Investigating Perceptual Congruence Between Data and Display Dimensions in Sonification

Jamie Ferguson & Stephen A. Brewster
Glasgow Interactive Systems Section
University of Glasgow
Glasgow, Scotland
j.ferguson.4@research.gla.ac.uk
stephen.brewster@glasgow.ac.uk
mig.dcs.gla.ac.uk

ABSTRACT
The relationships between sounds and their perceived meaning and connotations are complex, making auditory perception an important factor to consider when designing sonification systems. Listeners often have a mental model of how a data variable should sound during sonification and this model is not considered in most data:sound mappings. This can lead to mappings that are difficult to use and can cause confusion. To investigate this issue, we conducted a magnitude estimation experiment to map how roughness, noise and pitch relate to the perceived magnitude of stress, error and danger. These parameters were chosen due to previous findings which suggest perceptual congruency between these auditory sensations and conceptual variables. Results from this experiment show that polarity and scaling preference are dependent on the data:sound mapping. This work provides polarity and scaling values that may be directly utilised by sonification designers to improve auditory displays in areas such as accessible and mobile computing, process-monitoring and biofeedback.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation: User Interfaces: Auditory (non-speech) Feedback.

Author Keywords
Sonification, Auditory Display, Mental Models, Psychoacoustics.

INTRODUCTION
Sonification - "the use of non-speech audio to convey information" [28] - is an effective interaction modality in a wide range of contexts: for people with visual impairments [25], for data analysis [11], real-time monitoring [37] or interaction with devices that have little screen space [6]. In a parameter-mapping sonification system, data values are used to manipulate an acoustic parameter, such as frequency (pitch) or tempo, which facilitate the communication of the data value [20]. Many of these data:sound mappings fail to account for the listener's mental model of how the data value should sound during sonification - a mental model being defined as "a representation of some domain or situation that supports understanding, reasoning, and prediction" [17]. Instead of attempting to map data variables to acoustic parameters in a way that complements these mental models, designers commonly use well known auditory parameters such as pitch, loudness or panning [12]. In defaulting to these common auditory parameters, the opportunity to present a mapping that is complementary to the listener's expectations of the acoustic qualities of the data value being sonified is lost. This can result in reduced performance [43] and may lead to confusion, annoyance and fatigue, due to the mapping conflicting with the listener's mental model of how the data value should sound [45].

Pitch is unequivocally the most used acoustic parameter in sonification mappings [12] and has even been described as the "Hello World" of sonification [19]. This is to be expected, as pitch is one of the most salient characteristics of a musical sound and sonification designers commonly employ musical structures and aesthetics [20]. Regardless of the positive attributes of pitch as an acoustic parameter, there is a strong need to expand the parameter space for sonification. Each acoustic parameter is single-use - once pitch has been mapped to a data parameter it cannot be used again. The range of acoustic parameters available to sonification designers is limited, and only a small subset of these parameters has been empirically evaluated in a sonification context.

As stated by Walker [38], "the most successful representation of conceptual data depends on the most appropriate display dimension being used, and in the right way". It has been found that at a perceptual level, there is a measurable reduction in stimulus-response compatibility (how natural a response feels based on a stimulus [26]) when a poorly designed pitch display is used [40]. More dangerously, aircraft pilots and nuclear power plant controllers have turned off audible warning signals due to the sound design being unpleasant, or the information the alarm conveys being inaccurate [33].

ACM ISBN .
DOI:

Exploring acoustic parameters beyond the commonly used ones such as pitch, tempo, loudness, or panning, is a potential solution to the problem of perceptual congruency in sonification (i.e. what the listener perceives is complimentary with how they expect a data value to sound when it is sonified). *Psychoacoustic parameters* such as *fluctuation strength*, *roughness*, *sharpness* and *loudness* [13] are a reliable method of understanding the subjective qualities of sound - so much so that they have been found to show a higher correlation with human perception than physiological measurements [9]. These parameters are responsible for a substantial part of the affect a sound has on a listener - combined, they can model the relative degree of noise annoyance [10], and roughness independently has been found to be a primary component in both natural and artificial alarm sounds [2]. Previous work found that data:sound mappings in which the acoustic parameters were based on psychoacoustic parameters were effective in the context of an astronomical image quality assessment task [14]. However, this work was limited in scope, focusing on a single use-case - there is no research that provides a number of different contexts in which psychoacoustic parameters may be more effective than acoustic parameters traditionally used in sonification such as pitch and tempo.

There is little theory or evidence to guide sonification designers in the most effective data:sound mapping to use in particular contexts and there is even less research on evaluating the perceptual reactions of listeners to acoustic parameters in sonification, beyond pitch, tempo and loudness [38, 39, 43]. In this work, we utilise proven relationships between a number of psychoacoustic parameters and conceptual variables, such as the relationship between perceived danger and auditory roughness [2], and propose new ones in an attempt to expand the number of effective sonification parameters and to design data:sound mappings that are complementary to the listener's mental model of how a data variable should sound during sonification. The experiment described here uses magnitude estimation to map how the magnitudes of acoustic parameters relate to the perceived magnitudes of data parameters. Magnitude estimation has been shown as an effective method of establishing if data:sound mappings are conflicting with how the listener expects the data variable to sound like [38].

RELATED WORK

Since sonification has been recognised as a field of research in its own right, there have been a number of attempts to establish guidelines for effective sonification design. Early in the development of the field, Kramer [27] discussed some general principles for designing auditory displays. In a more extensive work, Barrass [3] proposed a number of "golden rules" for sonification design. Based on these two works, Anderson [1] identified some of the crucial elements that future research in this area must address - the mapping of data to parameters being one of these. These works provide a thorough view of what is required to provide a framework for sonification designers, however they are all theoretical and there remains little experimental evidence on which to base sonification design choices.

Perceptual Factors in Sonification Mappings

In Kramer's original principles for sonification design [27], a number of perceptual factors were established that may be practically implemented in a sonification system. Among these factors were so-called *affective associations* and *metaphorical associations*. These notions began to explore how important human perception is in sonification design, especially affective associations, which were described as "the association of feelings about data (if such feelings exist) with feelings aroused by changes in the sound". An example of such an association was given in the context of an ecologist: to such a researcher, data indicating a decrease in rain acidity would generally be described as undesirable, and therefore may cause a subtle negative affect. Therefore, a sonification mapping that may utilise this affect could be described as: an increase in "auditory ugliness" = an increase in an undesirable data variable. Kramer's affective associations have not been empirically tested, however the data:sound mappings described in this paper were motivated by the same inference of the importance of perceptual factors in sonification mapping design.

Although there are a number of previous studies investigating the information transmission capacity and accuracy of various data:sound pairings [5, 32], Walker & Kramer's study (originally presented in 1996 [41], published in 2005 [43]) was the first to directly investigate the importance of perceptual congruence between data and sound parameters in sonification. They investigated the use of a number of commonly used acoustic parameters (pitch, onset, loudness and tempo) to convey common data variables (temperature, pressure, size and rate) in a simple process-monitoring task. The mappings were grouped into four ensembles: *Intuitive*, *Okay*, *Bad* and *Random* based on how "intuitive" or "natural" the sound designers believed the mapping to be. They found the mappings that the sound designers believed to be optimal, e.g. temperature:pitch, did not result in either the most accurate responses or the fastest response time - in fact, the mappings in the *Bad* ensemble led to the fastest response time and *Random* led to the best performance. Some *post hoc* explanations were posited for these results that clearly indicated listener's mental models of data:sound relationships were key. Slower acoustic changes (such as slow tempos and onset) were representative of larger objects (i.e. due to inertia) and in some cases, the listener's mental model of how a data value "ought to sound" was reflected in their polarity choice - for example, increasing size was best represented by decreasing pitch.

Both of these possible explanations for this study's surprising results indicate that the listener's mental model of how they expect a data value should be sonically represented is key to its success. This study clearly shows the need to empirically test data:sound mappings, however this only answers part of the question. This study only investigated four parameters and they were chosen as they are common in sonification. There remains a number of parameters outside of the common set that need to be explored - many of which may provide more perceptually congruent data:sound mappings, and would increase the palette of parameters usable by sonification designers.

Magnitude Estimation

Walker expanded their earlier work [43] to further investigate the psychology relating to data:sound mappings [38]. In this research, magnitude estimation was validated as a tool for aiding the design of data:sound mappings - by measuring the perceived magnitude of a conceptual data dimension (e.g. temperature) when displayed by an auditory dimension (e.g. pitch). Another follow-up study replicated the experiment with a number of additional data:sound mappings [39].

Magnitude estimation is a standard method of psychophysical scaling, which maps the relationship between the magnitude of a sensory stimulus and its associated perceived intensity [36], resulting in a power function between the actual stimulus magnitude and the perceived magnitude. Today, magnitude estimation is a frequently used tool in studying the perception of hearing in a wide array of contexts including ergonomics [22] and medicine [24]. Walker's results [38, 39] showed that magnitude estimation provides reliable measures for both the scaling function (what increase is perceived in temperature for a 100Hz increase in pitch?) and polarity (is increasing pitch perceived as increasing temperature?) for a data:sound mapping. Based on these measures, it is therefore possible to determine if a mapping is suitably congruent with a listener's mental model of the data:sound relationship. Walker used the polarity as a predictor of the "naturalness of a mapping" - if a given polarity obtained a majority of all responses by participants in a block it was predicted to be a "good" polarity choice and it can therefore be predicted that the mapping itself is effective. We therefore used this method for the study presented in this paper.

Psychoacoustics and Sonification Parameters

Traditional psychoacoustics, the study of "the relations between sound stimuli and auditory perception" [49] , has provided a solid grounding for sonification research by contributing experimental methodologies, vocabulary and empirical insights into the perception of a number of fundamental auditory parameters such as pitch, loudness and roughness. Walker & Kramer's chapter in *Ecological Psychoacoustics* [42] provides a thorough overview of how work from the history of psychoacoustics relates to sonification.

Ferguson *et al.* [15] proposed the use of psychoacoustic parameters for sonification and discussed potential benefits, such as how they "may help in rendering displays that auditoraly describe their data intuitively and implicitly", however no experimental assessment was conducted. Neuhoff and Heller [30] proposed a method of auditory display that used data:sound mappings based on a "pre-existing cognitive structure in which to interpret the sonified changes in a variable". They suggested the use of real-world sound sources and events such as the sound of a car engine as acoustic parameters, as opposed to "low-level" acoustic dimensions such as frequency or intensity.

In a previous study, we successfully used psychoacoustic parameters to display the degree of focus of an astronomical image [14]. In this study, the acoustic parameters used in a sonification were the psychoacoustic sensation of *roughness* - the subjective perception of amplitude or frequency modulation of a sound; a signal that was broadband noise at 0 % and a pure clean tone at 100 % - this utilised a potential association of noise with uncertainty and a clear tone with clarity; and a redundant combination of roughness and noise. The results from this study indicated that the combination of these two parameters and the noise/clean signal independently performed at a similar level to the same task being carried out visually - suggesting that the listener's mental model of these data:sound mappings has some similarities to the visual equivalent - in this case, the blur of an image being related to the roughness or noise content of a sound.

Mental Models of Psychoacoustic Parameters

As shown in [14, 38, 43, 45], the mental models or "expectancies" a listener has of how a data value should sound during sonification is an important factor in the effectiveness of the auditory display. Psychoacoustic parameters such as roughness are a strong indicator of a listener's perceived quality of a sound [49]. Such an observed link between an acoustic attribute(s) and a variable, such as roughness and quality in this case, may be utilised to convey data parameters that fit semantically with the variable, for example roughness conveying uncertainty. Furthermore, roughness has been shown to be an integral component of how humans perceive danger [2], therefore it can be suggested that there a listener may expect a data variable relating to danger to sound rough.

Likewise, the effectiveness of the noise:focus mapping in [14] suggests that listeners expect a blurry image to sound noisy when it is sonified, whereas a crisp, focused image should sound clear. Based on these two examples, it is possible that these auditory parameters could convey semantically similar variables effectively, for example stress or error instead of danger, and quality or correctness instead of blur.

EXPERIMENT

Motivated by this literature, an experiment was conducted to investigate similar potential mental models between a conceptual data variable and an acoustic parameter used to convey it during sonification. In this experiment, magnitude estimation was used to evaluate the perception of the magnitude of three data variables: *stress*, *error* and *danger* based on four auditory parameters: *roughness*, *noise*, *combined roughness & noise* and *pitch*.

These data variables were chosen both on the basis of the work discussed in the previous section and because they are semantically broad, allowing for various use-cases to utilise this study's findings. Such scenarios could include:

- Biofeedback of physiological measures to reduce stress/anxiety such as the *Sonic Respiration* system developed by Harris *et al.* [18].
- Sonifying error values in non-visual graph presentation, such as in the *TableVis* system developed by Kildal & Brewster [25].
- Monitoring or alarm systems where multiple levels of danger, urgency or severity must be conveyed.

The experiment investigated how the perceived magnitude of the data variable changed as the magnitude of the auditory stimuli changed and in which polarities the participants perceived this mapping to be (i.e. a positive mapping being increasing pitch = increasing danger). Polarity information is a useful indication of when a data:sound mapping is conflicting with the participants' mental models [38]: agreement in the polarities between participants suggests correlation between their mental models and the data:sound mapping. Furthermore, magnitude estimation produces the slope values between each data and acoustic dimension, allowing future designers to scale their mappings appropriately.

Participants

Sixteen participants took part in the study (11 female, 5 male; mean age = 25 years, SD = 5.1 years). All were university students and staff. All participants reported no uncorrected vision impairment, no hearing impairment and no music/sound related neurological condition such as amusia [31].

Design

We conducted a magnitude estimation experiment, based on the design originally suggested by Stevens [34] and further described by Marks & Gescheider [29]. Applying this design to a sonification context was based on previous experiments by Walker [38] and Walker et al. [44].

Twelve conditions were examined in which the single independent variable in each condition was the acoustic parameter used to convey the data variables: error, danger and stress. Roughness, noise, and a combination of both were chosen as acoustic parameters. Pitch was also included to compare to a more traditional sonification mapping and also to investigate the effect of musicality. Most participants will be familiar with changes in pitch from listening to music, therefore we were interested in the effect on the polarities perceived with pitch compared to the other parameters that are non-musical.

Error, danger and stress were chosen as data variables, due to previous studies' suggestions that the perception of danger and error can be dependent on acoustic qualities [2, 14] and that they are a general set of potential use cases that are semantically similar - all of these are generally considered negative variables, as are the acoustic parameters of roughness and noise. The experiment used a within-subjects design. Each condition was a pairing of a data variable and acoustic parameter, dependent measures collected were the participants estimations of the magnitude of the data variable.

Stimuli

Roughness, noise, and a combination of both were chosen as acoustic parameters, due to the affect of roughness on the perception of danger [2] and noise/combination of roughness and noise due to their affect on the perception of image focus [14] (we suggest that error may be used as a more general variable than image focus).

Ten stimuli were used for each condition, each stimulus was 2 seconds in length. Brewster [7] showed that information encoded in sounds can be obtained from a stimulus between 1 and 2 seconds. Each stimulus had an amplitude envelope with a 0.2 second linear ramp onset (attack) and offset (release). An amplitude envelope was included in the sound design, as an abrupt start or stop of a sound can be perceived as unpleasant [4]. All stimuli were created in the Supercollider programming language [1]. The roughness, noise and combined sounds were identical to the ones used in our previous study [14], as we wanted to utilise the effect between these sounds and the perception of image focus found in that study, to investigate a similar effect for the perception of error. Pitch stimuli were based on the frequencies used by Walker [38].

Roughness

100 % sinusoidally amplitude modulated 1000 Hz pure-tone with modulation frequencies of 0, 2, 4, 7, 11, 16, 23, 34, 49 and 70 Hz.

Noise

This condition consisted of a 1000 Hz pure tone and broadband white noise, with the pure tone beginning at 100 % amplitude and noise at 0 %, respectively declining and ascending in 10 % increments.

Combined Roughness & Noise

Direct pairing of corresponding roughness and noise stimuli. Adjusting more than one auditory parameter simultaneously has been found to increase information transmission capacity [32] and have a greater effect on perceived urgency than independent parameters. Furthermore, the combined noise and roughness condition used in [14] showed slightly increased performance over the individual parameters, therefore we included a combined condition in this experiment.

Pitch

Pure tone with frequencies of 100, 200, 400, 800, 1000, 1400, 1800, 2400 and 3200 Hz. Loudness for each stimulus was normalised using Supercollider's built in library for basic psychoacoustic amplitude compensation: AmpComp.

Procedure

Before the first block of trials, the experimenter read allowed the following text (adapted from [38] and [29]) and the participant read along on-screen.

> *You will hear a series of sounds, one at a time, in random order. Your task is to indicate what magnitude of the variable the sounds seem to represent, by assigning numbers to them. For the first sound, assign it any number of your choosing that represents a value of the variable (e.g. stress). Then, for each of the remaining sounds, estimate its "stress", relative to your subjective impression of the first sound.*
>
> *For example in the case of stress as the variable, if the second sound seems to represent a stress level that is 10 times as "stressful" as the first, then assign it a number that is 10 times bigger than the first number. If the sound seems to represent a stress level that is one-fifth as "stressful", assign it a number that is one-fifth as large as the first number, and so on. You can use any range of numbers, fractions, or decimals that seem appropriate, so long as they are greater than zero.*

[1] http://supercollider.github.io

Here the first sound is used as a comparison value, it is the middle level of all the stimuli (level 5 of 10). The participant indicated the comparison magnitude once at the beginning of the block. In each trial, participants were presented with the comparison sound with its user-defined value simultaneously displayed on-screen and then after a one second pause, one of the nine remaining stimuli were presented. The participant was then asked to respond with a subjective value (i.e. the magnitude of stress they perceived). In a block of 27 trials, each stimuli was presented twice in random order and after a brief break, the next condition was presented, with new instructions that presented the mapping to be used. The order in which each block was presented was also randomised.

Results

Consistent with Walker's findings [38], for a given data:sound mapping most participants responded in a consistent polarity (either positive or negative), such that, for example, increasing pitch represented increasing danger. Some participants responded in an erratic polarity, therefore it was necessary to group the data into *positive*, *negative* or *no* polarity. This was done by calculating the Pearson coefficient for each participant in each block between the log of the responses and the log of the actual stimulus value. A participant's data were considered to have *no* polarity and therefore not included in further analysis if the absolute coefficient for that block did not reach statistical significance. Therefore, with 9 stimuli presented 3 times ($df = 27$), data sets with a correlation coefficient of less than $r_{critical} = 0.367$, $p = <0.05$) were removed. Data reaching statistical significance were grouped into polarities based on the sign of their correlation coefficient (positive or negative). Table 1 shows the number of participants who's data remained in the analysis of each condition, compared to the number of participants who originally carried out the experiment.

	Error	*Stress*	*Danger*
Noise	15 / 16	14 / 16	15 / 16
Roughness	12 / 16	11 / 16	13 / 16
Combined	13 / 16	13 / 16	14 / 16
Pitch	14 / 16	15 / 16	16 / 16

Table 1: Proportion of participants whose data was of a high enough collinearity to be used in analysis.

As the scale of the responses could vary widely between each participant, the geometric mean was calculated for all responses, for each stimulus in a given block (as per [35, 38]). A multivariate ANOVA found an overall effect that manipulating acoustic parameters lead to a significant change in the perception of the magnitude of a data variable ($F_{3,186} = 24.6$, $p < 0.001$).

To better understand the perceptions of each data:sound pairing, we investigated individual results for each pairing. A simple linear regression was carried out for each variable:sound pairing, using the logs of both the geometric means of responses for each stimulus and the actual stimulus values (as per [38]). The slope value of this regression indicates how much the perceived value changes based on changes in the stimulus. Table 2 shows the results of this analysis and the distributions of polarity choices.

The results showed that data for all data:sound pairings were statistically significant ($p < 0.05$). A majority of participants responded in a positive polarity for all data:sound pairings. This is consistent with our original motivations based on the effect found between roughness and danger [2] and noise and image focus [14]. No participants responded in a negative polarity in the pitch:stress condition and only one participant responded in a negative polarity in the pitch:danger condition. Over all conditions, five participants responded in a negative polarity, with four participants responding in a negative polarity in at least three conditions and one participant responding in a negative polarity in one condition.

Discussion & Design Guidelines

For the roughness, noise and combined conditions, it is reasonable that the majority of participants responded in a positive polarity, as the stimuli at level 0 for all of these conditions was an unmodulated pure tone. Participants may have perceived this clean tone as "no" error/stress/danger, due to the absence of any additional acoustic attribute in the sound. Furthermore, at the opposite end of the stimuli levels, the roughness condition at level 10 consisted of a very dissonant sound (roughness being a key part of the perception of dissonance [23]) and the noise and combined conditions at level 10 consisted purely of white noise. Therefore, participants may have associated these sounds with error/stress/danger, as these attributes of sound are normally considered unpleasant, just as error, stress and danger are generally considered unpleasant variables. The ubiquity of the responses for pitch mapping being in a positive polarity may be explained by higher pitches resulting in a higher sense of urgency, such as an infant crying [48]. These results suggest that for data sets relating to "unpleasant" or "undesirable" qualities such as the ones used in this study, the acoustic parameters used in this experiment may provide auditory displays that are consistent with listener's mental models of similar variables.

Only five participants' responses in any condition were in a negative polarity - with all but one of these participants repeatedly responding in a negative polarity. This suggests that a listener who perceives a data:sound mapping in a negative polarity, may be more likely to perceive a negative polarity mapping for any data:sound mapping in which that particular sound parameter is used. Furthermore, all but one of the participants that selected a negative polarity in any condition also responded with no polarity in at least one condition. This may suggest that these participants may not be representative of the consensus of polarity for a given mapping, as they responded with inconsistent polarities over the entirety of the experiment.

The results show that how a listener perceives the magnitude of a data variable varies greatly depending on the data:sound mapping used; participants heard the same sounds yet produced unique magnitude estimation slopes. This further underlines the findings by Walker [38] and Walker & Kramer [43] that the mapping topology used in a sonification system has a significant effect on how the user interacts with it. Using a magnitude

		Error	*Stress*	*Danger*
Noise				
$+^{ve}$	slope	**0.74**	**0.76**	**0.61**
		$n = 12$	$n = 10$	$n = 13$
		$r^2 = 0.99$	$r^2 = 0.99$	$r^2 = 0.99$
		$p < 0.001$	$p = < 0.001$	$p = < 0.001$
$-^{ve}$	slope	**-0.74**	**-0.8**	**-1.2**
		$n = 3$	$n = 4$	$n = 2$
		$r^2 = 0.49$	$r^2 = 0.66$	$r^2 = 0.5$
		$p = 0.02$	$p = 0.005$	$p = 0.02$
Roughness				
$+^{ve}$	slope	**0.56**	**0.68**	**0.53**
		$n = 11$	$n = 10$	$n = 12$
		$r^2 = 0.92$	$r^2 = 0.99$	$r^2 = 0.86$
		$p < 0.001$	$p = < 0.001$	$p = < 0.001$
$-^{ve}$	slope	**-0.62**	**-0.5**	**-0.67**
		$n = 1$	$n = 1$	$n = 1$
		$r^2 = 0.74$	$r^2 = 0.66$	$r^2 = 0.74$
		$p = 0.002$	$p = 0.005$	$p = 0.002$
Combined				
$+^{ve}$	slope	**0.89**	**1.01**	**0.7**
		$n = 11$	$n = 12$	$n = 11$
		$r^2 = 0.99$	$r^2 = 0.98$	$r^2 = 0.99$
		$p < 0.001$	$p = < 0.001$	$p = < 0.001$
$-^{ve}$	slope	**-1.34**	**-1.55**	**-0.96**
		$n = 2$	$n = 1$	$n = 3$
		$r^2 = 0.44$	$r^2 = 0.6$	$r^2 = 0.8$
		$p = 0.03$	$p = 0.008$	$p < 0.001$
Pitch				
$+^{ve}$	slope	**1.21**	**1.08**	**1.07**
		$n = 12$	$n = 15$	$n = 15$
		$r^2 = 0.99$	$r^2 = 0.99$	$r^2 = 0.97$
		$p < 0.001$	$p = < 0.001$	$p = < 0.001$
$-^{ve}$	slope	**-0.35**	***none***	**-0.67**
		$n = 2$	$n = 0$	$n = 1$
		$r^2 = 0.65$	*none*	$r^2 = 0.5$
		$p = 0.005$	*none*	$p = 0.02$

Table 2: Summary of results including slope value (bold). Underlined values indicate the most popular polarities.

estimation paradigm to estimate the polarities and slopes, like the one used in this study and Walker's initial studies [38] can point towards mappings that are consistent with the listener's mental model of how the data value should sound.

Design Guidelines

Roughness, noise and pitch map to "undesirable" variables

In all conditions, the most popular polarities were positive (i.e. an increase in the acoustic attribute was perceived as an increase in the data variable). Results from this study suggest that using increasing roughness, noise or pitch to represent an increase in similar "undesirable" data variables to the ones used here, may be a mapping that is aligned with a listener's expectations of how that data variable should sound when it is sonified.

An unmodulated signal can convey a minimum value

For the roughness, noise and combined conditions, the level 0 stimulus was an unmodulated tone (i.e. no roughness or no noise applied to it) and responses for all data:sound pairings for these conditions tended toward a positive polarity. This suggests that an unmodulated tone can convey an absence or a very low value of an "undesirable" data variable such as stress or danger. This is reflective of the findings in previous studies [14].

FUTURE WORK

This study focused on a limited set of unidimensional data-sound pairings, however one of the main goals of sonification research is to find ways of displaying complex, multidimensional data sets. Therefore, one area that we plan to study is the use of more complex signals as the carrier for roughness and noise and how these more complex signals compare to a listener's mental model of how they expect the data variables to sound. Fitch & Kramer [16] developed the variable of "piggy-back" parameters in which multiple acoustic parameters of a single auditory stream can carry information about multiple variables. This notion may be used to begin to investigate the perceptual factors involved in multidimensional data-sound mappings.

Furthermore, an example of such a complex signal could be a listener's own music. This could be used to create more aesthetically pleasing sonifications - an important considerations in making usable auditory displays (see [21]). Modulating a listener's own music has been shown to successfully convey information [18, 46], therefore another area of future work is to investigate what mappings of data variables to music-modulation parameters are optimal in terms of information transmission, aesthetics/comfort and correlation with listeners' mental model and expectations.

Another area we plan to expand on from this work is to investigate user's mental models and expectations of conceptual data variables when multimodal stimuli are used to convey them. For example, structured vibrations have been found to be able to convey complex information [8] and vibrotactile systems such as the *Shoogle* prototype [47] have utilised mental models of the physical world in their design. Further study into effectively utilising user's mental models to map data to vibrotactile parameters may not only improve haptic feedback, but increase the information carrying capacity of auditory displays by combining modalities.

Only a small subset of possible data:sound mappings were investigated in this study - focusing on investigating a potential link between sound parameters that are generally considered undesirable in music and data variables that share connotations of being negative, or undesirable. As the results of this study and previous experiments have shown [14, 38, 39], the congruency of a data:sound mapping with a listener's mental model of how the data variable should sound is key to the effectiveness of the sonification. Therefore, further study should be carried out on other data:sound mappings. Each context a sonification designer is creating an auditory display for will be unique, however an expansion of the current experimental work on data:sound mappings such as this study and [38, 39] will aid designers in choosing mappings that are usable for their purpose.

CONCLUSIONS

We investigated the effects of acoustic parameter choice on a number of data:sound mappings. We presented a study in which magnitude estimation was used to map how the perceived magnitude of a data variable changed based on a change in an acoustic parameter. Polarities and scales were derived for each mapping. We found that the acoustic parameter used to convey a data variable had a significant effect on the listener's perception of the magnitude of that data variable. This suggests that designing a data:sound mapping which is congruent with the listener's mental model of how they expect the data value to sound in a sonification system is key to successful mappings and sonifications.

ACKNOWLEDGMENTS

This research was funded by the EPSRC. We would like to thank David Borchers for his helpful statistics suggestions in the development of this work.

REFERENCES

1. Janet Anderson. 2005. Creating an empirical framework for sonification design. In *ICAD 2005, Proceedings of the International Conference on Auditory Display*. Georgia Institute of Technology.
2. Luc H Arnal, Adeen Flinker, Andreas Kleinschmidt, Anne-Lise Giraud, and David Poeppel. 2015. Human screams occupy a privileged niche in the communication soundscape. *Current Biology* 25, 15 (2015), 2051–2056.
3. Stephen Barrass. 1998. Some golden rules for designing auditory displays. *in Csound Textbook, B. Vercoe and R. Boulanger, eds., MIT Press, Cambridge, Mass* (1998).
4. Penny Bergman, Anders Sköld, Daniel Västfjäll, and Niklas Fransson. 2009. Perceptual and emotional categorization of sound. *The Journal of the Acoustical Society of America* 126, 6 (2009), 3156–3167.
5. Sara Bly. 1982. Presenting information in sound. In *Proceedings of the 1982 conference on Human factors in computing systems*. ACM, 371–375.
6. Stephen Brewster. 2002. Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing* 6, 3 (2002), 188–205.
7. Stephen Anthony Brewster. 1994. *Providing a structured method for integrating non-speech audio into human-computer interfaces*. Ph.D. Dissertation. University of York York, UK.
8. Lorna M Brown, Stephen A Brewster, and Helen C Purchase. 2006. Multidimensional tactons for non-visual information presentation in mobile devices. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. ACM, 231–238.
9. Michael Czaplik, Rolf Rossaint, Julia Kaliciak, Andreas Follmann, Sabine Kirfel, Roman Scharrer, Martin Guski, Michael Vorländer, Gernot Marx, and Mark Coburn. 2016. Psychoacoustic analysis of noise and the application of earplugs in an ICU: A randomised controlled clinical trial. *European Journal of Anaesthesiology (EJA)* 33, 1 (2016), 14–21.
10. Guo-Qing Di, Xing-Wang Chen, Kai Song, Bing Zhou, and Chun-Ming Pei. 2016. Improvement of Zwicker's psychoacoustic annoyance model aiming at tonal noises. *Applied Acoustics* 105 (2016), 164–170.
11. Wanda L. Diaz-Merced, Robert M. Candey, John C. Mannone, David Fields, and Emilio Rodriguez. 2008. Sonification for the Analysis of Plasma Bubbles at 21 MHz. *Sun and Geosphere* 3, 1 (2008), 42–54.
12. Gaël Dubus and Roberto Bresin. 2013. A systematic review of mapping strategies for the sonification of physical quantities. *PloS one* 8, 12 (2013), e82491.
13. Hugo Fastl. 1998. Psychoacoustics and sound quality metrics. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Vol. 1998. Institute of Noise Control Engineering, 3–10.
14. Jamie Ferguson and Stephen A Brewster. 2017. Evaluation of Psychoacoustic Sound Parameters for Sonification. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction-ICMI 2017*. ACM Press, 120–127.
15. Sam Ferguson, Densil Cabrera, Kirsty Beilharz, and Hong-Jun Song. 2006. Using psychoacoustical models for information sonification. In *ICAD 2006, Proceedings of the International Conference on Auditory Display*. Georgia Institute of Technology.
16. W Tecumseh Fitch and Gregory Kramer. 1994. Sonifying the body electric: Superiority of an Auditory Over a Visual Display in a Complex, Multivariate System. In *Santa Fe Institute Studies in the Science of Complexity*, Vol. 18. Addison-Wesley Publishing Co., 307–307.
17. D. Gentner. 2001. Mental Models, Psychology of. In *International Encyclopedia of the Social & Behavioral Sciences*. Pergamon, 9683 – 9687.

18. Jason Harris, Sarah Vance, Odair Fernandes, Avinash Parnandi, and Ricardo Gutierrez-Osuna. 2014. Sonic respiration: controlling respiration rate through auditory biofeedback. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2383–2388.

19. Christoph Henkelmann. 2007. Improving the aesthetic quality of realtime motion data sonification. *Computer Graphics Technical Report CG-2007-4. University of Bonn* (2007).

20. Thomas Hermann, Andy Hunt, and John G Neuhoff. 2011. *The Sonification Handbook*. Logos Verlag Berlin.

21. Bennett Hogg and Paul Vickers. 2006. Sonification abstraite/sonification concrete: An'aesthetic persepctive space'for classifying auditory displays in the ars musica domain. Georgia Institute of Technology.

22. Yu Huang and Weikang Jiang. 2017. The effect of exposure duration on the subjective discomfort of aircraft cabin noise. *Ergonomics* 60, 1 (2017), 18–25.

23. Phil N Johnson-Laird, Olivia E Kang, and Yuan Chang Leong. 2012. On musical dissonance. *Music Perception: An Interdisciplinary Journal* 30, 1 (2012), 19–35.

24. Alan Kan, Ruth Y Litovsky, and Matthew J Goupell. 2015. Effects of interaural pitch-matching and auditory image centering on binaural sensitivity in cochlear-implant users. *Ear and hearing* 36, 3 (2015), e62.

25. Johan Kildal and Stephen A Brewster. 2006. Non-visual overviews of complex data sets. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*. ACM, 947–952.

26. Sylvan Kornblum, Thierry Hasbroucq, and Allen Osman. 1990. Dimensional overlap: cognitive basis for stimulus-response compatibility–a model and taxonomy. *Psychological review* 97, 2 (1990), 253.

27. Gregory Kramer. 1994. Some organizing principles for representing data with sound. *Auditory Display-Sonification, Audification, and Auditory Interfaces* (1994), 185–221.

28. G Kramer, BN Walker, T Bonebright, P Cook, J Flowers, N Miner, J Neuhoff, R Bargar, S Barrass, J Berger, and others. 1999. The sonification report: Status of the field and research agenda. report prepared for the national science foundation by members of the international community for auditory display. *International Community for Auditory Display (ICAD), Santa Fe, NM* (1999).

29. Lawrence E Marks and George A Gescheider. 2002. Psychophysical scaling. *Stevens' handbook of experimental psychology* (2002).

30. John G Neuhoff and Laurie M Heller. 2005. One small step: Sound sources and events as the basis for auditory graphs.

31. JMS Pearce. 2005. Selected observations on amusia. *European neurology* 54, 3 (2005), 145–148.

32. Irwin Pollack and Lawrence Ficks. 1954. Information of elementary multidimensional auditory displays. *The Journal of the Acoustical Society of America* 26, 2 (1954), 155–158.

33. Robert D Sorkin. 1988. Why are people turning off our alarms? *The Journal of the Acoustical Society of America* 84, 3 (1988), 1107–1108.

34. Stanley S Stevens. 1955. The measurement of loudness. *The Journal of the Acoustical Society of America* 27, 5 (1955), 815–829.

35. Paul Strohmeier and Kasper Hornbæk. 2017. Generating Haptic Textures with a Vibrotactile Actuator. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 4994–5005.

36. Robert Teghtsoonian, SS Stevens, and Geraldine Stevens. 1975. Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects. *The American Journal of Psychology* 88, 4 (1975), 677.

37. A Väljamäe, T Steffert, S Holland, X Marimon, R Benitez, S Mealla, A Oliveira, and S Jordà. 2013. A review of real-time EEG sonification research. In *ICAD 2013-Proceedings of the International Conference on Auditory Display*. Georgia Institute of Technology.

38. Bruce N. Walker. 2002. Magnitude Estimation of Conceptual Data Dimensions for Use in Sonification. *Journal of Experimental Psychology: Applied* 8, 4 (2002), 211–221.

39. Bruce N Walker. 2007. Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Applied Cognitive Psychology* 21, 5 (2007), 579–599.

40. Bruce N Walker and Addie Ehrenstein. 2000. Pitch and pitch change interact in auditory displays. *Journal of Experimental Psychology: Applied* 6, 1 (2000), 15.

41. Bruce N Walker and Gregory Kramer. 1996. Mappings and Metaphors in Auditory Displays: An Experimental Assessment. In *ICAD 1996, Proceedings of the International Conference on Auditory Display*. Georgia Institute of Technology.

42. Bruce N Walker and Gregory Kramer. 2004. Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. *Ecological psychoacoustics* (2004), 150–175.

43. Bruce N Walker and Gregory Kramer. 2005. Mappings and metaphors in auditory displays: An experimental assessment. *ACM Transactions on Applied Perception (TAP)* 2, 4 (2005), 407–412.

44. Bruce N Walker, Gregory Kramer, and David M Lane. 2000. Psychophysical scaling of sonification mappings. In *ICAD 2000, Proceedings of the International Conference on Auditory Display*. Georgia Institute of Technology.

45. Bruce N Walker and David M Lane. 1996. Psychophysical Scaling of Sonification Mappings: A Comparison of Visually Impaired and Sighted Listeners. In *ICAD 2001, Proceedings of the International Conference on Auditory Display*. Georgia Institute of Technology.

46. Nigel Warren, Matt Jones, Steve Jones, and David Bainbridge. 2005. Navigation via continuously adapted music. In *CHI'05 extended abstracts on Human factors in computing systems*. ACM, 1849–1852.

47. John Williamson, Roderick Murray-Smith, and Stephen Hughes. 2007. Shoogle: excitatory multimodal interaction on mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 121–124.

48. Philip Sanford Zeskind and Victoria Collins. 1987. Pitch of infant crying and caregiver responses in a natural setting. *Infant Behavior and Development* 10, 4 (1987), 501–504.

49. Eberhard Zwicker and Hugo Fastl. 2013. *Psychoacoustics: Facts and models*. Vol. 22. Springer Science & Business Media.