University of Salford MANCHESTER

Acoustics Research Centre

A.J.TorijaMartinez@salford.ac.uk / A.S.Elliott@salford.ac.uk
http://www.acoustics.salford.ac.uk

ACOUS05235: Generation and Analysis of Artificial Warning Sounds for Electric Scooters

Dr Antonio J Torija Martinez & Dr Andrew Elliott
University of Salford, UK.

Client: emTransit B.V.
Address: Burgerweeshuispad 101,
1076 ER Amsterdam

Tel: +49 1520 7017095
Email: sebastian.schlebusch@ridedott.com
Report reference ACOUS05235
Report Issue Date: August 27th, 2021

A. Elliott
Approved: ……………………
Dr. Andrew Elliott

Approved: ………………………
Dr. Antonio J Torija Martinez

Research Team

This report has been prepared by:

- Dr Antonio J Torija Martinez
- Dr Andrew Elliott
- Dr Lara Harris
- Dr Zuzanna Podwinska
- Mr Connor Welham
- Mr Rory Nicholls
- Mr Bernard Steer
- Dr Duncan Williams

Executive Summary

- emTransit B.V (Dott) is a European mobility operator currently operating over 30,000 electric scooters in Belgium, France, Germany, Italy, Poland and now the UK. The company aims to expand its UK operations and **has recently won a tender for the Transport for London e-scooter trials**.
- Dott scooters is looking to mitigate potential safety hazards to pedestrians with the use of an **Acoustics Vehicle Alerting System (AVAS)** for distinct e-scooter category.
- This report presents the work carried out by Salford's Acoustics Research Centre (ARC) to create a stand-alone device to generate warning sounds for Dott's e-scooters.
- This report includes:
  - Sound generation process
  - Analysis of the warning sounds
  - Explanation of the implementation of a subjective experiment
  - Conclusions and recommendations for next steps, including how to optimise the sound generation system on the scooter, and how to continue the research for designing optimal warning sounds to maximise vehicle noticeability without increasing noise annoyance.
- Key outputs are:
  - A system (including hardware and software) has been developed to **generate in real time a warning sound, according to the scooter's operating conditions** (e.g., vehicle speed).
  - A laboratory study has been carried out to gauge pedestrian awareness of an approaching e-scooter with and without a warning sound added. Preliminary results suggest that **a significant benefit, in terms of vehicle noticeability, is observed with the addition of a warning sound.** Of the sounds tested, the addition of a broadband sound with modulated tones seems to be the most effective sound increasing vehicle noticeability.
- The development of technologies, innovations, goods and services within the Clean Growth sector, for instance **sustainable and inclusive micro-mobility**, is in strategic alignment with the University of Salford.

Table of Contents

1. Introduction

emTransit B.V (Dott) is looking to mitigate the potential safety hazards to pedestrians, with specific focus on visually impaired people, through the creation of an Acoustic Vehicle Alerting System (AVAS) for distinct e-scooter vehicle category. Therefore, Dott Scooters have a requirement to produce warning sounds on their range of electric scooters starting with model Okai O3. The warning sounds are required to meet specific requirements in terms for acoustic features to improve noticeability. Some of these requirements are in terms of one third octave band sound levels (see UNECE regulation R138.01[1]) but there is some freedom to meet these requirements with a brand specific sound.

Of primary importance is that the warning sounds are effective (i.e., increase noticeability of the vehicle) but at the same time it is also preferred that the generated sounds do not cause annoyance, or ideally, that they are perceived as positive.

The University of Salford's Acoustics Research Centre (ARC) has carried out a three-month project with the following objectives:

1. Development of a standalone hardware to generate warning sounds.
2. Development of a framework for synthesising warning sounds with different acoustic features (adapted to specific requirements).
3. Acoustic characterisation of warning sounds generated (sound level, frequency spectrum and directivity).
4. Investigation of noticeability of warning sounds using state-of-the-art sound quality metrics.
5. Development of a framework to investigate noticeability of warning sounds in context, via subjective experiments.

This project has been developed in closed collaboration with colleagues of Dott and Royal National Institute of Blind People (RNIB).

This report presents a proof of concept that is intended to inform a longer-term plan to implement an AVAS on future e-scooters as part of an integrated system. For the purposes of this project, the hardware system developed for producing warning sounds can be retrofitted in existing e-scooters.

Using off the shelf components the standalone electronic device has been assembled (in consultation with Dott). The electronic device is able to read the operational state of the e-scooter

---

[1] United Nations Economic Commission for Europe (UNECE). (2017). Regulation No. 138 of the Economic Commission for Europe of the United Nations (UNECE) — Uniform Provisions Concerning the Approval of Quiet Road Transport Vehicles with Regard to Their Reduced Audibility. ECE/TRANS/WP.29/GRB/2012/6, United Nations Economic Commission for Europe, Geneva, Switzerland.

(i.e., motor supply current or voltage, wheel/motor rotational speed, or motor vibration) and convert these quantities into warning sounds that represent the point of use.[2]

A subjective experiment has also been carried out where demo sounds produced in context (e.g., once heard in typical urban soundscapes) have been presented to a sample of participants. This experimental setup allows:

- The understanding of what acoustic signals improve vehicle awareness.
- The definition of minimum requirements to develop an AVAS for e-scooters considering a diversity of urban environments.
- The identification of the least annoying effective warning sounds.

This research, informing the development of technologies, innovations, good and services within the Clean Growth sector, such as micro-mobility, is in strategic alignment with the University of Salford.

2. Generation of Warning Sounds

2.1 Monitoring of vehicle operating conditions

In order to generate warning sounds that are linked to use/user behaviour it is necessary to monitor somehow the speed and rate of change of speed of the scooter in real time. This can be achieved using a number of different approaches including a tachometer (e.g., reed switch or hall effect sensor), GPS positioning, acceleration spectra (as measured on the motor hub) or voltage/current monitoring of the motor drive circuit. Most of these signals (excluding GPS) can be amplified and replayed using a simple circuit to create a basic speed related warning signal. It is also possible to enhance these signals as a specifically designed audio output to give a more noticeable and/or desirable sound; for example, using a microcomputer installed on the scooter to apply transformations to the speed related signal in the auditory domain. Such microcomputers can also be used (with attachments) to obtain the signals from the scooter.

The advantage of measuring wheel/scooter speed is that it allows a signal to be generated that is the same for all scooters even if they use motor drives with different vibro-acoustic signatures. The drawback is that the sound is not native to the device and it may be more appropriate in this case to generate a purely synthetic sound. However, synthesised sounds might be more difficult to reinterpret by listeners (e.g., pedestrians) as a vehicle warning sound until the sound becomes familiar (as has become the case with large goods vehicle reversing alarms). The same applies to the use of GPS for this purpose – moreover, GPS provides a less optimal inconstant data feed.

---

[2] The standalone electronic device has been configured so that it can produce multiple different warning sounds.

The use of an accelerometer on the scooters motor hub is potentially a good way to monitor use in real time. An accelerometer captures the native sound of the scooter which can be replayed in its raw unaffected state or can be enhanced via filters and effects to give a more desirable/noticeable sound. Alternatively, the frequency of operation (identified from a frequency spectrum of the operational accelerations) could also be used to infer scooter speed making the device similar to a tachometer for the generation of entirely synthetic sounds.

A further possible approach is to measure the current and/or voltage of the motor drive circuit. This data can be used in the same way as that of the accelerometer described above. The advantage of using this approach is that the voltage/current would not be susceptible to slight fluctuations in acceleration of the scooter caused by a rough or bumpy driving surface. The drawback is that obtaining this data cleanly (I.e., without significant measurement noise) for instantaneous replay can be challenging.

2.2 Sound synthesis

As mentioned above, the warning sound generated by the scooter should first and foremost be noticeable for reasons of pedestrian safety, but it should ideally also be pleasant, or at the very least inoffensive, whilst adding the minimum amount of additional unnecessary environmental noise. Note that environmental noise is regarded as a significant health concern, see for example reports from the world health organisation[3].

Although widely accepted as a significant challenge, the generation of warning sounds for moving vehicles is still in its infancy – the difficulty being to strike a balance between noticeability, sound quality and sound level. At the current time there is no widely accepted approach and different vehicle manufacturers are developing warning sounds that meet legislative criteria (where available) with widely varying levels and characteristics. The choice of sound currently appears to be one that meets legislative requirements minimum levels with a "brand" specific sound. For consumer vehicles this has long since been the case, for example the exhaust sound of a car can easily be tuned for refinement or to create the perception of power. For electric vehicles the sound design process is similar but rather than physical devices filtering the drive's output sound, it is instead an electric circuit or computer code.

Significant research, beyond the scope of the current project, is required to inform the selection of an 'ideal' warning sound for scooters. We can however make some qualitative observations of what might be an appropriate signal to be generated, based on what is currently familiar. For example, the noise radiated by a consumer vehicle such as a car is typically broadband noise caused by

---

[3] https://www.euro.who.int/en/health-topics/environment-and-health/noise/publications/2018/environmental-noise-guidelines-for-the-european-region-2018

tyre/road interaction which varies in frequency range depending on vehicle speed (e.g., imagine low speed rumble to high-speed hiss or swoosh). On top of this is typically also some tonal noise generated by the vehicle drive, where the frequency of the sound indicates the speed of the engine and the rate of change of the frequency indicates acceleration or deceleration rate, this is another important cue because it is a good indicator of user behaviour. Since both sounds are familiar it seems sensible as a starting point to consider each in isolation and in combination for the assessment of noticeability and annoyance/preference.

An alternative sound type which could also be considered is impulsive sounds such as clicks or beeps (e.g. like the sounds employed by a reversing heavy vehicle). Such sounds are potentially ideal because they are easier to localise than tonal and broadband sources. The drawback however is that these sounds are likely to cause greater annoyance and in situations where multiple scooters are present at one time it could potentially create a more confusing sound environment. Furthermore, if indicator (turning direction) sounds from scooters is also to be made audible there is potentially an argument that impulsive sounds should be reserved for this purpose rather than being a constant feature. Again, as mentioned above, much more research is required in this area to inform the selection or rejection of sounds of this nature.

In this study we have chosen, through discussions with the client (Dott) and the RNIB, to investigate the noticeability and sound quality of broadband, tonal and modulated sounds. Therefore, we have not included impulsive sounds at this stage. This selection was made due to the timescales and budget available for the research. We have however included a further sound which may provide a happy medium between tonal and impulsive: modulated tonal sound. A subjective test (see Section 4) will be used to evaluate these sounds in terms of noticeability and preference. A sound quality analysis using standard metrics is also employed (see Section 3). Described below is the method of generation for the broadband, tonal and modulated sounds used in the study.

The warning sounds are synthesised within a python environment, version 3.8.10. All sound synthesis involves taking various measurements from the scooter, the measurements currently available for use are accelerometer data, current measurements, and the rotational speed of the wheels. In the up-to-date implementation, only the accelerometer data is used. When trialling sounds, the current data was found to have only low-frequency components and was largely inaudible after filtering the noise from the signal and as a result was not considered as an audible element for any further implementation.

While creating the sounds and the initial script, the process was not run in real time. Instead, it was opted to use an already recorded dataset, and simply run that through the processes mentioned in the following paragraph. This allowed for all effects and data-handling to be tested thoroughly and to ensure the project could be worked on away from the scooter.

The signal chain for the synthesis process is as follows: The accelerometer data is fed into python and run through a 10th order digital low-pass Butterworth filter from the scipy signal library, with a critical frequency of 0.5 Hz to remove any unnecessary high-frequency data and potential noise present in the signal, the signal does have some resonances at higher frequencies, so it is important to remove these as quickly as possible. From here the only other libraries used are simpleaudio for audio output, and numpy for array manipulation.

Following this, the signal is run through a half-wave rectification equation, this involves summing the input signal with the absolute value of the input and dividing by 2. This is used to increase the upper-harmonic content of the accelerometer data, and helps to add clarity back to the signal after being filtered.

The final stages involve adding delay-based effects to the accelerometer data. These include a standard delay algorithm, which simply feeds the input back into the data after a predetermined number of samples, and a flanger effect which feeds a modulated-delayed signal back into the data also. The flanger primarily acts as a way to give the signal movement, as it can sound very static without any treatment.

The data is then normalised for audibility, and is outputted using the “simpleaudio” library. All of the above is achieved offline, using an array of previously recorded data. See Figure 1 for the overall flow of signal generation.

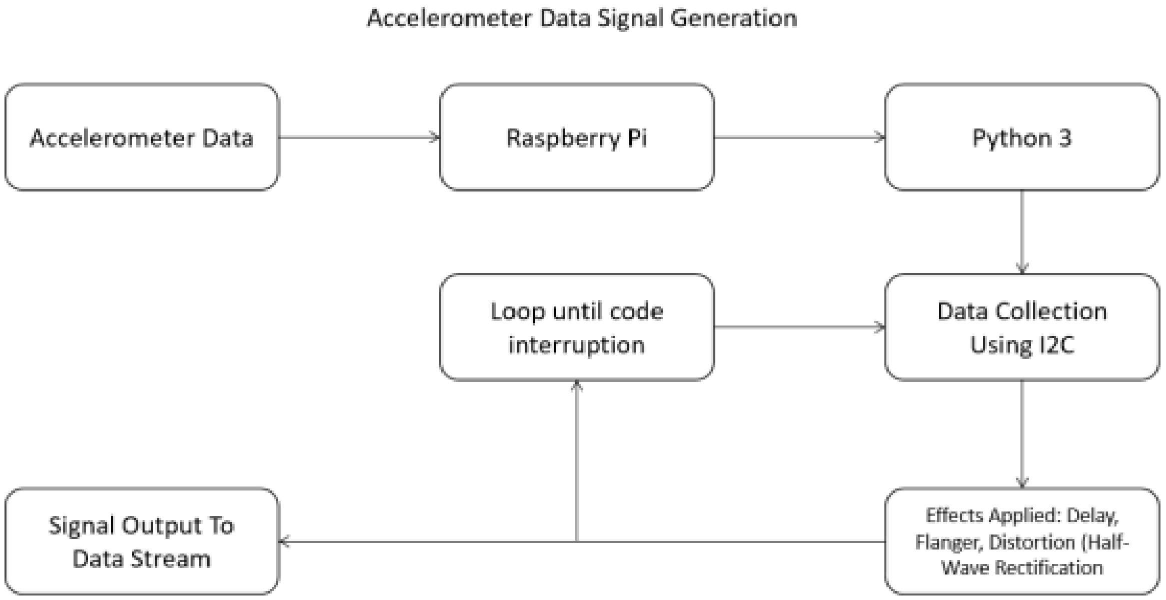


Figure 1***: Offline signal generation using accelerometer data, specifics of each block are explained in the above text.***

In terms of running the data through the scooter in real-time, the method implemented involves using the python library from the manufacturers for communication. This may also be achieved using the “SMBus” library too. The accelerometer is connected to the Raspberry Pi using the I2C protocol, this involves wiring up the accelerometer to the Rasberry Pi’s GPIO pins. The connections used are: +5V, Ground, SCL, and SDA. From here the manufacturers library for the accelerometer is used to create an object which may be used to call data from. Following this, the audio stream is initialised and runs through a callback method, which outputs the collects the data from the accelerometer in real time. The current implementation uses the data collected to modify the frequency of a sine-wave. Some issues which may arise from this are if the scooter had a sudden impact, it would be audible through the accelerometer data. Therefore, in future work, some limit should be added to the magnitude of the accelerometer data. Including this, an if statement is used to ensure the idle sound of the scooter is audible. This is done by taking the data received and setting a minimum value for output. Due to the nature of the output array being two separate channels, the final stage acts as a mixer, to ensure that the sub-woofer does not overshadow the beam-steered output. As a result, two multipliers are given to the data to ensure different levels from the two speaker arrays. This was done by ear and may not be applicable depending on different speaker placements.

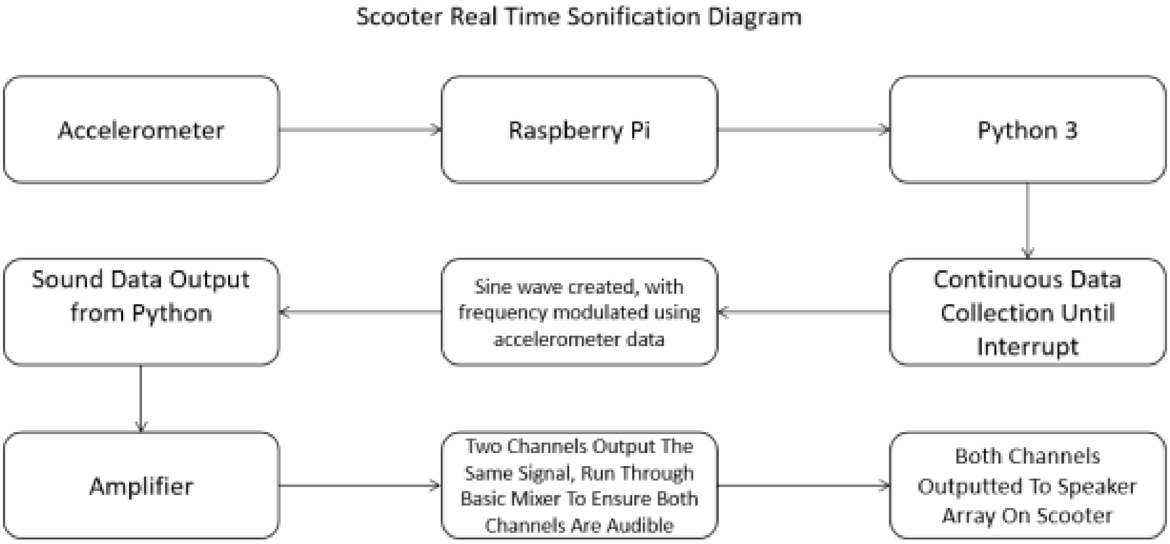


Figure 2*: Real Time Scooter Implementation for sonification of data.*

One final note on the real time implementation, the connection from Rasberry Pi to Amplifier has been done using the headphone output jack from the raspberry pi. This therefore bypasses any requirements for Bluetooth to output the audio. Figure 2 displays the flow diagram for the real time generation of warning sounds.

There are a number of timbre elements which are available, but not yet implemented fully. These include a white-noise signal; which would have a band-pass filter sweeping upwards as the speed of the scooter increases, Including this, some form of amplitude modulation given via data recorded from the scooter should be implemented.

For the subjective listening test, sound was synthesised using a DAW. These initial stages of sound synthesis were used to better understand what types of sound would be preferable for a warning system without having to worry about coding each element initially. These sounds were designed with the limitations of python in mind. Therefore, only simple waveforms and basic delay-based effects were used when needed. Including this, it allowed testing of signals which may be created, given more time on the project. These include standard waveforms which increase in pitch as the scooter increases in speed. The sounds were separated into three distinct categories: broadband,

broadband with tone, broadband with modulated tone. In designing these sounds, the broadband signals typically had a low-pass or band-pass filter with a cut-off which increased as vehicle speed increases. The tones were typically simple waves such as sine, square or sawtooth. This was done to keep python implementation as quick as possible. Finally, amplitude modulation was achieved through using the tachometer data. As this was a direct reading of the rotational speed of the wheel, this led to the frequency of the amplitude modulation to increase as the scooters speed increases. Figure 3 displays the sound generation process for the warning sounds used in the subjective experiment.

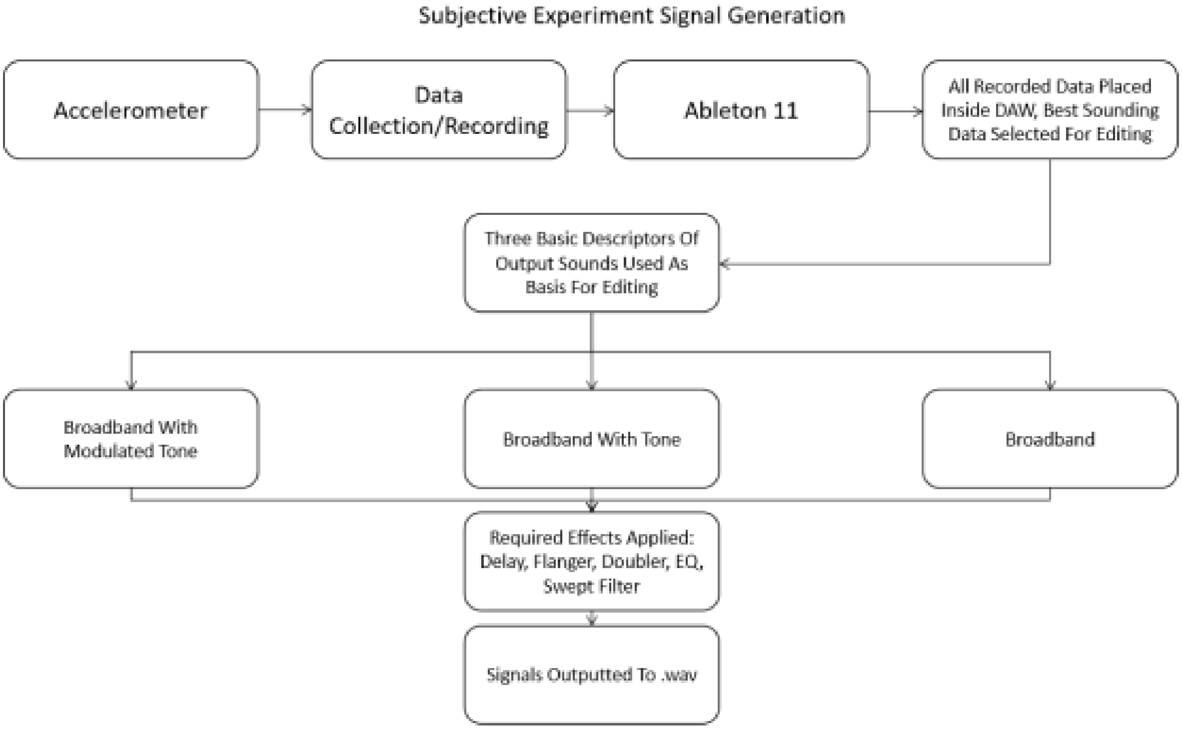


Figure 3*: Subjective listening test signal generation. All samples used in testing were created through Ableton 11. This was done to allow a greater spread of distinct sounds in a quicker time frame.*

2.3 Standalone hardware to sounds generation

Described in the previous section is a means of monitoring scooter use (speed and rate of change of speed; acceleration/rate of acceleration, or jerk) and how that information can be reinterpreted as a warning sound. Once these signals (warning sounds) have been transformed into warning sounds there is a further requirement to replay them using an audio system embedded on the scooter. The means of amplification and radiation of warning sounds is discussed in this section of the report. This section also describes the validation measurements performed to assess the utility of the prototype system.

Note that the work carried out here is a feasibility study investigating how to implement warning sounds within the limitations of package space available on a typical scooter; arriving at a cost-effective solution suitable for widespread implementation would require further study. For example, a warning sound device must be small, low cost, robust, reliable and weather resistant. Such aspects are not within the scope of the feasibility study.

Shown in Figure 4 are images of the components used to create the audible warning system's' output. The amplifier used is a two-channel device which requires power from an 8 to 24V supply; this can be plugged directly into the output of the sound generation module or can be connected via Bluetooth for wireless operation (with an additional component – Bluetooth receiver). The amplifier has a peak output of 2x55W at 8 Ohms, the amplifier is capable of outputting audio at 24bits and 192kHz sampling rate.[4]

Figure 4 also shows a small microphone array to be used on the stem or handlebars of the scooter. Using 5 miniature loudspeakers in a row a broadside array is made to achieve a figure of eight radiation pattern (broadside array). When enclosed as shown in Figure 4 (middle) and when attached to the scooter shielded behind the handlebars the array radiates sound predominantly in the forward direction above 1000Hz. The small broadside loudspeakers have a maximum output of 88dB, with a coil resistance of 8 Ohms and a rated power of 0.7 Watts.[5]

By using multiple drivers, a greater volume displacement (higher volume level) can be achieved in a small package along with the benefit of a directional sound output. This device alone provides sufficient sound output for a scooter warning system with levels exceeding 50dBA at 1.5m in the frequency range above 1000Hz. Note that the maximum output level achievable is not necessarily the recommended one (the ultimate aim should be to provide a solution with some headroom with respect to sound level without distortion).

Figure 4: Images left to right – Amplifier evaluation board, broadside microphone array and mini-

---

[4] The technical document for the amplifier can be found at: https://docs.rs-online.com/c53a/0900766b8152dcc1.pdf

[5] The technical document for the small forward-facing speakers can be found at: https://docs.rs-online.com/5e44/0900766b8157fdaa.pdf

loudspeaker (named mini-sub in the report).

For reasons of sound quality, a further device was also constructed as shown in Figure 4 (far right image). This additional unit may not be necessary in practice, but the two-channel amplifier allows its addition at little expense. Such a device could be installed on the underside of the scooter utilising free space in the battery compartment or side panels below the rear lights. The downside of this extra hardware is a small weight penalty and the unit cost (plus the extra amplifier channel). The advantage of its inclusion is an increased sound power output in the low frequency range where the micro speaker array is more limited. Its inclusion would extend the low frequency sound output level and potentially create a warmer more powerful sound. This may be important for perception of the brand. If this is not of interest the additional device should not be required.

Shown in Figure 5 below is a photograph of the scooter in the semi anechoic chamber. This lab condition represents a case where the scooter would be operating in an outdoor environment over a reflecting plane (floor/ground) such as a large car park but without any significant reflections from nearby obstacles. It is standard practice to use such a facility to test the sound level output and direction of sound radiation from equipment that is generally operable on a reflective surface. Testing conducted in this facility provides a measure of the radiated sound level (at a measured distance) which can be converted to a sound power. Using this facility, it is also possible to accurately determine the direction of the measured radiated sound as a function of angle and frequency.

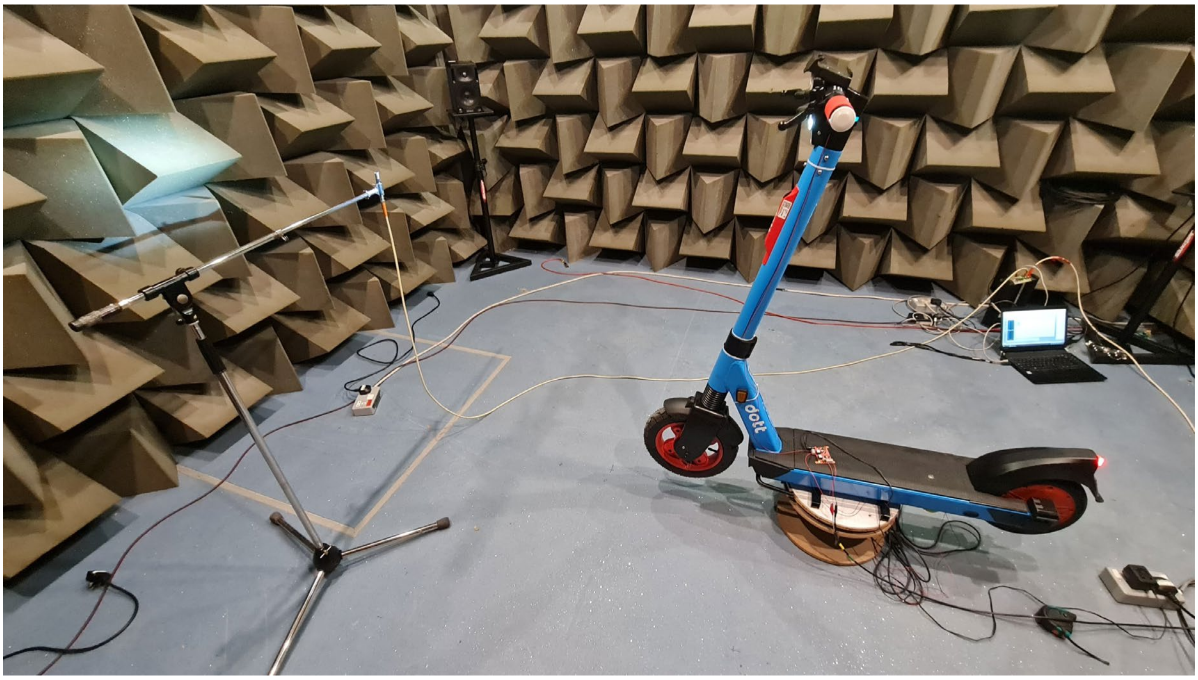
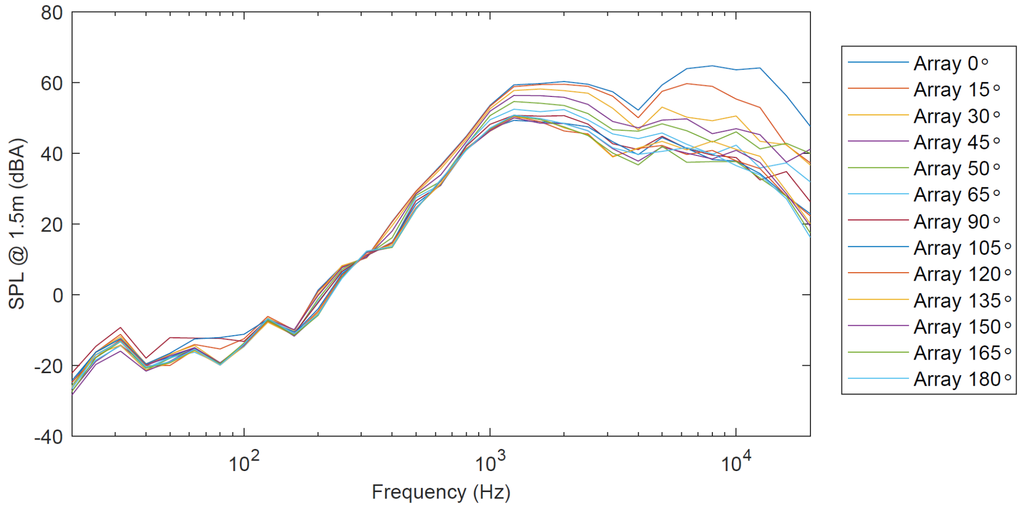


Figure 5: Dott scooter instrumented with forward firing microphone array and mini-sub for testing of sound level and directivity in the semi-anechoic chamber. The scooter is mounted on a rotating platform to allow small, measured rotations relative to a fixed microphone position.

Under test in the semi anechoic chamber the equipped scooter was rotated relative to a fixed microphone whilst measuring the A-Weighted sound level in one third octave bands. The frequency response function: sound pressure divided by voltage to amplifier (system frequency response) was also recorded. For the forward firing microphone array this was done in 15-degree increments and for the mini-sub 30-degree increments. Figure 6 below shows the relative levels of the two devices for different angles of reception at the receiver position. For the microphone array the sound level presented at the receiver microphone is 10-20dBA greater in the direction of travel when compared to the rear. The mini-sub situated beneath the scooter deck produces a higher sound level in the low frequency range and radiates in all directions as intended. As mentioned previously – the addition of such a unit is optional (and only required if the perception of brand sound is an important factor).



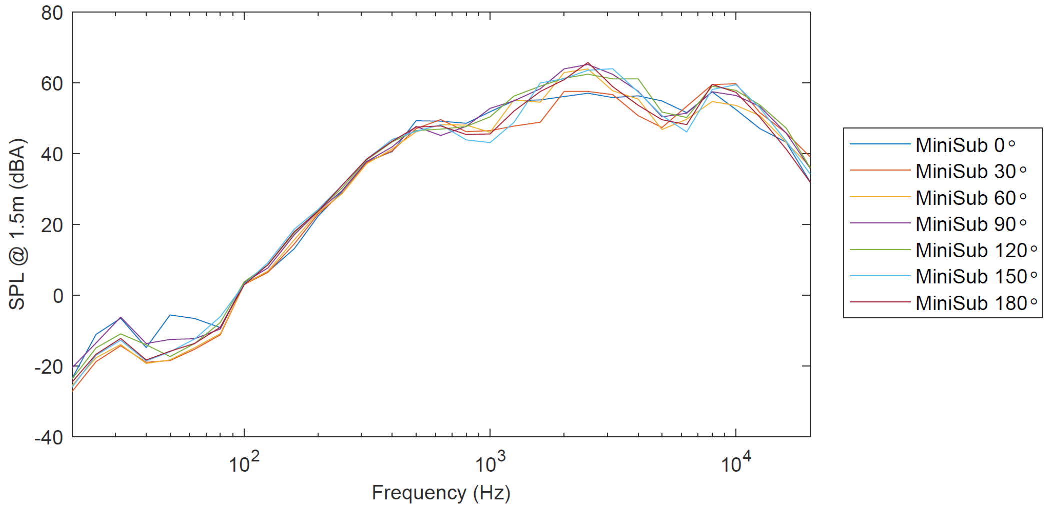


Figure 6: Upper plot – Maximum sound level output of the microphone array/amplifier combination for 15 degree increments of angle relative to 0 degrees which corresponds to the scooter facing directly towards the microphone in its direction of travel. The lower figure is the corresponding maximum sound output level from the mini-sub place below the scooter deck. Both plots are for the scooter at a 1.5m distance from the mid-point (approximate rider position) of the scooter.

3. Analysis and Assessment of Warning Sounds Generated

3.1 *Warning sounds vs. state-of-the-art regulation*

European and United States regulations such as 'United Nations Economic Commission for Europe (UNECE) 138.01'[6] and 'Federal Motor Vehicle Safety Standard (FMVSS) No. 141'[7] set the minimum requirements for AVAS in hybrid electric vehicles. To increase noticeability of hybrid electric vehicles, these regulations require AVAS to produce discrete tones as specific frequency ranges, e.g., 2 nonadjacent one-third octave bands from 315 Hz to 3150 Hz. UNECE 138.01 also requires pitch shifting (as a function of vehicle speed). Once the minimum set by these regulations are met, manufacturers can design their own warning acoustic signals to increase vehicle awareness for pedestrians.

A well-designed set of psychoacoustic features can substantially increase the noticeability of approaching vehicles.[8] Some specific features such as roughness, level variations or frequency shifts can lead to an early vehicle detection by pedestrians. Adding other features such as amplitude modulations can also increase noticeability and ease of localisation.

During this project, a series of audio signals were generated, including a variety of some of these psychoacoustic features. Below there is a list of the warning sounds better rated by Dott, RNIB and ARC's research team:

- Acceler_Square: Accelerometer Data with a square wave as the tonal element.

---

[6] United Nations Economic Commission for Europe (UNECE). (2017). Regulation No. 138 of the Economic Commission for Europe of the United Nations (UNECE) — Uniform Provisions Concerning the Approval of Quiet Road Transport Vehicles with Regard to Their Reduced Audibility. ECE/TRANS/WP.29/GRB/2012/6, United Nations Economic Commission for Europe, Geneva, Switzerland.

[7] Federal Motor Vehicle Safety Standard (FMVSS). (2019). Minimum Sound Requirements for Hybrid and Electric Vehicles. Docket No. NHTSA-2019-0085, Federal Motor Vehicle Safety Standard No. 141, Notice of proposed rulemaking, National Highway Traffic Safety Administration (NHTSA), US Department of Transportation, Washington, DC.

[8] A. Fiebig. Electric vehicles get alert signals to be heard by pedestrians: Benefits and drawbacks. Acoustics Today, 16(4), 2020. https://doi.org/10.1121/AT.2020.16.4.20

- AccelerCH2_Doubler: Accelerometer data with delay-based doubler effect, giving the sound a sense of space.
- AccelerCH2_Doubler_and_Tacho_Phaser: Accelerometer data with the doubler effect applied as above, this signal also includes the tacho data as an impulse with a phaser effect to add movement.
- AccelerCH2_Flanger: Accelerometer Data used with a flanger effect as a form of modulation.
- Broadband_LowPass: Low Pass filtered pink noise as the broadband sound, LP filter used to keep low frequency rumble throughout the filter sweep.
- Broadband_SawtoothSideChained_Acceler: Broadband sound with a sawtooth wav modulated by the amplitude of the Tacho.
- Broadband_SineLowPitch_Acceler: Notch filtered pink noise as the broadband sound with a pitch modulated low frequency sinewave as the tonal element, including the accelerometer data as a further tonal element.
- Broadband_Square: Notch filtered pink noise as the broadband sound with a pitch modulated Square wave as the tonal element.
- Broadband_Square_Acceler: Notch filtered pink noise as the broadband sound with a pitch modulated Square wave as the tonal element, including the accelerometer data as a further tonal element.
- BroadbandLowPass_Acceler: Low Pass filtered pink noise as the broadband sound, Accelerometer data used as tonal element.
- Current_Cleaner This is the current data with all high-frequency elements filtered out. This is done to remove the noise present in the signal, however this also removes much of the clarity and gain from the audio.
- TonePitched_Acceler: Sine wave with increasing pitch used as tonal element, with accelerometer data as secondary tonal element. Amplitude modulation happens at the top speeds of the scooter.
- TonePitched_Acceler_Broadband: Sine wave with increasing pitch used as tonal element, with accelerometer data as secondary tonal element. Broadband given by notch filtered white noise. Amplitude modulation happens at the top speeds of the scooter.

As a first step to understand the effect of each psychoacoustic feature on vehicle noticeability, a fundamental research was carried out in a subjective experiment (see Section 4). Three sounds were synthesised to be added as warning sounds to the e-scooter in the simulated scenarios presented to the participants of the experiment: (i) broadband sound, broadband plus tonal sound and broadband plus modulated tonal sound. Figure 7 shows the frequency spectra of these three synthesised sounds. A low-pass filter (cut-off frequency = 1000 Hz) was applied to synthesised the broadband sound. The broadband plus tonal sound is mainly composed by a fundamental frequency at 120 Hz and a series of its harmonics. The broadband plus modulated tonal sound has the same fundamental frequency and harmonics. For this sound there are other frequency components (as seen in Figure 2) consequence of the amplitude modulation. As described in Section 2, the amplitude modulation was achieved using the tachometer data[9]. The modulation rate in the

---

[9] Note that the modulation rate was variable, as it is a function of the tachometer data.

modulated tone sound ranges between ~2.7-3.9 Hz. This has been calculated by looking at the time between peaks and converting it into a frequency using 1/T.

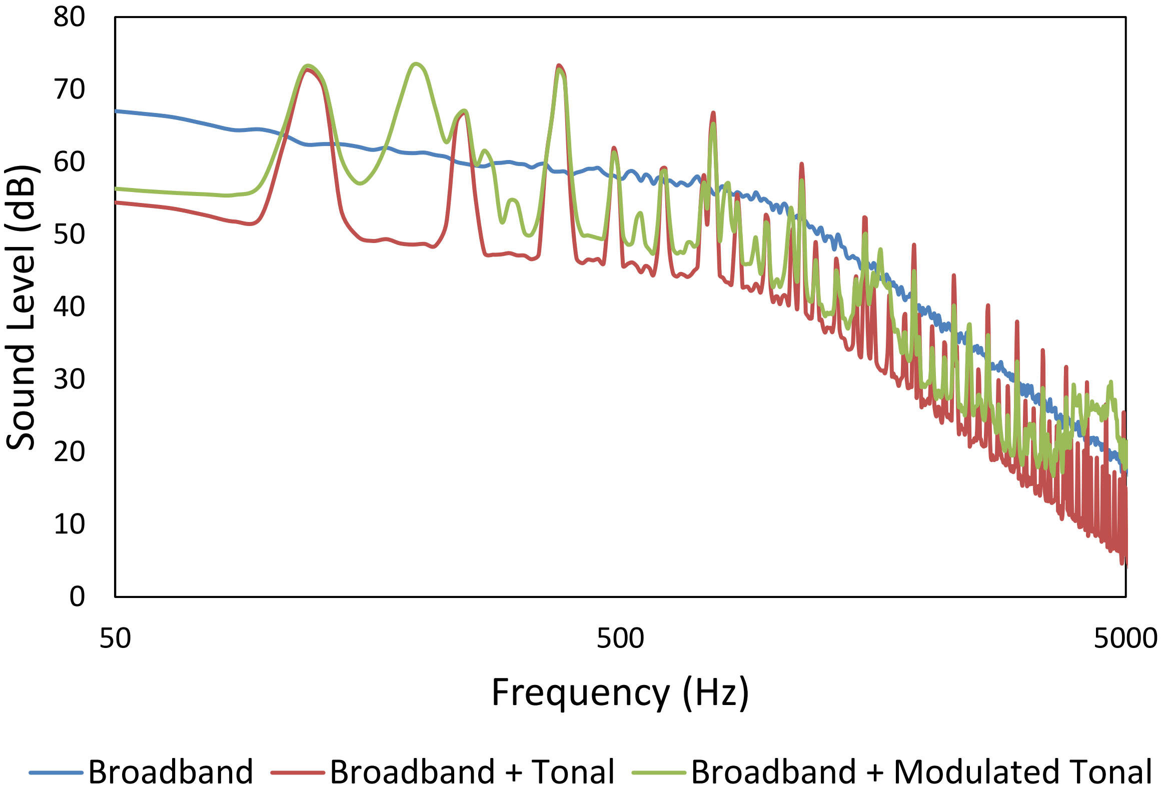


Figure 7: Frequency spectra of the three sounds synthesised for the subjective experiment.

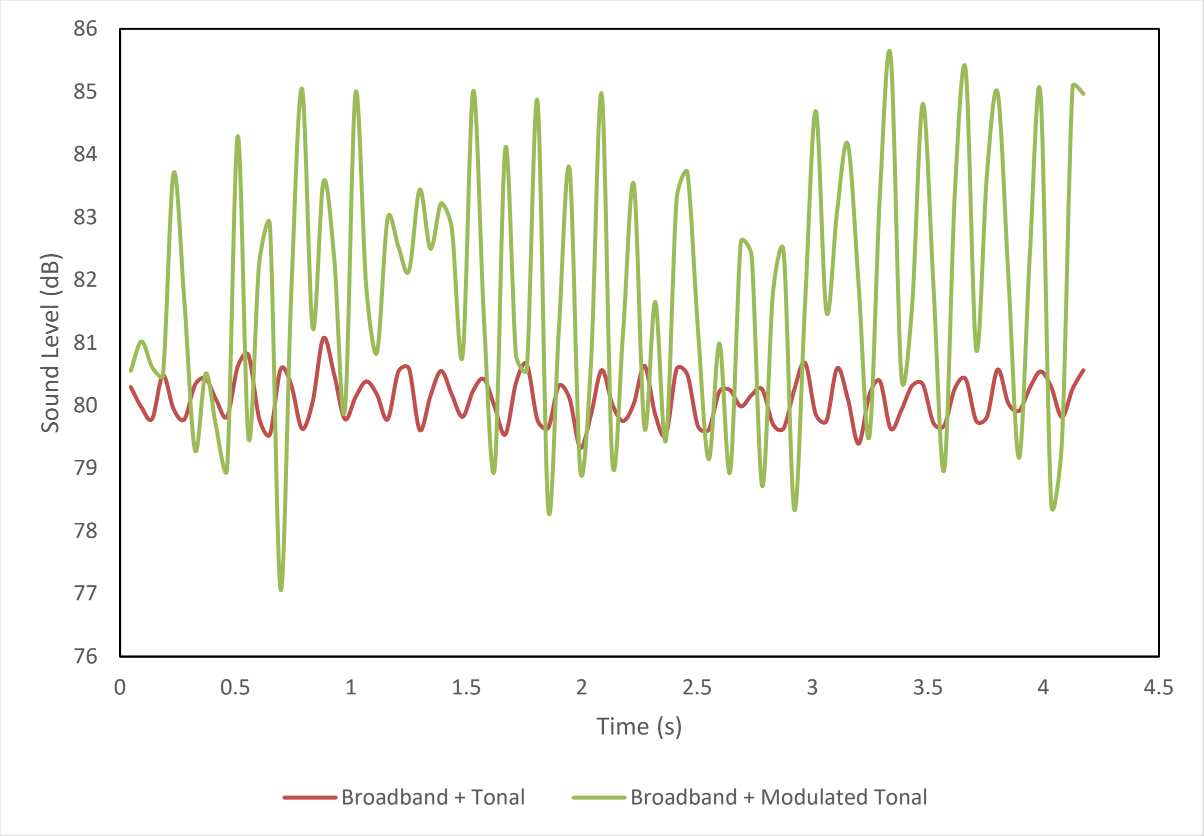


Figure 8: Time history (Sound Pressure Level vs. Time) of the broadband plus tonal and broadband plus modulated tonal sounds.

The temporal differences between the broadband plus tonal sound and the broadband plus modulated tonal sound are observed in Figure 8.

Following guidance for the European and US minimum overall sound pressure level requirements for hybrid electric vehicles, the A-weighted sound pressure level (at the received position) for these three synthesised sounds during the subjective experiment was set to 48 dBA.[10]

---

[10] Note that the sound pressure level produced by the e-scooter (without any added sound) at the receiver position was 46 dBA. This sound pressure level is lower than the minimum requirement for hybrid cars in Europe (i.e., 50 dBA), but it was deemed appropriate considering the difference in size between an hybrid car and a scooter.

3.2 Psychoacoustic assessment of warning sounds

The warning sounds developed underwent a psychoacoustic analysis, utilising industry standard calculation software (HEAD Acoustics ArtemiS Suite).

Sound Quality Metrics (SQMs) are good indicators of how the human auditory system reacts to different features of sound.[11] From these SQMs, and the results of the subjective experiment, it can be deduced which synthesised sounds may be appropriate for warning pedestrians of danger, while also mitigating rider and pedestrian perceived annoyance. The SQMs calculated were:

- Loudness: This metric provides an accurate representation of the sensation of sound intensity.
- Tonality: This metric describes the perceptual effects of the presence of spectral irregularities or discrete tones.
- Sharpness: It describes the perceptual effects of spectral imbalance of the sound towards the high frequency region.
- Roughness: It describes how rapid temporal fluctuations of the sound level are perceived.
- Fluctuation Strength: It describes how slow temporal fluctuations of the sound level are perceived.

Psychoacoustic annoyance models combined the contribution of a variety of SQMs to annoyance. The Zwicker's psychoacoustic annoyance model, accounting for the relation between annoyance and hearing sensations loudness (*N*), sharpness (*S*), fluctuation strength (*F*) and roughness (*R*) is given by

$$PA = N_5\left(1 + \sqrt{w_S^2 + w_{FR}^2}\right) \tag{1}$$

where

N_5 is the 5th percentile of the loudness

$$w_S = \{(S - 1.75) \cdot 0.25\lg(N_5 + 10),\ S > 1.75;\ 0,\ S \leq 1.75\,\} \tag{2}$$

$$W_{FR} = \frac{2.18}{N_5^{0.4}}(0.4F + 0.6R) \tag{3}$$

Zwicker and Fastl also developed an empirical method to estimate sensory pleasantness (expressed in terms of relative values). This method is based on relative values of sharpness (S), roughness (R), tonality (T) and loudness (N):

---

[11] Zwicker, E. and Fastl, H. (1999). "Psychoacoustics – facts and models." Berlin: Springer-Verlag.

$$\frac{P}{P_0} = e^{-0.7R/R_0} e^{-1.08S/S_0} \left(1.24 - e^{-2.43T/T_0}\right) e^{(-0.023N/N_0)^2} \tag{4}$$

Table 1 shows the value of the SQMs and Zwicker's psychoacoustic annoyance and sensory pleasantness calculated for the warning sounds synthesised, better rated by Dott, RNIB and ARC's research team. As can be seen in Table 1, the most preferred sound by Dott (i.e., Current_Cleaner) has the lowest value of Zwicker's psychoacoustic annoyance and the highest value of Zwicker's sensory pleasantness. The most preferred sound by Salford's research team (i.e., AccelerCH2_Flanger) has intermediate values for these metrics (according to the range of values of the list of sounds). These results suggest that the use of both Zwicker's psychoacoustic annoyance and sensory pleasantness can be very useful to inform the design of warning sounds with the target to achieve a balance between noticeability and annoyance.

Table 1*: Value of the Sound Quality Metrics (SQMs), Zwicker's psychoacoustic annoyance and sensory pleasantness for the warning sounds generated (better ranked by Dott, RNIB and ARC's research team).*

Warning Sound Generated	**Zwicker's Sensory Pleasantness**	**Zwicker's Psychoacoustic Annoyance**	**Loudness (sone)**	**Sharpness (acum)**	**Tonality (tu)**	**Roughness (asper)**	**Fluctuation Strength (vacil)**
Current_Cleaner (0.00 - 25.27 s)	0.53	9.04	7.51	0.35	0.30	0.34	0.011
TonePitched_Acceler (0.00 - 24.49 s)	0.21	11.78	10.20	0.92	0.68	0.29	0.017
Broadband_LowPass (0.00 - 24.38 s)	0.19	14.73	9.57	0.64	0.03	1.01	0.010
TonePitched_Acceler_Broadband (0.00 - 24.49 s)	0.17	12.86	10.80	1.00	0.54	0.37	0.015
AccelerCH2_Doubler (0.00 - 20.28 s)	0.11	15.69	10.60	1.02	0.28	0.93	0.016
AccelerCH2_Flanger (0.00 - 24.38 s)	0.10	14.10	10.80	1.20	0.43	0.59	0.021
Broadband_SineLowPitch_Acceler (0.00 - 24.38 s)	0.08	17.07	11.50	1.20	0.35	0.98	0.010
BroadbandLowPass_Acceler (0.00 - 24.38 s)	0.07	16.07	10.70	1.26	0.21	0.98	0.012
AccelerCH2_Doubler_and_Tacho_Phaser (0.00 - 20.02 s)	0.05	20.20	11.10	1.23	0.23	1.61	0.048
Acceler_Square (0.00 - 24.38 s)	0.04	16.55	11.20	1.66	0.64	0.95	0.012
Broadband_SawtoothSideChained_Acceler (0.00 - 24.38 s)	0.03	18.07	11.00	1.59	0.26	1.27	0.018
Broadband_Square_Acceler (0.00 - 24.38 s)	0.03	18.12	11.30	1.60	0.32	1.21	0.010
Broadband_Square (0.00 - 24.38 s)	0.02	17.50	10.70	1.78	0.32	1.25	0.007

The SQM scores are shaded to make it easier for the reader to compare values and see the range within each column, that is, within each feature and across the group of sounds evaluated. Lower scores are indicated by darker shading, except for *Sensory Pleasantness* and *Annoyance*, where it is reasonable to make a subjective judgement about the direction of scores. Higher values for these, indicating more pleasant / less annoying sounds, are shaded in green, and scores suggesting less pleasant / more annoying sounds are shaded in red.

Grey is used for the other features as it is harder to make a judgement for each one in isolation about whether sounds generating higher scores for these should be classed as better or worse from a subjective perspective. In any case, the darker values for the SQMs Loudness, Roughness and Tonality should produce higher values of noticeability. An optimum compromise between noticeability and annoyance requires a balance of the separate metrics, but further research is required to determine the maximum acceptable threshold for each of them when combined into one alert sound.

4. Subjective Experiment

4.1 Experiment Stimuli: Soundscape and e-scooter recordings

The sound and video stimuli were recorded in two separate locations: MediaCity and Peel Park in Salford. The first was an open urban area, including pedestrians, cyclists, hosptitality noise and music. The second was a quieter park area, with a play area, less pedestrian and cyclist activity, foliage noise and distant road traffic noise. These two locations were selected to test the warning sounds in two opposite locations in terms of levels of activity (i.e., distractions) and background noise. The 360-degree video stimuli were recorded using a Insta360 Pro 2 - 8K Professional 360 Camera[12]. The ambisonic audio stimuli were recorded using a Soundfield ST450 microphone, with a Zoom F8n Field Recorder. The ambisonic microphone was placed directly beneath the 360 camera to ensure the audio matched the video as best as possible without the microphone being seen by the 360 camera. A series of scooter pass byes were recorded, with the scooter operating at constant maximum speed (i.e., 15 mph), and approaching the camera (i.e., the simulated pedestrian) from behind at different angles. In the subjective experiment, the recorded sound and video stimuli were used in conjunction with the developed warning sounds to mimic scooter pass-by events, where the scooter would be generating a warning sound, as well as when they would not be generating a warning sound. Stimuli were also used without scooters included (i.e., no vehicles passing by or other vehicles passing by such a bicycle).

---

[12] See https://www.insta360.com/product/insta360-pro2

4.2 Experimental setup

The audio-visual scenes were presented to the participants via the Oculus Quest 2 VR headset, using a Focusrite Scarlett 2i4 audio interface and Beyerdynamic DT 250 headphones. The VR visual scenes and the recorded Ambisonic audio were synchronised and 20-second-long clips were selected, some of which included an e-scooter pass-by. Audio and video were combined in Unity and the Ambisonic recordings were decoded using the Steam Audio plugin, which provides binaural spatialisation with head-related transfer functions using head-tracking from the VR headset. In scenes with an added e-scooter sound, an audio object was created in Unity which followed the movement path of the e-scooter in the recording. The audio object was also spatialised with Steam Audio.

4.3 Experimental protocol

Each experimental session consisted of two parts. For the first part of the experiment, which studied noticeability of the e-scooter warning sounds, the participants were sat in a room wearing a VR headset and headphones and responded using the Oculus Quest 2 controller. In each experimental trial, they were shown one of the 360° video scenes and a short text excerpt, taken from the DeepMind Q&A Dataset (see Figure 9). They were asked to read the text and told they would be asked a question about it afterwards. At the same time, they were asked to press a button on the controller as soon as they detected a moving hazard, which was defined as anything that could potentially cause harm to the person if they were really in the situation displayed in the video. When the video was finished, a question about the text was displayed in front of the participant, with 4 possible answers, and the participant chose the answer they thought was correct using the VR controller. They were instructed to try to respond as accurately as possible, but to guess if they did not know the answer. This task was included to focus the attention of participants on something other than potential hazards in the scene. The intention was to create a distraction and increase cognitive load, thereby increasing the need for a more effective alert.

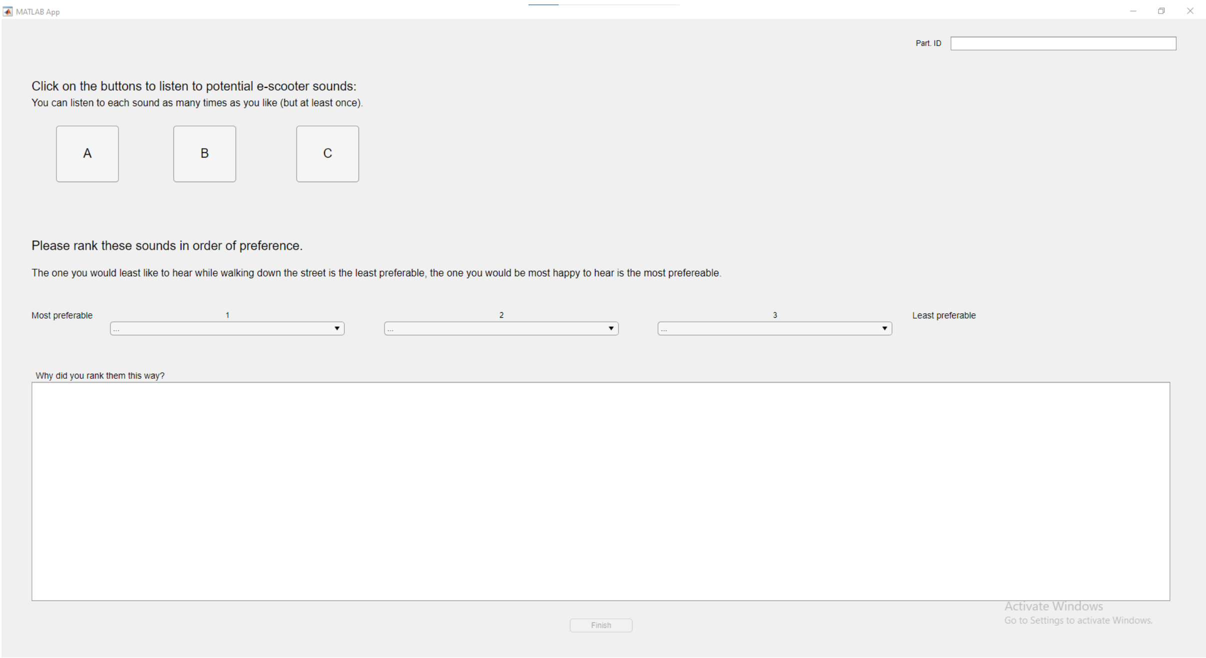
After they responded, a transition scene was displayed which asked them to press a button on the controller when they were ready to continue. Participants were told they could take a break at this point if needed. A short practice session was provided at the beginning to familiarise participants with the task and ensure that they were able to read the text clearly, and that the VR headset was comfortable and secure.

To the unaccustomed eye, a man toting 120 books while riding a stubborn donkey would seem nothing short of a circus spectacle. But for hundreds of children in the rural villages of Colombia, Luis Soriano is far from a clown. He is a man with a mission to save rural children from illiteracy.
"There was a time when many people thought that I was going crazy," said Soriano, a native of La Gloria, Colombia. "They'd yell, 'Carnival season is over.' ... Now I've overcome that."
Soriano, 38, is a primary school teacher who spends his free time operating a "biblioburro," a mobile library on donkeys that offers reading education for hundreds of children living in what he describes as "abandoned regions" in the Colombian state of Magdalena.

Figure 9. Example video scene in the noticeability experiment

Each session consisted of 20 trials. The independent variables studied in the experiment were environment (Peel Park and MediaCity) and warning sound (broadband sound, broadband plus tonal sound, and broadband plus modulated tonal sound). All participants were shown 3 video scenes from each environment, once with a warning sound, and once without. Which video scene was matched with which warning sound was randomised for each participant. Additionally, they were shown 4 video scenes from each environment which did not have an e-scooter in them – some with a bicycle pass-by, and some without any kind of moving hazard. These were included to make the task less predictable. The order of presentation of the trials was randomised for each participant, and the same video scene was never presented twice in a row.

The second part of the experiment was to study participants' preference for the three warning sounds. Before they started this part, participants were debriefed about the purpose of the experiment and told that we are studying e-scooter sounds. They were shown a user interface written in MATLAB, as shown in Figure 10. It included three buttons which allowed the participants to listen to each of the three sounds as many times as they wished. Then, they were asked to rank them from the most preferable to the least preferable, in terms of which sound they would most like to hear as a pedestrian, or which was the most pleasant. Finally, they were asked to leave a short comment justifying their choice (this was optional).



Click on the buttons to listen to potential e-scooter sounds:
You can listen to each sound as many times as you like (but at least once).

Please rank these sounds in order of preference.

The one you would least like to hear while walking down the street is the least preferable, the one you would be most happy to hear is the most prefereable.

Why did you rank them this way?

Figure 10. Experimental interface used to study participants' preference for the three warning sounds

4.4 Participants

Participants were recruited from within and outside the university. They were told that to take part they should have self-reported normal hearing. They were also advised that if they required glasses to read then the maximum frame size that would fit inside the VR headset was 13.5 cm wide by 4.5 cm tall.

In total, 15 people completed both parts of the experiment. They were asked to submit some demographic data at the start of the session; questions were asked about factors that it was suspected might influence performance in the task.

Handedness:

All 15 people were right-handed; there is therefore no reason to suspect that response time from any individual was affected by them having to perform the task with a controller designed for their non-dominant hand (the experiment was programmed to respond to triggers from only the right-hand Oculus controller).

Hearing / Visual impairment:

All but one person reported having no hearing impairment as far as they were aware (self-reported normal hearing). The one person that said they had a hearing impairment did not give any details.

Five people reported having a visual impairment; one of these did not leave any details, but the rest said their impairment was minor and corrected with glasses or contact lenses e.g., short-sighted. Two participants needed glasses to read clearly but were able to wear them during the experiment without any issues.

Participant 10 was the one person that reported having both a visual and hearing impairment but did not give any further details; they were also the only person who did not respond (click) in any trials at all which contained scooters. We do not have sufficient information to determine the cause of this performance e.g., whether the person could not detect the scooter due to reduced audio-visual acuity, or didn't understand the task.

Native language:

Participants were asked about their native language as the experiment required reading text and answering a question about it. This task could therefore have been more challenging (distracting) for people whose native language is not English. Most participants (9 of the 15) were not native British English speakers, with one other reporting that they were bilingual (English / Italian). When asked for details, 3 people did not leave any; otherwise, the native languages reported were: Portuguese, White Spanish, Mandarin Chinese, Estonian, Persian (Farsi), and Arabic.

Age and gender:

Age bracket and gender were not controlled during recruitment but were monitored as a check that the sample was not skewed towards any particular demographic. Typical age bracket options were presented: 18-24, 25-34, 35-44, 45-54, 55-64, and 65+. Nobody was in the top two categories. Most people were aged between 18 (the minimum allowed) and 44, but within that the sample was fairly evenly distributed by age and gender (male/ female; nobody checked the gender option 'other').

Samples for Noticeability and Preference Analysis

Data from all 15 participants was analysed for the Preference task. For the Noticeability analysis, data was excluded from anyone clicking in less than 50% of trials as this meant they were not performing the task as expected. These were participants 2, 9 and 10. Everyone else responded (clicked) in 100% of trials containing scooters. Figure 11 illustrates the differences in distribution for age between the samples used in the Noticeability and Preference analysis; Figure 12 and Figure 13 show equivalent information for gender and native language, as percentages of the sample size (12 and 15 for Noticeability and Preference respectively). It can be seen that excluding data from 3 participants for the Noticeability analysis changed the distribution across age brackets slightly (a greater proportion falling into the 25-34 range), but distributions for native language (British English / Other) and Gender (Male / Female) remained similar to those present in the Preference group.

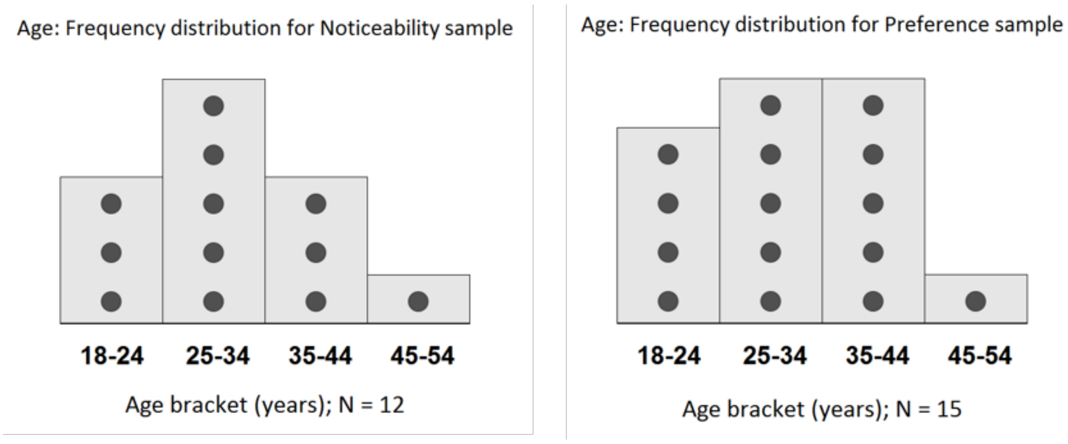


Figure 11: Age distribution of participants. The Noticeability sample featured a relatively higher proportion of people in the 25-34 bracket than the Preference equivalent.

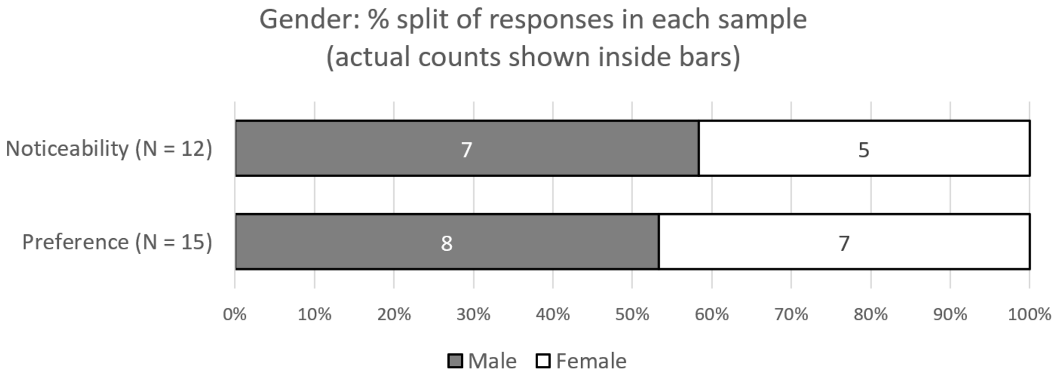


Figure 12: Gender distribution of participants (an 'other' gender option was offered but nobody selected it).

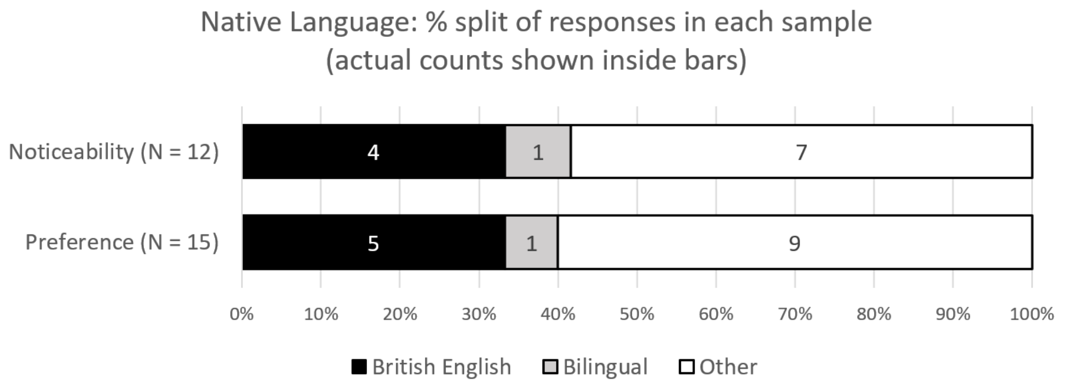


Figure 13: Native language distribution. In both analyses, native British English speakers formed about one third of the overall sample.

4.5 Results and discussion

Noticeability results

Response times were recorded for each scene containing an e-scooter pass-by, calculated from the beginning of a scene until button press. If a participant pressed the button more than once during a scene, the response closest in time to when the scooter passes them was chosen for analysis.

Because in each video scene the scooter pass-by was at a different time, the data is best analysed by comparing responses to the same scene with and without a warning sound. Figure 14 shows the benefit of introducing the three warning sounds, calculated as the difference in response time between the same video scene with and without a warning sound, for each participant, in the two different environments. Positive values mean that response time was faster with the warning sound, negative – that it was slower. From the plot, it appears that the broadband sound does not provide any benefit to response times, however, both broadband plus tones and broadband plus modulated tones might decrease response times slightly.

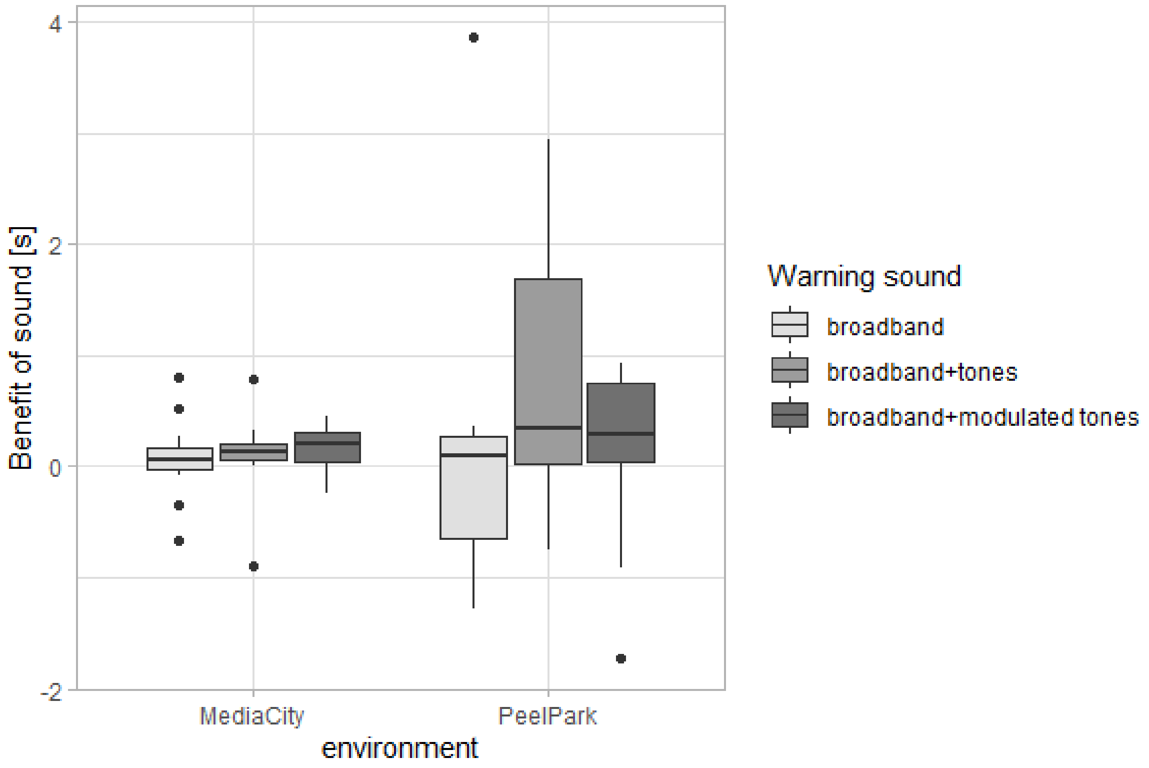


Figure 14. Boxplots of the benefit of introducing a warning sound, calculated for each participant. The horizontal lines show the median, and the shaded boxes show the 25th and the 75th percentile of the data.

To test if introducing any of the warning sound provides a decrease in response times, a mixed-effects linear model was fit to the data using the R package lme4. A mixed-effects model allows to include grouping variables (random effects), which in this case were: *Participant* and video *Scene*. By using *Scene* as a random effect, the model can take into account that each scene has a different 'baseline' response time and assign a different intercept to each scene. This is particularly important here, as the scooter pass-by was by design at different time point in different scenes. The dependent variable in the model was response time (*RT*), and the independent variables were *Sound* (with levels: no sound/ broadband/ broadband+tones/ broadband+modulated tones) and *Environment* (MediaCity/ Peel Park), and the interaction between the two variables, to test if any of the sounds works better in either environment. The model definition was thus:

$$RT \sim Sound + \ Environment + Sound{:}Environment \ + \ (1 \mid Scene) \ + \ (1 \mid Participant) \quad (5)$$

The contribution of each variable was then tested with the Anova function from the 'car' package in R. The Table 2 shows the results of this analysis.

Table 2: Results of the ANOVA analysis of the contribution of each variable to response time.

Variable	F	Df	Df.res	p-value
Sound	**2.97**	**3**	**121.01**	**0.035**
Environment	0.03	1	4	0.879
Sound - Environment	1.56	3	121.01	0.202

The results show that *Sound* is a significant predictor of response time (p=0.03). Neither the type of environment, nor the interaction between Sound and *Environment* were statistically significant.

To find out which warning sounds show a significant difference in response times compared to no sound, contrasts are calculated using the 'emmeans' package in R. The results are shown in Table 3.

Table 3: Results of the contrast analysis with the contribution of each sound tested to the response time.

contrast	estimate	SE	df	t.ratio	p.value
broadband – no_sound	-0.09	0.16	124	-0.53	0.882
broadband+tones – no_sound	-0.18	0.17	124	-1.07	0.570
broadband+modulated_tones – no_sound	**-0.48**	**0.16**	**124**	**-2.91**	**0.012**

In fact, only the warning sound with modulated tones showed a statistically significant difference from having no warning sound. On average, introducing the modulated warning sound decreased response times by 0.48 seconds.

Preference results

Figure 15 shows the distribution of ranks given to the three different sounds, where rank 1 is the most preferred sound, and rank 3 is the least preferred sound.

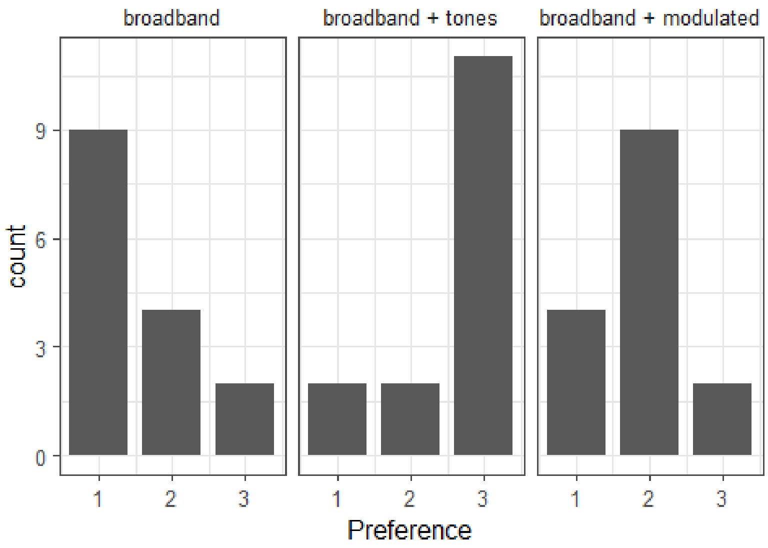


Figure 15. Distribution of preferences for the three warning sounds.

As can be seen from Figure 15, the participants most often chose the broadband sound as their most preferred sound, and the broadband plus tones sound as the least preferred one. The sound with broadband and modulated tones was most often ranked in the middle.

In the free input field, participants characterised the **broadband** sound as “more relaxing”, “more soothing”, “least annoying – continuous and not tonal”, “almost like the sea” and said it “sounds more like traffic noise and a stream rushing”.

The **broadband plus tones** sound was described as a “buzzing sound that is very irritating”, “deafening as it emitted a continuous loud hum”, “annoying”, “most annoying as it was strongly tonal”, “too annoying”. Two participants expressed they would not want to listen to it regularly by writing: “I could not listen to it for long periods of time as it was not pleasant to listen to” and “very grating on the ears and I would definitely not want to listen to that every day!”.

Finally, the **broadband plus modulated tones** sound had the most mixed responses. Participants described it as “reasonably ok”, “a bit annoying”, “in the middle - it was annoying because of the variation” and wrote: “I’m still not very comfortable with [it], but that seems was better than the rest of choices”. One participant, however, thought it was “really distinctive and very annoying”. A few participants also commented that it “sounded futuristic”, “sounds like something from Tron which is ok but not entirely what I'd want to hear in the environment” and “[it] is a bit scary as it sounds like a beam but it is the one I would choose”.

Discussion

The results detailed above show that in controlled experimental conditions, a significant benefit of introducing a warning sound to an e-scooter was demonstrated. In particular, the broadband sound with modulated tones decreased detection time of the moving hazard by 0.48 seconds. To put the number in context, as the scooter was moving at 15 mph, this translates to noticing it at a distance 3.2 meters further away than when there is no scooter. It is worth pointing out that in the experiment, the scooter approach was only a few seconds long, and a larger benefit might be possible to measure if these approaches were longer.

The warning sound which performed best in terms of noticeability – broadband with modulated tones – was also ranked as second in the preference task, which shows that a good compromise between noticeability and annoyance can be found with well-designed sounds, and that amplitude modulation can be a way of achieving this.

The statistical tests did not show any interaction between the sound and environment, even though these types of effects would be expected given the different background noise levels in both environments. However, it is worth noting here that because of the relatively small number of participants, the statistical analysis is likely underpowered, and a relationship between type of sound and environment might be uncovered with sufficient sample size.

The absolute values of the detection time advantage (with simulated sounds compared to no additional sound) should be treated with some caution as this is likely to have been influenced by the experimental design and context; for example, a change in attentional state or direction of approach would likely alter results. Taking this into account, it is concluded from the results of this preliminary subjective experiment that a sound can be designed which alerts people sooner to the presence of a scooter, even in relatively complex noisy urban environments when their attention is focused on another task. It has been shown that an alert sound can be effective without an unacceptable increase in overall sound pressure level or annoyance to people nearby.

5. Recommendations for Further Work

This section presents a non-exhaustive list of recommendations for further work. These recommendations are grouped into two main areas: (i) generation of warning sounds, and (ii) performance assessment of warning sounds to increase vehicle noticeability.

- Generation of warning sounds: The current system consists of hardware (e.g., Raspberry Pi, microphone array) and software (i.e., coding for generating audio files). This system monitors vehicle operating conditions, and then generates an audio signal in real-time. Therefore, the system is able to produce a sound of varying characteristics (e.g., pitch shifting) as a function of vehicle operation (e.g., vehicle speed). Further work is proposed to optimise the system:
    - Increase robustness and integration of the developed system. As yet robustness, pricing of components and other challenges relating to stresses imposed during use have not been considered in this study.
    - Further implementation of more audio features for the real-time signal generation. Including much of the elements used in the offline signal generation such as delay-

    based effects, filtering, amplitude modulation, and other signal components such as white noise and impulsive sounds.
  - Expand the system so the warning sounds are optimised for the specific soundscape. In other works, a microphone could record the existing background noise in the specific area of operation; the system could take this data and then generate a warning sound with a SPL according to the background noise to ensure noticeability. This is likely a challenging task, as the system must work in real-time.
  - Further research is required to achieve an optimised solution when a set of suitable candidate sounds have been selected for a real-world trial.
- Assessment of warning sound to increase vehicle noticeability: During this project a subjective experiment has been carried out when a series of participants have assessed the noticeability of the scooter with under four different conditions: (i) no added sound), (ii) added broadband sound, (iii) added broadband plus tonal sound, and (iv) added broadband plus modulated tonal sound. This experiment has provided useful information for the basic understanding of the contribution of each of these acoustic features to the vehicle awareness. Also, a series of Sound Quality Metrics have been implemented to analyse the psychoacoustic annoyance and sensory pleasantness for a series of candidate warning sounds generated. Further work is proposed below:
  - Optimise the warning sounds generated for a better balance between noticeability and annoyance.
  - Carry out a comprehensive assessment of sounds with different acoustic features for vehicle awareness under controlled conditions (in the lab).
  - An extensive virtual experiment would be required for further controlled lab testing. This would likely be in virtual reality with full spatial (3D) audio reproduced over loudspeakers, in an immersive environment possibly like the Octave facility at the University of Salford. This permits participants to move around in the space. Whether using this or a VR headset as was used in this study, it would be necessary to simulate a wider range of scenarios; this includes environment, number of hazards (including multiple scooters), and different directions of approach, taking into account the directivity of proposed hardware solutions. Full simulation rather than real filming would permit creation of more hazardous situations than it was possible to create in this study. This study would require a larger pool of participants than it was possible to test in these initial experiments.
  - In collaboration with RNIB, design a series of subjective experiments optimised for partially sighted or blind people.
  - Once candidate warning sounds are designed and agreed with interested parties (i.e., Dott and RNIB), carry out field trials with the warning sounds system implemented in Dott scooters. The subjective experiments carried out at Salford have been (or will be) designed to have a substantial ecological validity (e.g., using 3D audio and VR). However, it is almost impossible to include all the complexity of usual conditions in typical urban environments. Therefore, a more holistic assessment of the warning sounds system is proposed to be done during a series of field trials in representative urban scenarios.

In the short term, it has been agreed with Dott the development of a roadmap for a holistic assessment of the warning sounds within the London (and other cities) urban system. This is to be developed and presented to Transport for London. To gather useful data for the assessment of the warning sounds for vehicle awareness, it is proposed the following:

- Agree with Dott and RNIB the planning of the field trials.
- Salford's ARC to advise on additional resources needed to equip 2 scooters with the system developed for the generation of warning sounds.
- RNIB to assess and provide feedback for the warning sounds used during the subjective experiment at Salford (i.e., including warning sounds and existing soundscape in two different locations).
- Salford's ARC to advise on additional resources needed to advise on the survey methodology to conduct the warning sounds assessment in live environments like London.
- It is expected that such an exercise would generate qualitative data that could be used to aid the design of a controlled laboratory experiment already outlined; the feedback from this exercise would be essential to ensure that the subsequent research is not only fully accessible for all participants, but features the range of scenarios and hazards of highest priority to the blind and partially-sighted community.

5.1 Proposal for Next Steps and Estimate of Associated Costs

During the existing project the following tasks have been completed:

- Generation of sounds and demonstration of implementation on scooter
- VR pilot study with analysis/report and feedback from stakeholders

Following discussions with Dott the following next steps have been agreed and a cost estimate is provided below as a starting point for further discussions. In section 5.2 potential routes to funding are highlighted.

Next steps:

1. Sound optimisation and subjective testing (cost estimate depending on scope £50,000)
    a. Generation of broader range of function related sounds
    b. Evaluation using sound quality metrics
    c. Screening process using simplified subjective test
    d. Decision on candidate sounds
    e. Modify VR experiment based on feedback from RNIB to include an additional scenario and develop distraction test based on interaction with a mobile device (TBA with Dott/RNIB)
    f. Full scale VR test with 50 participants
2. Integration (cost estimate depending on scope £25,000)
    a. Liaise with Dott engineers to find effective solution to on board sound generation
    b. Provide hardware requirements

    c. Provide code for sound generation (based on 1.f) to be functional where possible with existing hardware on e-scooter
    d. Note that code will be provided in Python format – functional on a device with computing power equivalent to a Raspberry Pi 4.
3. Field trials (cost estimate depending on scope £15,000)
    a. Support Dott to equip 2 scooters with functional sounds from the integration phase
    b. Provide outline recommendations for mobile phone app allowing scooter users to detail their experience – e.g. near misses/sound quality
    c. Development of questionnaire to assess reactions from the general public in terms of added awareness and annoyance
    d. NOTE- Attendance at these field trials by University of Salford is optional with additional associated cost (not included)
    e. NOTE – Integration into app to be carried out by Dott's existing app developers (not included)
4. Meetings/reporting (cost estimate depending on scope £5,000)

5.2 Routes to funding

This three-month project has served to set the foundations for a large-scale project. A proof of concept has been developed, and a fundamental research has been carried out to understand the noticeability of different acoustic features.

There is the objective for Dott to conduct a series of field trials in London (and potentially other cities) with the system developed for generating warning sounds. That would require covering the expenses for the Salford's ARC team, in terms for staff time and resources. Due to the short time scales for carrying out that work, it is assuming that these expenses could be directly covered by Dott (with some potential contribution from the University of Salford via the HEIF fund).

During the regular discussions with Dott and RNIB, the different parties have expressed an interest in engaging in a larger knowledge exchange project, to develop an AVAS integrated system, and to expand the research about warning sounds to increase vehicle awareness both in the lab under highly controlled conditions and in the field.

It has been agreed that UKRI funding is not optimal for this larger scale project, due to the timelines of such funding avenues. It is considered that Innovate UK type funding is more appropriate. Different options are considered at this state, including:

- Innovate UK Knowledge Transfer Partnerships (KTP)
- Higher Education Innovation Fund (HEIF) Knowledge Exchange
- iCase PhDs

Appendix A presents the University of Salford Engagement Model for industry, and appendix B displays a comparison between different options for applying for funding.

6. Summary

Sound generation:

- A standalone system for the generation of warning sounds have been developed. The system, consisting of both hardware and software, generates a warning sound as a function of the scooter's operating conditions (i.e., vehicle speed).
- A real-time implementation has been developed using a Raspberry Pi computer, within a Python environment. Appendix C presents the pseudocode for warning sound generation and scooter implementation.
- A system has been developed for the amplification and subsequent radiation of sound according to a voltage input at line level.
- This device can be connected to any module that generates a sound signature (e.g., a warning sound system) and this may be wired or transmitted (e.g., by Bluetooth).
- The microphone array fitted below the handlebars of the scooter provides a strong output with a maximum output level exceeding 50dBA in the direction of travel above 1000Hz (at 1.5m from the rider position).
- The array radiates predominantly in the forward direction as requested by RNIB.
- The unit has been tested in the laboratory and a front to rear bias of around 10dBA at 1000Hz was observed. 20dBA at 5000Hz. This is likely to be a desirable feature.
- To put this in context – humans perceive a reduction of 10dBA to be approximately a halving of loudness. At 1000Hz the sound level radiated by the loudspeaker array would therefore appear twice as loud for an approaching scooter when compared to one which has passed by.
- Further research is required to achieve an optimised solution when a suitable candidate sound has been selected for a real-world trial. Due to budget and time constraints, this study has not considered the pricing of components and other challenges relating to stresses imposed during use.
- Reproducing a generated sound at a sufficient sound level and focusing it on a specific direction appears relatively straightforward but there will be additional costs to the scooter manufacturer. This is fully achievable but with associated costs.

Noticeability and preference of demo warning sounds:

- A laboratory study has been carried out to gauge pedestrian awareness of an approaching scooter without and with added warning sounds.
- The broadband plus modulated tones sound has been found to decrease the detection time of the approaching scooter by 0.48 seconds (compared to the scooter without any added

sound). With the scooter moving at 15 mph, this translates to noticing it at a distance 3.2 meters further away than when there is no scooter.

- The broadband plus modulated tones sound performed best in terms of noticeability and was also ranked as second in terms of preference. Amplitude modulation seems a very efficient acoustic feature to increase vehicle noticeability.
- This indicates that a good compromise between noticeability and annoyance can be achieved with a well-designed warning sound.
- Further research is needed to design warning sounds with an optimal balance between noticeability and annoyance.

Appendix A: University of Salford Engagement Model

Academic Year September 2021 to August 2022

1. Definitions

1.1 Live Briefs

A short, assessed, project of between 1-4 weeks offered to undergraduate and postgraduate students within their existing studies. Typically take the form of a bespoke project e.g., create a new logo, devise a marketing brand, create a software app etc. The external organisation will create a project brief to submit to the University, then students will typically remain based in the University with some visits to the organisation’s premises to work on the project, with direction from the lecturers. A live brief will be unpaid.

1.2 Masters Project

A Masters project is an opportunity for postgraduate students to complete their studies by undertaking an internship (1-3 months) or placement (6 months). The brief is set by the external organisation, and the student works on-site or remotely to meet its objectives over the period of the agreed project. A three-month project is paid or unpaid depending on the nature of the project, and a six-month project is paid.

1.3 Student Project

A project undertaken outside of formal teaching, therefore not assessed by the University. The student undertakes the project in addition to their studies to obtain real world experience. Projects are agreed between the student and the external organisation. The project is typically undertaken at the organisation's premises. The project can vary in length from 1 week to 3 months.

1.4 Year in Industry

Undergraduate student with two complete years of study that will work with an external organisation for 12 months on a paid contract. The year will form part of the student's studies and be assessed, so the organisation will need to remain in contact with the placement tutor to monitor and assess the student's progress.

1.5 *Industrial Cooperative Awards in Science & Technology (iCASE)*

Industrial CASE studentships are allocated directly to a limited number of businesses. The company takes the lead in defining a student project. The University of Salford, in partnership with the company recruit a suitable eligible candidate. Projects must be in the engineering or physical sciences and are jointly supervised by the academic and industrial partner.

1.6 *Knowledge Transfer Partnership (KTP)*

A KTP enables a business to innovate and grow and bring in new skills and the latest academic thinking to deliver a specific, strategic innovation project through a knowledge-based partnership with the University.

KTPs typically run for 2years and are part funded by InnovateUK (for more information click here).

2. Key Dates

Career Fairs		
Due to COVID-19, Career Fairs will be 'virtual'		
Salford Business School	Wednesday November 4th	10:00 to 16:00
School of Arts & Media	Wednesday November 11th	10:00 to 16:00
School of Health & Society	Wednesday November 18th	13:00 to 19:00
School of Science, Engineering & Environment	Wednesday November 25th	10:00 to 16:00

Due to COVID-19 Career Fairs will be 'virtual' and run over 6hours. The Virtual Career Fairs will allow exhibitors to have a variety of digital content as well as the opportunity to have one-to-one appointments or webinars during the event. Click Here to access the Expressions of Interest Form

University Open Days, 2020 (on-line)
October 10th (Undergraduate)
November 18th (Postgraduate)

3. Student Availability

Live Briefs	**Trimester 1**	**Trimester 2**	**Trimester 3**
Proposal submission	July 1st to Aug 30th	Oct 1st to Jan 10th	Mar 1st to May 1st
Start/End	Oct 5th / Dec 18th	Feb 1st / May 28th	June 7th to Sept 13th

Masters Project	**Trimester 1**	**Trimester 2**	**Trimester 3**
Application Date	July 1st to Aug 30th	Oct 1st to Jan 10th	March 1st to May 1st
Shortlisting	Minimum of 6weeks before project start date		

Start Date	Oct 5th / Dec 18th	Feb 3rd to May 28th	June 7th to Sept 13th

Student Project	**Trimester 3**
Application Date	Nov 1st to March 1st
Shortlisting / interviews	Minimum of 6weeks before project start date
Start Date	June 1st

Year in Industry	
Application Date	Nov 1st to May 31st
Shortlisting / interviews	Minimum of 6weeks before project start date
Start Date	June 1st to Sept 1st

Proposal submission: period when industry/business partners submit ideas/proposals to the University for the University to incorporate the project within the relevant course/s and enable the University to promote it to students.

Start/End Date: period when students are available to undertake the Live Brief/Project/etc.

Shortlisting / interviews: period set aside for the industry/business partner to undertake their recruitment process.

3.1 *Availability timetable*

Sept	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Trimester 1					Trimester 2				Trimester 3		
🟦	🟩	🟩	🟩	🟥	🟩	🟩	🟩	🟥	🟨	🟨	🟨

Key

🟥	Examination period – no access to students
🟨	Students available to start Year in Industry. Also, Postgraduate/Masters projects, work placements, live briefs, etc.
🟦	Student induction/registration.
🟩	Peak of student availability

Appendix B: Comparison Between Funding Opportunities

	Knowledge Transfer Partnership (KTP)	**HEIF Knowledge Exchange**	**iCase**	**Industrial Masters by Research**
Project duration	12 to 36 months	12 to 18 months	42 months	12 months
Eligibility	All private sector industries; High growth SME in key market sector *or Large company;* Clear business case with significant increase in profit.	All industries; all sectors; SME or Large organisation; Clear business case.	All industries; all sectors; SME or Large organisation; Cutting-edge research relevant to the organisations' priorities and objectives	All industries; all sectors; SME or Large organisation; Cutting-edge research relevant to the organisations' priorities and objectives
Delivery mechanism	KTP Associate jointly recruited; Associate on fixed term contract (employed by University); Associate based at company premises	KE Affiliate jointly recruited; Affiliate on fixed term contract (employed by University); Affiliate based at company premises	Student jointly recruited; University based with no less than 4 months spent at company	Student jointly recruited; University based with no less than 4 months spent at company
Restrictions	Company partner must have UK presence 5+ employees 2 years audited accounts Significant knowledge transfer required	Company partner must have UK presence Good credit rating	Company partner must have UK presence Good credit rating	Company partner must have UK presence Good credit rating
Cost of project	2 year KTP on average costs £125,000	12 month project on average costs £64,000	42 month project on average costs £102,000	12 month project on average costs £30,000

Grant contribution	Up to 67%	Up to 60%	£35,722	25%
Typical Company contribution	SME £20,625 Large £31,250 per annum	SME £25,750 Large £35,300 per annum	£66,341 + travel and consumables Over 3 ½ years	£22,750 per annum
Academic commitment	Equivalent of ½ day per week by academic team	Equivalent of ½ day per week by academic team	100 hours per year	100 hours per year
REF returnable	Yes – all income	Yes – company contribution only	Yes – company contribution only	Yes – company contribution only
Academic benefit	REF returnable income; Publications Case studies; Potential for Higher Degree registration for Associate; Industrial collaboration	REF returnable income; Case studies; Industrial collaboration	REF returnable income; Publications; Case studies; PhD completion within 4 years; Industrial collaboration; Research and teaching reflecting the needs of the economy.	REF returnable income; Case studies; Industrial collaboration
Submissions	6 times per year- linked to sector competitions	Anytime	Anytime – linked to student registrations	Anytime – linked to student registrations

Appendix C: Pseudocode for Signal Generation and Scooter Implementation

3.1 Pseudo-code for offline python signal generation:

Libraries used:

Scipy.signal

Numpy

Simpleaudio as sa

```
Function Playdata(input):
    Play = sa.playbuffer(input,1,2,44100)
    Play.wait_done()
    Play.stop()

Function Delay(input, fs, delay_time):
    filter = zeros(delay_time*fs+1)
    filter( delay_time*fs) = 0.7
    Output = convolve(input,filter)

Function Flanger(input):
    Lfo = Sawtooth(2*pi*Lfo_Freq)
    Index = [number_of_samples-Lfo_amp*Lfo]
```

```
For i in range(length(input))
Output = input[I] + input[index[i]]

Function softclip(input):
Output = arctan(input)

Function HWR(input, mix):
Hwr = (Input+ abs(Input)) / 2
Output = mixval * Hwr + (1-mixval) * Input

Function LowPass(Input,Fc):
b,a =LP_Filter(Order,Fc)
Signal_Filtered = Apply_Filter(b,a,input)
Output = Signal_Filtered/max(abs(Signal_Filtered))

Data = Input.Read('AccelerometerData.csv')
b,a = signal.butter(10,0.5)

Filt_Data = filter(Data,b,a)

Delay_Dat = Delay(Filt_Data,1)

Flange_Dat = Flanger(Delay_Dat)

Datout = Flange_Dat*32767/max(abs(Flange_Dat)) #normalisation of data for 16 bit integer values
```

```
Playdata(Datout)
```

3.2 Pseudo-code for on-scooter signal generation:

```
 Import Accelerometer_lib
Import time
Import numpy as np

BLOCK_SIZE = 1024
 i2c = open(i2c_board)
sensor = Accelerometer_lib.Accelerometer_name (i2c, address)

Function callback:
Windowing = hanningwindow(blocksize)
Data = Sensor.read()
Sig = sinewave(2*pi*f*(Data))
Window_Data = Sig*windowing
Output(Channel1) = Window_Data*gain1
Output(Channel2) = Window_Data*gain2

Stream.Open(SampleRate = 44100, blocksize = BLOCKSIZE, callback)
Try:
        While True:
Sleep(2000ms)
```

```
Except KeyboardInterrupt:

Pass
```