REVISITING PULSE RATE, FREQUENCY AND PERCEIVED URGENCY: HAVE RELATIONSHIPS CHANGED AND WHY?

Christian Gonzalez, Bridget A. Lewis, and Carryl L. Baldwin

George Mason University
4440 University Drive
Fairfax, VA
cgonza12@gmu.edu

ABSTRACT

Research in psychophysics and auditory warnings during the early 1990's created much of the theoretical groundwork for auditory alert design today. The main goal of this series of experiments was to reevaluate key auditory parameters (pulse rate and fundamental frequency) that have been shown to exhibit psychophysical relationships with perceived urgency in an updated context. Our results suggest that the relationship between pulse rate and perceived urgency may have weakened since the early 1990's, but the relationship between frequency and perceived urgency remains relatively stable. However, the relationship between pulse rate and perceived urgency was more reliable across multiple study manipulations relative to the relationship between frequency and perceived urgency. Based on its robustness across variable acoustic contexts, auditory alert designers wishing to convey a range of urgency levels may be more successful utilizing pulse rate rather than frequency.

1. INTRODUCTION

Sounds can capture people's attention no matter where they are looking. This makes the auditory modality well-suited for signaling events of varying criticality during visually demanding tasks like driving. The auditory environment within consumer and commercial vehicles is quickly becoming more heavily loaded with safety, communication and navigation technologies. This ongoing increase of in-vehicle technology requires manufacturers to develop auditory alerts that convey varying levels of urgency. Because sound is used to convey many different meanings, it becomes imperative to consider how various auditory parameters may impact perceptions of urgency and how appropriately matched these parameters are to the hazard levels they connote.

In 1993, Hellier, Edworthy and Dennis [1] demonstrated that Stevens' Power Law exponents [2] could be used to quantify the relationship between changes in auditory parameters and changes in perceived urgency. In their seminal paper, the authors had participants produce line ratings to represent the perceived urgency of sounds that varied systematically in different auditory parameters. These ratings were then used to create psychophysical functions (summarized by a Stevens' Power Law exponent) describing the relationship between each parameter and perceived urgency.

The power law exponents they identified successfully predicted the perceived urgency ratings of a new set of stimuli. Their results demonstrated that a) it is possible to assess the relationship between different auditory parameters and perceived urgency using psychophysical methods, and b) some parameters have a stronger relationship with urgency than others.

The impact of their work on auditory warning design has been substantial. This paper, as well as several other related papers [3–5], has served as the basis for the urgency mapping literature. This work has focused on systematically manipulating auditory parameters in a context neutral format. In a limited number of studies, urgency mapping has been examined within the context of driving [6–8].

However, in the nearly 20 years since the publication of Hellier et al.'s [1] original article, the prevalence and diversity of auditory alerts has increased dramatically. Given the increasingly complex soundscape, perceptions of various auditory parameters may have changed or may be influenced by concurrent changes in other dimensions. For example, increases in frequency may not seem as urgent if pulse rates are changing at the same time. We sought to examine these issues in the current investigation.

Stevens [9] has demonstrated that psychophysical judgments are relative to the set of stimuli being presented. As in-vehicle alerts can vary greatly in the information they represent, sets of alerts within vehicles will need to be heterogeneous [10], [11] in order for drivers to discriminate between the different intended meanings. Allowing participants to rate all levels of each parameter within the same experiment may help us better understand how a heterogeneous alert environment impacts ratings of perceived urgency.

1.1 The present studies

The primary goal of the present investigation was to replicate the basic psychophysical relationships between key auditory parameters (namely pulse rate and frequency) observed by

Hellier et al. [1]. Secondly, we wished to examine the impact of auditory context (presenting several different parameter changes within a single experiment) on perceived urgency by utilizing a within-subjects design. And finally, we sought to examine the impact of presenting the sounds within a driving context. Note that we made no attempt to simulate an actual driving context, but rather merely asked participants to consider how urgent the sounds presented would seem if heard while driving.

Psychophysical relationships between stimuli and subjective ratings have been shown to be consistent across cultures [12], [13], participants [14] and samples, [15–17] (see [18] for review) thus allowing for comparison across the four experiments presented here. The following studies systematically manipulated pulse rate, fundamental frequency and intensity because all three have been shown to relate to changes in perceived urgency [1], [3], [10]. We examined a number of methodological changes that impact the coherence between our results and those observed previously by Hellier et al. [1]. In general, we expected to find that increases in pulse rate and frequency would lead to higher ratings of perceived urgency. Based on Hellier et al.'s [1] results, we expected pulse rate to exhibit the strongest relationship with perceived urgency. We hypothesized that rating multiple types of stimuli at once could potentially cause participants to calibrate their ratings of perceived urgency based on a single, highly urgent sounding parameter.

2. METHOD (EXPERIMENT1)

2.1 Participants

Twenty-six graduate and undergraduate George Mason University students aged 18 to 25 (mean = 20.08; 12 female) voluntarily participated for class credit. All participants indicated they had normal hearing. A unique sample of participants was used for each experiment.

2.2 Design

A within-subjects designed was utilized. Each participant experienced and gave subjective ratings for all magnitude levels of all alerts. Each alert was presented three times within the experiment and all alerts were completely randomized. The average rating of each alert was used for analysis to mitigate any order effects.

2.3 Materials

2.3.1. Equipment

Alerts were presented in a sound attenuated laboratory on an Optiplex 745 Dell PC with a SoundMAX Integrated Digital HD Audio Driver Analog Device sound card. All alerts were presented through a pair of Sennheiser HD-280 stereo headphones. There was no evidence of intensity disparity between the left and right channel.

A MATLAB based program was written to present alerts as well as collect subjective ratings of urgency, annoyance and acceptability using a digital slider. The range of the slider included values between 0-100 and allowed participants to see their current rating. The program also allowed participants to adjust each rating until they felt it accurately reflected their perceptions before submitting.

2.3.2. Stimuli

A total of 21 stimuli were created, seven for each of the three auditory parameters that were investigated: fundamental frequency, intensity and pulse rate. Experiments 1 and 2 used stimuli that varied in all three parameters (21 total), but Experiments 3 and 4 used only stimuli that varied in pulse rate and fundamental frequency (14 total). Frequency and pulse rate alerts were created following the specifications of Hellier et al. [1], whereby varying durations of silence separated several standard "basic" pulses. The basic pulse used, based on the pulse-burst principles described by Patterson [19], was a 200 millisecond (ms) sine wave (20 ms on/offset) with a fundamental frequency (F_0) of 300 Hz and 15 harmonic components. Each alert was then made up of parametric variations of the basic pulse and varying durations of silence. Only one alert parameter was manipulated at a time while all other parameters were held constant to the basic pulse as described above. Unless intensity was being specifically manipulated, the basic pulse was presented at 75 dBA. This methodology ensured that our stimuli matched those used by Hellier et al. [1] exactly.

Table 1 provides a description of the stimuli used in the four experiments. The bolding within the columns indicates specifically what parameter changed in each stimulus. The seven fundamental frequency alerts consisted of six basic pulses of the same F_0 played in succession. There was no silence between the pulses and each alert had a total duration of 1200 ms. The 20 ms on/offset allowed the pulses to be discernible without the need for silence between pulses.

The seven intensity alerts varied in a similar fashion. Each alert consisted of six basic pulses with an F_0 of 300 Hz played in succession. Total duration for each alert was 1200 ms. Again, the on/offset allowed the pulses to be discerned without silence between each pulse. Using a Brüel & Kjær Sound Level Meter, we verified the intensity of each stimulus. Decibel measurements were taken from the individual pulses rather than the entire alert to avoid including the decreasing intensity of the onset and offset in our measurement.

The seven pulse rate alerts consisted of between four and twelve basic pulses (F_0 = 300 Hz) the inter-pulse interval (IPI) - or silence between pulses - varied from 475 to 9 ms. Pulse-to-pulse duration is defined as the duration from the start of one pulse to the start of the next pulse (pulse duration + IPI). The total time of each pulse rate alert approached, but did not exceed 2500 ms so each alert varied slightly in total duration. Pulse rate was derived via a formula based on one previously used by Hellier et al. [1]:

$$\text{2500ms/pulse-to-pulse time} \tag{1}$$

2500 ms represents the total approximate duration of each stimulus. 2500 ms was used as the total duration to standardize the rates for all pulse rate stimuli although the total true durations varied slightly. For example, a stimulus with a pulse rate of 3.69 would consist of four basic pulses of 200 ms each

	Level	# of Pulses	FFq (Hz)	IPI (ms)	Length (ms)	Intensity (dBA)
	1	4	300	**475**	2225	75
	2	5	300	**302**	2210	75
	3	6	300	**238**	2389	75
Pulse Rate	4	8	300	**118**	2430	75
	5	9	300	**60**	2280	75
	6	10	300	**50**	2450	75
	7	12	300	**9**	2499	75
	1	6	**210**	0	1200	75
	2	6	**250**	0	1200	75
	3	6	**260**	0	1200	75
Fundamental Frequency	4	6	**320**	0	1200	75
	5	6	**440**	0	1200	75
	6	6	**500**	0	1200	75
	7	6	**680**	0	1200	75
	1	6	300	0	1200	**66**
	2	6	300	0	1200	**69**
	3	6	300	0	1200	**72**
Intensity	4	6	300	0	1200	**75**
	5	6	300	0	1200	**78**
	6	6	300	0	1200	**81**
	7	6	300	0	1200	**84**

Table 1: Stimuli for Experiments 1 - 4. Experiments 1 and 2 used all 21 stimuli. Experiments 3 and 4 used only pulse rate and frequency stimuli.

separated by 475 ms of silence. Because following the last pulse was simply 275 ms of silence the total true duration of this alert is 2225 ms rather than 2500 ms.

2.4 Procedure

After completing an informed consent form, participants were told they would be presented with a variety of auditory alerts, which they would then rate on urgency, annoyance and acceptability. They were asked to imagine these alerts were presented in a driving context, but we did not specify in what capacity (e.g. collision warning, navigation, or communication etc.). We operationally defined acceptability as “How likely you would be to purchase a vehicle with this type of alert.”

Participants then completed a brief practice with non-experimental auditory alerts to familiarize themselves with the rating slider. During the actual experiment, participants received a fixation cross on a black screen for 500 ms, then the auditory alert and a black screen, then three separate rating screens for urgency, annoyance and acceptability. The rating screen order was consistent throughout the experiment. The experiment took approximately 30 minutes to complete.

3. RESULTS AND DISCUSSION (EXPERIMENT 1)

Though annoyance and acceptability results were analyzed in Experiments 1 - 4, they will not be discussed here. The goal of this paper was to re-examine Hellier et al.’s [1] findings, which pertained only to ratings of perceived urgency. See [18 *under review*] for a closer examination of urgency and annoyance ratings.

Results were analyzed according to Hellier et al.’s [1] specifications. Exponents were calculated for each parameter according Stevens’ [2] methodology. All raw values were log transformed and their geometric means were taken. All parameter values were also log transformed. This allowed us to create a log-log plot of perceived urgency ratings as a function of changes in each parameter. The slope of the best-fit line plotted through these points is the exponent used in Stevens’ Power Law:

$$P = kS^n \tag{2}$$

Where P is the perceived urgency rating of the physical stimulus (S), k is a constant and n is the exponent found using empirical data. Smaller exponents are related to smaller changes in perceived urgency as the stimulus changes whereas larger exponents (generally greater than 1) are related to larger changes in perceived urgency relative to stimulus changes. Similar to Hellier et al.’s [1] findings, this experiment also found a large portion of the variance could be accounted for by the slope of a best fit line (see Table 2). This supports Hellier et al.'s [1] assertion that it is possible to systematically quantify changes in perceived urgency with relation to changes in parameter level. The percent variance explained in Table 2 is in reference to the variance explained among the seven log-transformed mean values of each parameter level, not the variance explained among all participants' ratings.

As illustrated in Table 2, we found that intensity produced the largest exponent (n = 3.8) while pulse rate (n = .52) and frequency (n = .54) produced similar, much smaller exponents. Our pulse rate exponent was nearly 60% smaller than Hellier et al.'s [1] (see Table 6) indicating a weaker relationship with urgency than expected. However, our frequency exponent was slightly larger than Hellier et al.'s [1] findings.

Level	Pulse Rate	Fundamental Frequency	Intensity
	Mean rating value (0-100) and standard deviation		
1	46.61 (21.5)	46.90 (19.4)	40.70 (22.5)
2	55.59 (17.6)	53.54 (18.5)	54.33 (20.7)
3	64.90 (16.3)	55.06 (16.8)	62.82 (19.6)
4	71.97 (15.4)	70.70 (14.8)	69.18 (14.6)
5	75.11 (13.6)	69.19 (18.8)	74.23 (10.81)
6	72.75 (14.3)	74.13 (15.6)	78.90 (10.7)
7	78.61 (12.4)	73.52 (19.3)	84.55 (9.4)
Exponent	0.51	0.54	3.8
% Variance accounted for	0.91	0.75	0.92

Table 2: Experiment 1 - Effects of Three Auditory Parameters on Perceived Urgency.

4. METHOD (EXPERIMENT 2)

4.1 Introduction

Experiment 2 was very similar to Experiment 1 with the exception that we provided an additional visual cue to better connote a driving context for participant ratings.

4.2 Participants

Thirty-one graduate and undergraduate George Mason University students aged 18 to 25 (mean = 19.5; 22 female) voluntarily participated for class credit.

4.3 Procedure

Experiment 2 followed the exact specifications of Experiment 1, except that instead of a black screen with a fixation cross, participants saw a generic car dashboard on the screen.

5. RESULTS (EXPERIMENT 2)

Results were examined using the exact procedure described for Experiment 1. Again, we found intensity alerts produced the largest exponent (n = 2.6) by far (see Table 3), though 30% smaller than Experiment 1. Pulse rate alerts produced a similar exponent as Experiment 1 (n = .47) while frequency alerts produced a much smaller exponent (n = .29). These findings still differ greatly from Hellier et al.'s [1] results (Table 6) and may suggest that the relationship between frequency and perceived urgency may be more sensitive to even small changes in context than pulse rate. The fact that intensity produced such a large exponent could be indicative participants calibrating their ratings of perceived urgency for pulse rate and frequency alerts. Pulse rate and frequency could have been perceived as less urgent in the context of another seemingly much more urgent sounding alert (intensity alerts). While the relationship between intensity and perceived urgency is seemingly quite strong, if we wish to maintain guidelines established by Patterson [17] [i.e. warnings should be presented at least 15 decibels (dB) above ambient background noise], intensity would likely not be a feasible parameter to manipulate in a noisy vehicle. For Experiment 3, we eliminated intensity from our manipulations and examined the impact of experiencing changes in pulse rate followed by frequency on ratings of perceived urgency.

Level	Pulse Rate	Fundamental Frequency	Intensity
	Mean rating value (0-100) and standard deviation		
1	52.31 (27.9)	60.93 (23.8)	53.22 (26.1)
2	57.72 (26.4)	64.10 (22.8)	60.81 (23.6)
3	64.61 (22)	65.19 (22.5)	66.28 (20.7)
4	72.47 (21.8)	71.31 (18.1)	70.76 (17.7)
5	74.54 (21.4)	70.47 (19.9)	74.49 (16.9)
6	74.39 (18.6)	73.63 (17.7)	80.57 (11.3)
7	77.32 (18.3)	75.11 (18.4)	84.31 (12.8)
Exponent	0.47	0.28	2.6
% Variance accounted for	0.94	0.87	0.97

Table 3: Experiment 2 - Effects of Three Auditory Parameters on Perceived Urgency.

6. METHOD (EXPERIMENT 3)

6.1 Introduction

Based on the results of Experiment 2, we split Experiment 3 into blocks. Block 1 always consisted of only pulse rate alerts and Block 2 always consisted of only frequency alerts. Block order was not manipulated and participants were not made aware of the block changes. This was done to mitigate any rating calibration effects discussed in Experiment 2. Also, this more closely mirrors Hellier et al. [1] where they ran four individual smaller experiments to collect ratings of urgency.

6.2 Participants

Thirty graduate and undergraduate George Mason University students aged 18 to 29 (mean = 20.52; 6 female) voluntarily participated for class credit. All participants indicated they had normal hearing.

6.3 Equipment

Due to our inability to closely reproduce Hellier et al.'s [1] findings, we decided to change our rating scale to more accurately mimic the paper and pencil methodology they used. Instead of a digital slider, participants provided ratings via an on-screen line draw. Participants could draw a straight, horizontal line anywhere on the rating screen using the mouse. The maximum possible length was the equivalent of 394 millimeters (the maximum possible line length in Hellier et al.'s [1] study). Participants were not given feedback on the magnitude of their ratings. The length of the line was recorded in pixels then converted to mm to more accurately reflect the data and scale used in Hellier et al. [1].

This change in scale coupled with a change in parameters investigated is not an ideal manipulation. However, [9] has demonstrated that, in general, relationships between stimuli and ratings are independent of the rating scale. Thus, we combined the manipulations in order to constrain the number of experiments in this series.

6.4 Procedure

Other than the changes noted above, the procedures were identical to Experiment 1 and 2. We verified that participants understood that the length of the line reflected the magnitude of their rating, such that longer lines meant alerts were more urgent, more annoying, and more acceptable and vice versa. The total experiment time was reduced to 20 minutes.

7. RESULTS AND DISCUSSION (EXPERIMENT 3)

The change in rating scale from Experiments 1 and 2 to Experiment 3 necessitated a slightly lengthier transformation in order to compare across studies. Ratings were converted from mm on the screen to a percentage of the total possible mm rating they could have given, thus allowing for comparison between mm ratings and slider ratings out of 100. These converted percentage values were then log transformed to derive the exponents. The methods used in Experiment 3 resulted in a much larger exponent relative to Experiments 1 and 2 for pulse rate alerts (n = .77). However, frequency alerts produced a non-significant exponent (n = .10). Once again we were unable to closely replicate Hellier et al.'s 1993 [1] findings. The dramatic change in the exponent for frequency seemed likely due to an order effect. Participants may have calibrated their ratings of frequency relative to the block in which pulse rate was manipulated. The increase in the observed exponent for pulse rate in Experiment 3 supports our previous suspicion that exposure to sounds varying in intensity (Experiments 1 and 2) resulted in some rating compression. The methodology in Experiment 3 mirrors Hellier et al. [1] in stimuli, rating method and (pseudo) between-subjects design more so than of the previous studies. Although our pulse rate exponent is closer to their original findings, it is still over 40% smaller. This may indicate that even under nearly identical conditions, the relationship between pulse rate and perceived urgency has changed over the last 20 years.

8. METHOD (EXPERIMENT 4)

8.1 Introduction

In order to examine the potential order effects from Experiment 3, we ran Experiment 4 with a reversed block order where Block 1 consisted of only changes in frequency and Block 2 consisted of only changes in pulse rate.

Level	Pulse Rate	Fundamental Frequency
	Mean rating value (%) and standard deviation	
1	29.08 (16.7)	31.50 (20.1)
2	34.09 (19.3)	34.89 (22.3)
3	38.58 (20.2)	31.54 (22.7)
4	47.66 (22.3)	36.12 (20.5)
5	52.37 (23.5)	37.93 (24.5)
6	56.31 (24.4)	37.91 (30.2)
7	64.20 (26.1)	43.05 (32.4)
Exponent	0.47	0.28
% Variance accounted for	0.94	0.87

Table 4: Experiment 3 - Effects of Three Auditory Parameters on Perceived Urgency.

8.2 Participants

Ten graduate and undergraduate George Mason University students aged 18 to 22 (mean = 19.12; 13 female) voluntarily participated for class credit. All participants indicated they had normal hearing.

9. RESULTS AND DISCUSSION (EXPERIMENT 4)

Results were analyzed using the same procedures employed in Experiment 3. We observed a smaller exponent for manipulations of pulse rate (n = .50) and a much larger exponent for frequency (n = .38) indicating there may have been some block order effects on ratings of perceived urgency. However, our exponent for frequency matched Hellier et al.'s [1] findings almost exactly. This suggests that changes in frequency may still have a similar relationship with perceived urgency though only under specific and homogenous conditions.

10. GENERAL RESULTS

Table 6 summarizes exponent values, effect sizes and 95% Confidence Intervals (CIs) for pulse rate and frequency across all four studies and Hellier et al. [1]. In order to investigate differences across studies we utilized 95% CIs of the exponents derived from the log-log regression plots. Though 95% CIs were not reported in their original article, Hellier et al. [1] did provide raw data from their experiments. This allowed us to analyze their data and identify the CIs for pulse rate and frequency exponents. We converted their raw millimeter values to percentages and then log transformed them, similar to Experiments 3 and 4, allowing for comparison on a 0-100 scale across all experiments.

Level	Pulse Rate	Fundamental Frequency
	Mean rating value (%) and standard deviation	
1	34.46 (18.5)	32.27 (16.6)
2	32.86 (8.2)	35.93 (18.6)
3	46.19 (20.9)	38.81 (22.2)
4	53.29 (22.5)	37.53 (22.1)
5	51.11 (17.1)	45.50 (25)
6	50.22 (23.43)	48.29 (24.4)
7	58.23 (20.7)	53.19 (29)
Exponent	0.50	0.38
% Variance accounted for	0.87	0.88

Table 5: Experiment 4 - Effects of Two Auditory Parameters on Perceived Urgency.

Because of the variation in samples and methodologies we encourage caution when interpreting the CIs across the four experiments and Hellier et al. [1]. Finding statistically significant differences was not the ultimate goal for this series rather exploration of previous relationships. Table 6 also includes R^2 values reported in Hellier et al. [1] for comparison purposes. (R^2 values are equivalent to the percent of variance accounted for).

10.1 Pulse rate across experiments

We found the only experiment that did not fall within the 95% CI of another was Experiment 3. This experiment produced the largest exponent for pulse rate (n = .77) and falls outside of the CI of Experiment 2. Experiment 3 also approached the upper limits of Experiment 1 and 4's CIs. However, overall, exponents from the four experiments remained quite similar. In comparison none of the exponents from the four experiments fell within the 95% CI of Hellier et al.'s [1] exponent value, indicating that the relationship between pulse rate and perceived urgency may have weakened over time.

Figure 1 shows a log-log plot of pulse rate on the x-axis and average perceived urgency rating on the y-axis. The mean rating values for each level of pulse rate are shown with a line of best fit for each experiment and Hellier et al. [1]. The exponents reported in Table 6 reflect the slope of each line. This figure highlights the similarity in slopes across all four experiments. The y-intercepts for Experiments 3 and 4 differ from Experiments 1 and 2 because of the difference in scales between the two sets of studies. However, a general slope characteristic is maintained by the four experiments illustrating the results reported in Table 6. In comparison, the slope of the line of best fit from Hellier et al. [1] appears much steeper demonstrating the large difference in exponents.

10.2 Frequency across studies

We found a much larger range of exponents for frequency across all four experiments. We also found that no single experiment fell outside the 95% CI of any other. However, when looking at the exponent values we see a much larger spread for frequency than pulse rate. We also see consistently smaller R^2 values for frequency across experiments. Only three of the four experiments found a relationship between changes in frequency and perceived urgency. However, all four experiments fall within the 95% CI of Hellier et al.'s [1] exponent value. In addition to producing the same exponent, Experiment 4 also produced upper and lower CI bounds similar to Hellier et al. [1]. This may indicate that the relationship between frequency and perceived urgency has changed less than pulse rate over time. Figure 2 shows a log-log plot similar to Figure 1, but with fundamental frequency on the x-axis. Figure 2 illustrates the variation in slopes across all 4 experiments as well as the greater variance of mean data points compared to Figure 1. The similarity in slopes of Experiment 4 and Hellier et al. [1] is also evident though the y-intercept values differ due to scale differences.

Parameter	Exp.	n	R^2	*p*	95% Lower	95% Higher
Pulse Rate	1	.51	.91	0	.33	.71
	2	.47	.94	0	.33	.62
	3	.77	.99	0	.70	.85
	4	.50	.87	0	.28	.73
	Hellier et al. [1]	1.35	.98	0	1.16	1.54
Frequency	1	.54	.75	.01	0.18	.90
	2	.28	.87	0	0.16	.42
	3	.10	.15	.40	NS	NS
	4	.38	.88	0	0.23	.55
	Hellier et al. [1]	.38	.93	0	.29	.54

Table 6:Summary results Experiment 1 – 4 and Hellier et al. [1]

11. GENERAL DISCUSSION

Our findings indicate that across various procedural and methodological manipulations and within homogenous and heterogeneous alert sets, pulse rate exhibits a relatively robust relationship with perceived urgency. However, the magnitude of this relationship may have weakened since Hellier et al.'s 1993 [1] experiment 20 years ago. Though it is difficult to systematically evaluate the role that increased technology and

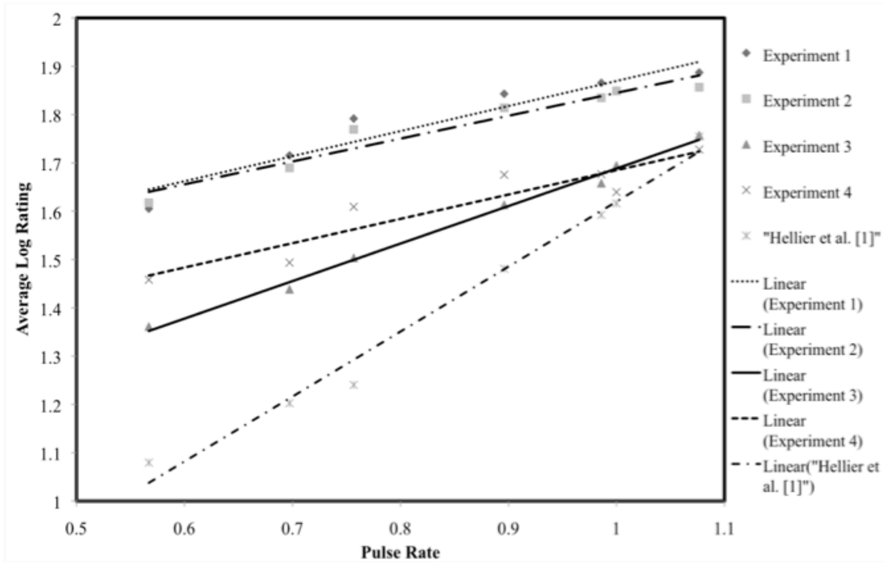

Figure 1: Log-log plot of pulse rate and ratings of perceived urgency across four experiments.

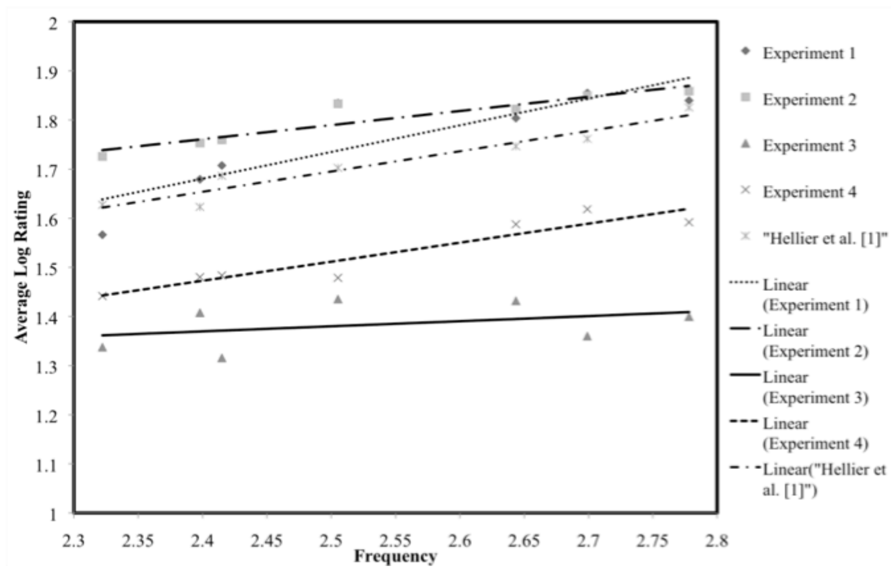

Figure 2: Log-log plot of frequency and ratings of perceived urgency across four experiments.

sound exposure plays in this weakening, it seems plausible that sensitivity to changing pulse rates has decreased with a general increase in exposure to technology. The stability of pulse rate over multiple studies is in agreement with Patterson's [19] finding that temporal patterns are the critical structural difference when distinguishing between sounds. Furthermore, the robustness of pulse rate across manipulations may also be explained by the ability of the auditory system to distinguish minute changes in timing in concert with other highly variable parameters, as exhibited by the role of temporal characteristics in perception of phonemic changes resulting from coarticulation and changes in Voice Onset Time on the millisecond level [21], [22].

The relationship between frequency and perceived urgency seems to be conditional on the presence of other alerts against which it may be compared. Furthermore, rating order also seems to have a large impact on the relationship between frequency and perceived urgency. Hellier et al. [1] suggested a similar unreliability in the frequency exponent they reported claiming the metathetic nature of frequency [23] as a potential explanation. Previous research [24] has also shown the ability to retain pitch decays over time and is subject to interference from other tones [25]. This may make frequency less than ideal for conveying multiple levels of urgency especially in a heterogeneous environment. Surprisingly, while the relationship of pulse rate and perceived urgency seems to have weakened over time, frequency, under homogenous conditions (Experiment 4), was the only parameter that produced a similar power law exponent to those reported by Hellier et al. [1].

Though different samples of participants were used for each experiment, both pulse rate and frequency ratings were subject to the same potential influence of individual differences. As [26] has shown, individual variation in power law exponents is indeed random and pooling subject data results in reliable exponents over time and across samples. Barring that, there is still a chance that discrepancy in methods represents some of the variation in power law exponents found between studies. However, across manipulations and ostensibly different samples, pulse rate maintained a relatively stable exponent clearly different from Hellier et al.'s [1] results.

The two main findings applicable to auditory alert designers are: 1) Alerts within a heterogeneous set (similar to what may be found in vehicles or other complex auditory environments) may have different relationships with perceived urgency than those same alerts in a homogenous set. 2) When it is critical to convey a specific level of urgency aurally, pulse rate may be the most reliable and robust parameter to manipulate. However, due to the apparent weakening of the magnitude of the relationship between pulse rate and perceived urgency, increased levels of pulse rate may be needed to convey the same level of urgency achieved 20 years ago.

Together, the present series of experiments examined some of the many methodological factors that may impact the relationship between auditory parameters and perceptions of urgency. The relationship between pulse rate and perceived urgency appears to have changed over time, but it remains well-suited for use in the design of effective in-vehicle alerts and alarms. In the future, we plan to extend this work into higher fidelity simulations where we can evaluate the impact of more realistic driving contexts on ratings of perceived urgency.

12. ACKNOWLEDGEMENTS

The authors would like to thank Stephanie M. Pratt and Daniel M. Roberts for making this work possible through their assistance in coding and developing the various study iterations.

13. REFERENCES

[1] E. Hellier, J. Edworthy, and I. Dennis, "Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 35, no. 4, pp. 693–706, 1993.

[2] S. S. Stevens, "On the psychophysical law.," *Psychological review*, vol. 64, no. 3, p. 153, 1957.

[3] E. Hellier and J. Edworthy, "On using psychophysical techniques to achieve urgency mapping in auditory warnings," *Applied Ergonomics*, vol. 30, no. 2, pp. 167–171, 1999.

[4] J. Edworthy, S. Loxley, and I. Dennis, "Improving auditory warning design: Relationship between warning sound parameters and perceived urgency," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 33, no. 2, pp. 205–231, 1991.

[5] E. Hellier, J. Edworthy, and I. Dennis, "A comparison of different techniques for scaling perceived urgency," *Ergonomics*, vol. 38, no. 4, pp. 659–670, 1995.

[6] C. L. Baldwin, "Verbal collision avoidance messages during simulated driving: perceived urgency, alerting effectiveness and annoyance," *Ergonomics*, vol. 54, no. 4, pp. 328–337, Apr. 2011.

[7] C. L. Baldwin and J. F. May, "Loudness interacts with semantics in auditory warnings to impact rear-end collisions," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 14, no. 1, pp. 36–42, Jan. 2011.

[8] D. C. Marshall, J. D. Lee, and P. A. Austria, "Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 49, no. 1, pp. 145–157, 2007.

[9] S. S. Stevens, "On the operation known as judgment," *American Scientist*, vol. 54, no. 4, pp. 385–401, 1966.

[10] J. Edworthy, E. Hellier, K. Titchener, A. Naweed, and R. Roels, "Heterogeneity in auditory alarm sets makes them easier to learn," *International Journal of Industrial Ergonomics*, vol. 41, no. 2, pp. 136–146, Mar. 2011.

[11] M. A. Nees and B. N. Walker, "Auditory Displays for In-Vehicle Technologies," *Reviews of Human Factors and Ergonomics*, vol. 7, no. 1, pp. 58–99, Aug. 2011.

[12] E. C. Haas and J. Edworthy, "Designing urgency into auditory warnings using pitch, speed and loudness," *Computing & Control Engineering Journal*, vol. 7, no. 4, pp. 193–198, 1996.

[13] C. Ellen and J. G. Casali, "Perceived urgency of and response time to multi-tone and frequency-modulated warning signals in broadband noise," *Ergonomics*, vol. 38, no. 11, pp. 2313–2326, 1995.

[14] Chillery, J. A. and Collister, J. B., "Possible confusions amongst a set of auditory warning signals developed for helicopters," Farnborough: Royal Aircraft Establishment, Technical Memo FS(F) 655, 1986.

[15] B. N. Walker, "Consistency of magnitude estimations with conceptual data dimensions used for sonification," *Applied Cognitive Psychology*, vol. 21, no. 5, pp. 579–599, Jul. 2007.

[16] J. Edworthy and S. Newstead, "On the Stability of the Arousal Strength of Warning Signal Words," 1999.

[17] M. Teghtsoonian, "Children's scales of length and loudness: A developmental application of cross-modal matching," *Journal of Experimental Child Psychology*, vol. 30, no. 2, pp. 290–307, Oct. 1980.

[18] E. Hellier and J. Edworthy, "Subjective rating scales: scientific measures of perceived urgency?," *Ergonomics*, vol. 45, no. 14, pp. 1011–1014, 2002.

[19] R. D. Patterson, *Guidelines for auditory warning systems on civil aircraft*. Civil Aviation Authority, 1982.

[20] Gonzalez, C., Lewis, B. A., Pratt, S. M., Roberts, D. M., and Baldwin, C. L., "Perceived urgency and annoyance of auditory alerts in a driving context. (under review)," in *Proceedings from HFES '12*, Boston, MA, 2012.

[21] L. Lisker and A. S. Abramson, "A cross-language study of voicing in initial stops: acoustical measurements," *Word*, vol. 20, no. 3, pp. 384–422, 1964.

[22] W. Strange, R. R. Verbrugge, D. P. Shankweiler, and T. R. Edman, "Consonant environment specifies vowel identity," *The Journal of the Acoustical Society of America*, vol. 60, no. 1, pp. 213–224, 1976.

[23] S. S. Stevens and E. H. Galanter, "Ratio scales and category scales for a dozen perceptual continua.," *Journal of Experimental Psychology*, vol. 54, no. 6, p. 377, 1957.

[24] D. Deutsch, "Delayed pitch comparisons and the principle of proximity," *Perception & Psychophysics*, vol. 23, no. 3, pp. 227 – 230, 1978.

[25] D. Deutsch, "The processing of pitch combinations," *The psychology of music*, pp. 271–316, 1982.

[26] M. Teghtsoonian and R. Teghtsoonian, "How repeatable are Stevens's power law exponents for individual subjects?," *Attention, Perception, & Psychophysics*, vol. 10, no. 3, pp. 147–149, 1971.