USING SEMANTIC DIFFERENTIAL SCALES TO ASSESS THE SUBJECTIVE PERCEPTION OF AUDITORY WARNING SIGNALS

Joana Vieira
Ergonomics Lab
Faculty of Human Kinetics
CIAUD, Lisbon School of Architecture,
Universidade de Lisboa

ALGORITMI, University of Minho, Braga

CCG - Centre for Computer Graphics
Guimarães, Portugal

joana.vieira@ccg.pt

Jorge Almeida Santos
School of Psychology
University of Minho
Braga, Portugal

ALGORITMI, University of Minho, Braga

CCG -Centre for Computer Graphics
Guimarães, Portugal

jorge.a.santos@psi.uminho.pt

Paulo Noriega
Ergonomics Lab
Faculty of Human Kinetics
CIAUD - Lisbon School of Architecture
Universidade de Lisboa,
Portugal

pnoriega@campus.ul.pt

ABSTRACT

The relationship between physical acoustic parameters and the subjective responses they evoke is important to assess in audio alarm design. While the perception of urgency has been thoroughly investigated, the perception of other variables such as pleasantness, negativeness and irritability has not. To characterize the psychological correlates of variables such as frequency, speed, rhythm and onset, twenty-six participants evaluated fifty-four audio warning signals according to six different semantic differential scales. Regression analysis showed that speed predicted mostly the perception of urgency, preoccupation and negativity; frequency predicted the perception of pleasantness and irritability; and rhythm affected the perception of urgency. No correlation was found with onset and offset times. These findings are important to human-centred design recommendations for auditory warning signals.

1. INTRODUCTION

The study of the psychological correlates of physical parameters motivated early psychophysical research. This was considered a tool to better study and understand the mind [1]. Early studies focused on sensory thresholds of humans, associating the human response to the systematic variation of a physical stimulus. Nowadays, this interest in human response is broader. Could we know more than sensory responses? Could similar methods be used to comprehend the relation between physical parameters and affective responses?

Several experimental methodologies attempt to understand the association of physical parameters with subjective perceptions and evaluations by humans. Mostly derived from these early works, it is common to have controlled laboratorial set-ups to understand how certain emotional states can be triggered. This happens because there is consensus and robustness in what a culturally similar group of participants finds *pleasant*, *attractive*, or *annoying*.

For instance, semantic profiling stemmed from the wine tasting industry and is currently being applied in other areas such as acoustics [2], [3]. Here, the evaluators can taste and compare several samples of wines and then verbally create a vocabulary describing the perceptual differences between the wines. Later, consensus is achieved among all gathered vocabularies. Another technique, Kansei Engineering [4], originated in the automotive industry in Japan intending to quantitatively connect affective responses of the customers to physical design specifications. The evaluation method pairs representative samples of the product under evaluation with representative words usually presented in a semantic differential scale (a scale between two polar adjectives).

In the auditory modality, the semantic differential scale method is used to understand which variations in which acoustic parameters should be implemented in order to trigger the appropriate affective, attentional or motor response. While the method is commonly applied in alarm design (e.g.: trendsons [5]), disciplines such as sound design for products [6] or music theory [7] are also interested in knowing exactly which acoustic structure originates which affective response.

In the auditory alarm design context, early work by Roy D. Patterson [8], [9], Judy Edworthy and Elizabeth Hellier [10]–[12] has set the fundamental work grounds to understand the perception of urgency. However, not all auditory warning signals are associated with urgent events, and thus the same work needs to be made to comprehend which acoustic parameters might trigger, for instance, irritability, preoccupation, unpleasantness or others - depending on their context and adequate response. This knowledge will allow designing more appropriate audio alarms or warning signals for environments heavily populated with alarming sounds, such as control rooms, intensive care units or operating theatres.

The purpose of this study is to use a semantic differential scale methodology to understand which acoustic parameters of simple computer-generated sounds have an effect on perceived urgency, pleasantness, irritability, preoccupation, speed, and positiveness. Its specific aim is to create a predictive model that indicates which acoustic parameters (spectral or temporal) activates the subjective perceptions mentioned above. In the future, these findings will be used for the design of auditory warning signals from medical devices.

2. METHOD

The selected methodology was based on previous studies of Kansei Engineering [13] and semantic differential scales applied to psychoacoustic studies [10], [11], [14].

2.1. Selection of representative pairs of words

When using semantic differential scales, it is of extreme importance to select pairs of words that can adequately describe the object under evaluation [4]. For this, in a pilot study, people were asked to suggest words they associated with artificial sounds, in all possible contexts. Any words, adjectives or not, were accepted. Examples of sounds were referred, such as sounds from household devices, electronics, sounds from inside the vehicle, or alarms from queuing services. In total, 183 words were suggested that described sensations, emotions and perceptions evoked by sounds.
The most frequent words were *shrieking, loud, alert, irritating, deafening, confusing, noisy, pleasant, short,* and *sweet*. Other words related with a) physical properties (*low, short, long, fast, vibrant, synchronous, slow, repetitive, harmonious*); b) positive feelings (*relief, calm, curiosity, fresh, gentle, positive, relaxing, pleasant, melodic, peaceful, soft*); c) negative feelings (*boring, anxious, unpleasant, strident, fiddly, nervous, stressful, irritating, intrusive, angry, frustrating, penetrating*); d) other words (*critical, strong, important, order, respect, safety, attention, artificial*).
All were grouped considering similitude of meaning and frequency. This resulted in 11 words and corresponding negation. Then, five human factors and acoustics researchers selected the most fitted pairs to describe artificial sounds/alarms, resulting in 6 pairs. The final six pairs of words are in Table 1.

Table 1: *Pairs of Words used for evaluation.*

1	Not very - very Urgent
2	Unpleasant-Pleasant
3	Not very - Very Irritating
4	Not very - Very Preoccupying
5	Slow - Fast
6	Negative - Positive

All pairs of words were presented in an analog visual scale, ranging from 0 to 100 mm without numbers (*Figure 2)*

2.2. Selection of acoustic parameters

This phase consisted in selecting the acoustic parameters to be manipulated, so in the evaluation phase they could be paired with the chosen pairs of words. Two types of parameters were selected: spectral and temporal characteristics of sound. Four acoustic parameters were analysed in the present study:

1) *Frequency*: referred to by Hertz (Hz) where 1 Hz is one cycle per second.
2) *Amplitude Envelope*: the shape of a waveform's intensity throughout time. Rise (onset) and fall (offset) times were edited in milliseconds (ms).
3) *Speed*: determined by the inter-pulse interval with faster bursts possessing shorter inter-pulse intervals.
4) *Rhythm*: regular occurrence of an auditory event in time. This occurrence can have a given pattern that can be cyclic, thus having periodicity.

A total of three levels were defined for Frequency, Speed and Onset. Rhythm had two levels. The objective was to have three different levels of priority, similarly to an emergency signal (level 1 in table 2), a warning signal (level 2) and an information notification (level 3). Table 2 depicts all levels for each parameter.
Values and directionality of the variations were established after literature and international standards on the design of audio warning signals, detailed in the following sections.

Table 2*: Levels of variation in each acoustic parameter*

	1	**2**	**3**
F0 Frequency	2500 Hz	1500 Hz	500 Hz
Speed	x4	x2	x1
Rhythm	Regular	Regular	Regular
	Syncopated 0	Syncopated 5	Syncopated 10
Onset	Regular	Slow onset	Slow offset

2.2.1. *Frequency*

The fundamental frequency of a signal should depend on the purpose and context of the signal. Whether it is an emergency or an information signal, or whether it is to be used in a public or private space. For instance, Begault and Godfroy [15] proposes a range between 300 Hz – 1000 Hz for NASA's crew exploration vehicles, while ISO 7731 [16] for danger signals proposes frequency components in the 500 Hz to 2 500 Hz range. Specifically for medical devices, IEC 60601- 1-8:2012 [17] proposes a frequency range between 500 Hz and 3 000 Hz. Because the aim of this study is to help in the design of medical devices' audio alerts, three levels of the range suggested by [17] standard were chosen as a fundamental frequency.
All agree the auditory signals should have several harmonics. Begault and Godfroy [15] state that "*there should be four or more harmonically related spectral contents*"; IEC 60601- 1-8:2012 [17] and ISO 7731 [16] also propose four or more harmonics to improve spatial localization and signal audibility.
For this study, three levels of frequency (F0) were chosen: 2500 Hz, 1500 Hz, and 500 Hz. All had four harmonics.

2.2.2. *Speed*

ISO 7731 [16] recommends the temporal distribution of the signal to be pulsating rather than continuous in time; Patterson, Edworthy, and Lower [9] mention speed as the main variable for the perception of priority. ANSI/ASA S3.41 [18] recommends a temporal pattern of three 1-s pulses with 1.5s silence; ISO 9703-1:1994 [19] (this standard has been withdrawn) proposed multiple pulses with an interval between of 0.15- 0.5 s, depending on the priority of the alarm. Similarly, IEC 60601- 1-8:2012 [17] proposes three different pulse duration patterns according to high, medium or low priority of the alarm, respectively 75 ms to 200 ms (high) and 125 ms to 250 ms (medium and low), but only mentions the interpulse interval should be "*speeding up > regular/slowing*".

For this experiment, the strategy adopted by Edworthy, Loxley, and Dennis [10] was applied by creating three levels of speed with a systematic relationship: the faster speed was twice the speed of the moderate one, which was twice the speed of the slower one. The temporal distribution of the "pulse + silence" was repeated three times when the speed was x1 (slow) and x2 (moderate), and it was repeated five times when speed was in x4 (fast). However, the silence duration differed according to speed. In x1 it had 1 s, in x2 it had 0.5 s and in x4 it had 0.25 s.

2.2.3. *Rhythm*

The standard IEC 60601- 1-8:2012 [17] suggests syncopated or "off-beat" rhythms for higher priority alarms and regular rhythms for medium and low priority alarms. Edworthy, Loxley, and Dennis [10] have found the inverse relation with syncopated rhythms being perceived as less urgent than a regular one.
For this study, the stimuli rhythm was based on the syncopation index of Fitch and Rosenfeld [20] (index 0, 5 and 10), and all stimuli were tested both with regular and syncopated rhythm.

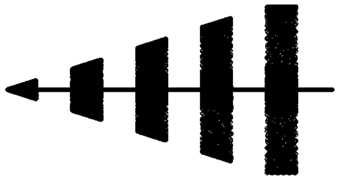
2.2.4. *Onset – Offset*

The rise and fall time of an auditory warning is defined in IEC 60601- 1-8:2012 [17] as "*the interval over which the pulse increases from 10 % to 90 % of its maximum amplitude*". While initially this standard proposed a rise time of 10 to 20% of the stimuli's total duration, a 2012 amendment changed this to allow for rise times of up to 40% of the total duration. Due to hardware constraints, this rise time should not be less than 10-ms long. The manipulation of rise times provide, according to the standard, more psychoacoustic cues of greater urgency, where rapid rise times are perceived as more urgent than slow rise times. Edworthy, Loxley, and Dennis [10] found that a regular 20 ms onset was considered more urgent than a pulse with a slower onset.
In this study, stimuli had either a slow onset (180 ms; offset of 20 ms), a regular onset and offset (20 ms) or a slow offset (180 ms; onset of 20 ms).

2.3. Auditory Stimuli

In order to test all variables, a combination of all parameters was performed, generating 54 stimuli (3 Frequency x 3 Speed x 2 Rhythm x 3 Onset/Offset). All audio stimuli were generated in R Studio using Seewave [21] and TuneR [22] packages. A modular approach as first proposed by Patterson [8] and used in Edworthy, Loxley, and Dennis [10] was applied, where pulses were firstly created and then grouped into longer bursts of sound, which were then intercalated with periods of silence to form the full warning. All pulses were 200-ms long. Figure 1depicts two warning signals.

a)



b)

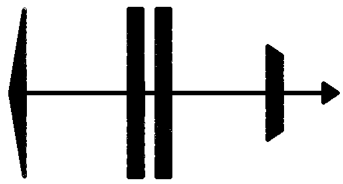


Figure 1: *Depictions of a) Stimuli with 180-ms onset, regular rhythm, speed level 4 (1500 Hz); b) Stimuli with 20-ms onset, 180-ms offset, syncopated rhythm, speed level 1 (500 Hz))*

2.4. Participants

Twenty-six participants took part in the study (17 female, 9 male), from 20 to 50 years (M=33, SD=10), all with normal hearing and most (22) without formal musical education. Data collections were carried out in two geographical locations in order to gather a higher number of participants, using the same equipment.

2.5. Apparatus

The study took place in a quiet room, where the participant was seated in front of a display and made the sound evaluation using a computer mouse by clicking on the visual analog scale. Participants used AKG Pro Audio K271 MKII headphones and all stimuli were presented using PsychoPy [23] software running on a Lenovo G500s with a 3rd generation Intel® Core™ i7-3612 processor and a Conexant Audio HD. Audio stimuli were presented in 77 dB SPL.

2.6. Procedure

Participants were welcomed and explained the main objective of the study, which consisted in evaluating several sounds according to a set of properties. They sat in front of a screen and placed the headphones. There was one participant per experimental session. After signing an informed consent and answering demographic questions, the instructions were given by the experimenter. These referred that after presenting a sound, an adjective was going to be presented, and participants should evaluate that sound according to that adjective. There were a total of six adjectives, and participants were told they should evaluate how much the sound was pleasant or unpleasant, irritating or not, preoccupying or not, slow or fast, urgent or not urgent and, finally, negative or positive (Table 3).
Participants were told they could only make the evaluation after hearing the entire sound once, which could last between 2 to 5 seconds. Participants could navigate with the mouse on the line of the scale, but after clicking with the mouse, it could not be changed. The scale was a continuous 100-mm scale. Before starting the experiment, all participants went through a training phase with the same scales and four different sounds (from [17]).

Table 3: *Descriptors per pair of words. A sheet with this information was always near the participant*

Pair of words	Description
Unpleasant Pleasant	I dislike the sound and it bothers me// I like the sound and it does not bother me
Not Irritating Very Irritating	The sound does not make me feel irritated and impatient// The sound makes me feel irritated and impatient
Not Preoccupying Very Preoccupying	The sound does not make me feel worried and alarmed// The sound makes me feel worried and alarmed
Slow Fast	The sound has a slow pace// The sound has a fast pace
Not Urgent Very Urgent	The sound communicates a need that may not be immediate// The sound communicates an immediate need
Negative Positive	The sound communicates a negative information// The sound communicates a positive information

After, the experimental session began, the screen displayed one pair of words at the time (Figure 2). The presentation of sound files was randomized, as well as the presentation of the pairs of words.

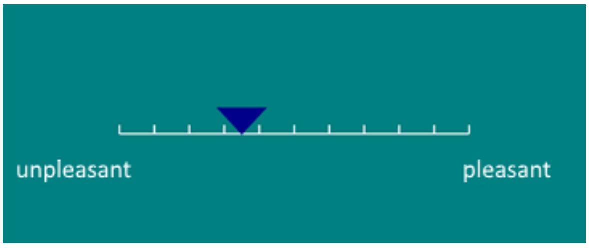


Figure 2: *Image of the evaluation interface with the semantic differential scale under evaluation*

Due to the great number of stimuli to be evaluated, there were two intervals, which had the length the participant preferred. The total procedure lasted between 30 to 40 minutes. Each participant evaluated each stimuli once for each pair of words. In total, each participant made 324 evaluations (54 stimuli x 6 pairs of words).

3. RESULTS

Because participants did not repeat the evaluation, it was important to assess the degree of agreement between participants as raters of a given stimuli. Outliers were removed from the sample using Tukey's method due to its independency from data distribution. This method ignores the mean and standard deviation, which are influenced by the outliers, by using an inter-quartile range approach (above and below the 1.5*IQR).

3.1. Inter-participant concordance

Kendall's W (also known as Kendall's coefficient of concordance) is a non-parametric statistic and can be used for assessing agreement among raters. Kendall's W ranges from 0 (no agreement) to 1 (complete agreement). The value of Kendall's W was calculated per pair of words to verify if the stimuli were rated in more or less the same order per participant. The results are in *Table 4*. All tests revealed a significant value of Kendall's W. As expected, because it was the most objective adjective, the pair "Slow-Fast" obtained the highest value of concordance, followed by "Not Urgent – Urgent".

Table 4: Values of Kendall's W

	Kendall's W
Slow – Fast	0.70 ***
Not Urgent - Very Urgent	0.61 ***
Not Irritating - Very Irritating	0.45 ***
Not Preoccupying -Very Preoccupying	0.42 ***
Unpleasant – Pleasant	0.36 ***
Negative – Positive	0.12 ***

*** *Significant ($p < .001$) ** ($p < .01$) * ($p < .05$)*

This analysis only shows consistency, and does not reveal the nature of the classification made by the participants. For this purpose, correlational (Table 5) and linear regression analysis were performed after all data was pooled and averaged.

Table 5: *Correlations between the four acoustic parameters and the six pairs of words*

	Frequency	Speed	Rhythm	Onset
Irritating	0.69 ***	0.13	0.27 *	0.04
Positive	-0.27 *	-0.70 ***	-0.18	0.02
Pleasant	-0.88 ***	-0.26	-0.14	0.04
Preoccupying	0.15	0.80 ***	0.31 *	-0.03
Urgent	0.11	0.83 ***	0.37 *	0.00
Fast	0.14	0.78 ***	0.41 ***	-0.01

	Irritating	Positive	Pleasant	Preoccupying	Urgent	Fast
Irritating	---					
Positive	-0.41 ***	---				
Pleasant	-0.82 ***	0.53 ***	---			
Preoccupying	0.40 ***	-0.88 ***	-0.48 ***	---		
Urgent	0.38 *	-0.86 ***	-0.45 ***	0.96 ***	---	
Fast	0.40 *	-0.85 ***	-0.47 ***	0.97 ***	0.98 ***	---

*** *Significant ($p < .001$) ** ($p < .01$) * ($p < .05$)*

The significant correlations found with *Frequency* were with Irritating ($r(52) = .69$, $p < .001$), Positive ($r(52) = -.27$, $p < .05$) and Pleasant ($r(52) = -.88$, $p < .001$); with *Speed*, the stronger correlations were with Positive ($r(52) = -.70$, $p < .001$), Preoccupying (r(52) = .80, p < .001), Urgent (r(52) = .83, p < .001), and Fast (r(52) = .78, p < .001); with *Rhythm* were Irritating (r(52) = .27, p < .05), Preoccupying (r(52) = .31, p < .05), Urgent (r(52) = .37, p < .05), and with Fast (r(52) = .41, p < .001). No correlations were found with the acoustic parameter Onset-Offset. For this reason, this variable will not be used in further analysis.

Additionally, it can be seen that the pair of words Irritating correlated significantly with all other words, negatively with Positive and Pleasant. The pair of words Positive and Pleasant were negatively correlated with Preoccupying, Urgent and Fast. And Preoccupying was correlated with Urgent, and Fast.

Following this, all relations between acoustic parameters and pairs of words were explored using linear regression models (Table 6-9).

3.1.1. Frequency

The *Frequency* (Table 6) variable had three levels, and each level increased the perception of unpleasantness of our participants, with 500 Hz (B= 53.96, F = 100.2, R^2 = .80, p < .001) 1500 Hz (B= -15.55, p < .001) and 2500 Hz (B= -23.77, p < .001). A similar pattern was found regarding the perception of irritableness, with 500 Hz (B= 39.11, F = 29.43, R^2 = .54, p < .001) 1500 Hz (B= 19.82, p < .001) and 2500 Hz (β= 24.75, p < .001). No significant relations with Frequency were observed among the other pairs of words.

Table 6: *Results of linear regression by levels of Frequency (500 Hz, 1500 Hz and 2500 Hz). N = 54. 95% Confidence Interval (only* R^2 *> 0.5 are depicted)*

FREQUENCY	Unpleasant – Pleasant		Not Irritating - Very Irritating	
	B	**CI**	**B**	**CI**
Intercept (500)	53.96 ***	51.54 – 56.38	39.11 ***	34.26 – 43.96
1500	-15.55 ***	-18.97 – -12.12	19.82 ***	12.96 – 26.68
2500	-23.77 ***	-27.19 – -20.34	24.75 ***	12.96 – 26.68
F	100.2		29.43	
R^2	.797		.536	

**** Significant (p < .001) ** (p <.01) * (p < .05)*

These two regression models are plotted in Figure 3, with the y-axis depicting the 20-80 mm fraction of a 100-mm visual analog scale.

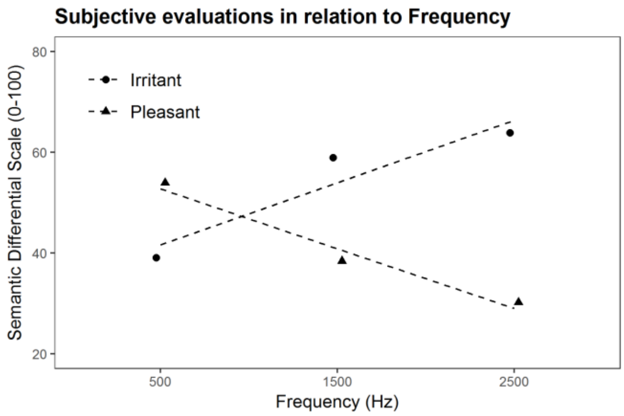


Figure 3*: Significant regressions for Frequency as predictor. Relationship between participant's evaluation of a sound as Irritable (*R^2 *= .54) or Pleasant (*R^2 *= .80) and three levels of increasing frequency.*

According to these results, the higher the sound's frequency, the more irritant and the less pleasant the sound is evaluated.

3.1.2. Speed

The *Speed (*Table 7*)* variable also had three levels and it was the variable which better explained the variance of four pairs of words. As the speed increased, so did the perception of Urgency (Speed x1 B= 31.70, p < .001, Speed x2 B= 23.82, p < .001 and Speed x4 B= 39.68, p < .001, F = 76.11, R^2 = .75): Preoccupation (Speed x1 B= 34.31, p < .001, Speed x2 B= 20.37, p < .001 and Speed x4 B= 30.78, p < .001, , F = 69.47, R^2 = .73) and Speed (Fast) (Speed x1 B= 32.95, p < .001, Speed x2 B= 27.33, p < .001 and Speed x4 B= 40.85, p < .001, , F = 63.9, R^2 = .72). The inverse pattern was observed in the Negative-Positive pair of words, with the perception of positiveness decreasing as the speed increased (Speed x1 B= 50.60, p < .001, Speed x2 B= -8.75, p < .001 and Speed x4 B= -11.95, p < .001, F = 37.9, R^2 = .60). No significant relations with Speed were observed among the other pairs of words.

Table 7: *Results of linear regression by levels of Speed (x1, x2, x4). N = 54, 95% Confidence Interval (only* R^2 *> 0.5 are depicted)*

SPEED	Not Preoccupying-Very Preoccupying		Slow – Fast		Not Urgent – Very Urgent		Negative – Positive	
	B	**CI**	**B**	**CI**	**B**	**CI**	**B**	**CI**
Interc. (x1)	34.3 ***	30.5 – 38.1	33.0 ***	27.7 – 38.2	31.7 ***	27.1 – 36.3	50.6 ***	48.6 – 52.6
x2	20.3 ***	15.0 – 25.7	27.3 ***	19.9 – 34.7	23.8 ***	17.3 – 30.3	-8.8 ***	-11.6 - -5.9
x4	30.7 ***	25.4 – 36.1	40.9 ***	33.5 – 48.2	39.7 ***	33.2 – 46.2	-12.0 ***	-14.8 - - 9.1
F	69.47		63.9		76.11		37.9	
R^2	.731		.715		.749		.598	

**** Significant (p < .001) ** (p <.01) * (p < .05*

The four regression models are plotted in Figure 4, with the y-axis depicting the 20-80 mm fraction of a 100-mm visual analog scale.

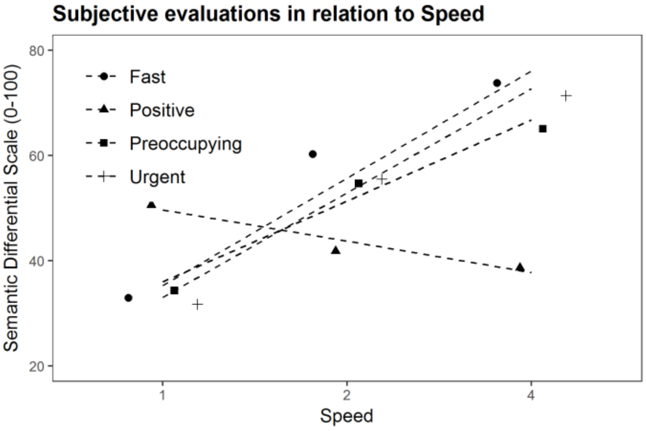


Figure 4: *Significant regressions for Speed as predictor. Relationship between participant's evaluation of a sound as Fast (R^2=.71), Positive (R^2=.60), Preoccupying (R^2=.73), or Urgent (R^2=.75) and three levels of increasing speed*

According to these results, the higher the sound's speed, the more urgent, fast and preoccupying and the less positive it is evaluated.

3.1.3. Rhythm

The *Rhythm (Table 8)* variable had two levels and the perception of speed (word fast) (Regular B= 47.47, $p < .001$, Syncopated B= 16.41, $p < .001$, F=10.4, $R^2 = .17$) and urgency (Regular B= 45.89, $p < .001$, Syncopated B= 13.95, $p < .001$, F=8.26, $R^2 = .14$) increased when the rhythm was syncopated.

Table 8: Results of linear regression by levels of Rhythm (Regular, Syncopated). N = 54, 95% Confidence Interval (only $R^2 > 0.5$ *are depicted)*

RHYTHM	Slow – Fast		Not Urgent – Very Urgent	
	B	CI	B	CI
Intercept (Regular)	47.47 ***	40.25 – 54.69	45.89 ***	39.00 – 52.78
Syncopated	16.41 **	6.20 – 26.62	13.95 **	4.21 – 23.69
F	10.4		8.26	
R^2	.167		.137	

**** Significant (p < .001) ** (p <.01) * (p < .05)*

Having significant regression coefficients means the *Rhythm* is correlated with both subjective perceptions, nevertheless, the model does not account for the variability found among the data.

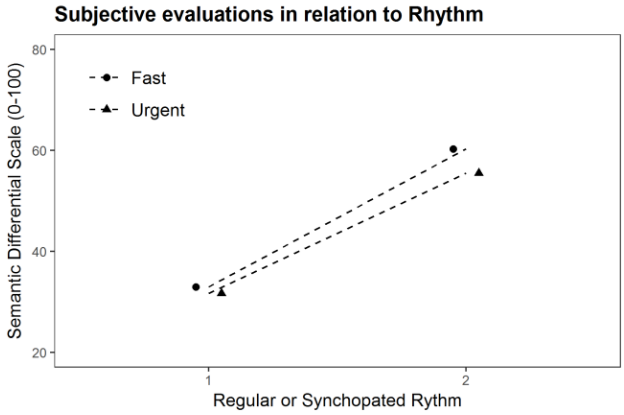


Figure 5: *Significant regressions for Rhythm as predictor. Relationship between participant's evaluation of a sound as Fast (R^2=.17) and Urgent (R^2=.14) and two levels of Rhythm (regular or syncopated).*

4. DISCUSSION

The obtained results demonstrate how the semantic differential scale methodology is robust and useful for the analysis of relations between subjective perceptions and physical acoustic parameters. Firstly, although some pairs of words were extremely subjective, like *unpleasant – pleasant*, and vague, like *negative-positive*, there was consistency among participants, revealed in the significant values of Kendall's *W* in all pairs of words. This was an important result, as it allows to somewhat balance an obvious limitation of this study, which was the lack of repetitions of the evaluation sessions. At first, one could expect large inter personal variability regarding such subjective perceptions, but these observations serve as an addition to the strengths of this simple method. It is important to add that during the data collection phase, some participants had informally mentioned they had trouble in classifying a sound as negative or positive, even though they had the definition sheet nearby. It is then somewhat surprising to understand that, although difficult, the classification was congruent among raters, later relating significantly to the manipulation of the *Speed* variable.

Regarding the associations between subjective and physical variables, with *Frequency*, it was observed that the subjective perceptions in which it had more effect were *Pleasantness* and *Irritability*. Again, although apparently a very personal evaluation, most participants found high-frequency audio signals as unpleasant and irritant. This is an important result that confirms that an alarm, to essentially fit its purpose of communicating an urgent event, does not need to increase its frequency. In fact, it should not, as it only affects the negative affective perception of the signal

Also importantly, and in agreement with Patterson's suggestions and standard norms, *Speed* is the variable which mostly affects an alarm's perception of urgency, communication of preoccupation or that "something" negative is happening. In applied settings, it is important to bear in mind that an increase in these subjective perceptions should be made via inter-pulse interval.

Rhythm obtained results also aligned with the [17] standard, with participants evaluating as significantly more urgent those auditory signals with syncopation than those with regular rhythm. However, the association found was not robust, and no more elations can be made. One explanation can be that the irregularity of the rhythm might have been affected by the slow onsets and offsets, not allowing to hear the full structure of the auditory signal.

Contrary to the literature and standards, the onsets and offsets of the auditory signals had no effect on the perception of any pair of words. In the future, the variations of this parameter should be more numerous, and evaluations should consider this manipulation only. This would allow clarifying the effect this parameter has without interacting with other manipulations.
With this study, it was possible to understand which acoustic features trigger what affective state when designing for auditory warning signals. For instance, that a signal to be understood as urgent should have shorter and irregular inter-pulse intervals, preferably with lower frequencies. However, these sound design recommendations must co-exist with other requirements such as the ability to localize audio warning signals in an open space, and the ability to recognize it among other devices with similar spectral and temporal patterns.

5. CONCLUSION

A study was performed to better understand the psychological correlates of acoustic parameters. Fifty-four stimuli were created manipulating frequency, speed, rhythm and onset and offset times. Twenty six participants listened to each stimuli six times, each time considering a different pair of words presented in a visual analog scale. These words were selected among more than a 100 sound-related words. The applied methodology consisted in using semantic differential scales. The findings allowed to consolidate this method as a good evaluator of subjective perceptions. Results have demonstrated that the acoustic features which most contribute to the perception of these states in audio stimuli are frequency (pleasantness and irritability) and speed (urgency, preoccupation and negativity). Rhythm also affected the perception of urgency, although to a lesser extent, with irregular rhythms obtaining higher ratings for the perception of urgency.
This was the first study intending to use a human-centred approach to the design of auditory warning signals. After these fundamental associations between acoustic parameters and subjective perception have been established, the next step will be to apply them in the design of better auditory warning signals for medical devices.

6. ACKNOWLEDGMENTS

This work was supported by grant no. POCI-01-0145-FEDER-031943, co-financed by COMPETE2020 under the PT2020 programme, and supported by FEDER

7. REFERENCES

[1] G. A. Gescheider, *Psychophysics: The Fundamentals*, Third Edit. London: Lawrence Erlbaum Associates, 1997.

[2] T. Lokki, H. Vertanen, A. Kuusinen, J. Pätynen, and S. Tervo, “Auditorium acoustics assessment with sensory evaluation methods,” *Proc. Int. Symp. Room Acoust. Melbourne, Aust.*, no. August, pp. 1–10, 2010.

[3] T. Lokki, “Tasting music like wine: Sensory evaluation of concert halls,” *Phys. Today*, vol. 67, no. 1, pp. 27–32, 2014.

[4] S. Ishihara, “Psychological Methods of Kansei Engineering,” in *Kansei/Affective Engineering*, M. Nagamachi, Ed. Boca Raton: CRC Press, 2011, pp. 31–38.

[5] J. Edworthy, E. Hellier, K. Aldrich, and S. Loxley, “Designing Trend-Monitoring Sounds for Helicopters: Methodological Issues and an Application,” *J. Exp. Psychol. Appl.*, vol. 10, no. 4, pp. 203–218, 2004.

[6] E. Özcan-Vieira, *Product Sounds - Fundamentals & Applications*. 2008.

[7] H. Von Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music.* Longmans, Green, 1912.

[8] R. D. Patterson, “Guidelines for Auditory Warning Systems on Civil Aircraft,” Eindhoven, The Netherlands, 1982.

[9] R. Patterson, J. Edworthy, and M. Lower, “Alarm sounds for medical equipment in intensive care areas and operating theatres,” London, 1986.

[10] J. Edworthy, S. Loxley, and I. Dennis, “Improving auditory warning design: relationship between warning sound parameters and perceived urgency.,” *Hum. Factors*, vol. 33, no. 2, pp. 205–231, 1991.

[11] 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.

[12] J. Edworthy, E. Hellier, and R. Hards, “The semantic associations of acoustic parameters commonly used in the design of auditory information and warning signals.,” *Ergonomics*, vol. 38, no. 11, pp. 2341–2361, 1995.

[13] J. Vieira, J. M. A. Osório, S. Mouta, P. Delgado, A. Portinha, J. F. Meireles, and J. A. Santos, “Kansei engineering as a tool for the design of in-vehicle rubber keypads,” *Appl. Ergon.*, vol. 61, 2017.

[14] E. Hellier and J. Edworthy, “On using psychophysical techniques to achieve urgency mapping in auditory warnings,” *Appl. Ergon.*, vol. 30, no. 2, pp. 167–171, 1999.

[15] D. Begault and M. Godfroy, “Auditory Alarm Design for NASA CEV Applications,” in *13th International Conference on Auditory Display*, 2007, pp. 131–138.

[16] ISO, *ISO 7731 - Ergonomics - Danger signals for public and work areas — Auditory danger signals*, vol. 2003. 2003.

[17] AAMI, *ANSI/AAMI/ IEC 60601- 1-8:2006 & A1:2012 MEDICAL ELECTRICAL EQUIPMENT – Part 1-8: General requirements for basic safety and essential performance – Collateral Standard: General requirements, tests and guidance for alarm systems in medical electrical equip*. 2013.

[18] A. S. of America, *ANSI/ASA S3.41 Audible Emergency Evacuation Signal*. 2015.

[19] I. O. for Standardization, *ISO 9703-1:1994 - Anaesthesia and respiratory care alarm signals*. 1994.

[20] W. T. Fitch and A. J. Rosenfeld, “Perception and Production of Syncopated Rhythms,” *Music Percept. An Interdiscip. J.*, vol. 25, no. 1, pp. 43–58, Sep. 2007.

[21] J. Sueur, T. Aubin, and C. Simonis, “SEEWAVE, a free modular tool for sound analysis and synthesis,” *Bioacoustics*, vol. 18, no. 2, pp. 213–226, Jan. 2008.

[22] U. Ligges, S. Krey, O. Mersmann, and S. Schnackenberg, “tuneR: Analysis of music,” 2016.

[23] J. W. Peirce, “PsychoPy—Psychophysics software in Python,” *J. Neurosci. Methods*, vol. 162, no. 1–2, pp. 8–13, May 2007.