ORIGINAL ARTICLE

Evaluating use of the Doppler Effect to Enhance Auditory Alerts

Stuart Cunningham[a] and Iain McGregor[b]

[a]Centre for Advanced Computational Science, Manchester Metropolitan University, Manchester M1 5GD, UK; [b]School of Computing, Edinburgh Napier University, Edinburgh EH10 5DT, UK

ARTICLE HISTORY
Compiled January 29, 2021

ABSTRACT
Auditory alerts are an essential part of many multi-modal interaction scenarios, particularly in safety and mission critical settings, such as hospitals and transportation. A variety of strategies can be employed in the design of auditory alerts, often orienting manipulation of volume and pitch parameters. However, manipulations by applying a Doppler effect are under-investigated. A perceptual listening test is conducted (n=100) using multiple alert sounds that are subjected to a variety of volume, pitch and Doppler manipulations, with the unaltered sounds serving as a benchmark. Applying a mixed methods approach consisting of inferential statistics and thematic analysis, it is found that decreases in volume and a Doppler simulation of a sound moving away reduce importance and urgency, increase safety, are harder to detect, and are perceived as being more distant in perceptions of auditory alerts. Further, increases in volume and a Doppler simulation of a sound approaching are effective in communicating safety, whilst pitch manipulations were much less effective. Further work is required to provide wider, ecologically valid, verification of these findings, particularly as to how listener detection of Doppler and volume manipulations can be improved.

KEYWORDS
Auditory alerts; Doppler effect; listening test; mixed-methods; sound design.

1. Introduction

As a moving sound source passes a stationary listener it will be perceived as either rising or falling in pitch, due to the object's relative position (Neuhoff & McBeath, 1996). This effect was formalised by Doppler in 1842 in relation to the observed colour of distant stars, and how the corresponding wavelengths of light could be used to calculate relative movement (Doppler & Studnica, 1903). Sound sources that are discerned as approaching are experienced as more negative and regarded as more important than those that are stationary or receding (Tajadura-Jiménez, Väljamäe, Asutay, & Västfjäll, 2010). This effect is more pronounced the louder a sound is and whether it is artificial or natural in nature, unpleasantness also has a greater effect. The effect can be mathematically described in (1) for the perceived frequency f' of a sound source

CONTACT Stuart Cunningham. Email: s.cunningham@mmu.ac.uk

approaching or moving away from a static listener

$$f' = f\frac{c}{c \pm v} \tag{1}$$

where f is the frequency of the sound emitted at the source, c denotes the speed of sound in air and v is the velocity of the sound source.

Surprisingly the longer a sound takes to increase in amplitude the more threatening it is, although this could be due to fast attack sounds being perceived as quieter than slow attack (Tajadura-Jiménez et al., 2010). Rising intensity of a sound (15 dB) associated with an object's movement towards to a listener has been shown to decrease reaction times and increase alertness (Bach et al., 2008). There is considered to be an overestimation of the sound pressure level changes for approaching sounds rather than receding, which is accompanied by the perception that advancing auditory events are closer than withdrawing ones (Neuhoff, 2001).

This study aims to explore whether applying the Doppler effect to auditory cues using Digital Signal Processing (DSP) could alter an alert's perceived importance, level of danger, urgency, ease of detection and proximity.

2. Background

2.1. Audio Alerts

Pitch and amplitude changes have been used successfully to convey the perceived urgency of aviation alerts, with higher and louder being more urgent. Changes which occurred over 20 ms were utilised to make the sound events less startling, as well as providing a sense of spatial front-back (z-axis) movement (Patterson, 1990). Frequency affects perceived loudness of a sound source (Fletcher & Munson, 1933), and louder sounds will generally be perceived as higher pitched than quieter sounds, which inevitably affects the perceived timbre (Melara & Marks, 1990). Other characteristics of sounds that give rise to their perception as alarms were shown to be: an amplitude peak-to-total-time-ratio of 70% or greater, inter-burst intervals $\leq$ 125 ms, a minimum of three harmonics being present, and a base frequency $\geq$ 1000 Hz (Singer, Lerner, Baldwin, & Traube, 2015).

Audible alarms are commonly set to an amplitude that will make them 15 – 25 dB above the pre-existing noise floor. The main reason is to ensure that no other sound events mask the alert. There is also the need to ensure that the auditory event is not too loud in order to prevent it being permanently switched off, attended to prior to the issue being solved, emotionally disturbing other users who might not be able to react to the sound, and most importantly not cause the startle effect, which could dramatically slow down reaction times (Edworthy, 1994).

Experimental work into the perception of urgency, as a function of various manipulations to the sound stimuli, has been undertaken with a group of human participants. The sounds used in these studies were subject to a range of acoustic manipulations such as fundamental frequency, harmonic content and complexity, amplitude envelope, and temporal and melodic characteristics. It has been shown that manipulations of acoustic parameters, particularly when used in combination, lead to a range of predictable and readily detected outcomes in the perception of their urgency (Edworthy, 1994). In particular, speed (the duration of the interval between pulses) was shown to provide the greatest practical economy, in terms of the extent of manipulation required to allow participants to differentiate between urgencies of large magnitudes (Hellier, Edworthy, & Dennis, 1993).

There is also an established trend that the more serious the event an alert sound represents the higher its pitch (Sinclair, 2012). In contrast it has been proposed that modulation of lower frequencies is more effective for listener detection of alert sounds with both younger and older users, specifically a fast modulation around 500 Hz (Huey, Buckley, & Lerner, 1996). There is also a clear requirement that an alert sound should not reach pressure levels where users' hearing might be damaged, as discovered in the work relating to cordless telephone ringing sounds reported by Orchik, Schumaier, Shea, and Moretz Jr (1985).

Habituation, desensitisation or inattentional deafness are ongoing issues in environments where multiple, similar, auditory alerts repeatedly occur (Cabrera, Ferguson, & Laing, 2005; Causse, Imbert, Giraudet, Jouffrais, & Tremblay, 2016). Masking is also an issue, with sometimes less serious auditory alerts preventing essential sounds from being attended to (Hasanain, Boyd, Edworthy, & Bolton, 2017), although any form of masking has a significant effect upon accurate interpretation, and the similarity of alert cues can lead to error rates of around 30% (Bolton, Zheng, Li, Edworthy, & Boyd, 2019).

Edworthy et al. (2017) studied the efficacy of four sets of alternate audio alert designs for recognizability and localisation in medical equipment. A total of 194 participants were involved in evaluating the sounds, with each being presented with normalised stimuli of between 75 to 80 dB(A), which were presented via headphones. Each participant was allocated to one of the four sets of sound design sets under evaluation. Their work found that designs making use of auditory icons, word rhythms and simple acoustic metaphors outperformed the alert sounds specified in the associated IEC 60601-1-8:2006 standard, with auditory icons producing the strongest outcomes.

Understandably, a great deal of work on the topic of auditory alerts has emerged in fields where the visual attention of a user is otherwise engaged, such as when driving a vehicle or other form of transportation, and for pedestrians (sighted and visually impaired) who may be in the vicinity of vehicles with very low noise emissions, such as electric cars (Begault, Anderson, & McClain, 2003; Chamard & Roussarie, 2012; Fagerlönn, 2011; Han & Lee, 2017; Kim, Emerson, Naghshineh, Pliskow, & Myers, 2012; Singer et al., 2015; Yasui, 2018, 2019). In particular, Kim et al. (2012) noted the potential difficulty that visually impaired pedestrians might have in determining the direction (such as turning a corner) of a vehicle equipped with an artificial alert sound, which indicated the vehicle was accelerating. Chamard and Roussarie (2012) also undertook work in this field, utilising sounds that are directly influenced by the vehicle's characteristics (direction, speed, and so on).

The use of spatial movement of alert sounds, such as spatial jitter, has been examined and shown to lower the threshold of detection by listeners by either 7.8 or 13.4 dB, depending upon conditions, against a simulated ambient noise floor. These results were obtained in a relatively small study simulating alerts for aircraft pilots and using HRTF convolution techniques to simulate the spatial motion of the alert sounds presented via headphones (Begault et al., 2003). Such findings suggest that aural cues that have acoustic characteristics, indicative of motion, may improve the sensitivity of listeners to alerts of this type.

2.2. The Doppler Effect

In real world environments, the Doppler effect is a naturally occurring phenomenon for the human listener and will be perceived in situations where sonic cues may be interpreted as alerts. For instance, one might consider the sound of an emergency services vehicle approaching and passing by a stationary listener, where the pitch, as well as amplitude, of the siren will increase as the vehicle approaches and decrease as it moves away from the listener. The same is true for any sound emitting object to a greater or lesser extent. Thus, it is reasonable to postulate that human cognitive processing of sounds that exhibit the Doppler effect will provoke a range of cognitively driven or acute stress responses, such as one of *fight-or-flight* (Cannon, 1916). Such an acute stress response is especially likely in the case of a sound with a Doppler effect that rises in pitch, emulating the rapid approach of an object, for evolutionary reasons.

The perceptions and properties of the Doppler effect have been studied previously by Oechslin, Neukom, and Bennett (2008), who performed a study with twelve human listeners, reporting as having normal hearing, in a controlled acoustic environment (an anechoic chamber). Participants were blindfolded and asked to provide a binary rating for a series of sounds. The ratings were either that a sound was moving towards or away from the participant. A deception or placebo device was deployed in the study in that the participants, prior to being blindfolded, were able to see the speaker setup with the speaker positioned on a rail and told that it could move towards or away from them. In reality, the speaker was static and DSP used to produce pitch changes in the sounds presented. Crucially, the amplitude of the stimuli was kept static so as to isolate the effect of only the Doppler pitch change. The sounds presented fell into three categories: sirens, pink-noise and sine tones with 100 trials of each category presented.

Although the number of listeners used by Oechslin et al. (2008), were relatively small, the study found significant differences in several key areas. Most relevant to our own work is the phenomenon that participants were able to detect approaching siren sounds far more accurately than sounds that were moving away, suggesting a greater sensitivity to alert sounds of an approaching nature, and providing evidence that supports the hypothesis that human listeners are particularly sensitive to sounds exhibiting a Doppler effect.

These existing works in the field demonstrate the potential and value of investigating multiple sound design types and strategies in devising effective audio alerts. However, it is notable that the use of the Doppler effect is under-examined in the field, leaving significant scope for its efficacy to be explored, leading to the primary research activities that are subsequently reported in this article.

3. Method

3.1. Participants

One hundred Edinburgh Napier University members of staff and students took part in the study. Of the participants 39% were academic staff, 29% support staff, and 32% students. Each participant believed themselves to have normal hearing for their age. This screening ensured that audio clips could be accurately perceived. Ages ranged from 18 to 68 with a median of 42, 50% of participants were male, 49% female and 1% non-binary. Everyone was recruited via email or face to face. Participation was voluntary and no one was compensated in any way for taking part in the study.

3.2. Materials

Participants listened to 70 short mono 16 bit 44.1 kHz wav audio clips that were randomly embedded within PowerPoint presentations running on a 13-inch MacBook Pro. The alert sounds were selected at random from a variety of online sources to achieve a wide range. These were played back through a pair of Genelec 8030A loudspeakers, at an RMS of 60.7 dB, with a peak of 92.3 dB (A weighting). The clips varied in duration from 0.7 s to 14 s, with a median of 2.65 s (see Table 1). Printed A4 paper forms were used to capture participants' responses, with pens being provided for convenience.

Table 1. Auditory alerts: level of importance, duration, amplitude, pattern, pitch and envelope.

Description	Importance	Duration	RMS	Peak	Pattern	Pitch	Envelope
Email alert	Low	1.6s	60.7 dBA	79.8 dBA	3 x 0.14s pulses, 1.2s pulse, no gaps	Descending 589, 392.5, 293.9 Hz	Inverse sawtooth
Emergency Alert System	High	14.0s	60.7 dBA	75.4 dBA	3 x 0.9s pulses, 0.9s gaps, 8.4s pulse	949.5 – 2994 Hz (short), 855 – 963 Hz (long)	Square
Emergency vehicle	Medium	2.8s	60.7 dBA	75.8 dBA	8 x 0.3s pulses, no gaps	482 - 4535 Hz Repeated Ascending then Descending	Inverse sawtooth
Fire alarm	High	1.1s	60.7 dBA	81.3 dBA	25 x 0.4s pulses, no gaps	0.87, 1.9, 3.2, 4.8, 6.6, 8.7, 10.8 kHz continuous pitches	Inverse sawtooth
Forward collision warning	Medium	3.0s	60.7 dBA	73 dBA	15 x 0.12s pulses, 0.08s gaps	1.3 kHz continuous pitch	Sawtooth
Medical equipment alarm	High	4.7s	60.7 dBA	75.6 dBA	10 x 0.15s pulses, 0.12s, 0.17s, 1.6s gaps	261 – 2355 Hz (5 bands) continuous pitch	Square
Smoke alarm	High	1.4s	60.7 dBA	75.5 dBA	4 x 0.3s pulses, 0.4s gaps	3.4 kHz continuous pitch	Sine
SMS alert	Low	0.7s	60.7 dBA	81.2 dBA	0.1s, 0.1s, 0.4s pulses, no gaps	294 Hz, 440.8 Hz, 591 Hz, rising pitch	Sin
Telephone ring	Low	2.5s	60.7 dBA	81.3 dBA	2 x 0.6s pulses, 1.3s decay	100 Hz – 13 kHz (9 bands) continuous	Sine
Truck backup beeps	Medium	3.7s	60.7 dBA	74.7 dBA	4 x 0.5s pulses, 0.4s gaps	1.2 – 5.8 kHz (4 bands) continuous	Sine

3.3. Design

Ten discrete sound clips were chosen to represent a range of auditory alerts with different levels of importance (see Table 1). Auditory alerts were sourced from government, manufacturers and sound effects libraries, and were utilised under fair use/dealing conditions. Seven versions of each clip were created: *Doppler down*, *Doppler up*, *Normal*, *Pitch down*, *Pitch up*, *Volume down* and *Volume up*, in order to identify which parameters (independent variables of pitch/amplitude) had the greatest effect. All of the modifications were made using a Waves Doppler plug-in (see Figure 1). This allowed separate control of gain and pitch, both of which were set to 100% when used, which represented the correct shift based upon the duration of the clip. Panning, Air Damp and Reverb were all bypassed so as not to provide additional auditory cues. To illustrate, a description of each of the seven modifications for the medical equipment alarm can be observed in Table 2 , which describes the parameters of the manipulations.

Table 2. Medical equipment high alarm amplitude and pitch modifications.

Modification	RMS	Peak	Amplitude	Pitch
Original	60.7 dBA	74.7 dBA	100%	100%
Pitch up	60.7 dBA	76.2 dBA	Logarithmic 99.9 – 102.6%	Logarithmic 100 – 105.94%
Pitch down	60.7 dBA	76.1 dBA	Logarithmic 102.15 – 100.15%	Logarithmic 100 - 93.93%
Volume up	60.7 dBA	84.2 dBA	Logarithmic 66.4 – 138.69%	100%
Volume down	60.7 dBA	87.3 dBA	Logarithmic 162.85 – 54.09%	100%
Doppler up	60.7 dBA	84.6 dBA	Logarithmic 67.31 – 141.84%	Logarithmic 100 – 105.94%
Doppler down	60.7 dBA	87.2 dBA	Logarithmic 159.03 – 54.33%	Logarithmic 100 - 93.93%

Five-point scales were utilised to capture the responses of participants using the antonyms: *Important / Unimportant*, *Safe / Dangerous*, *Urgent / Trivial*, *Difficult to detect / Easy to detect* and *Close / Distant*. The survey finished with a question relating to the level of confidence with the responses given, again on a five-point scale, ranging from *No Confidence* to *Strong Confidence*. Followed by an area to add any further comments.

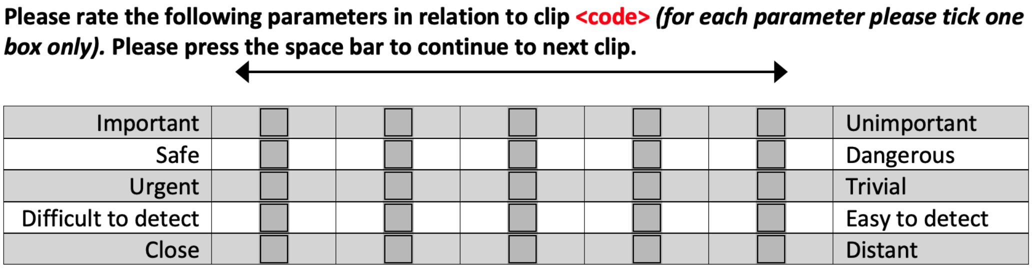
Three potentially similar parameters were identified for exploration *Importance*, *Danger* and *Urgency*. *Important / Unimportant* was sourced from Wille et al. (2017) who found a consistency with regards to this scale for context free interpretation of auditory stimuli. *Safe/Dangerous* came from Edworthy, Hellier, Aldrich, and Loxley (2004) who utilised it as part of research into Helicopter monitoring sounds. *Urgency* is not normally measured using antonyms but was included in this manner in order to maintain consistency for participants. A more common approach is to use a scale running from least urgent to most urgent (Reynolds, Rayo, Fitzgerald, Abdel-Rasoul, & Moffatt-Bruce, 2019). Related concepts were chosen in order to attempt to more fully understand the effects of the modifications upon the auditory alerts. *Danger*, *importance* and *urgency*, whilst similar are not identical concepts, as an alert can still indicate a high level of urgency, but be safe, or be important but not urgent. Ease of detection was chosen to monitor the clarity of the modifications, and proximity to gauge the extent of any distancing, commonly associated with the Doppler effect.

Figure 1. Waves Audio Ltd. Doppler Plug-in (Ltd., 2020).

3.4. Procedure

Participants sat at a table in front of a laptop and a pair of loudspeakers and did not receive any training prior to taking part. They first read a participant information sheet, then a privacy notice. If participants were happy to proceed, they signed an informed consent form.

After entering the participant information on a printed A4 form, each person listened to the 70 randomised audio clips, which played only once, within a PowerPoint presentation via the laptop, that they were able to control. Each participant listened to a uniquely randomised sequence of wav files. Every clip was rated on a printed form using a five-point scale for perceived Importance, Level of danger, Urgency, Ease of detection and Proximity (see figure 2), with each parameter being considered at the same time. The order of the dimensions on the printed response sheet did not change, in order to minimise potential confusion. Participants took anywhere between 13 and 43 minutes, with the average being close to 30 minutes, as they could proceed at their own pace. After entering the participant information, each person listened to the first audio clip, which played only once, and rated the randomised sequence of clips using the five-point scales (Importance, Level of danger, Urgency, Ease of detection and Proximity). This took participants anywhere between 13 and 43 minutes, with the average being close to 30 minutes, as they could proceed at their own pace.


Please rate the following parameters in relation to clip <code> *(for each parameter please tick one box only).* Please press the space bar to continue to next clip.

Important	☐	☐	☐	☐	☐	Unimportant
Safe	☐	☐	☐	☐	☐	Dangerous
Urgent	☐	☐	☐	☐	☐	Trivial
Difficult to detect	☐	☐	☐	☐	☐	Easy to detect
Close	☐	☐	☐	☐	☐	Distant

Figure 2. Printed scales and instructions for participant rating of randomised sequence of audio clips.

After the experiment was completed each participant was provided with a verbal explanation of the underlying principles being tested. All of the responses were subsequently transferred to an Excel spreadsheet for analysis, with the open response comments being coded using Nvivo 12 in order to perform a thematic analysis.

4. Results

4.1. Participant Ratings

Figures 3 through to Figure 7 summarise the participants' responses across all of the ten auditory alerts and their modifications for each of the dependent variables. The *Volume up* and *Doppler up* versions achieve being perceived as the most urgent (Figure 5) and most important (Figure 3) as well as being the most dangerous (Figure 4). This effect is mirrored for the *Doppler down* and *Volume down*, which are perceived as being the least urgent, important, and safest of all versions. These two sounds are also indicated to be the most difficult effects to detect (Figure 6), which may go some way to explain their relatively poor efficacy in conveying the other characteristics assessed. Evaluation of the close/distant measurement shows that the *Doppler down/up* and *Volume down/up*, as one might intuitively expect, contribute to the greatest differentiation of a sound source being near to, or far away from, the listener's position (Figure 7).

Whilst participants rated the sounds using five-point ordinal scales, we employed parametric statistical analysis. This is common practice in the analysis of subjective listening test data (Počta & Beerends, 2015; Villegas, Stegenborg-Andersen, Zacharov, & Ramsgaard, 2016) and forms part of a series of international recommendations (ITU-R, 2015, 2019), which highlight benefits of statistical sensitivity and power.

An alpha level of 0.05 was used to determine statistical significance in inferential

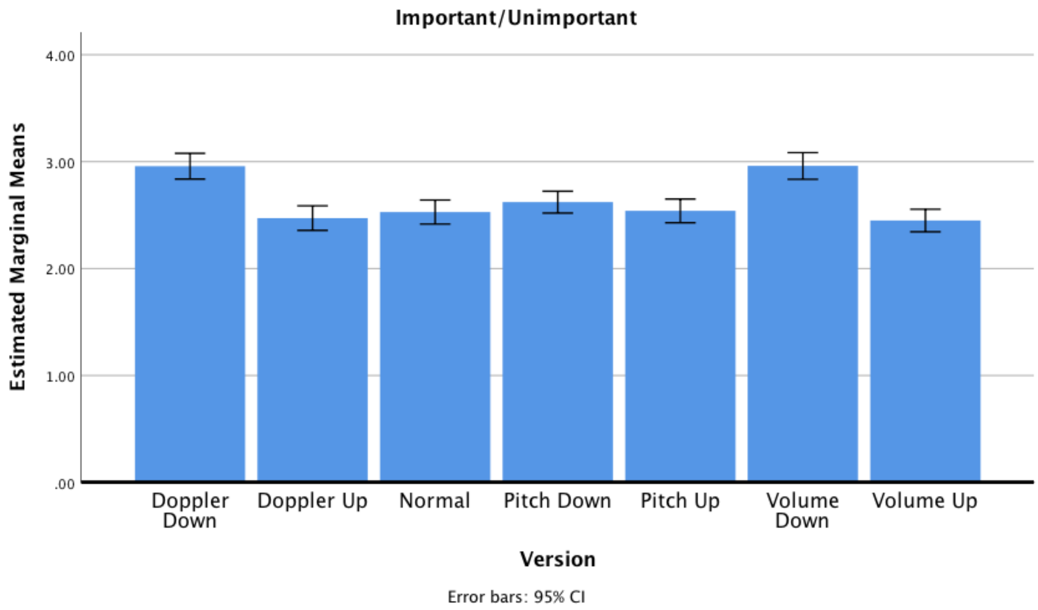


Figure 3. Unimportant / Important Participant Ratings (all sounds)

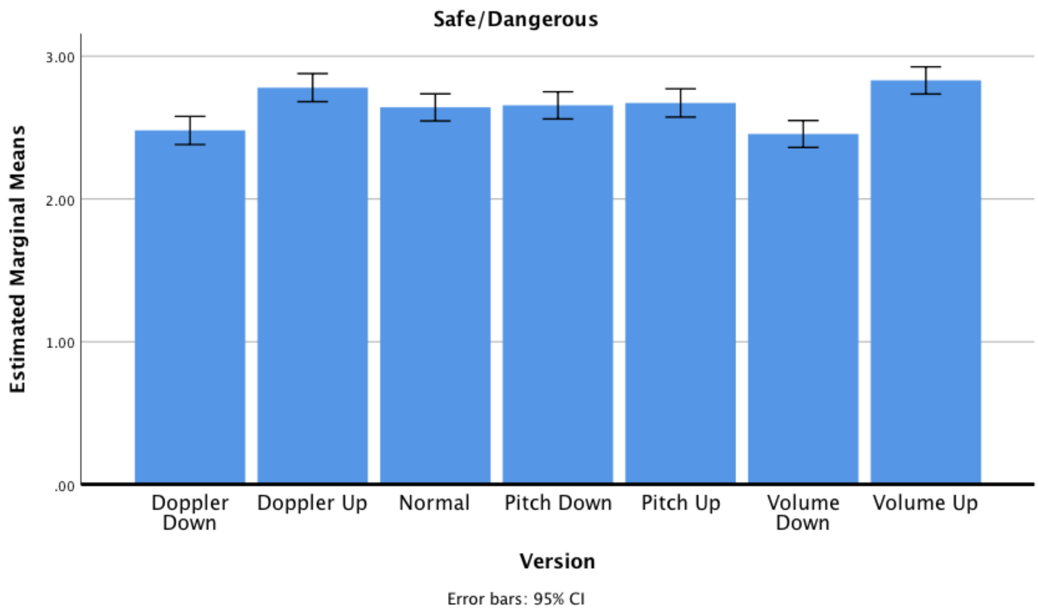


Figure 4. Safe / Dangerous Participant Ratings (all sounds)

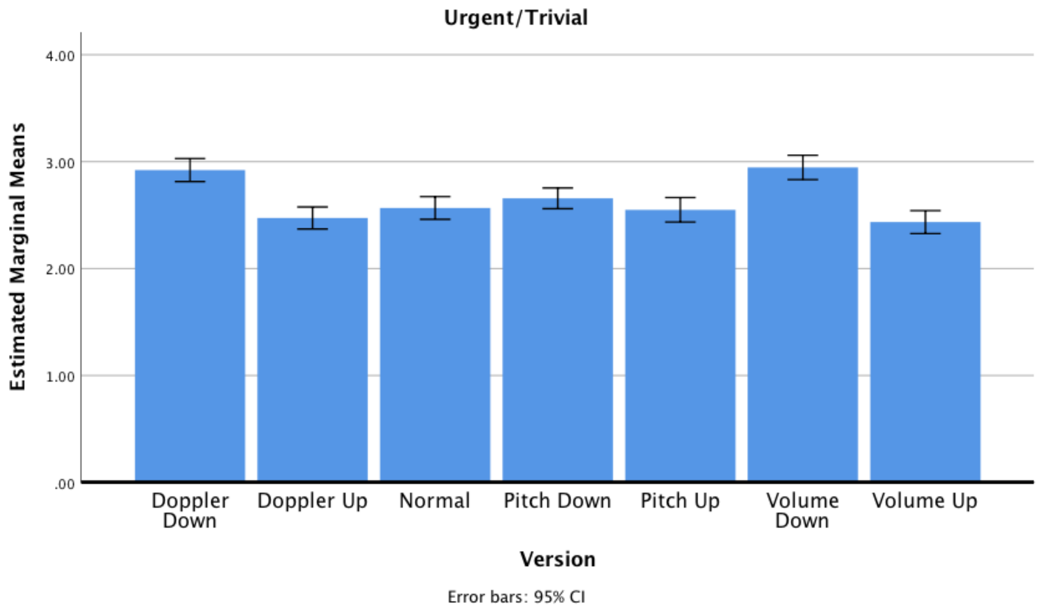


Figure 5. Urgent / Trivial Participant Ratings (all sounds)

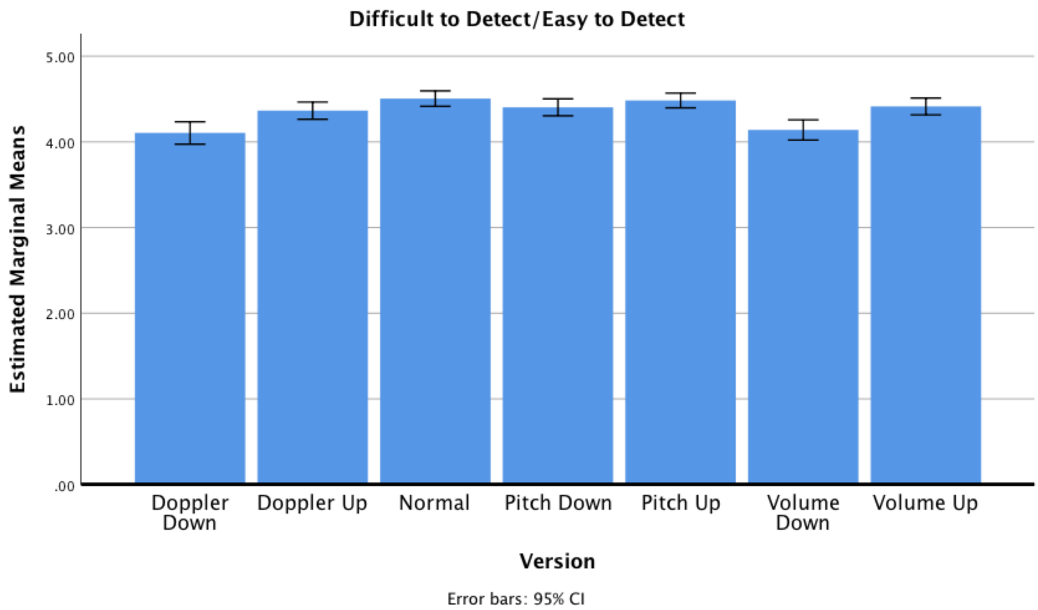


Figure 6. Difficult / Easy to Detect Participant Ratings (all sounds)

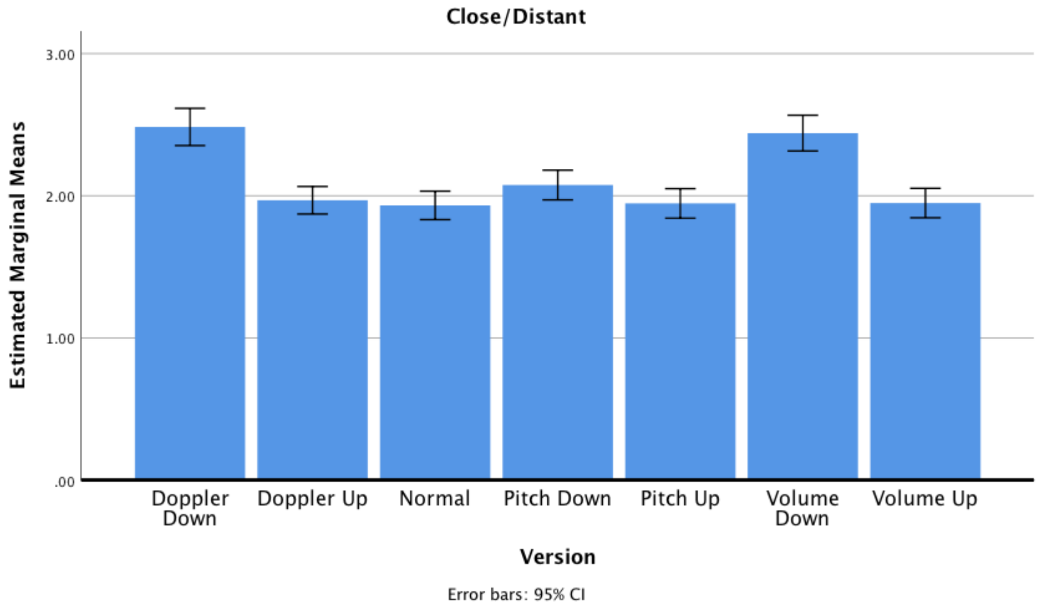


Figure 7. Close / Distant Participant Ratings (all sounds)

tests. To investigate any potential relationship between the dependent variables, a set of Pearson product-moment correlations were run using the ratings participants provided for the normal condition of all of the ten alert sounds. The normal version was chosen due to its purpose as a control condition in the study.

There was a strong, positive correlation between Important/Unimportant and Urgent/Trivial ($r = 0.821, n = 1000, p < 0.001$). A moderate, negative correlation between Important/Unimportant and Safe/Dangerous ($r = -0.560, n = 1000, p < 0.001$). A weak, negative correlation between Important/Unimportant and Difficult to detect/Easy to detect ($r = -0.121, n = 1000, p < 0.001$). There was a moderate, negative correlation between Safe/Dangerous and Urgent/Trivial ($r = -0.571, n = 1000, p < 0.001$). A weak, positive correlation between Safe/Dangerous and Close/Distant ($r = 0.143, n = 1000, p < 0.001$). There was a weak, negative correlation between Urgent/Trivial and Difficult to detect/Easy to detect ($r = -0.101, n = 1000, p = 0.001$). There was a weak-to-moderate, negative correlation between Difficult to detect/Easy to detect and Close/Distant ($r = -0.349, n = 1000, p < 0.001$).

These correlations may provide an indication that the Important/Unimportant, Safe/Dangerous and Urgent/Trivial variables, in particular, may have been measures of a phenomenon that participants considered to be similar. However, since these dependent variables were selected from sources in the literature, they are retained separately for in-depth analysis in this work and the phenomenon may be worthy of future investigations.

Subsequently, a two-way MANOVA with repeated measures was used to investigate the effects of the two independent variables: *alert* (Email alert; Emergency Alert System; Emergency vehicle; Fire alarm; Forward collision warning; Medical equipment high alarm; Smoke alarm; SMS alert; Telephone ring; and Truck backup beeps); and *modification* (Doppler down; Doppler up; Normal; Pitch down; Pitch up; Volume down; and Volume up) upon the five dependent variables: *Important / Unimportant; Safe / Dangerous; Urgent / Trivial; Difficult to Detect / Easy to Detect; and Close /*

Distant.

The two-way MANOVA revealed statistically significant main effects in *alert*, $F(45, 54) = 46.232, p < .0005$; Wilk's $\Lambda = 0.250$, partial $\eta^2 = .975$ and a statistically significant difference in *modification*, $F(30, 69) = 9.616, p < .0005$; Wilk's $\Lambda = 0.193$, partial $\eta^2 = .807$ upon listener perceptions, as measured by the five dependent variables.

To investigate further, participant ratings for all alerts were combined so as to expose the overall differences between the modifications in comparison to the normal, or baseline, version. This is specifically dealt with by performing post-hoc analysis, using the Bonferroni correction, to identify statistically significant differences between the normal version of each alert sound and versions with volume, pitch and Doppler modifications, as shown in Table 3 and the differences are presented graphically in Figure 8.

Table 3. Post-hoc analysis of alerts with respect to participant ratings showing significant modifications when compared to the normal condition.

Dependent variable	**Significant Modifications (Mean Difference and p-value)**
Important/Unimportant	Doppler down ($MD = 0.429, p < 0.001$) Volume down ($MD = 0.431, p < 0.001$)
Safe/Dangerous	Doppler down ($MD = -0.162, p < 0.001$) Doppler up ($MD = 0.138, p = 0.007$) Volume down ($MD = -0.186, p < 0.001$) Volume up ($MD = 0.189, p < 0.001$)
Urgent/Trivial	Doppler down ($MD = 0.355, p < 0.001$) Volume down ($MD = 0.379, p < 0.001$) Volume up ($MD = -0.132, p = 0.006$)
Difficult to detect/Easy to detect	Doppler down ($MD = -0.401, p < 0.001$) Doppler up ($MD = -0.140, p < 0.001$) Pitch down ($MD = -0.102, p = 0.001$) Volume down ($MD = -0.366, p < 0.001$) Volume up ($MD = -0.092, p = 0.003$)
Close/Distant	Doppler down ($MD = 0.552, p < 0.001$) Pitch down ($MD = 0.143, p = 0.004$) Volume down ($MD = 0.508, p < 0.001$)

Overall, when looking at these post-hoc results it is possible to see broad trends in terms of the *Doppler down* (orange) and *Volume down* (green) modifications having the greatest effect upon the participants' interpretation of the clips. Applying a *Doppler down* or *Volume down* modification made alerts seem less important, safer, less urgent, less easy to detect and less close. *Pitch down* had a considerably lesser effect, and only significant for two of the dependent variables. *Doppler up* (grey) was significant for being less safe, and less easy to detect, *Volume up* (blue) was similarly perceived as being less safe, more urgent, and less easy to detect. Use of a *Pitch down* modification had the effect of making alerts seem less easy to detect and less close. The *Pitch up* modification made no significant difference to any of the responses.

To explore the individual alert sounds, one-way MANOVA with repeated measures tests were employed to examine the impact of the alert modifications upon participants' ratings on the five dependent variables. The one-way MANOVA tests revealed statistically significant main effects for the *Emergency Alert System*, $F(30, 70) = 3.927, p < .001$; Wilk's $\Lambda = 0.373$, partial $\eta^2 = 0.627$; *Emergency vehicle*, $F(30, 70) = 4.384, p < .001$; Wilk's $\Lambda = 0.347$, partial $\eta^2 = .653$; *Fire alarm*, $F(30, 70) = 5.216, p < .001$; Wilk's $\Lambda = 0.309$, partial $\eta^2 = .691$; *forward collision warning*, $F(30, 70) = 4.471, p < .001$; Wilk's $\Lambda = 0.343$, partial $\eta^2 = .657$; *Medical equipment alarm*, $F(30, 70) = 6.793, p < .001$; Wilk's $\Lambda = 0.253$, partial $\eta^2 = .747$; *Smoke*

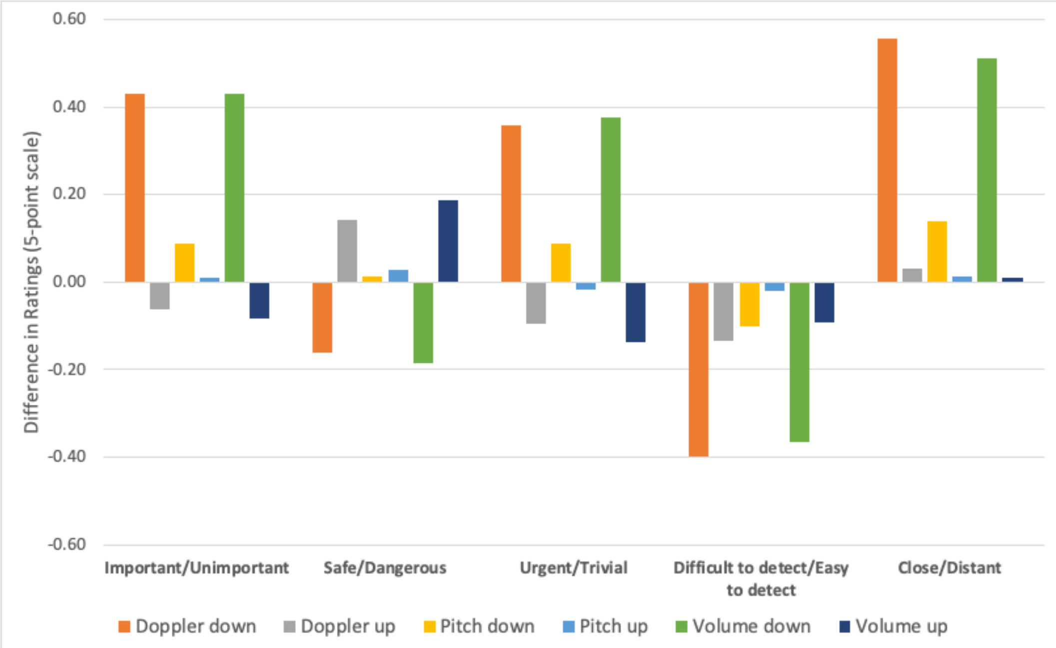


Figure 8. Combined results of modifications compared to their normal (unaltered) versions for all alerts.

alarm, $F(30, 70) = 5.716, p < .001$; Wilk's $\Lambda = 0.290$, partial $\eta^2 = .710$; *SMS alert*, $F(30, 70) = 1.697, p = .036$; Wilk's $\Lambda = 0.579$, partial $\eta^2 = .421$; *Telephone ring*, $F(30, 70) = 3.762, p < .001$; Wilk's $\Lambda = 0.383$, partial $\eta^2 = .617$; and *Truck backup beeps*, $F(30, 70) = 7.519, p < .001$; Wilk's $\Lambda = 0.237$, partial $\eta^2 = .763$. There was no statistically significant main effect for the *Email alert*.

For the auditory alerts that exhibited significant main effects, post-hoc analysis using the Bonferroni correction was performed to determine where significant differences were achieved between the normal version of the alert and the modifications. This analysis is summarised in Table 4.

Considering the significant differences noted in the alert sounds individually (Table 4), the prevalence of the *Doppler down* and *Volume down* modifications is reinforced as influencing the dependent variables and were the modifications most commonly found to influence participant ratings.

Where *Doppler down* and *Volume down* impacted upon the Important/Unimportant variable, it is seen that they had a largely consistent effect of making alerts less important.

The *Doppler down* modification significantly increased safety in the emergency alert system, forward collision warning, smoke alarm, and truck backup beeps alert sounds, but decreased safety in the telephone ring.

Volume down make it more difficult to detect all alert sounds, with the exception of the email alert (no significant differences due to any modification) and the SMS alert. A similar situation is observed in the *Doppler down* modification, noting that it did not impact detection in the forward collision warning and that a *Doppler up* modification had a nearly identical effect to *Doppler down* in the fire alarm alert.

Both *Doppler down* and *Volume down* were consistent in producing an effect of making alerts more trivial, specifically the emergency alert system, forward collision warning, smoke alarm, and truck backup beeps. *Doppler down* was the only modifi-

Table 4. Post-hoc analysis of each auditory alert with respect to participant ratings showing significant modifications when compared to the normal condition.

Auditory Alert	**Dependent Variable - Significant Modifications (Mean Difference and p-value)**
Emergency Alert System	Important/Unimportant - Volume down ($MD = 0.390, p = 0.01$) Safe/Dangerous - Doppler down ($MD = -0.340, p = 0.027$) Safe/Dangerous - Volume down ($MD = -0.380, p = 0.017$) Urgent/Trivial - Doppler down ($MD = 0.350, p = 0.039$) Urgent/Trivial - Volume down ($MD = 0.360, p = 0.048$) Difficult/Easy to Detect - Doppler down ($MD = -0.440, p = 0.002$) Difficult/Easy to Detect - Volume down ($MD = -0.560, p < 0.001$) Close/Distant - Doppler down ($MD = 0.850, p < 0.001$) Close/Distant - Volume down ($MD = 0.880, p < 0.001$)
Emergency vehicle	Difficult/Easy to Detect - Doppler down ($MD = -0.390, p = 0.001$) Difficult/Easy to Detect - Volume down ($MD = -0.340, p < 0.001$) Close/Distant - Doppler down ($MD = 0.860, p < 0.001$) Close/Distant - Volume down ($MD = 0.900, p < 0.001$)
Fire alarm	Important/Unimportant - Doppler down ($MD = 0.420, p = 0.011$) Important/Unimportant - Volume down ($MD = 0.840, p < 0.001$) Urgent/Trivial - Volume down ($MD = 0.580, p < 0.001$) Difficult/Easy to Detect - Doppler down ($MD = -0.480, p = 0.001$) Difficult/Easy to Detect - Doppler up ($MD = -0.480, p < 0.001$) Difficult/Easy to Detect - Volume down ($MD = -0.550, p < 0.001$)
Forward collision warning	Important/Unimportant - Doppler down ($MD = 0.850, p < 0.001$) Important/Unimportant - Volume down ($MD = 0.850, p < 0.001$) Safe/Dangerous - Doppler down ($MD = -0.510, p < 0.001$) Safe/Dangerous - Volume down ($MD = -0.480, p = 0.001$) Urgent/Trivial - Doppler down ($MD = 0.740, p < 0.001$) Urgent/Trivial - Volume down ($MD = 0.860, p = 0.001$) Difficult/Easy to Detect - Volume down ($MD = -0.360, p = 0.018$)
Medical equipment alarm	Important/Unimportant - Doppler down ($MD = 0.515, p = 0.001$) Important/Unimportant - Doppler up ($MD = -0.354, p = 0.039$) Urgent/Trivial - Doppler down ($MD = 0.465, p = 0.004$) Difficult/Easy to Detect - Doppler down ($MD = -0.687, p < 0.001$) Difficult/Easy to Detect - Volume down ($MD = -0.545, p < 0.001$) Close/Distant - Doppler down ($MD = 1.111, p < 0.001$) Close/Distant - Volume down ($MD = 0.899, p < 0.001$)
Smoke alarm	Important/Unimportant - Doppler down ($MD = 0.960, p < 0.001$) Important/Unimportant - Volume down ($MD = 0.680, p < 0.001$) Safe/Dangerous - Doppler down ($MD = -0.780, p < 0.001$) Safe/Dangerous - Volume down ($MD = -0.580, p = 0.001$) Urgent/Trivial - Doppler down ($MD = 0.900, p < 0.001$) Urgent/Trivial - Volume down ($MD = 0.840, p = 0.001$) Difficult/Easy to Detect - Doppler down ($MD = -0.570, p < 0.001$) Difficult/Easy to Detect - Volume down ($MD = -0.440, p < 0.001$) Close/Distant - Doppler down ($MD = 0.470, p = 0.016$)
SMS alert	Important/Unimportant - Pitch down ($MD = 0.430, p = 0.006$) Difficult/Easy to Detect - Doppler down ($MD = -0.300, p = 0.007$) Difficult/Easy to Detect - Pitch down ($MD = -0.310, p < 0.017$)
Telephone ring	Safe/Dangerous - Doppler down ($MD = 0.360, p = 0.006$) Safe/Dangerous - Doppler up ($MD = 0.480, p < 0.001$) Safe/Dangerous - Volume up ($MD = 0.310, p = 0.004$) Urgent/Trivial - Volume up ($MD = -0.260, p = 0.045$) Difficult/Easy to Detect - Doppler down ($MD = -0.310, p = 0.002$) Difficult/Easy to Detect - Volume down ($MD = -0.230, p = 0.021$) Close/Distant - Doppler down ($MD = 0.500, p < 0.001$) Close/Distant - Volume down ($MD = 0.450, p < 0.001$)
Truck backup beeps	Important/Unimportant - Doppler down ($MD = 0.650, p < 0.001$) Important/Unimportant - Volume down ($MD = 0.830, p < 0.001$) Safe/Dangerous - Doppler down ($MD = -0.450, p < 0.001$) Safe/Dangerous - Volume up ($MD = -0.500, p < 0.001$) Urgent/Trivial - Doppler down ($MD = 0.490, p < 0.001$) Urgent/Trivial - Volume down ($MD = 0.620, p < 0.001$) Difficult/Easy to Detect - Doppler down ($MD = -0.430, p = 0.002$) Difficult/Easy to Detect - Volume down ($MD = -0.520, p = 0.001$)

cation to achieve this in the medical equipment alarm and *Volume down* in the fire alarm sounds.

The other modifications appeared only in a small number of alert types as presenting a significant effect. For example, *Pitch down* was successful only in the SMS alert in making it less important and harder to detect. *Volume up* had a contradictory effect on two sounds, as it made the telephone ring more dangerous but the truck backup beeps less dangerous. It also made the telephone ring more urgent. *Doppler up* made the medical equipment alarm more important and increased the safety of the telephone ring.

4.2. Participant Comments

A total of 36 participants provided information in the comments section of the response form. Following an initial inspection of the comments, 33 usable responses were available with three being removed as the participants reported that they had no additional comments to leave.

The comments remaining were subjected to a qualitative analysis following the process of thematic analysis (Braun & Clarke, 2006; Guest, MacQueen, & Namey, 2011). This provided an opportunity to gain an alternate perspective into the experiences and perceptions of the participants with the sounds in the study. This is a natural and suitable approach since the study being reported is largely dependent upon the subjective experience of its participants. It was anticipated that additional, descriptive information about the qualities of the sounds might be highlighted using timbral descriptions, as well as the five-point scale antonyms, as well as reflection upon the sounds and the testing process itself. Thematic analysis and qualitative techniques are found to provide useful insight in audio listening studies (Francombe et al., 2018; Ratcliffe, Gatersleben, & Sowden, 2013) and can be used to complement and triangulate with quantitative data (Cunningham & McGregor, 2019).

The comments provided were read initially to form a broad set of initial themes. During this process, it became clear that the main terms used by participants related to the antonyms used to identify the bipolar points on the rating scales. As such, rather than duplicate and potentially over analyse these terms, high-level themes were established for each of the five scales used in the ratings exercise and the set of comments were appraised again using this scheme, whilst also including any further themes that were encountered.

The complete set of themes, their definition, number of references, and exemplar statements from participants (with participant numbers shown in parentheses) are shown in Table 5.

The scale receiving the most responses was that of Close / Distant. In addition to comments provided about the perception of distance contributing to other qualities, such as urgency, a number of participants noted that they found it a challenging task to respond to this question. This seemed to be a particular issue when the proximity characteristics were compounded with the perception of other qualities changing in the sound, giving explanations such as:

- *Difficult to state whether close or distant when volume occasionally varies from low to high or vice versa* (P82);
- *Further the tests continue, there seems to be a correlation between increasing volume and proximity, increasing pitch and danger, whilst constant pitches and volume indicate stability.* (P36);

Table 5. Summary of Themes from Participant Comments

Theme	Definition	n	Example Response
Close / Distant	Scale used in rating task	10	*A perceived change in distance I felt cause a drop in urgency.* (P18)
Safe / Dangerous	Scale used in rating task	7	*Generally, sounds imitating phone rings are not important or dangerous.* (P61)
Urgent / Trivial	Scale used in rating task	5	*The sound which increased in volume made a significant increase in the urgency.* (P05)
Important / Unimportant	Scale used in rating task	4	*Sound clips which had a 'flatlining' section felt most important each time.* (P46)
Easy / Difficult to Detect	Scale used in rating task	2	*I also wasn't quite sure the parameters between easy + difficult to detect.* (P64)
Testing Process	Issues relating to the listeners' experience of participating in the listening study	15	*I feel that some of the clips were quite similar and therefore concerned that I may not have been consistent in my ratings.* (P83)
Context	Information about the environment, situation or prior knowledge of the listener to place the sound into context	14	*I felt my reactions would strongly depend on a real-life context. Also familiar "false-alarm" sounds influenced my responses.* (P50)
Sound Perception	Perception of the sound including affective response	4	*Some of the sounds with missing pitches were quite abrasive.* (P16)
Attention	Ability of the sound to obtain the attention of the listener	1	*Sounds that changed in volume/tone/intensity or were unfamiliar were quite attention-grabbing.* (P75)

- *For some of the sounds with missing pitches over time, I had different opinions for the close/distant question between the start and the end of the sound* (P16);
- *Found the question regarding distance/proximity a little difficult especially when 'alarm' was getting louder/quieter... Interesting!* (P81).

Several of the comments received in the remaining themes derived from the scales in the test overlapped with comments that were provided by participants in the *Context* and *Testing Process* and themes and thus they are discussed in a holistic manner here.

To deal with the mixture of the responses as they related to the theme of *Context*, in this subset of participant responses there was an overwhelming sense that the ratings provided in several of the scales would be dependent upon what else was happening around the listener at the time that they heard the sound. It was also highlighted that this may, or may not, also be dependent upon prior knowledge or experience that the participants had with the sounds. This is best exemplified by a number of the responses provided:

- *One last observation is that the level of urgency, importance and danger vary on the listener's viewpoint. If for instance the alarm was meant for a specific action or not.* (P02);
- *Some of these are subjective, as if you recognise the sounds you can relate them to specific things (e.g. one sounded like 'fully charged iPhone' so I related to it as that.* (P37);
- *It was difficult to not think about the "CONTEXT" around the sound, instead of the sound itself. For example: Why is the phone ringing?* (P40);
- *Conditioned by knowledge that certain sounds are normally used for general announcements, modem failures etc. Siren sounds are always associated with urgency, but not necessarily for you.* (P51);
- *Some sounds were part of my life so I used my usual reaction to them in my answers. Some were familiar but I could not remember what for. One was unknown so I gauged on my impressions.* (P54);

- *I found myself doubting whether SAFE meant did I feel safe when I heard it, in that a police siren can reasure [sic] you because you know it might mean help is at hand - or is it dangerous because it is on its way to potentially nasty incident.* (P64).;
- *Found myself considering the social/cultural context which I relate to the sounds.* (P92).

Relating the sounds to the concept of a scenario emerged as a notable element that could be considered in later work and clearly one that participants were seeking to help them to provide as appropriate a ratings as possible for each sound version that had been presented.

In terms of the *Testing Process* theme, a number of notable points were raised, which were generally oriented around the concepts of listening testing methodology, reducing listener fatigue, greater explanation about the meaning of the scales, and consistency in performance during the test itself. The broad issues relating to testing methodology are best exemplified by the following comments:

- *As a participant I have been struggling with certain sounds, in specific at identifying what/where they have been generated from.* (P02);
- *For future testing possibly have a couple of pretest sounds to help establish persons sense of close/distant perspective.* (P20);
- *Listener fatigue sets in towards the lat 20 clips or so as well.* (P36);
- *Gets fatiguing towards the end (last quarter)* (P38);
- *I found it really interesting and surprisingly difficult, since some tracks were almost the same but I expereienced [sic] different feelings.* (P95).

Finally, as a general outcome related to the testing process, some participants reported that they had concerns about their ability to perform consistently. Reasoning for this was sometimes provided, although not in all cases. However, given the previous set of comment that highlighted issues around context and listener fatigue, it may be reasonable to assume that these are factors that played a part. The most notable comments received in terms of listener consistency included:

- *Mostly I was torn between whether I had answered consistently.* (P64);
- *I feel that some of the clips were quite similar and therefore concerned that I may not have been consistent in my ratings.* (P83);
- *I found myself analysing what I thought the sounds were. And my response was as much about the meaning of the sound, and what I associated it with, as the volume or 'urgency' of the sound. Not sure how consistent I was...* (P85);
- *I think I may have switched my assessment of whether a sound was "safe" or "dangerous" midway through because I started thinking about whether I would consider myself safer after hearing a warning sound like a police siren or if I would consider myself in danger because said siren means something not great is going on if I need the police to rescue me.* (P98).

Overall, these responses have provided useful additional context to the quantitative results. Some of the most useful aspects highlighted are in relation to the listening test itself and some of the challenges and issues that participants reported facing.

For example, drawing especially upon the theme of context, the issue of meaning or association of the sounds with objects or events used was touched upon by several participants. It was clear that a number of participants were drawing upon their past experiences of such alert sounds and that these were influencing their perception of some of the characteristics being measured. For example: *"Generally, sounds imitating phone rings are not important or dangerous"* (P61). Whilst the experimental controls adopted for our study deliberately removed a situational context, disassociating past experiences of sound alerts is harder to achieve, likely resulting in the sounds not being evaluated exclusively on their acoustic properties, except, of course, in the case of participants hearing a sound for the first time.

Highlighting the issues around providing meaningful context and greater clarity about the factors being evaluated will be extremely useful in further research activities and in the design of more experiments, which is discussed further in the next section of this article.

5. Conclusions and Future Work

One, perhaps not altogether surprising result from this study is in terms of dynamics. It is commonly believed that loudness affects perceived urgency, as does speed and pitch (Haas & Edworthy, 1996), whereas the clips in this study with the loudest peaks in terms of SPL were not necessarily reported as the most urgent. Both *Doppler down* and *Volume down* had higher mean peak values of 87.2 and 86.9 dBA respectively than the *Normal* (77.4 dBA), *Doppler up* (82.6 dBA) and *Volume up* (81.6 dBA) clips, the latter three all having a higher perceived urgency, despite all having an identical RMS of 60.7 dBA. Neuhoff (2001), writes about sounds which appear to be coming towards a listener potentially being perceived as louder than those receding.

Patterson (1990) is often cited for stating that auditory alerts with fast attacks are generally perceived as more dangerous than those with slower attack, and there is certainly a general convention for this approach when urgency is desired to be communicated. In this study *Volume up* had the greatest effect on perceived danger followed by *Doppler up*, both of which had much slower attacks than *Volume down* and *Doppler down*, which were reported as having the highest level of safety, with fast attacks.

Pitch changes on their own had comparatively little effect on the alerts, especially with regards to *Pitch up*. *Pitch down* had slightly more impact, but not as much as *Doppler down* or *Volume down*. Higher pitches are considered to be more effective for auditory alerts, and this can be seen partially by the small reduction in safety and importance when the pitch falls. However, in this study there was no evidence that increasing the pitch increased the importance or danger, which may be due to the pitches of the original auditory alerts already being within the most effective frequencies for their purpose.

This study has suggested that DSP could be used for current alert sounds to alter their perceived importance, safety, and urgency, but there is an effect upon their ease of detection and perceived proximity. The major advantage is that existing cues could be augmented without end-users having to learn new meanings, as the Doppler effect and the associated pitch and volume changes are inherent in everyday life, making them easy to understand. There are two immediate scenarios, for fire alarms and medical alerts, although other applications might also be suitable.

In large buildings there can be multiple routes to fire exits, and signage can be confusing for those not familiar with the layout, especially if there are smoke or flames. There is almost always a legal requirement to have multiple heat and smoke sensors located throughout a building, and these can have built-in sounders or sirens. Three distinct states could be used to indicate the level of danger, and therefore provide guidance about which route to take through a building with multiple routes of egress. If a sensor is detecting smoke or excessive heat processing could be applied to raise pitch, volume or both to indicate aurally that the alarm represents an active area of danger. This would allow inhabitants to choose a safer route, should one be available. If a fire alarm is being triggered centrally the amplitude and pitch could remain as they are. Alarms located next to fire exits, where there is no smoke or excessive heat could have *Pitch down*, *Volume down* or both applied.

In hospitals it is common practice to group patients with similar conditions together in wards, or adjacent rooms. This leads to similar if not identical machinery to be used to monitor patients, which can lead to confusion with medical practitioners needing to quickly identify which patient needs assistance. Both higher and lower pitched sounds are generally easier to spatially locate than those in the mid-range due to interaural phase differences (IPD) between approximately 80 - 770 Hz, as well as interaural level differences (ILD) above 1.6 kHz, which are believed to be the most effective for lateral orientation (Blauert, 1997). The small pitch changes associated with the Doppler effect might provide additional spatial cues to aid quick sound source identification. The variation in amplitude could also be helpful, in that with current alerts the highest setting often masks other sound sources. In this scenario there would be a cross over as the highest state would cycle through from silent and ramp up to maximum volume, whereas the lowest state would start off at maximum volume and be reduced to silence. This would help ensure that each state, High, Mid or Low would all potentially be clearly audible at some point during the cycles. This would also help assist with the issue of medical practitioners not wanting any new sounds, considering that the sonic environment in hospitals is already over-saturated (Edworthy et al., 2017).

The next stage of the research is to establish the extent of any pitch or volume changes required as more subtle or more extreme modifications might be better suited. For this experiment settings were based upon 100% representation of the pitch and amplitude changes which would occur naturally at moderate temperatures at sea level if each sound source were either to move from a standing start or to a complete stop for the duration of the clip. The standard Waves plugin can range from 0 – 200% for each parameter, and DSP could easily be developed to extend this range, should it prove effective. A simple experiment where auditory cues are played concurrently on spatially separated speakers could be used to query participants about which cue was the most important or dangerous. This approach was not taken for the initial study in order to explore whether or not there was an effect to explore further.

The issue of ecological validity, scenario, and the context where sounds are played was highlighted strongly by participants in the qualitative responses and subsequent analysis provided. Presenting the sound stimuli to the participants without context and in a highly controlled environment was a deliberate and purposeful choice on our part at this stage in the research, so as to eliminate any possible distractions and masking effects upon the stimuli themselves. However, as our qualitative analysis has highlighted, participants were often relying upon this information to make sense of the sounds being presented, as exemplified by the field of semiotics (Jekosch, 1999; Ostwald, 2019). On this point, an important finding from research into the monitoring sounds of helicopters (Edworthy et al., 2004) was that researchers should be aware of how participants in experiments may perceive sounds both acoustically and in terms of the perceived meaning given the situation or context that the listener finds themselves in. This could be especially true where there is potential for direct mapping of underlying parameters in the phenomenon to acoustic attributes to be perceived or real. For example, an audible electrocardiogram directly sonifies human heartbeats by adjusting the interval between pulses, but both extremely rapid pulses and extremely distanced pulses would cause a clinician to identify these conditions as requiring urgent attention.

A natural next stage, therefore, in the expansion of the findings presented here, would be to place the sounds created in this study into situational contexts or scenarios that could be simulated in some way and potentially providing participants with a pretext and primer on the sounds and objects to be encountered. Such an approach could be accompanied by requesting participants to report their initial perceptions or experiences with the sounds under investigation, prior to the experiment commencing. This could then be done simply, by mixing these sounds with a background of ambient recordings from a set of environments or expanded into a small computer game, requiring participants to make decisions based upon the sound stimuli presented. The use of a game-based approach in particular would provide the opportunity for metrics to be recorded, such as reaction time and visual attention duration, that do not rely upon participants having to self-report these features using a scale. Following the feedback from participants with regards to the listening test itself, and being mindful of comments relating to listener fatigue, it may be prudent to reduce the number of sounds being evaluated in each experimental session to provide more focused tasks. Finally, 3D computer game activities would also permit the simulation of spatial audio processing, allowing the sounds to be presented in an immersive soundscape, closely mimicking that of the real-world, whilst also providing the affordances of a controlled, laboratory-based study.

References

Bach, D. R., Schächinger, H., Neuhoff, J. G., Esposito, F., Salle, F. D., Lehmann, C., ... Seifritz, E. (2008). Rising sound intensity: an intrinsic warning cue activating the amygdala. *Cerebral Cortex*, *18*(1), 145–150.

Begault, D. R., Anderson, M. R., & McClain, B. U. (2003). Spatially-modulated auditory alerts..

Blauert, J. (1997). *Spatial hearing: the psychophysics of human sound localization.* MIT press.

Bolton, M. L., Zheng, X., Li, M., Edworthy, J. R., & Boyd, A. D. (2019). An experimental validation of masking in iec 60601-1-8: 2006-compliant alarm sounds. *Human factors*, 0018720819862911.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77–101.

Cabrera, D., Ferguson, S., & Laing, G. (2005). Development of auditory alerts for air traffic control consoles. In *Audio engineering society convention 119.*

Cannon, W. B. (1916). *Bodily changes in pain, hunger, fear, and rage: An account of recent researches into the function of emotional excitement.* D. Appleton.

Causse, M., Imbert, J.-P., Giraudet, L., Jouffrais, C., & Tremblay, S. (2016). The role of cognitive and perceptual loads in inattentional deafness. *Frontiers in human neuroscience*, *10*, 344.

Chamard, J.-C., & Roussarie, V. (2012). Design of electric or hybrid vehicle alert sound system for pedestrian..

Cunningham, S., & McGregor, I. (2019). Subjective evaluation of music compressed with the acer codec compared to aac, mp3, and uncompressed pcm. *International Journal of Digital Multimedia Broadcasting*, *2019*.

Doppler, C., & Studnica, F. J. (1903). Ueber das farbige licht der doppelsterne und einiger anderer gestirne des himmels. *Prag, K. Bohm gesellschaft der wissenschaften, 1903..*

Edworthy, J. (1994). The design and implementation of non-verbal auditory warnings. *Applied ergonomics*, *25*(4), 202–210.

Edworthy, J., Hellier, E., Aldrich, K., & Loxley, S. (2004). Designing trend-monitoring sounds for helicopters: methodological issues and an application. *Journal of Experimental Psychology: Applied*, *10*(4), 203.

Edworthy, J., Reid, S., McDougall, S., Edworthy, J., Hall, S., Bennett, D., ... Pye, E. (2017). The recognizability and localizability of auditory alarms: Setting global medical device standards. *Human factors*, *59*(7), 1108–1127.

Fagerlönn, J. (2011). Urgent alarms in trucks: effects on annoyance and subsequent driving performance. *IET Intelligent Transport Systems*, *5*(4), 252–258.

Fletcher, H., & Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, *12*(4), 377–430.

Francombe, J., Woodcock, J., Hughes, R. J., Mason, R., Franck, A., Pike, C. R., ... others (2018). Qualitative evaluation of media device orchestration for immersive spatial audio reproduction. *Journal of the Audio Engineering Society*, *66*(6), 414–429.

Guest, G., MacQueen, K. M., & Namey, E. E. (2011). *Applied thematic analysis.* Sage Publications.

Haas, E. C., & Edworthy, J. (1996). Designing urgency into auditory warnings using pitch, speed and loudness. *Computing & Control Engineering Journal*, *7*(4), 193–198.

Han, M. U., & Lee, S. K. (2017). Detectability evaluation for alert sound in an electric vehicle. *Transactions of the Korean Society of Mechanical Engineers A*, *41*(10), 923–929.

Hasanain, B., Boyd, A. D., Edworthy, J., & Bolton, M. L. (2017). A formal approach to discovering simultaneous additive masking between auditory medical alarms. *Applied ergonomics*, *58*, 500–514.

Hellier, E. J., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human factors*, *35*(4), 693–706.

Huey, R. W., Buckley, D. S., & Lerner, N. D. (1996). Audible performance of smoke alarm sounds. *International journal of industrial ergonomics*, *18*(1), 61–69.

ITU-R, R. (2015). Bs. 1116-3,". *Methods for the subjective assesment of small impairments in audio systems," International Telecommunication Union-Radiocommunication Sector*.

ITU-R, R. (2019). Bs.1284-2,". *General methods for the subjective assessment of sound quality," International Telecommunication Union-Radiocommunication Sector*.

Jekosch, U. (1999). Meaning in the context of sound quality assessment. *Acta Acustica united with acustica*, *85*(5), 681–684.

Kim, D. S., Emerson, R. W., Naghshineh, K., Pliskow, J., & Myers, K. (2012). Vehicle surge detection and pathway discrimination by pedestrians who are blind: Effect of adding an alert sound to hybrid electric vehicles on performance. *British journal of visual impairment*, *30*(2), 61–78.

Ltd., W. A. (2020). *Doppler effect plugin.* Retrieved from `https://www.waves.com/plugins/doppler`

Melara, R. D., & Marks, L. E. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & psychophysics*, *48*(2), 169–178.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, *13*(2), 87–110.

Neuhoff, J. G., & McBeath, M. K. (1996). The doppler illusion: The influence of dynamic intensity change on perceived pitch. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 970.

Oechslin, M., Neukom, M., & Bennett, G. (2008). The doppler effect—an evolutionary critical cue for the perception of the direction of moving sound sources. In *2008 international conference on audio, language and image processing* (pp. 676–679).

Orchik, D. J., Schumaier, D. R., Shea, J. J., & Moretz Jr, W. H. (1985). Intensity and frequency of sound levels from cordless telephones: A pediatric alert. *Clinical pediatrics*, *24*(12), 686–690.

Ostwald, P. F. (2019). *The semiotics of human sound* (Vol. 36). Walter de Gruyter GmbH & Co KG.

Patterson, R. D. (1990). Auditory warning sounds in the work environment. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *327*(1241), 485–492.

Počta, P., & Beerends, J. G. (2015). Subjective and objective assessment of perceived audio quality of current digital audio broadcasting systems and web-casting applications. *IEEE Transactions on Broadcasting*, *61*(3), 407–415.

Ratcliffe, E., Gatersleben, B., & Sowden, P. T. (2013). Bird sounds and their contributions to perceived attention restoration and stress recovery. *Journal of Environmental Psychology*, *36*, 221–228.

Reynolds, M. E., Rayo, M. F., Fitzgerald, M., Abdel-Rasoul, M., & Moffatt-Bruce, S. D. (2019). Guidance for custom alarm design: You don't have to "crank it up to 11". In *Proceedings of the international symposium on human factors and ergonomics in health care* (Vol. 8, pp. 142–145).

Sinclair, P. (2012). Living with alarms: the audio environment in an intensive care unit. *AI & society*, *27*(2), 269–276.

Singer, J., Lerner, N., Baldwin, C., & Traube, E. (2015). Auditory alerts in vehicles: Effects of alert characteristics and ambient noise conditions on perceived meaning and detectability. In *Conference on the* (Vol. 1, pp. 1–14).

Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., & Västfjäll, D. (2010). Embodied auditory perception: The emotional impact of approaching and receding sound sources. *Emotion*, *10*(2), 216.

Villegas, J., Stegenborg-Andersen, T., Zacharov, N., & Ramsgaard, J. (2016). Effect of presentation method modifications on standardized listening tests. In *Audio engineering society convention 141*.

Wille, M., Theis, S., Rasche, P., Bröhl, C., Kummer, R., & Mertens, A. (2017). What does beep mean? : context free interpretation of short sinus wave stimuli. In *Proceedings of the human factors and ergonomics society europe chapter 2016 annual conference : Human factors and user needs in transport, control, and the workplace / edited by dick de waard, antonella toffetti, rebecca wiczorek, andreas sonderegger, stefan röttger, petr bouchner, thomas franke, stephen fairclough, matthijs noordzij, and karel brookhuis* (p. 173-184). Retrieved from `https://publications.rwth-aachen.de/record/726441`

Yasui, N. (2018). Effect of amplitude fluctuation on detectability of alert sound for electric and hybrid vehicle in an actual environment. In *Inter-noise and noise-con congress and conference proceedings* (Vol. 258, pp. 2217–2225).

Yasui, N. (2019). Subjective evaluations of detectability of alert sound for electric and hybrid electric vehicle under actual environment. In *Inter-noise and noise-con congress and conference proceedings* (Vol. 259, pp. 6357–6364).

6. Author Biographies

Dr Stuart Cunningham is Senior Lecturer at Manchester Metropolitan University (UK). He holds BSc and MSc awards from the University of Paisley (UK) and a PhD from the University of Wales (UK). His research interests are in sonic interaction, affective computing, and digital media technologies to improve wellbeing.

Dr Iain McGregor is the Programme leader for both the undergraduate and online postgraduate Sound Design degrees at Edinburgh Napier University. He specialises in listening studies and runs the University's auralisation suite. Current research projects include Auditory interfaces, Foley, Human Robot Interfaces, Internet of Things, Sonification and Video Game Audio.