PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Analysing the effectiveness of approaches to auralisation for applications in environmental acoustics

Alex SOUTHERN[(1)], Frank STEVENS[(2)], Damian MURPHY[(2)]

[(1)]AECOM Ltd., United Kingdom, alex.southern@aecom.com

[(2)]AudioLab, University of York, Heslington, frank.stevens@york.ac.uk / damian.murphy@york.ac.uk

Abstract

Auralisation is becoming more commonplace as a means of communicating and disseminating plans for changes in our environmental landscape due to major infrastructure projects such as road, rail and airport developments, where noise can be a significant issue. This paper presents an approach to this challenge in the context of a major public consultation for a UK road project as well as commenting on different means for facilitating presentation of, and engagement with, both audio and visual material. Although such methods are generally used for information only, every effort is made to ensure that the final result is both plausible and comparable to existing on-site conditions, and calibrated accordingly, with a view to supporting other more standardised methods and metrics. However, it is also possible to develop more quantitative measures as to how effective and successful such auralisations are in terms of conveying a representation of a given scene, and any planned for interventions. In this work the Self Assessment Manikin and soundscape classification descriptors are used in a series of listening tests to better understand the approach that has been adopted for these types of projects.

Keywords: Sound, Insulation, Transmission

1 INTRODUCTION

This paper presents the results of a listening study that uses the Self Assessment Manikin and soundscape classification descriptors, as presented previously in (14)(15), for road traffic auralisations (16) as typically presented at public consultations as part of the UK planning consent process. The motivation for the study is to understand if these methods for rating and categorising auralisations, or *sound demonstrations*, using simple and quick to administer surveys might provide a way to enhance the planning and decision making processes when considering the potential noise impacts of a proposed development. Note that the use of *sound demonstration* is preferred as a less technical and more accessible term when dealing with and presenting auralisations to non-technical stakeholders and user groups. We also analyse the results using the Perceived Noise Impact Rating (R_{PNI}) (12) to better understand the general applicability of this metric for quantifying attitudes towards noise impact.

The remainder of the paper is organised as follows: first we provide some background to the consultation process and the adoption of auralisations for stakeholder engagement; Section 3 presents the method and technique for the survey and is followed by the results and conclusions.

2 BACKGROUND

Nationally Significant Infrastructure Projects (NSIPs) are developments considered by the UK Government to be of significant national importance such that the developer must apply to the Planning Inspectorate for a Development Consent Order (DCO) to receive permission to undertake the work (3), rather than the more usual application for local planning permission. The process of applying for a DCO is set out in the UK Planning Act 2008 and has 6 stages: Pre-application; Acceptance; Pre-examination; Examination; Decision and Post-Decision. The first stage, Pre-application, is relevant to this work as there is a statutory requirement for the developer to consult the public, and other relevant concerned bodies, and provide an opportunity for them to comment and potentially influence the developing proposals. Prior to this statutory consultation it is not uncommon for developers to undertake an additional non-statutory consultation where design options are presented and comments from the public and other bodies are invited.

The government provides relevant guidance on consultation principles (1)(2) and included in this advice are suggestions to, "avoid lengthy documents", and provide, "enough information to ensure those consulted understand the issues and can give informed responses". The guidance also suggests that consideration should be given to tailor the consultation to the needs and preferences of different groups or demographics, such as older people or younger people which might not respond to traditional consultation methods.

In relation to communicating noise impacts, the traditional approach is to present noise level predictions for any proposed changes in the form of sound-level/noise contours that present the relative change from a baseline. This usually requires an acoustic specialist to help interpret the resulting contour maps and so avoid misinterpretation of what changes in decibel quantities might mean. While such contour maps are well established as a means of conveying the results of noise level predictions, they do not help many stakeholders understand what the real-world impact might be in terms of changes in overall or frequency dependent sound level, and by extension what the resulting qualitative sound character might be. Furthermore, this limitation also makes the positive/negative impacts of any scheme, design or planned mitigation strategy equally difficult to communicate effectively.

Such consultations are seen as the best opportunity for local communities, local authorities and other stakeholders to influence proposed designs and so understand any potential impacts, positive and/or negative. Hence, in recent years, there has been a growing interest and demand for the use of auralisations - or sound demonstrations - at stakeholder engagement events on NSIPs to more effectively present, and enable engagement with, changes in noise levels, and by extension, the associated soundscape. Sound demonstrations presenting multiple acoustic contexts arising from various scheme options can therefore help to highlight the potential impact of the proposed scheme, particularly in relation to road, railway or airport developments, where changes in noise levels resulting from the scheme's completion are potentially significant and of wider public concern. By making these comparisons directly, stakeholders are able to arrive at a more informed opinion about how the development might impact on them and their community.

However, what is missing from any such consultation, given that the sound demonstrations are used to inform rather than to formally test, is any robust measure of the perceived impact of the resulting soundscape as presented to a listener. The next section presents the methods that will be used to gather and assess such perceptual data based on established methods from environmental soundscape studies.

3 METHOD

This section introduces the key methods used in the listening tests that follow, the metrics that will be derived and analysed, a description of the the auralisations used in the study, and the approach to administering the listening test survey itself.

3.1 The Self Assessment Manikin

The Self Assessment Manikin (SAM) is a method for measuring emotional responses to various stimuli developed by Bradley and Lang in 1994 (5). It was developed from factor analysis of a set of Semantic Differential (SD) pairs rating both aural (6) and visual stimuli (4) (using, respectively, the International Affective Digital Sounds database, or IADS, and the International Affective Picture System, or IAPS). The three factors developed for rating emotional response to a given stimuli are:

- **Valence:** How positive or negative the emotion is, ranging from unpleasant feelings to pleasant feelings of happiness.
- **Arousal:** How excited or apathetic the emotion is, ranging from sleepiness or boredom to frantic excite-

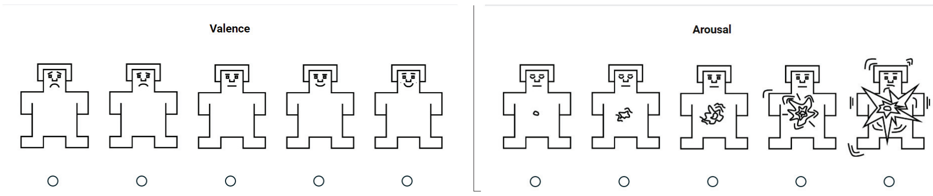

Figure 1. The Self-Assessment Manikin (SAM) as used in this study, after (5) previously used in (13).

ment.

- **Dominance:** The extent to which the emotion makes the subject feel they are in control of the situation, ranging from not at all in control to totally in control.

These results were used by Bradley and Lang to create the SAM itself as a set of pictorial representations of the three identified factors. The version of the SAM used in this experiment (as shown in Fig. 1) contains only the Valence and Arousal dimensions following results from previous studies showing these two factors of the SAM to be most useful in the context of environmental soundscape assessment (13).

3.2 Soundscape Classification Descriptors

A *soundscape* is usually considered as the aural equivalent of a landscape (10), but more recently the *ISO 12913-1:2014, Acoustics - Soundscape - Part 1: Definition and conceptual framework* (7) has provided a more precise definition, where soundscape is defined as the perception or understanding of an acoustic environment which is made up of sound sources. Typically, sound sources can be split into three categories - natural, human, and mechanical - and previous work has shown how the prevalence of sounds belonging to these categories might affect the subjective rating of a soundscape in terms of emotional experience, as well as how a soundscape is perceived as belonging to each of these three categories when used more generally as soundscape classification descriptors (14). The work showed that soundscapes rated as more mechanical in nature tend to exhibit lower valence and higher arousal in the SAM results, and that highly natural soundscapes will exhibit high valence and low arousal. We adopt these descriptors again in the survey for this work in order to gain further understanding of how they might provide a more robust and quantitative insight to the collective opinion of a surveyed population subjected to a series of prepared sound demonstrations.

3.3 Survey Approach

The survey was administered as an online set of 20 questions which required the subject to use over ear headphones set at a constant volume and undertaken in a relatively quiet space. The majority of the participants were audio or acoustics professionals from various parts of the world, including New Zealand, Australia, USA and Canada, but predominantly the UK. Each question required a 10s auralisation to be rated using the Valence and Arousal dimensions of the SAM (using a discrete selection on the 5-point scale of the SAM) and using the three soundscape classification descriptors (also using a discrete selection on a 5-point scale ranging from 'Not At All' through to 'Very Much'). The 10s audio clips represented five locations relating to a typical road traffic scheme in the UK *Without* and *With* the proposed scheme in place - that is, the soundscape as currently experienced, and the soundscape as potentially experienced with any proposed changes made. The participants were not aware of these locations, nor whether the examples were presented with or without scheme. The clips were presented in a random order for each participant. The results are presented in Section 4 and include the perceived noise impact rating which is explained below in section 3.4. The audio clips consist of a mix of rural countryside sounds with distant road traffic, busy A-roads (dual and single carriageway), and road junction approaches consisting of vehicles moving at different speeds and accelerating/decelerating. Visual stimuli did not accompany the audio clips to avoid biasing test subjects towards preference with or without scheme, and to

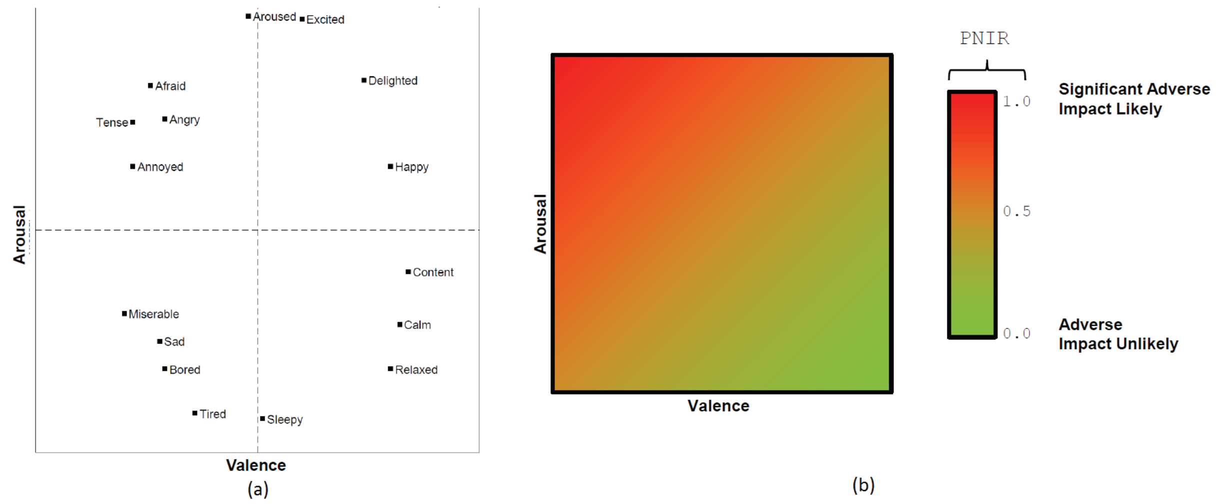

Figure 2. In (a) an example of a the circumplex model of affect defined by the valence and arousal dimensions of the SAM. A 2D graphical representation of emotion. In (b) a visual representation showing how R_{PNI} maps onto the the Circumplex Model of Affect.)

avoid any bias based on recognising the locations being presented.

3.4 The Rating of Perceived Noise Impact

Valence and arousal can be used to represent emotional state on the *Circumplex Model of Affect*, as shown in Figure 2(a) (9). Derived from this we also used the Rating of Perceived Noise Impact, R_{PNI}, in this study as first introduced in (12):

$$R_{PNI} = 1 - 0.5(1 - A + V) \qquad (1)$$

where parameters A and V are the Arousal and Valance scores from the SAM. R_{PNI} can therefore be determined for an individual or across a population, or subset or scenarios, providing the convenience of a single figure objective parameter, but with a relationship to the underlying subjective responses for a presented soundscape. R_{PNI} has been shown to be an appropriate measure of the subjective experience of a soundscape, and potentially able to identify differing emotional states not shown by arousal and valence results individually (13). Figure 2(b) presents a visual representation of how R_{PNI} maps onto the Circumplex Model of Affect from which it has been derived.

4 RESULTS

In the following the results are presented from 28 test participants for the Soundscape Categorisation, SAM and R_{PNI} for responses to all 20 questions for the 5 locations With (W) and Without (WO) the scheme in place in Figure 3, Figure 4 and Figure 5 respectively. For each rating scale the mean value is plotted with bars showing the standard error.

At Locations 1, 2 and 5 it can be observed that with the scheme in place the participants have rated the soundscape in a positive way - higher valence, lower arousal, lower R_{PNI}. This quantification supports the informal comments and opinions expressed by some of the participants after listening to the clips. Similarly, for location 3 there seems to be a slight benefit from the scheme, for example at this location the existing road is barely audible and becomes even less audible with the scheme in place. This situation is captured and confirmed by the survey results, the arousal is slightly lower and the valence is slightly higher with the scheme in place and this is also supported by the R_{PNI} which is lower with the scheme in place. It is worth recalling

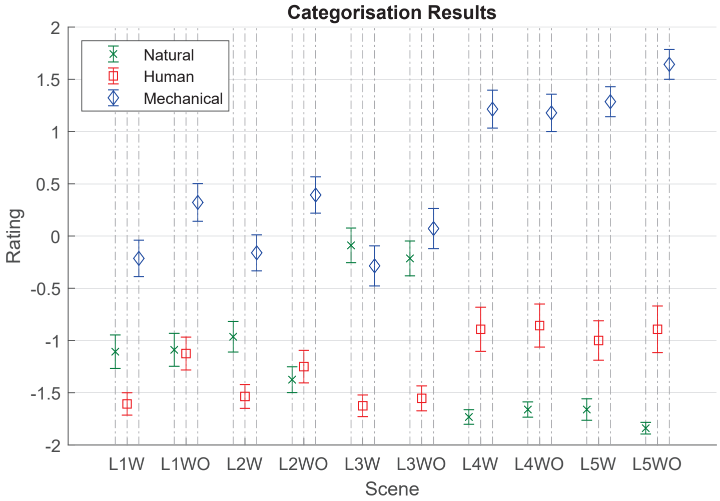

Figure 3. Results of the categorisation scoring for all 5 locations with (W) and without (WO) the scheme in place. The mean value is plotted with bars showing the standard error.

that the W and WO scheme scenarios are presented randomly and cannot be directly compared AB style by the participant. This is an example of the robustness of the survey method for capturing the opinion towards more subtle changes in the perceived soundscape.

The results for location 4 are mixed/negligible and again this corresponds with the informal comments of the participants. At location 4 the road traffic sound level does not differ between with and without scheme, this is because the busy local road heard is not affected by the scheme being in place and as a result masks any increase in road traffic resulting from the scheme. This means there is no change at this location, and it is not a pleasant place to be, both of which are apparent when looking at the categorisation results, which are the same in both with and without scheme across all of the results in Figure 3, Figure 4 and Figure 5 (a), Further to this L4 and L5W scored very similarly for valence and R_{PNI}, despite one being distant highspeed road traffic and the other being close proximity engine and tyre noise including acceleration and deceleration vehicle movements. If we put these in to the context of L5WO which was the soundscape that was rated as most mechanical and received the lowest valence scores, it suggests that while L4 and L5W are not idyllic the results could well be indicating that R_{PNI} in the region of 0.5 and 0.6 indicates an acoustic environment that is tolerable.

In addition, it can be observed that for this dataset that the mechanical and valence scores are inversely related. This suggests that the presence of mechanical sounds when being used to quantify vehicles and road noise might well be used to estimate the valence scores, and as such reduce the number of ratings that consultees need to do.

Overall, looking at the R_{PNI} scores, and calculating the differences between them With and Without scheme is one way to simply indicate objectively the mood of the surveyed population, see Figure 5 (c). This quantifies the most beneficial and contentious parts of a scheme and where people are less concerned about noise.

The test results were also tested for normality using a Shapiro-Wilk's test (11). This testing indicated that none of the results were normally distributed. As such, a Mann-Whitney U-test (8) has been used to evaluate the significance of the effect of the presence of each scheme at each of the five locations.

Figure 5 (b) shows the Mann-Whitney U-test results for the five locations, indicating where there is a significant difference between the results for each rating scale (i.e. valence, arousal, the three soundscape categories, and R_{PNI}) for each of the five locations, comparing the location with and without the scheme present in each case.

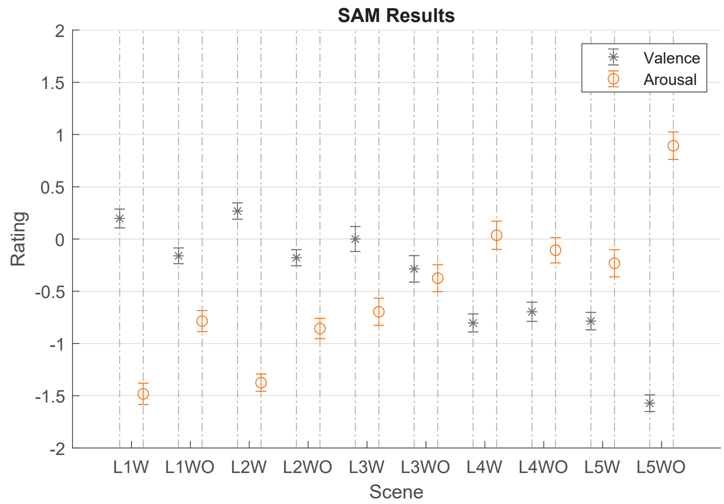

Figure 4. Results of the SAM scoring of Valence and Arousal for all 5 locations with (W) and without (WO) the scheme in place. The mean value is plotted with bars showing the standard error.

It is clear from these results L3 and L4 which as previously described have very little or no change, respectively have been quantified as such by the participants of the survey by looking at the Mann-Whitney U-test results.

5 CONCLUSIONS

In this paper we have presented results of a soundscape listening survey undertaken with the goal of quantifying the collective opinion of the soundscape survey participants. Currently the feedback, if any, from sound demonstrations delivered at public consultation does not tend to capture or quantify the opinions of the people listening to them. It could be helpful to the planning process if general mood and attitudes of the population could be quantified in some objective way and this work has explored this possibility.

The results have indicated that a combination of the soundscape categorisation rating and the valence and arousal dimensions of the self assessment manikin presented in the form of R_{PNI} could provide a means of quantifying subjective opinions about a scheme. This approach has potential value to the planning application process to help developers quantify where public opinion about noise is being driven by factors/attitudes other than noise, and to communicate in the planning application the locations where changes in road traffic noise level are subjectively negligible, adverse or beneficial. The relative magnitude of the planned changes can also be indicated in the form of differences in the perceived noise impact rating between with and without scheme in place.

The results also support the findings from previous work in (14) that the soundscapes rated as more mechanical in nature tend to exhibit lower valence and higher arousal in the SAM results, and that highly natural soundscapes will exhibit high valence and low arousal. It was observed that for the test soundscapes used in this study, which consisted of mainly natural and/or or road related noise, that the mechanical and valence scores might be inversely related and as such could mean one of them could be omitted when conducting surveys on soundscapes with a similar sound character.

The results in this work also support previous findings that the SAM is a powerful tool for soundscape analysis, and further to this the calculation of the perceived noise impact rating is a potential useful and informative way to quantify the attitudes of survey group towards planned changes to an environment and including the magnitude of those changes whether positive or negative.

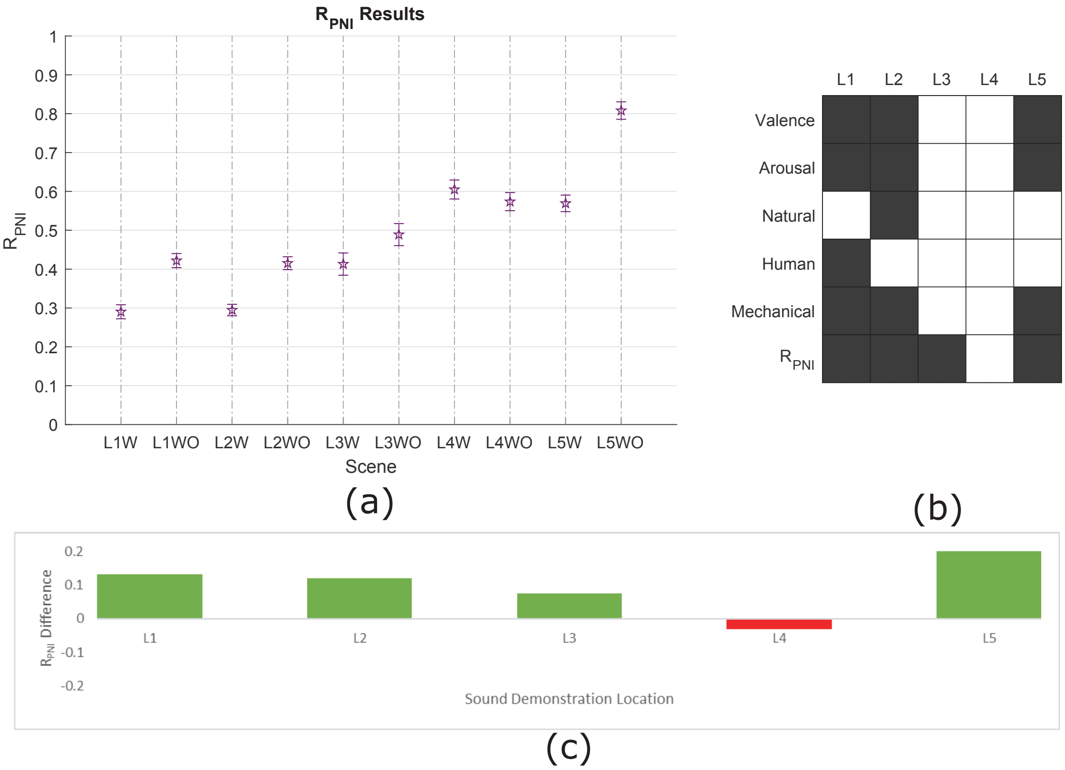

Figure 5. In (a) Results of the perceived noise impact rating, R_{PNI}, calculated from the SAM valence and arousal scores using equation 1 for all 5 locations with (W) and without (WO) the scheme in place. The mean value is plotted with bars showing the standard error. In (b) Mann-Whitney U-test results comparing the results for each rating scale for each of the 5 locations with and without the scheme in place. A dark square indicates a significant difference at 95% confidence ($p < 0.05$). In (c) The difference between the With and Without results for the Perceived Noise Impact Rating, a positive difference indicates a positive attitude towards the potential change.

Further work should include new surveys for other schemes to help confirm the findings in this work, and would ideally involve members of the public, unlike this work in which acoustics and audio specialists were the test participants.

REFERENCES

[1] Consultation principles: guidance, The Cabinet Office, Last updated 19 March 2018. https://www.gov.uk/government/publications/consultation-principles-guidance. (Accessed: 2019-05-29).

[2] Planning Act 2008: Guidance on the pre-application processs, Department for Communities and Local Government, March 2015. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/418009/150326_Pre-Application_Guidance.pdf. Accessed: 2019-05-29.

[3] The Planning Inspectorate planning process. https://infrastructure.planninginspectorate.gov.uk/application-process/the-process/. Accessed: 2019-05-29.

[4] M. Bradley, B. Cuthbert, and P. Lang. Picture media and emotion: Effects of a sustained affective context. *Psychophysiology*, 33(6):662–670, 1996.

[5] M. Bradley and P. Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.

[6] M. Bradley and P. J. Lang. *The International affective digitized sounds (IADS): stimuli, instruction manual and affective ratings*. NIMH Center for the Study of Emotion and Attention, 1999.

[7] Acoustics – Soundscape – Part 1: Definition and conceptual framework. Standard, International Organization for Standardization, Geneva, CH, Sept. 2014.

[8] H. Mann and D. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.

[9] J. A. Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980.

[10] R. M. Schafer. *The soundscape: Our sonic environment and the tuning of the world*. Simon and Schuster, 1993.

[11] S. Shapiro and M. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, pages 591–611, 1965.

[12] A. Southern, F. Stevens, and D. Murphy. Sounding out smart cities: Auralization and soundscape monitoring for environmental sound design. *The Journal of the Acoustical Society of America*, 141(5):3880–3880, 2017.

[13] F. Stevens. *Strategies for Environmental Sound Measurement, Modelling, and Evaluation*. PhD thesis, University of York, 2018.

[14] F. Stevens, D. T. Murphy, and S. L. Smith. Soundscape categorisation and the self-assessment manikin. In *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, Edinburgh, UK, 2017.

[15] F. Stevens, D. T. Murphy, and S. L. Smith. Soundscape auralisation and visualisation: A cross-modal approach to soundscape evaluation. In *DAFx 2018*, Aveiro, Portugal, 9 2018.

[16] M. Vorländer. *Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. Springer Science & Business Media, 2007.