eVADER: A Perceptual Approach to Finding Minimum Warning Sound Requirements for Quiet Cars.

Ryan Robart, Etienne Parizet, Jean-Christophe Chamard, Karl Janssens, Fabio Biancardi, Joseph Schlittenlacher, Wolfgang Ellermeier, Perceval Pondrom, James Cockram, Paul Speed-Andrews, et al.

eVADER: A Perceptual Approach to Finding Minimum Warning Sound Requirements for Quiet Cars.

Ryan Robart[1], Etienne Parizet[1], Jean Christophe Chamard[2], Karl Janssens[3], Fabio Bianciardi[3], Josef Schlittenlacher[4], Wolfgang Ellermeier[4], Perceval Pondrom[4], James Cockram[5], Paul Speed-Andrews[5], and Gemma Hatton[5]

[1] *INSA-Lyon, 69100 Villeurbanne, France, Email: ryan.robart@insa-lyon.fr*
[2] *PSA Peugeot Citroën*
[3] *LMS International*
[4] *Technische Universität Darmstadt*
[5] *Nissan*

Introduction

The danger of quiet cars, such as hybrid and electric vehicles, to pedestrians has become an important issue for public policy [1], car manufacturers as well as the scientific community. The eVADER project (Electric Vehicle Alert for Detection and Emergency Response) has taken a progressive approach to this problem by using psychoacoustics to try to establish some basic guidlines for adding sound to such vehicles. Project contributors include: car manufacturers; technological , material, & engineering companies; academic laboratories; and the European Blind Union, which is a non-profit organization dedicated to the safety and quality of life of the visually-impaired people of Europe. Requirements for potential warning sounds included the following premises:

1. Sounds should be easily detected by pedestrians despite a range of background noise.
2. Sounds should add as little noise as possible to the environment in the interest of avoiding noise pollution and annoyance.
3. Sounds should convey information regarding vehicle dynamics, such as speed, to pedestrians (especially the visually-impaired).

The research presented here focused specifically on premise 1, with special consideration given to the loudness of the warning sounds as required by premise 2. Another experiment which will focus directly on annoyance (premise 2), is planned but has not yet begun. Experiments focused on premis 3 are currently underway and will not be included in this report. Past research has used a localization paradigm to test the detectability of hybrid and internal combustion cars by measuring listener reaction times to binaural recordings of such vehicles [2]. A similar method was adopted in this experiment with some important modifications. These modifications are associated with creating a virtual environment that seemed as natural as possible to listeners while gathering data. Also, in the interest of testing prototypical sounds, these sounds had to be designed according to a quantifiable structure, and then added to the virtual environment.

The primary goal of this experiment was to test if some distinct features of potential warning sounds might be heard more easily than others if the overall level is held constant among all sounds. In order to derive the appropriate features to test, it was necessary to review related literature from various areas of study. It is beyond the scope of this document to completely review such research. Still it should be noted that the psychoacoustic research regarding listener sensitivity and masking was considered [3]. Perceptual research regarding auditory scene analysis was also taken into account [4]. Perhaps most importantly, the literature focused on alarm sound design and perception was also instructive in our endeavour [5-12]. As a result, the following three features were selected for testing: tonal content; frequency detuning; and amplitude modulation.

Since it is likely that warning sounds composed of broadband noise would difficult to segregate from a noisy environment [3-4], all sounds were composed of a specified number of harmonic sinusoidal harmonic tones. Each of the proposed sound-features (factors) varied according to 3 levels, as displayed in table 1.

Factors (Sound Features)	Levels 1	2	3
Frequency Detuning:	Harmonic: no frequency detuning	Sinusoid: 2 highest harmonics (+/- 75 Hz) at 4 & 5 Hz	Sawtooth: 2 highest harmonics (+/- 75 Hz) at 4 & 5 Hz
Tonal Content: f0 = 300 Hz (pure tones)	3 harmonics	6 harmonics	9 harmonics
Amplitude Modulation:	Flat: no amplitude modulation	Periodic: All bands modulated between 20% and 100% SPL at 8 Hz	Irregular: 4 distinct envelopes of variable fs, intermittently applied

Table 1: Description of stimulus dimensions of for each level of each sound feature selected to be tested.

These 9 prototype sounds were layered onto a recording of an electric vehicle to emulate an electric vehicle with added sound emitting from a loudspeaker (EV+S). Adhering to the localization paradigm used by Robart & Rosenblum [2], recordings of an electric vehicle (EV) and a normal car (Diesel) were included as reference stimuli. Thus, there were a total of 11 stimuli used in the experiment. General predictions were that the Diesel should be detected earlier and more accurately than the EV. Furthermore it was predicted that there could be some ambiguity among the EV+S's if the timbre factors (frequency detuning, tonal content, and amplitude modulation) have varying impact on perceptibility.

Methods & Materials

Experiment Design

In a fractional repeated measures design [13-14] 11 stimuli were presented 8 times each, 4 repetitions from each direction. A fractional design was chosen primarily for two reasons. The first was the fact that all combinations of factors (3 x 3) would require 27 distinct stimuli. Since each stimulus lasted 10.8 seconds, it was determined that the experiment would likely be uncomfortably long with 27 stimuli. The second reason for using a fractional design was due to the novelty of our application, and a lack of research that specifically designates a paradigm to manipulate the chosen factors (frequency detuning, tonal content, and amplitude modulation). As a result, Taguchi's method [13] was chosen as the best way to begin to choose the combinations of factors. Using an orthogonal array (table 2), we could designate which factor combinations to use as the sounds in the EV+S's.

Stimulus Code	Levels of Factors		
	Frequency Detuning	Tonal Content	Amplitude Moduation
111	1	1	1
122	1	2	2
133	1	3	3
212	2	1	2
223	2	2	3
231	2	3	1
313	3	1	3
321	3	2	1
332	3	3	2

Table 2: Orthogonal array used to choose factor X level combinations used for the design of stimuli. The stimulus code for each sound refers to the level combinations of each factor as shown above.

Sound Design

Sounds were constructed using a custom synthesizer (Max/MSP) and normalized to ensure that all sounds had the same overall level. As previously stated, binaural recordings of an electric and diesel cars were used as reference stimuli. These recordings consisted of a single car approaching and eventually passing directly in front of a dummy-head. The track was 60 meters and the dummy head was situated in the middle (30 meters). The speed of the cars was 20 km/h, which remained constant (figure 1). This scenario had been selected as one of the most dangerous situations for pedestrians in a previous eVADER workpackage [15].

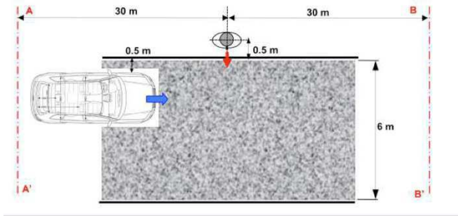

Figure 1: This image depicts illustrates both the recording procedure and the 'waiting to cross' scenario used in the experiment. In the recording procedure, a dummy-head was centrally located facing the track (red arrow).

The nine warning sounds were spatially synthesized so that they could be layered onto the recording of the electric vehicle. This was done to create the impression that the warning sounds were emitted from a loudspeaker mounted on the car. Recordings of a busy freeway and a rainstorm were mixed to create a realistic, and challenging background context. The level of the background set at approximately 65 dBA. This level was chosen to avoid making the task of hearing the electric vehicle too difficult (floor effects). The peak dBA levels for all stimuli is displayed in figure 2 .

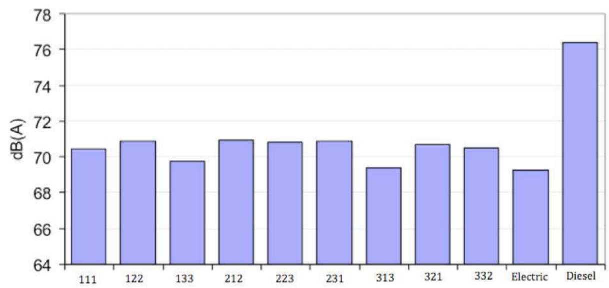

Figure 2: Peak levels of the 11 stimuli used in the experiment. The labels on the x axis describe the stimuli. The numbered stimuli, such as '111' etc., contain a coding that describe a certain combination of the levels of the three independent variables: frequency modulation; tonal content; amplitude modulation.

Participants & Procedure

Ninety-one participants (aged 20-72) participated in the experiment and 33 were visually impaired. The experiment was conducted in three separate laboratories, but the procedure was maintained in every case. Participants were instructed to imagine that they wished to cross a somewhat busy street during a rainstorm. They were told that their task was to indicate as quickly and accurately as possible when they heard a car approaching them by pressing buttons on the keyboard. The space-bar indicated a leftward approach, while the enter key indicated a rightward approach. The participants were also told that in order to attain a realistic presentation, the cars would appear in a somewhat random fashion. That is, the onset of each stimulus was randomized between 1-20 seconds from the offset of the previous stimulus. So, sometimes the next trial would begin 1 second after the previous, or it could begin up to 20 seconds after the previous trial. Each participant received a unique presentation of stimulus order and inter-trial intervals. Participants were familiarized with all sounds and underwent a short training exercise before the actual experiment. The experiment was split into 2 blocks, 44 trials in each. After the first block, participants were given the option to take a short break if needed. The entire procedure lasted roughly 45-60 minutes.

Materials

The experiment was conducted in a quiet room, and sounds were delivered via Stax electrostatic headphones and amplifier system. The stimuli and background sounds were presented using Delphi software, which controlled the timing of events as well as recording data. Both reaction time and accuracy were measured as a result.

Results

Data from participants (7) were rejected if more than 40% of the stimuli were missed and/or erroneously localized.Between subject differences were quite high. In order to reduce the variance due to this, the reaction time data was centered on individual means to reduce variability (figure 3).

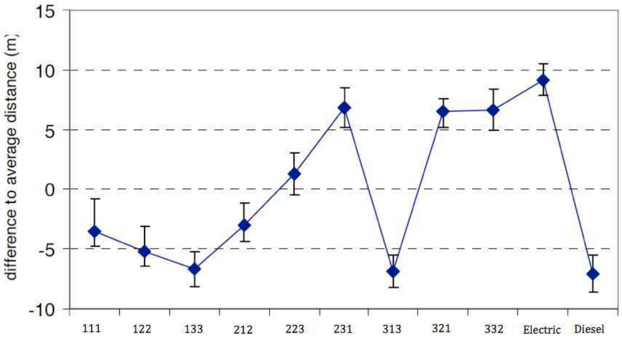

Figure 3: The diamonds indicate the mean reaction times for all participants for each stimulus. The standard error was made to be equal among the stimuli by centering the data as you can see by the error bars.

No differences were found between visually impaired and sighted people. Furthermore, there was no difference between direction of approach, or different ages. The fractional 3(frequency detuning) X 3(tonal content) X 3 (amplitude modulation) ANOVA produced main effects for all factors ($p < .05$). However, these results are difficult to interpret, because interactions cannot be accounted for in a fractional design. However, it seems that the factor of amplitude modulation had the most discernable impact on detectability (figure 4). In that, as the amplitude modulation increased from level 1 to level 3, listeners reliably detected the vehicle faster.

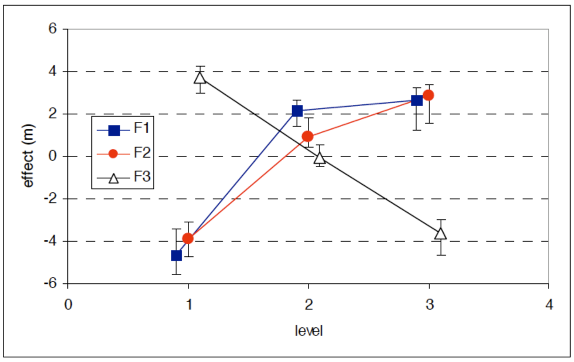

Figure 4: The interaction plot displays the differential effects of each factor at each level (faster is < 0, and slower is > 0). Zero may be thought of as the overall average reaction time. F1 refers to frequency detuning, F2 refers to tonal content and F3 refers to amplitude modulation.

Based on the data, it seems that different factor combinations are heard earlier than others. More specifically, the results suggest that the best warning sound should have a temporally irregular amplitude modulation structure (level 3); and contain few harmonics (level 1), and contain no frequency detuning (level 1). However, this must be concluded with caution as the interactions cannot be explained in a fractional design. Furthermore, a sound with all 3 of those features (*113*) was not tested as the *113* combination was not specified by the orthogonal table (figure 2). Still, it is evident that 2 of the EV+S's (*133*; *313*) were detected as early as the Diesel.

Accuracy was measured by analysing correct vs. incorrect detections of the direction of approach. Interestingly, *313* produced fewer errors than any other sound with exception to the Diesel. A post-hoc t-test showed that the difference in average errors (*313* vs. *Diesel)* was not significant $t_{(83)} = 1.34$ ($p = .18$). Still the Diesel produced twice as many errors as the *313* sound (figure 5).

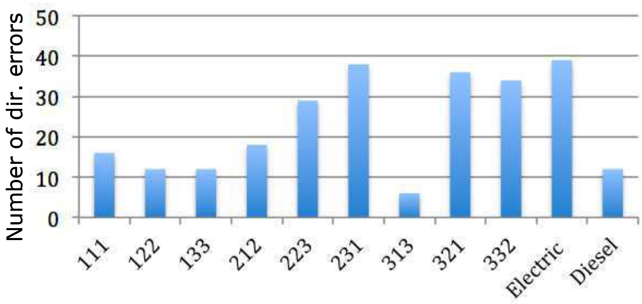

Figure 5: This graph represents all directional errors (no misses) for all 84 participants, separated according to stimuli. Stimulus *313* is the only sound that produced less (50% fewer) errors than the Diesel.

A miss occurred when the participant failed to detect a stimulus during a trial. Post-hoc t-tests showed that the EV produced more misses than all other stimuli ($p < .05$). There were no significant differences in misses among all other stimuli.

Conclusions

The general prediction that the EV would be heard later and less accurately than the Diesel was confirmed based on post-hoc comparisons of detection and accuracy. This can be seen a replication of previous work [1], and speaks to the strength of the general method. Regarding the EV+S's, the results indicate that the independent variables chosen (frequency detuning, tonal content, and amplitude modulation) have differential effects on the perceptibility of warning sounds. Based on the fractional ANOVA, and post-hoc comparisons regarding detection and accuracy, it was concluded that sound 313 is the best combination of the factors tested for use as a quiet warning sound in eVADER's prototype system.

Interestingly, it seems that our incremental manipulation of the factors among the different stimuli significantly effected the detectability of the vehicles. These results cannot be due to SPL dBA (see figure 2). Overall, the results suggest that, with careful sound design, warning sounds need only add 2 dBA to an electric vehicle to be at least as detectable as a normal car. This is likely due two factors:

1. temporally irregular amplitude modulation (level 3 amplitude modulation)
2. few harmonics (level 1 tonal content)

The results also suggest that the harmonic structure should not contain frequency detuning. But the best EV+S tested (*313*) did contain a saw-tooth detuning of the highest two harmonics, so it is unclear if this is true. Certainly, further tests need to be conducted to test the strength of the suggested combination(s), and perhaps more importantly, to find the constraints of potential manipulations.

In light of the recently proposed minimum requirements for sounds added to quiet cars [1], these results might seem surprising. For example, NHTSA recommends that warning sounds should be primarily composed of filtered, broadband noise. This could make figure-ground segregation difficult for listeners, especially if the warning sounds are emitted at low-levels. Furthermore, it is likely that adding broadband noise to cars would contribute to noise pollution. Moreover, while the current minimum requirements do not disallow amplitude modulation, it is certainly not required. Based on the results of this experiment, it is clear that there should be a requirement for a minimum amount of audible amplitude modulation. Future research should include perceptual tests of warning sounds based on the minimum requirements proposed by NHTSA. Hopefully, the research presented here will help increase the growing knowledge base of vehicle sound perception, as well as the use perceptual research when making such guidelines in the future.

Aknowledgements

The authors acknowledge the funding of the EC for the project eVADER (Electric Vehicle Alert for Detection and Emergency Response), grant agreement # 285095. The work published within reflects only the opinion of the authors and should not be affiliated with any other institution, public or private.

References

[1] U.S. Department of Transportation National Highway Traffic Safety Administration (2013). Minimum Sound Requirements for Hybrid and Electric Vehicles. *Docket Number NHTSA-2011-0100*.

[2] Robart, R. & Rosenblum, L.D. (2009). Are hybrid cars too quiet? *Proceedings oft he 157th meeting oft he Acoustical Society of America:* Paper 5aNS8.

[3] Moore, B.C.J (2003). *An Introduction to the Psychology of Hearing.* 5th ed. Academic Press, San Diego, Ca.

[4] Bregman, A, (1990). *Auditory scene analysis: The perceptual organization of sound.* CMA: The MIT Press.

[5] Misdariis, N., Cera, A., Levallois, E., Locqueteau, C. (2012). Do electric cars have to make noise? An emblematic opportunity for designing sounds and soundscapes. *Proceedings of the Nantes Acoustic Conference* 23-27 April 2012, Nantes, France.

[6] Edworthy, J., Hellier, E., Titchener, K., Naweed, A., Roels, R. (2011). Heterogeneity in auditory alarm sets makes them easier to learn. *International Journal of Industrial Ergonomics*, 41, 136-146.

[7] Parizet, E. (2011). Signaux d'alerte : analyse bibliographique. Technical Report to CETIM, 2011.

[8] Tabata, T., Konet, H., Kanuma, T. (2011). Development of Nissan Approaching Vehicle Sound for Pedestrians. *Proceedings of the InterNoise Conference,* Osaka Japan.

[9] Catchpole, K.R., Mckeown, J.D., Withington, D. J. (2004). Localizable auditory warning pulses. *Ergonomics*, *47* (7), 748 – 771.

[10] Edworthy, J. & Hards, R. (1997). Learning auditory warnings: The effects of sound type, verbal labeling and imagery on the identification of alarm sounds. *International Journal of Industrial Ergonomics*, *24*, 603-618.

[11] Edworthy J., Loxley S., Dennis I. (1991). Improving auditory warning design: relationship between warning sound parameters and perceived urgency. *Human Factors* , *33* (2), 205-231.

[12] Patterson, R.D. (1990). Auditory warning sounds in the work environment. *Philosophical Transcriptions of the Royal Society of London B 327*, 485-492.

[13] Taguchi G, Konishi S. (1987). *Taguchi methods. Orthogonal arrays and linear graphs. Tools for Quality Engineering.* Dearborn (MI): American Supplier Institute Press; 1987.

[14] Koehl, V., Parizet, E. (2006). Influence of structural variability upon sound perception: Usefulness of fractional factorial designs. *Applied Acoustics, 67*, 249–270.

[15] Garcia, J. & Baudet, G. (2012). Acoustic measurement of close to accident scenarios. *Deliverable of eVADER project.*