echevarria Sanchez is funded by the People Programme Marie Curie Actions of the European Union's Seventh Framework Programme FP7/2007e2013/under REA grant agreement n° 290110, SONORUS "Urban Sound Planner".

Kang Sun is funded by the Chinese Scholarship Council (CSC), the support of this organization is gratefully acknowledged.

REFERENCES

- [1] Echevarria Sanchez, G., Van Renterghem, T., Sun, K., De Coensel, B., Botteldooren, D. (2017), Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space , *submitted to Landscape and Urban Planning*
- [2] Sun, K., De Coensel, B., Echevarria Sanchez, G., Van Renterghem, T., Botteldooren, D. (2017) The interaction between personal attention focusing capability and noise sensitivity and the influence of sound source visibility and green elements visibility on road traffic noise annoyance, *in preparation*.
- [3] De Coensel, B., Botteldooren, D., De Muer, T., Berglund, B., Nilsson, M. E., & Lercher, P. (2009). A model for the perception of environmental sound based on notice-events. *The Journal of the Acoustical Society of America*, 126(2), 656-665.
- [4] Viollon, S., Lavandier, C., & Drake, C. (2002). Influence of visual setting on sound ratings in an urban environment. *Applied acoustics*, 63(5), 493-511.
- [5] Li, H. N., Chau, C. K., & Tang, S. K. (2010). Can surrounding greenery reduce noise annoyance at home?. *Science of the Total Environment*, 408(20), 4376-4384.
- [6] Van Renterghem, T., & Botteldooren, D. (2016). View on outdoor vegetation reduces noise annoyance for dwellers near busy roads. *Landscape and urban planning*, 148, 203-215.
- [7] Bangjun, Z., Lili, S., & Guoqing, D. (2003). The influence of the visibility of the source on the subjective annoyance due to its noise. *Applied Acoustics*, 64(12), 1205-1215.
- [8] Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of cognitive neuroscience*, 11(5), 473-490.
- [9] Van den Brink, R. L., Cohen, M. X., van der Burg, E., Talsma, D., Vissers, M. E., & Slagter, H. A. (2014). Subcortical, modality-specific pathways contribute to multisensory processing in humans. *Cerebral cortex*, 24(8), 2169-2177.
- [10] Macdonald, J. S., & Lavie, N. (2011). Visual perceptual load induces inattentive deafness. *Attention, Perception, & Psychophysics*, 73(6), 1780-1789.
- [11] Raveh, D., & Lavie, N. (2015). Load-induced inattentive deafness. *Attention, Perception, & Psychophysics*, 77(2), 483-492.
- [12] Altieri, N., Stevenson, R. A., Wallace, M. T., & Wenger, M. J. (2015). Learning to associate auditory and visual stimuli: behavioral and neural mechanisms. *Brain topography*, 28(3), 479-493.
- [13] Oldoni, D., De Coensel, B., Boes, M., Rademaker, M., De Baets, B., Van Renterghem, T., & Botteldooren, D. (2013). A computational model of auditory attention for use in soundscape research. *The Journal of the Acoustical Society of America*, 134(1), 852-861.
- [14] Maffei, L., Masullo, M., Aletta, F., & Di Gabriele, M. (2013). The influence of visual characteristics of barriers on railway noise perception. *Science of the Total Environment*, 445, 41-47.