

Exterior Sounds for Electric and Automated Vehicles: Loud is Effective

August 2023

Pavlo Bazilinskyy^{a,*}, Roberto Merino-Martinez^b, Elif Özcan^c, Dimitra Dodou^d, Joost de Winter^d

^aFuture Everyday, Department of Industrial Design, Eindhoven University of Technology, Groeneloper 3, 5612 AZ Eindhoven, The Netherlands

^bFaculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands

^cFaculty of Industrial Design, Delft University of Technology, Landbergstraat 15, 2628 CE Delft, The Netherlands

^dFaculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

Abstract

Exterior vehicle sounds have been introduced in electric vehicles and as external human-machine interfaces for automated vehicles. While previous research has studied the effect of exterior vehicle sounds on detectability and acceptance, the present study takes on a different approach by examining the efficacy of such vehicle sounds in deterring people from crossing the road. An online study was conducted in which 226 participants were presented with different types of synthetic sounds, including sounds of a combustion engine, pure tones, combined tones, and beeps. A vehicle moving in a straight trajectory at a constant velocity of 30 km/h was used, and no visual information was provided. Participants in the role of pedestrians were asked to hold down a key when they felt safe to cross. After each trial, they assessed whether the vehicle sound was easy to notice, whether it gave enough information to realize that a vehicle was approaching, and whether the sound was annoying. The results showed that sounds of higher modeled perceived loudness, such as continuous tones with high frequency, were the most effective in deterring participants from crossing the road. The tested intermittent beeps proved to be relatively ineffective, presumably because no valuable information could be derived during the inter-pulse intervals. Tire noise proved to be relatively effective in deterring participants from crossing, and was found to be the least annoying. These results may prove insightful for the improvement of synthetic exterior vehicle sounds.

Keywords: electric vehicle, synthetic sound, road safety, listening experiment, exterior vehicle sound

1. Introduction

More than 270,000 fatal pedestrian traffic accidents occur annually worldwide [1], the majority of which occur during road crossing [2]. Causes of pedestrian-vehicle accidents include underestimation of the crossing gap or the time needed to cross [3], low visibility [4], and visual obstruction of the approaching vehicle [5]. Augmenting the sounds emitted by vehicles could potentially help prevent unsafe crossing. Such solutions have been the topic of investigation for electric vehicles (EVs) as well as in the form of external human-machine interfaces (eHMIs) for automated vehicles (AVs).

1.1. Sound design for electric vehicles

The market penetration of EVs is increasing, and countries worldwide have announced plans to cease sales of new internal combustion engine vehicles (ICEVs) between 2025 and 2035 [6-10]. However, one of the issues with EVs is that the lack of combustion noise can make EVs so quiet that they may become unsafe for vulnerable road users (VRUs) with visual impairments [11, 12]. Therefore, in recent years, legislation in Europe [13, 14] and the United States [15] has been introduced to address the inherent quietness of EVs at low speeds. The law specifies that, up to 20 km/h in Europe and 30 km/h in the USA, EVs and hybrid vehicles must emit a minimum number of decibels, distributed across a minimum number of frequency bands. Beyond these speeds, the noise generated by the tires and aerodynamic drag is considered sufficient, making synthetic noise unnecessary. Furthermore, the legislation stipulates additional requirements. The sound must be dependent on the speed [14, 15], and within the European legislative framework, the sound must be continuous [13]. In Europe, there is even an allowance for the inclusion of specific preferred sounds selectable by the driver [13]. In summary, the current legislation aims to ensure safety through synthetic sounds but also provides some room for creativity, allowing automobile manufacturers to create vehicle-specific branding through the selection of different frequencies [13].

Previous studies have explored various types of artificial sounds for EVs, including engine sounds, hums, whistles, beeps, white noise, as well as music/melodious sounds [16-18]. The effect of spectral characteristics has been examined, encompassing aspects such as volume, pitch, frequency, and modulation of frequency and amplitude [12-22]. Additionally, the influence of background noise has been explored, including traffic noise in different conditions [12, 19, 23-27], crowded areas [24, 26], as

well as pink and white noise [23, 28]. Outcome measures included detectability measured with a response key [12, 18, 20, 22, 26-29] as well as subjective qualities such as perceived pleasantness [18, 27, 30], appropriateness [27], authenticity [22], alertability [22], powerfulness [18], irritability [22], and annoyance [28, 29]. The results showed that higher frequencies, particularly when background noise is present [19, 21], frequency modulation [20], and irregular amplitude modulation [12] resulted in better detection. Furthermore, higher frequencies [22], saw-tooth signals [22, 28], and louder signals [29] tend to be perceived as more annoying. Naturally, loudness also affects detectability [31].

Only a few studies have investigated the relationship between detectability and annoyance in EVs, and which sound characteristics may result in high detectability, yet low annoyance. In an experiment with 30 participants, Lee et al. [28] reported that sounds with amplitude and frequency modulation led to faster detection and lower perceived annoyance than saw-tooth signals. Petiot et al. [27] used an interactive genetic algorithm to develop sounds: 15 assessors rated sounds in terms of their detectability and pleasantness. The sounds were weighted combinations of four components (a thermic motor sound, a harmonic sound, and two types of filtered broadband noises). Two different filters were also applied to the final sound. In total, more than 70 parameters were adjustable, including the frequencies and amplitudes of the four components. The genetically evolved sound was found to result in statistically significantly higher ‘fitness’ (i.e., a combination of higher detectability and lower unpleasantness) than sounds developed by human designers instructed to develop sounds satisfying these two criteria. It was also found that including the sound of a motor is important for detectability.

The effect of sounds on participants’ willingness to cross has not been extensively investigated. An exception is the on-road study by Wall Emerson et al. [32], in which visually impaired participants were standing next to a road with approaching traffic and asked to indicate their willingness to cross by pressing and releasing a button. The authors reported that hybrid vehicles switching to internal combustion soon after the start accelerating from a stop were considerably more detectable than hybrid vehicles switching to internal combustion later (i.e., after having reached a higher speed), indicating the importance of engine sound as a cue of approaching traffic.

1.2. Sound design for automated vehicles

Next to EVs, auditory signals could be useful in the form of auditory external human-machine interfaces (eHMIs) that communicate the state or intention of automated vehicles (AVs) to VRUs. The majority of eHMIs developed so far are visual, but a number of auditory eHMIs have also been proposed (see [33] for an overview). Mahadevan et al. [34] used the verbal messages “I see you” and “cross” together with visual and tactile (via a mobile phone) feedback and found that a combination of modalities improved pedestrians’ awareness of the approaching AV compared to single modalities. Deb et al. [35] tested a horn, music, and the verbal message “safe to cross” and found that verbal messages were preferred over the abstract sounds tested. Music was tested by Florentine et al. [36], who noted that it helped draw pedestrians’ attention. Inspired by research on EV sounds, Moore et al. [37] proposed the use of a synthetic engine sound to indicate the intention of a driverless AV to stop in front of a pedestrian. The authors tested the concept with a hybrid vehicle in a Wizard-of-Oz naturalistic setting. The results showed that participants in the role of pedestrians rated the clarity of the AV’s intention higher in the presence of engine sound compared to a condition without sound.

While the mandatory exterior vehicle sounds for EVs are primarily intended to ensure that EVs are audible in the same manner as ICEVs, serving as a compensation for the intrinsic quietness of EVs, this implementation does not exploit the full potential of what can be achieved with sound. Just as visual eHMIs on AVs provide the opportunity to offer additional information that would otherwise be invisible to pedestrians [38], external vehicle sounds can be more than just mimics of traditional ICEVs. They could take various forms, such as loud tones or warning beeps, as investigated in this current work. Other possibilities that are not pursued in this study could include linking the sounds to the stopping or turning intentions of the AV or dynamically adapting the sounds to the traffic situation, with options such as directional emission or sounds that rely on pedestrian recognition by the AV.

1.3. Study aim

The aim of this study is to examine the effectiveness of various types of synthetic exterior vehicle sounds that could be used as auditory eHMI to inform pedestrians of an approaching vehicle. We evaluated the effect of pure tones, combined tones, intermittent tones, beeps, and engine/tire sounds—presented with and without background noise—on perceptual factors such as the extent to which the sound is easy to notice and annoying. Moreover, instead of asking participants to press a key as soon as they detected

the exterior vehicle sound, we asked them to hold a key when they would be willing to cross the road. We expect crossing deterrence to increase with loudness and tonal frequency, whereas annoyance is likely to increase with loudness [39-41]. Participants were recruited via crowdsourcing, resulting in a larger sample and higher statistical power compared to previous studies.

2. Method

2.1. Sound emission from a moving source

In this study, auditory stimuli were generated from a vehicle passing a stationary observer. In this section, we describe the assumptions made during the simulation of the passing vehicle. In the subsequent section, we detail the various artificial sounds that were employed to eventually produce the auditory stimuli that were presented to the participants in the experiment.

This study assumes a relatively simple sound source and observer geometry. A two-dimensional arrangement is employed, where the observer and the sound source are on the same plane, and there is no reflection of sound from buildings or other structures. Furthermore, the sound source is assumed to be a monopole, a theoretical point source that radiates sound evenly in all directions. Modeling a passing car as a monopole could be a reasonable simplification for studying how a stationary observer perceives the sound. When the observer's distance is considerably greater than the size of the car itself, the specifics of the car's acoustic near field can often be disregarded. Previous research on the design of external vehicle sounds for EVs also employed a monopole assumption (see [27, 28, 42]; although other methods such as multiple monopoles [43], directional emission [44], and physical model methods [45] have also been used for the simulation of exterior vehicle sounds).

The sound source was modeled to move at a constant velocity \mathbf{V} of 30 km/h (≈ 8.33 m/s, Mach number $M \approx 0.024$) in the positive x direction along a straight line at a distance $y_s = 3$ m from the observer (Figure 1). In Figure 1, θ is the angle between the velocity vector \mathbf{V} and the source position vector \mathbf{r} . The initial position of the sound source at $t = 0$ is defined at $(x_{s,0}, y_s)$ and at a general instant t as $(x_{s,0} + V \cdot t, y_s)$. In the generated auditory stimuli, the source moved from $x_{s,0} = -60$ m to a final position of $x_{s,f} = 53$ m. At the given speed of 30 km/h, this meant that the vehicle passed the observer after 7.2 s and that the audio clip concluded after 13.6 s.

The decision to use -60 m and $+53$ m was driven by the objective to keep the audio clips limited in duration (specifically, 13.6 s); this approach enabled participants to assess a substantial number of stimuli without causing the entire experiment to become overly lengthy. At a position of -60 meters, the sound volume was still low (see Figure 2), and at 53 m, where the vehicle had already passed 5 s prior, no additional valuable information could be obtained regarding the participants' inclination to cross.

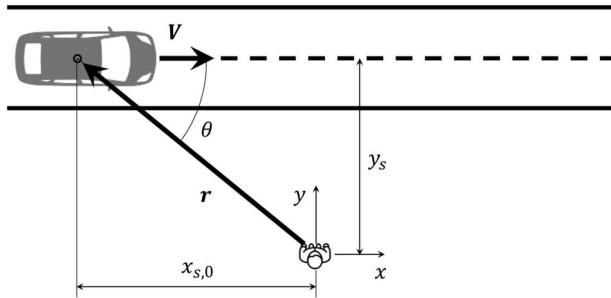


Figure 1. Schematic representation of the vehicle approaching a stationary pedestrian. In this figure, the x -axis and y -axis originate from the observer. The x -axis runs parallel to the road, and the y -axis is perpendicular to the road. In the audio clips, the car exhibits uniform motion from $(x, y) = (x_{s,0}, y_s) = (-60, 3)$ to $(x_{s,f}, y_s) = (53, 3)$.

For a stationary observer, the amplitude of a sound emitted by a moving source changes over time due to the relative motion of the source. Equation 1 gives the amplitude modulation factor [46, 47]. It describes the amplitude of the sound pressures observed by the pedestrian ($A(t_{obs})$)

relative to the amplitude of the sound pressures observed by the pedestrian if the vehicle would be standing still ($V = 0$) in front of the pedestrian at coordinate $(0, y_s)$ (A_0).

$$\frac{A(t_{obs})}{A_0} \approx \frac{y_s}{r(1 - M \cdot \cos \theta)^2} \quad (1)$$

in which M is the Mach number, defined as $V/c \approx \frac{8.33 \text{ m/s}}{343 \text{ m/s}} \approx 0.024$, where c is the speed of sound.

The sound signal is observed with a slight delay with respect to the emission time:

$$t_{obs} = t + r/c \quad (2)$$

Equation 1 describes that the observed sound amplitude increases as the vehicle gets closer to the pedestrian and decreases once the vehicle has passed the pedestrian. Furthermore, the effect of the source motion causes a frequency shift due to the Doppler effect, as described by Equation 3 [46].

$$f_{observed} = \frac{f_{emitted}}{1 - M \cdot \cos \theta} \quad (3)$$

2.2 Auditory stimuli

Thirty auditory stimuli were synthetically generated. These stimuli can be classified into four categories: (1) continuous pure tones at a single frequency, (2) intermittent pure tones (a 500-ms interval emitting followed by a 500-ms interval not emitting), (3) combined tones, and (4) double beeps. In addition, a stimulus with a diesel engine sound signal [48] was included as a baseline representing an ICEV. The selection of a diesel engine's sound was made because of its unique and characteristic noise, which sets it apart from other sounds such as those made by tires or wind. A stimulus with only tire noise [49] was also included to assess the performance of a quiet EV/AV.

The tonal sounds of categories (1), (2), and (3) were presented at four frequencies: 350 Hz, 500 Hz, 1000 Hz, and 2000 Hz. The tonal sounds of categories (1), (2), and (3) were presented at four frequencies: 350 Hz, 500 Hz, 1000 Hz, and 2000 Hz. It is known that the hearing of young adults is most sensitive in the range of 2000 Hz to 5000 Hz, with peak sensitivity around 3000 Hz [50]. We opted for lower frequencies, from 350 to 2000 Hz. Reasons for avoiding higher frequencies included potential issues with the directivity of the exterior loudspeakers, the possibility that atmospheric absorption might attenuate the sound emitted too much (necessitating very loud emission levels to remain effective), and the fact that the optimal frequency tends to reduce with biological age [51], making testing at 3000 Hz less relevant. For comparison, the (recommended) frequencies of sirens of emergency vehicles are also within the range of 400 Hz to 2000 Hz [51, 52, 53].

The combined tones of category (3) were the same as the continuous pure tones of category (1) but with two additional tones of lower amplitude at frequencies 90 Hz above and below the main tone. The combined tones were expected to be perceived as more annoying because of the addition of the extra tone at +90 Hz [54]. The double-beep signal (4), with each beep at a tone of 1800–1900 Hz, consisted of eight pairs of 240-ms beeps separated by a silent 100-ms interval within each pair and a 1000-ms interval between pairs. This stimulus was tested by Bazilinskyy et al. [55] (in a series of double-beep stimuli with 2000, 1000, 750, and 430 ms intervals). These authors found that shorter intervals between beeps led to a higher perceived urgency. For the current study, the stimulus with a medium (1000 ms) interval was selected.

The set of 15 signals (3 tonal sounds x 4 frequencies + double beeps + ICEV + tires) was offered with background noise (a recording of a quiet street [56]) and without, resulting in a total of 30 stimuli. In all cases with background noise, the tire noise sound was also added (for the tire noise stimulus, this means that it was offered once with and once without background noise). In the cases without background noise, no tire noise was added to test the pure effect of the synthetic sound. The sounds with background noise were presented at a lower volume than their counterparts without background noise (see Figure 2 for an illustration), with the purpose of making the background noise relatively prominent compared to the artificial tones that were part of the stimuli.

Figure 3 shows the spectrogram of the continuous pure tone at 2000 Hz, with and without background noise. From the spectrogram, the sound of the vehicle can be clearly distinguished. The change in frequency at around 7.2 s corresponds to the Doppler effect (see Section 2.2). The regular peaks between 3000 and 6000 Hz in the bottom spectrogram correspond to bird tweets in the background noise.

All stimuli were generated at a sampling rate of 44.1 kHz. The duration of each stimulus was 13.6 s. Table 1 provides an overview of the auditory stimuli used in the experiment. Appendix A delivers a full account of all auditory stimuli used in this study. There, we display the amplitude of the signal as a function of time and also provide the discrete Fourier transforms of these signals.

Table 1
Auditory stimuli included in the experiment

No	Characteristics
1	Pure tone, continuous, 350 Hz
2	Pure tone, continuous, 500 Hz
3	Pure tone, continuous, 1000 Hz
4	Pure tone, continuous, 2000 Hz
5	Pure tone, intermittent (13 x [500 ms emitting, 500 ms not emitting]), 350 Hz
6	Pure tone, intermittent (13 x [500 ms emitting, 500 ms not emitting]), 500 Hz
7	Pure tone, intermittent (13 x [500 ms emitting, 500 ms not emitting]), 1000 Hz
8	Pure tone, intermittent (13 x [500 ms emitting, 500 ms not emitting]), 2000 Hz
9	Combined tone, continuous, 350 Hz (± 90 Hz)
10	Combined tone, continuous, 500 Hz (± 90 Hz)
11	Combined tone, continuous, 1000 Hz (± 90 Hz)
12	Combined tone, continuous, 2000 Hz (± 90 Hz)
13	Double beeps (8 x [240 ms beep, 100 ms pause, 240 ms beep, 1000 ms pause]), 1800–1900 Hz
14	Diesel engine
15	Tires on asphalt

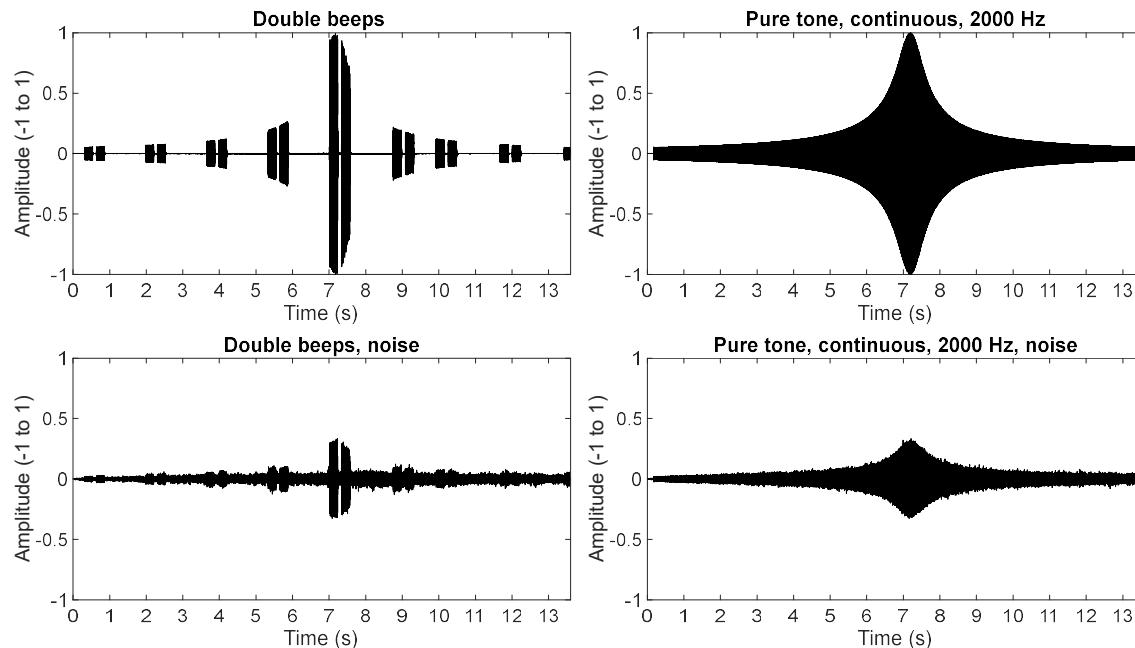


Figure 2. Sound signal as a function of the elapsed time for two example stimuli without background noise (top) and the same stimuli with background noise (bottom).

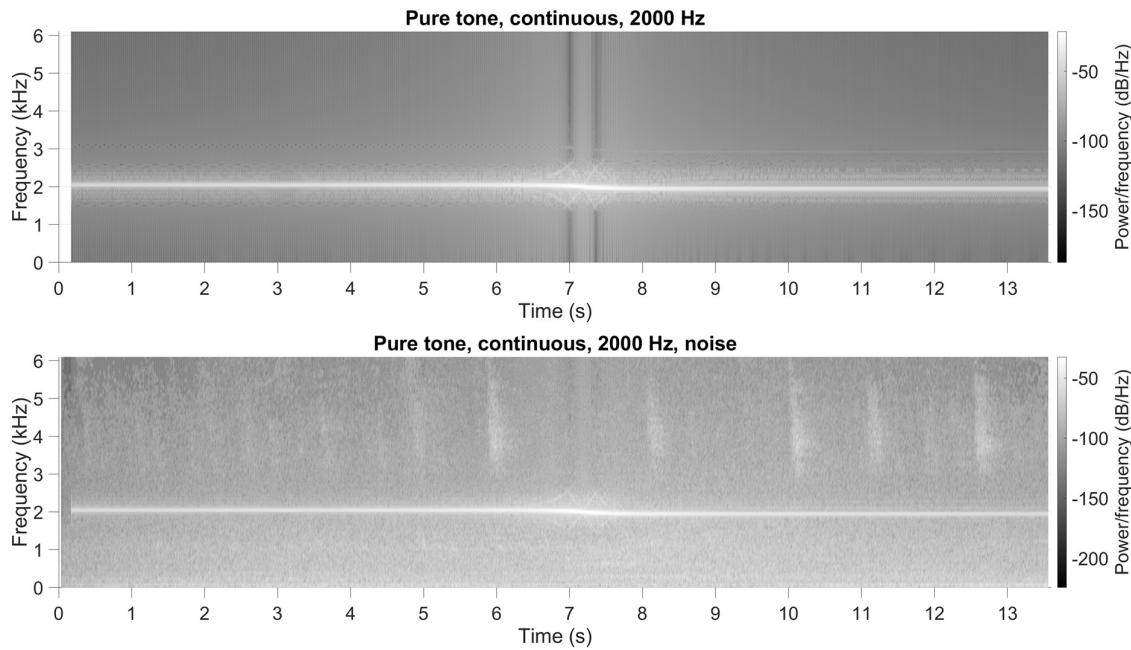


Figure 3. Spectrogram of pure continuous tone (2000 Hz) stimulus with and without background noise.

2.3. Participants

Participants subscribed to the study through the crowdsourcing service Appen (<https://appn.com>). They could become aware of this research by logging into one of the channel websites (e.g., <https://www.ysense.com>) where our study was presented on a list of other projects. We allowed contributors from all countries to participate. Participants were not allowed to complete the study more than once using the same worker ID. A payment of 0.50 USD was offered after the completion of the experiment.

In addition to the crowdsourced participants, a small number of participants were recruited among acquaintances to conduct the same study. Due to COVID-19 restrictions at the time of the study, these participants conducted the experiment online using their personal computers, rather than in the laboratory. They answered the same pre-experimental questions as the crowdsourced participants but did so using a Google Form questionnaire instead of Appen. The experiment itself was presented in the same online environment as that of the crowdsourced participants.

The research was approved by the Human Research Ethics Committee of the Delft University of Technology (reference number 1233).

2.4. Procedure

The study was presented in English. At the top of the webpage introducing our study, contact information was provided. Participants were informed that they could contact the investigators to ask questions about the study and that they had to be at least 18 years of age. Information about anonymity and the voluntary nature of participation was also provided.

Participants first provided demographic information about their age, gender, hearing problems, use of headphones at the moment, and driving experience. Next, they were asked to leave the questionnaire by clicking on a link that opened a webpage with the experiment and were presented with the following instructions:

"Imagine that you are a pedestrian standing on the side of the road. You will listen to 60 sounds of vehicles driving by you. When the sound is playing, press and HOLD 'F' when you feel safe to cross the road in front of the car. You can release the button and then press it again multiple times during the sound. After each sound, you will be asked to answer a few questions. After each 10 sounds you will be able to take a short break. Sometimes you will be asked to listen to a phrase and type what was said."

Please make sure that your audio is on. On the next page, you will listen to a song. When you will be listening to the song, adjust your volume level to be able to hear the song clearly. Do NOT change your volume level till the end of the experiment. Press 'C' to proceed." The song used to adjust the volume was instrumental and copyright-free, taken from [57]. On a scale from 0 to 1, the maximum amplitude of the digital sound from the music was 0.1 (compare this with Figure 2 and Figure A1 in the Appendix, which indicate that the maximum amplitude of the auditory stimuli without background noise is 1.0). Previous literature has shown that the preferred playback level for music listeners with headphones or earphones corresponds to a mean A-weighted ear canal sound pressure between 70 and 80 dBA, and a standard deviation of 7 to 10 dBA [58-60].

The experiment was created using a modified framework based on jsPsych [61] that was used in a previous study on the measurement of reaction times to auditory, visual, and multi-modal stimuli [62], as well as in studies investigating the willingness of pedestrians to cross in front of an automated vehicle, using the same keypress method as in this study [63, 64]. The sounds were pre-loaded before the start of the experiment to prevent delays during the experiment.

The participants had to respond to 60 sounds presented in blocks. Participants were randomly assigned to either listen to 30 auditory stimuli with background noise first, followed by 30 without background noise, or vice versa. Each stimulus was presented twice per block, in a random order that differed for each participant. Before each stimulus sound, the participants were instructed as follows: "*Start by HOLDING the 'F' key. Release the key when it becomes unsafe to cross; press again when safe to cross*". The instruction remained visible throughout the duration of the stimulus. After each stimulus, they were asked to rank the sound based on three criteria: (1) "easy to notice", (2) "gave me enough information to realize that a vehicle was approaching", and (3) "annoying". Each criterion was ranked by moving a slider between 0 and 10 (Figure 4). The participant could not proceed to the next stimulus before having moved all the sliders. Participants did not receive feedback on their responses.

Please rate the following statements based on the sound that you just listened to. Provide your answers by moving the sliders. You will not be able to continue before moving all sliders.

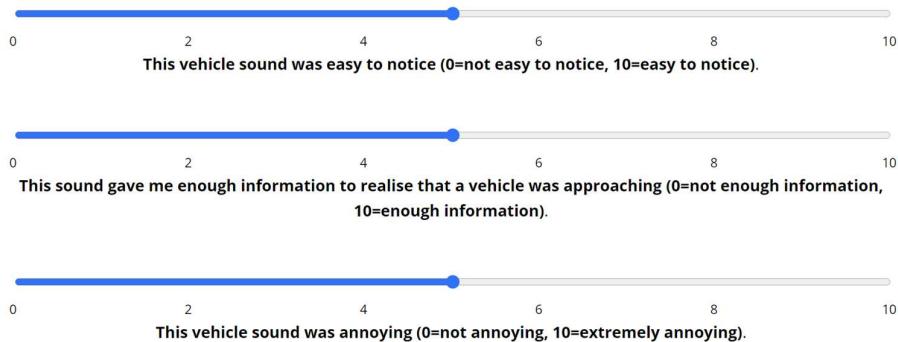


Figure 4. A depiction of the questions that the participants received after listening to each sound.

In order to ensure attentive participation, five test phrases were injected, randomly selected from one of the following six: (1) "Oranges are orange", (2) "Lemons are yellow", (3) "Cherries are red", (4) "Apples are green", (5) "Blackberries are black", and (6) "Grapes are blue". The test phrases were generated using the British English Amy female voice available at [65]. The participants had to type the test phrase they listened to. On a scale from 0 to 1, the maximum amplitude of the digital sound from the test phrases was 0.1, identical to the amplitude of the music that was used by the participants to adjust their volume settings. After the experiment, the response typed by the participant was automatically compared to the correct response to determine whether participants were able to hear the sound and were still attentive to the task.

It should be noted that the synthesized voice may have been difficult to understand because they were presented without context. The purpose of the test phrases was to ascertain whether participants had their sound enabled and maintained sufficient attentiveness throughout the experiment, not to evaluate

their proficiency in English. Consequently, we adopted a scoring method in which only the first three letters of the fruit or the first three letters of the color were considered as correct responses. For instance, for the test phrase “lemons are yellow”, responses containing the strings “lem” or “yel” were marked as correct. The comparison was not case sensitive.

After every ten trials, participants were presented with text indicating how many of the 60 sounds they had completed, for example: “*You have now completed 10 sounds out of 60. When ready press ‘C’ to proceed to the next batch.*”

At the end of the experiment, the participants were given a unique code. They had to enter the code on the questionnaire as proof that they completed the experiment in order to receive their remuneration.

2.5. Data analysis

For each auditory stimulus, a ‘crossing deterrence score’ was calculated based on the participants’ keypress behavior. The concept of such a score computation originates from previous research [63, 66–68], where the time intervals used for calculating the score corresponded to triggers like the moment the vehicle started to brake, the activation of an eHMI, or the vehicle coming to a halt. In our experiment, we expect participants to release the response key, indicating they feel less safe to cross, as the vehicle approaches, and to press it again after the vehicle has passed, i.e., after the 7.2-s mark. We defined the crossing deterrence score over the interval from 1.0 to 7.2 s, incorporating a 1-s start-up margin (see Figure 5 for justification). More specifically, the crossing deterrence score was calculated as 100% minus the percentage of keypresses in the 1.0 to 7.2-s interval, averaged across all trials and participants.

Moreover, the loudness for the 1.0–7.2 s interval was computed from the sound signal using ISO 532-1 [69] (Zwicker loudness; for stationary sounds: [70]; for time-varying sounds: [71]) and ISO-532-2 [72] (Moore-Glasberg method [73]). These methods determining acoustic loudness take into account characteristics of human hearing, such as the dependence of sound transmission through the middle ear on frequency [70]. Swart and Bekker [74] compared the Zwicker loudness with other psychoacoustic metrics and showed that it was suitable for distinguishing between similar vehicle sounds. In addition to the Zwicker loudness and Moore Glasberg methods, we also included Integrated Loudness (LUFS) as a loudness score [75, 76]. In addition to the Zwicker loudness and Moore-Glasberg methods, we also incorporated Integrated Loudness (LUFS) as a loudness score [75, 76]. Integrated loudness is considered a more versatile measure of perceived loudness; it is commonly applied to broadcasting and streaming services, where dynamic content spans a range from speech to music and sound effects [77].

Pearson correlation coefficients were computed between the crossing deterrence scores, loudness levels, and the scores of the three questions presented after each stimulus (“easy to notice”, “gave me enough information to realize that a vehicle was approaching”, and “annoying”), averaged across all participants.

3. Results

A total of 995 people participated between September 16 and 17, 2020. However, due to a data storage error, data for 420 participants was unavailable, reducing the effective sample size to 575. From this pool, we applied a filtering process, removing participants based on the following criteria: (1) self-reported non-compliance with reading the instructions, (2) being under 18 years of age, (3) completion of the study within an implausible time frame of 15 minutes, given that the total duration of the stimuli was 13.6 seconds multiplied by 60 audio clips (equaling 816 seconds), plus the required time to answer the questions, (4) self-reported hearing problems, and (5) incorrect answers to two or more of the test phrases. These five criteria led to the exclusion of 370 participants, 352 of whom were excluded due to errors in the test phrases, thus yielding a final sample size of 205.

The 205 participants had a mean age of 37.1 years ($SD = 11.4$ years, median: 36). Of the 205 participants, 140 were male, 63 were female, and 2 preferred not to respond. The countries that were most represented were Venezuela ($n = 77$), India ($n = 13$), Russia ($n = 11$), and the United States ($n = 11$). The participants took, on average, 48.0 minutes to complete the study ($SD = 19.8$ min, median = 43.3 min). Of the 205 participants, 84 confirmed the use of headphones at the time, while 120 did not, and one chose the option ‘I prefer not to respond’.

From the recruiting via acquaintances, 21 people participated between 4 December 2020 and 12 January 2021. This sample consisted of 14 males and 7 females. Collectively, they had an average age of 32.0 years ($SD = 11.1$, median: 30). None reported wearing headphones at the time of the study. The remainder of the analysis will be conducted for the 205 crowdsourced and 21 additionally recruited participants together ($n = 226$). This combined sample consisted of 154 males, 70 females, and 2 participants who indicated 'I prefer not to respond' to the gender question. The mean age for this combined group was 36.6 years ($SD = 11.5$, median = 34.5).

Figure 5 shows the corresponding keypress percentages for two example signals from 30 auditory samples. A substantial difference in keypresses can be distinguished, with the continuous tone being a stronger deterrent to cross (i.e., participants released the key earlier) compared to the beeping sound.

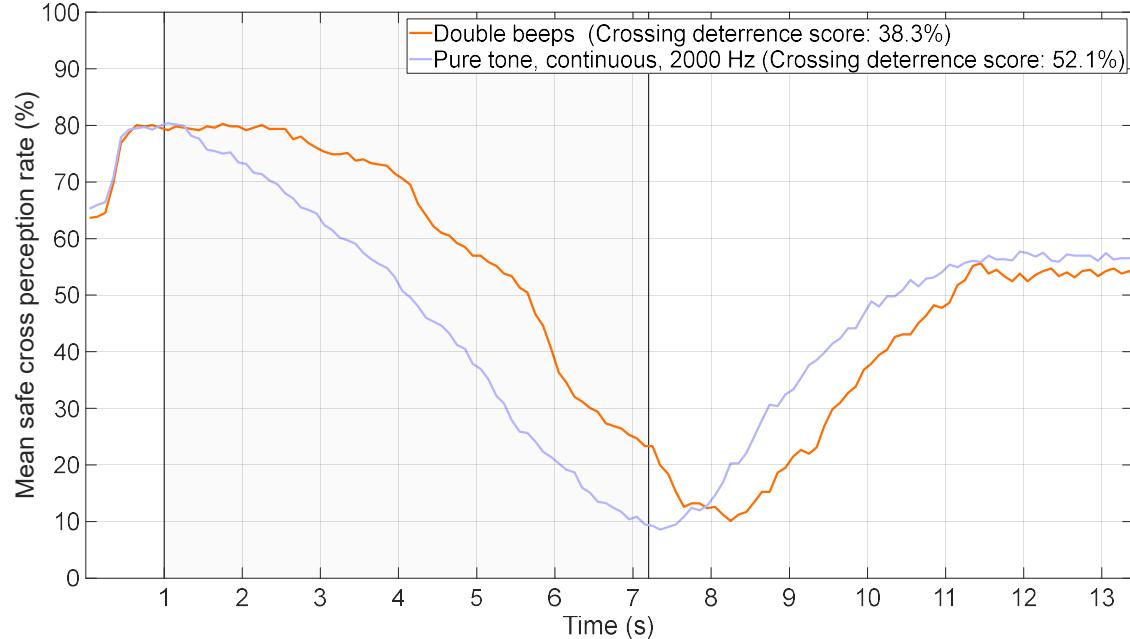


Figure 5. The mean safe cross perception rate (equivalent to the percentage of trials in which the 'F' key was pressed) plotted as a function of elapsed time for two selected samples. The legend shows the mean and SD of the crossing deterrence score. The gray background indicates the time interval across which the crossing deterrence score was computed.

Table 2 shows the correlation matrix between the crossing deterrence score, the three self-reported measures, and the loudness for the interval from 1.0 to 7.2 s. The scores were averaged over the 226 participants. One notable observation, illustrated in Figure 6, is that louder sounds generally had a higher crossing deterrence score. When sounds implicit to driving (such as tire noise and diesel engine noise) are excluded, the correlation between the Moore-Glasberg loudness and the crossing deterrence score, as depicted in Figure 6, is considerably stronger ($r = 0.86$, $n = 26$) compared to when these sounds are included ($r = 0.72$, $n = 30$).

Table 2
Means, standard deviations, and correlation coefficients of the variables for the auditory stimuli ($n = 30$).

	Mean	SD	1	2	3	4	5	6	7
1 Crossing deterrence score (%)	44.0	4.09							
2 Easy to notice (0–10)	7.20	0.49	0.38						
3 Gave enough information (0–10)	6.72	0.41	0.27	0.94					
4 Annoying (0–10)	5.34	1.21	0.66	0.79	0.60				
5 Zwicker loudness, stationary (-)	25.4	5.21	0.66	0.26	0.09	0.44			
6 Zwicker loudness, time-varying (-)	17.4	4.93	0.70	-0.12	-0.14	0.13	0.64		
7 Moore-Glasberg loudness (-)	25.0	6.42	0.72	-0.02	-0.08	0.20	0.79	0.81	
8 Integrated loudness (LUFS)	-21.5	6.41	0.45	0.60	0.39	0.75	0.57	0.09	0.04

Table 2 and Figure 7 show that sounds with higher crossing deterrence scores were generally also perceived as more annoying. However, driving-associated sounds, namely tire noise and the sound of a diesel engine, were found to be relatively non-annoying (see Figure 7). These sounds were also somewhat hard to notice (Figure 8) but still resulted in fairly high crossing deterrence (see Figures 6 & 7).

High tones yielded a relatively high crossing deterrence score, encouraging participants to release the key early, while lower tones scored substantially lower, as seen in Figures 6 and 7. Additionally, intermittent tones resulted in lower crossing deterrence than continuous tones.

Finally, Table 2 highlights the noteworthy finding that the four different measures of perceived loudness capture various aspects. The ISO-532 measures (Zwicker, Moore-Glasberg) primarily correlate with the crossing deterrence score, reflecting keypress behavior, while the integrated loudness seems to correlate more strongly with the subjective measures of 'easy to notice' and 'annoyance'.

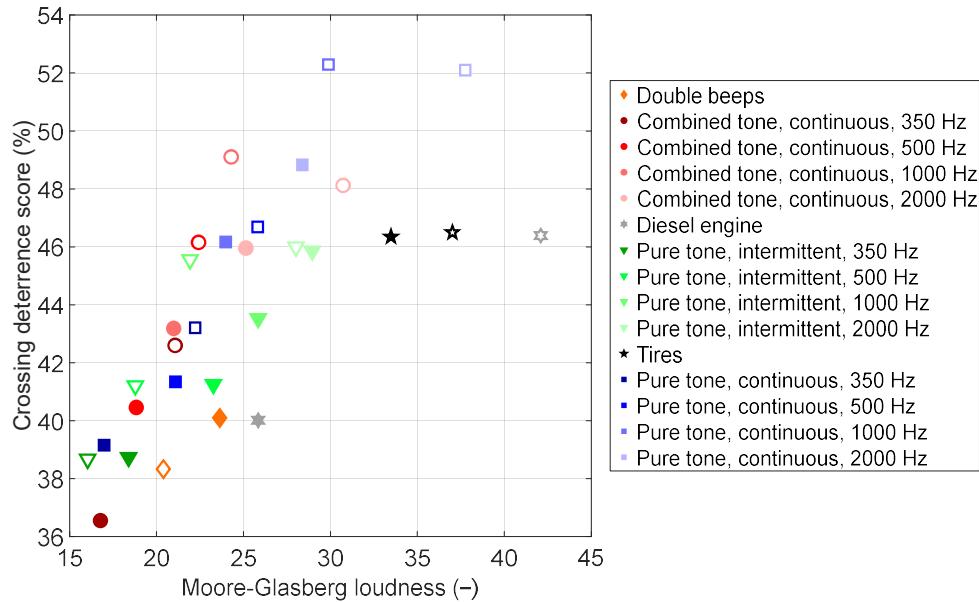


Figure 6. Scatter plot of the crossing deterrence score (based on keypress inputs) and the computed loudness score of the 30 stimuli. Filled markers pertain to sounds with added background noise; open markers, that is, with a white interior, relate to sounds without any added background noise.

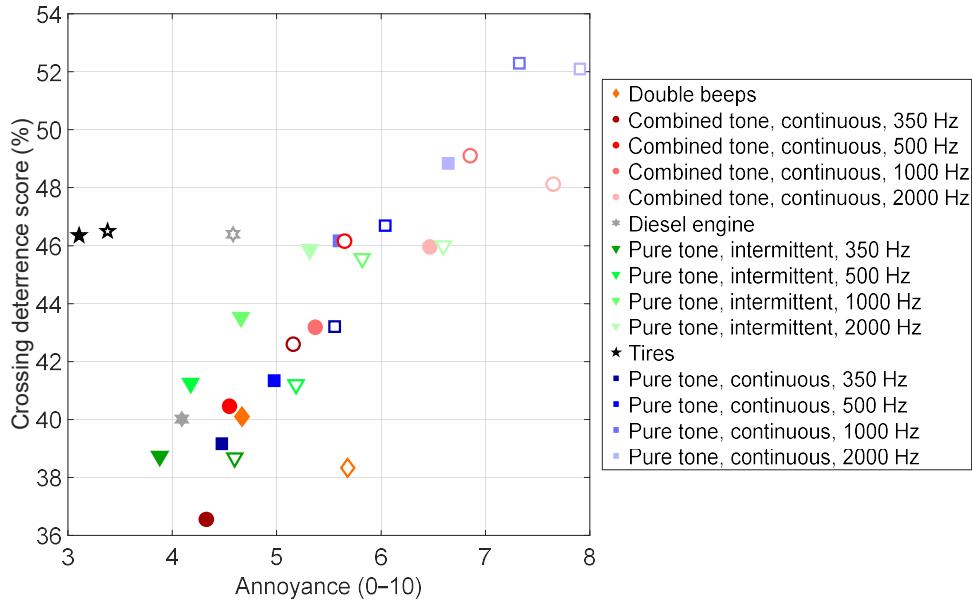


Figure 7. Scatter plot of the crossing deterrence score (based on keypress inputs) and the perceived annoyance of the 30 stimuli. Filled markers pertain to sounds with added background noise; open markers, that is, with a white interior, relate to sounds without any added background noise.

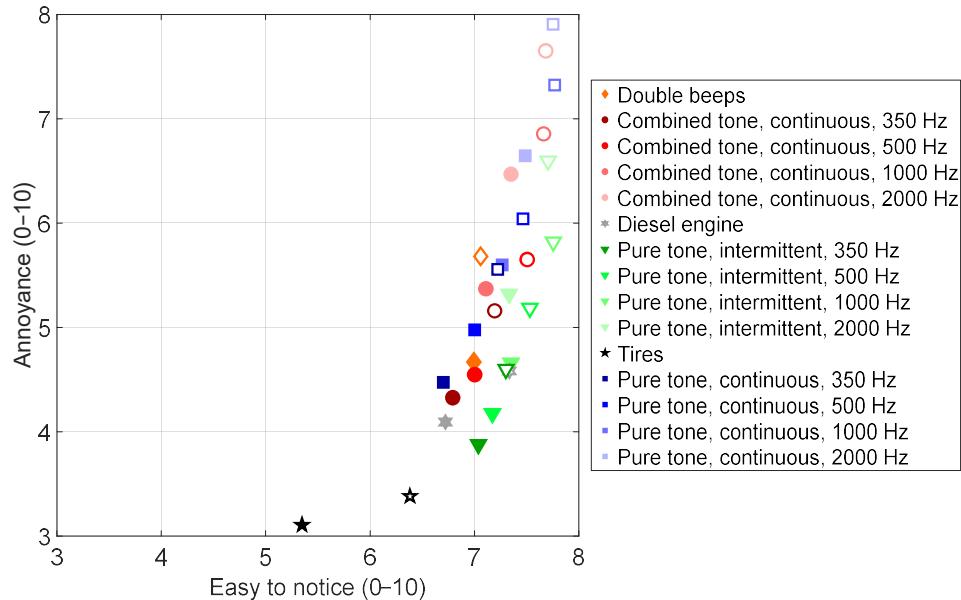


Figure 8. Scatter plot of the perceived annoyance and the perceived 'easy to notice' of the 30 stimuli. Filled markers pertain to sounds with added background noise; open markers, that is, with a white interior, relate to sounds without any added background noise.

4. Discussion

This online study tested how vehicle sounds, naturalistic ones, as well as tones and beeps, influence participants' crossing intentions. In order to isolate the effect of sound, participants were not presented with any visual information, which may be representative of situations in which visual information is unavailable due to visual impairment or visual obstruction.

A distinctive feature of our research is that we not only queried subjective qualities, as previous studies have done (e.g., [16, 21, 30]), but also procured objective data through a keypress. With regard to the

latter, a unique aspect is that we did not measure reaction times with the intention of assessing detectability (see, for example, [12, 18, 22, 26-29]), but instead asked participants to press and hold down the response key for as long as they believed it was safe to cross. This approach yields information equivalent to that obtained from reaction times, supplemented by the fact that participants can also integrate a decision threshold with respect to what they deem as a ‘safe crossing’. Furthermore, through our keypress method, we were able to depict the dynamics of crossing tendencies, as demonstrated in Figure 5, from which a crossing deterrence score was subsequently calculated.

The results showed that loud sounds were the most effective in discouraging participants from crossing the road. Our findings also indicated that across all 30 stimuli, the more effective sounds, i.e., those causing higher crossing deterrence, were also the more annoying sounds. This supports the “trade-off hypothesis of pleasantness and power” that was reported by Bisping [78] for in-vehicle sounds and according to which an increase in perceived pleasantness of a sound beyond a certain level has negative consequences on the perceived powerfulness of the car and vice versa (and see [79] for a similar negative correlation between valence and dominance/arousal of auditory stimuli). These results raise the question of whether current EVs, which tend to emit naturalistic sounds or pleasant tones [18], are optimally effective in ensuring the safety of pedestrians.

Beeps and intermittent sounds, particularly those of low tonal frequency, yielded relatively low crossing deterrence, even when there was no background noise. One possible explanation is that the inter-pulse intervals we used were relatively long (1000 and 500 ms for beeps and intermittent tones, respectively). Previous studies that have investigated shorter intervals found that perceived urgency increases as the beep rate increases, following Stevens’ power law [55, 80-82]. In practice, slow beeps are typically used for relatively slowly evolving situations, such as a truck reversing, whereas fast beeps indicate an approaching hazard [83]. Moreover, the duration of the beeps themselves was short, which inhibited speed estimation. At the same time, during the non-emitting intervals, particularly if these are long, no new information is conveyed, leading to a lag in information processing.

In evaluating our results, it must be taken into account that loudness can be defined in various ways, and that our different loudness scores did not exhibit particularly strong intercorrelations. The loudness scores correlated with objective outcomes (crossing deterrence score) and subjective outcomes (including ‘easy to notice’, ‘enough information’, and ‘annoyance’) to different extents. One possible explanation resides in the underlying purposes of the metrics, where the ISO-532 methods (Zwicker and Moore-Glasberg) are psychoacoustic models that take into account human hearing as a function of frequency content [69-73, 84], with the Moore and Glasberg method being a more recent alternative [72, 73, 85]. Integrated loudness represents a more practical industry standard used to assess a wide variety of audio content, including different types of music, in terms of loudness [75, 76, 86, 87]. These differences are also observable in our results where, for a given sound type (such as the pure tone), there exists a highly linear relationship between the Moore-Glasberg loudness and objective and subjective outcomes (see Figures 6 & 7), while different types of sounds do not fit this pattern, presumably because the ISO-532 methods were not originally intended or validated to compare sounds that are fundamentally distinct from one another [88].

A limitation of this study is that only a small range of relatively simple signals was tested. Moreover, only non-verbal stimuli were tested because, while verbal messages can have richer semantic content, non-verbal ones can convey information faster [89]. Nevertheless, a direct comparison of verbal and non-verbal sounds would deserve further investigation.

Another limitation is that we did not have control over the participants’ volume settings or the quality of their audio equipment, something that is particularly relevant given the fact that many participants were from low-income countries where laptops or PCs might be of inferior quality. However, we believe that the findings of our study, which concern *relative* comparisons between stimuli, are robust due to the substantial sample size. In our previous research using a similar crowdsourcing approach, we found that despite absolute differences in the outcomes between individual countries, and variations in age and gender among participants from these nations, the ranking of stimuli in comparison to each other remains consistent across countries [68, 90]. This means that the ordinal relationships are invariant, even though specific differences exist between the countries.

Lastly, the ecological validity of our experiment is limited because the sounds were tested in a computer environment and in a predictable scenario of an approaching vehicle. In real-world situations, road users

divide their attention over a large number audiovisual stimuli, some of them even distracting from the crossing task, which would potentially impact the detectability of auditory signals emitted from approaching cars.

Our study queried each participant about every sample, assessing to what extent the sound was easy to notice, to what extent it provided sufficient information that a vehicle was approaching, and to what extent it was annoying. The first two items exhibited a very strong correlation across the 30 samples and are thus nearly redundant. Accordingly, our three questions capture the two key features of sounds, specifically whether the sound is loud (powerful, potent) and whether it is annoying (unpleasant). Previous research, which examined a large quantity of samples from the automotive domain and encompassed a greater number of rating scales for these sounds, indicates a third dimension based on multivariate analysis. This dimension can be characterized as brightness or clarity and may also relate to the timbre of the sound [91, 92]. In other words, our three questions may provide a reasonable depiction of the subjective attributes of sounds but are not comprehensive; future research could include a larger number of diverse rating scales. A disadvantage of this, however, is that the experiment may become unduly lengthy for participants and may be perceived as monotonous. As a remedy, a between-subjects design could potentially be chosen, wherein not all participants would need to evaluate all sound samples.

It is important to note that our findings do not necessarily imply that electric vehicles or eHMIs for automated vehicles should be made louder (and, by extension, more annoying). Doing so is one of the possibilities, but not the only way of informing vulnerable road users. Another potential method to enhance detectability and possibly increase crossing deterrence is to reduce background noise, for instance, by mitigating various forms of noise pollution on streets. For example, it may be expected that with the increasing electrification of vehicle fleets and the corresponding decline of combustion engine vehicles, artificial engine sounds may be more readily heard. Other options would be to continue research into sounds that are easily detectable yet not annoying, a topic that has been explored before in the literature [27].

Our research provides insights regarding the design of deterrent sounds that are nonetheless not annoying. Research suggests that implicit communication, defined by signals emanating from the vehicle itself, such as speed and pedestrian-vehicle distance, is often sufficient, and that explicit forms of communication, like light-emitting eHMIs, hand gestures, or eye contact, are not frequently used or needed [93, 94]. Our research seems to present a similar picture when it comes to exterior vehicle sounds, where the sounds of the tires and engine alone yielded a relatively high crossing deterrence score, while being considered the least annoying. One possible explanation is that tire and engine sounds constitute noise across a broad frequency spectrum making them less annoying than an explicit tone. Another possible explanation is that tire and engine noises are familiar and naturally associated with a moving vehicle, whereas understanding the meaning of a beep or tone might require some learning, equivalent to the phenomenon that the reliance on visual eHMIs necessitates learning [95-97]. On the other hand, it should be noted that tire noise can also be confused with the sound of wind or other noises that are naturally present [18]. In conclusion, it is our belief that the quest for optimal exterior vehicle sounds—those that are effective without being annoying—resides in the domain of implicit communication, such as the use of tire or engine noise. These findings align with earlier suggestions made by [16, 30].

Finally, it must be noted, as indicated in the introduction, that the precise imitation of the sounds of an ICEV may not necessarily be required or optimal in this context. An interesting approach is employed by Mauder [98], who entirely refrained from imitating ICEV sounds. Instead, microphones were used to amplify the sound of the AV motors, aiming to provide a unique experience for both interior and exterior vehicle sounds. A sound resembling that of tires or engines could also be amplified and modified, an approach that is already being adopted by EV manufacturers to comply with legislation. This strategy could be further developed toward more intelligent or adaptive exterior vehicle sounds for AVs and EVs, a process that would necessitate the amendment of existing legislation.

To conclude, this study found a fairly strong relationship between the acoustic loudness of a sound and its effectiveness in preventing participants from crossing the road. The present findings provide a useful reminder that loudness may be the primary (yet not the only) factor to consider in exterior sound design for EVs and AVs. The study also suggests that intermittent beeps may need to be avoided as they may impede the pedestrian's ability to perceive the speed and distance of the approaching vehicle.

5. Supplementary material

The questionnaire used, the auditory stimuli, the raw data underlying Table 2, the song for volume adjustment, the test phrases, anonymized data, and MATLAB code are accessible at <https://www.dropbox.com/sh/qbho9jyea9z0wgd/AADNFSJCC4N-HewqEU8IXIFGa?dl=0>. These data will be published later in a certified repository.

Acknowledgment

The research of Pavlo Bazilinskyy and Joost de Winter is supported by Joost de Winter's grant 016. Vidi.178.047 (2018–2022; "How should automated vehicles communicate with other road users?"), which is financed by the Netherlands Organisation for Scientific Research (NWO).

References

- [1] World Health Organization. Pedestrian safety. A road safety manual for decision-makers and practitioners. http://apps.who.int/iris/bitstream/handle/10665/79753/9789241505352_eng.pdf;2013
- [2] DaSilva MP, Smith JD, Najm WG. Analysis of pedestrian crashes. Report No. DOT-VNTSC-NHTSA-02-02. John A. Volpe National Transportation Systems Center, Cambridge, MA; 2003
- [3] Habibovic A, Davidsson J. Causation mechanisms in car-to-vulnerable road user crashes: Implications for active safety systems. Accid Anal Prev 2012;49:493-500. <https://doi.org/10.1016/j.aap.2012.03.022>
- [4] Governors Highway Safety Association. Pedestrian traffic fatalities by state. 2021 Preliminary data (January – December). <https://www.ghsa.org/sites/default/files/2022-05/Pedestrian%20Traffic%20Fatalities%20by%20State%20-202021%20Preliminary%20Data%20%28January-December%29.pdf>; 2022
- [5] Bálint A, Labenski V, Köbe M, Vogl C, Stoll J, Schories L, et al. Use case definitions and initial safety-critical scenarios. Report No. D2.6. Project SAFE-UP. https://static1.squarespace.com/static/5efaed43294db25b18168717/t/627e752a8d7775630d2ea94a/1652454782434/SAFE-UP_D2_6_2Bcase%2Bdefinitions%2Band%2Binital%2Bsafety-critical%2Bscenarios_.pdf; 2021
- [6] European Parliament News. EU ban on the sale of new petrol and diesel cars from 2035 explained. <https://www.europarl.europa.eu/news/en/headlines/economy/20221019STO44572/eu-ban-on-sale-of-new-petrol-and-diesel-cars-from-2035-explained>; 2022
- [7] GOV.UK. Government takes historic step towards net-zero with end of sale of new petrol and diesel cars by 2030. [https://www.gov.uk/government/news/government-takes-historic-step-towards-netzero-with-end-of-sale-of-new-petrol-and-diesel-cars-by-2030](https://www.gov.uk/government/news/government-takes-historic-step-towards-net-zero-with-end-of-sale-of-new-petrol-and-diesel-cars-by-2030); 2020
- [8] The White House. Executive order on catalyzing clean energy industries and jobs through federal sustainability. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/12/08/executive-order-on-catalyzing-clean-energy-industries-and-jobs-through-federal-sustainability>; 2021
- [9] Government of Canada. Building a green economy: Government of Canada to require 100% of car and passenger truck sales be zero-emission by 2035 in Canada. <https://www.canada.ca/en/transport-canada/news/2021/06/building-a-green-economy-government-of-canada-to-require-100-of-car-and-passenger-truck-sales-be-zero-emission-by-2035-in-canada.html>; 2021
- [10] Roadmaps for Energy. The facts on: Norway's alleged plan to "prohibit sales of all vehicles using fossil fuels by 2025". <https://roadmapsforenergy.eu/norway-fossilfuel-car-ban>; 2023
- [11] Liu S, Fitzharris M, Oxley J, Edwards C .The impact of electric / hybrid vehicles and bicycles on pedestrians who are blind or have low vision. Accident Research Centre. Monash University; 2018
- [12] Parizet E, Ellermeier W, Robart R. Auditory warnings for electric vehicles: Detectability in normal-vision and visually-impaired listeners. Appl Acoust 2014;86:50–8. <https://doi.org/10.1016/j.apacoust.2014.05.006>
- [13] European Commission. Commission Delegated Regulation (EU) 2017/1576 of 26 June 2017 amending Regulation (EU) No 540/2014 of the European Parliament and of the Council as regards the Acoustic Vehicle Alerting System requirements for vehicle EU-type approval (Text with EEA relevance, §III.3.a(i)). <https://eur-lex.europa.eu/eli/reg/del/2017/1576/oj>; 2017
- [14] United Nations. Agreement concerning the adoption of harmonized technical United Nations regulations for wheeled vehicles, equipment and parts which can be fitted and/or be used on wheeled vehicles and the conditions for reciprocal recognition of approvals granted on the basis

- of these United Nations regulations (Revision 3, including the amendments which entered into force on 14 September 2017), Addendum 137: UN Regulation No. 138, §6.2.8.
<https://unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/2017/R138r1e.pdf>; 2017
- [15] Code of Federal Regulations; § 571.141 Standard No. 141; Minimum Sound Requirements for Hybrid and Electric Vehicles. <https://www.ecfr.gov/current/title-49/section-571.141>; 2023
 - [16] Wogalter MS, Lim RW, Nyeste PG. On the hazard of quiet vehicles to pedestrians and drivers. *Appl Ergon* 2014; 45:1306–12. <http://doi.org/10.1016/j.apergo.2013.08.002>
 - [17] Yamauchi K, Menzel D, Takada M, Nagahata K, Iwamiya SI, Fastl H. Psychoacoustic examination of feasible level of additional warning sound for quiet vehicles. *Acoust Sci Technol* 2015;36:120–5. <https://doi.org/10.1250/ast.36.120>
 - [18] Singh S, Payne SR, Jennings PA. Toward a methodology for assessing electric vehicle exterior sounds. *IEEE Trans Intell Transp Syst* 2014;15:1790–800. <https://doi.org/10.1109/tits.2014.2327062>
 - [19] Hsieh MC, Chen HJ, Tong ML, Yan CW. Effect of environmental noise, distance and warning sound on pedestrians' auditory detectability of electric vehicles. *Int J Environ Res Public Health* 2021;18:9290. <https://doi.org/10.3390/ijerph18179290>
 - [20] Fleury S, Jamet É, Roussarie V, Bosc L, Chamard JC. Effect of additional warning sounds on pedestrians' detection of electric vehicles: An ecological approach. *Accid Anal Prev* 2016;97:176–85. <https://doi.org/10.1016/j.aap.2016.09.002>
 - [21] Misdariis N, Cera A, Levallois E, Locqueteau C. Do electric cars have to make noise? An emblematic opportunity for designing sounds and soundscapes. Paper presented at Acoustics 2012, Nantes, France, 23–27 April 2012
 - [22] Frlić Sekulić N, Blagojević I, Popović V, Janković S. Development of pedestrian alert system for use in electric vehicles. *Tehnički Vjesnik* 2019;26:1614–19. <https://doi.org/10.17559/TV-20180730144329>
 - [23] Grosse J, Weber R, Van de Par S. Comparison of detection threshold measurements and modeling for approaching electric cars and conventional cars presented in traffic and pink noise. In: Proc Meetings Acoust 2013;19:050192. <https://doi.org/10.1121/1.4800370>
 - [24] Yamauchi K, Menzel D, Fastl H, Takada M, Nagahata K, Iwamiya SI. Cross-cultural study on feasible sound levels of possible warning signals for quiet vehicles. *INTER-NOISE NOISE-CON Congr Conf Proc* 2011;4:3236–41.
 - [25] Kournoutsos N, Cheer J. An environmentally adaptive warning sound system for electric vehicles. *INTER-NOISE NOISE-CON Congr Conf Proc* 2019;259:683–91.
 - [26] Poveda-Martinez P, Peral-Orts R, Campillo-Davo N, Nescolarde-Selva J, Lloret-Climent M, Ramis-Soriano J. Study of the effectiveness of electric vehicle warning sounds depending on the urban environment. *Appl Acoust* 2017;116:317–28. <https://doi.org/10.1016/j.apacoust.2016.10.003>
 - [27] Petiot JF, Legeay K, Lagrange M. Optimization of the sound of electric vehicles according to unpleasantness and detectability. *Proc Des Soc : Int Conf Eng Des* 2019 ;1 :3949–58. <https://doi.org/10.1017/dsi.2019.402>
 - [28] Lee SK, Lee SM, Shin T, Han M. Objective evaluation of the sound quality of the warning sound of electric vehicles with a consideration of the masking effect: Annoyance and detectability. *Int J Automot Technol* 2017;18:699–705. <https://doi.org/10.1007/s12239-017-0069-6>
 - [29] Jacobsen G, Song W, MacDonald E. Predicting detectability and annoyance of EV warning sounds using partial loudness. *INTER-NOISE NOISE-CON Congr Conf Proc* 2016;253:886–95.
 - [30] Petiot JF, Kristensen BG, Maier AM. How should an electric vehicle sound? User and expert perception. *Int Des Eng Tech Conf Comput Inform Eng Conf* 2013;55928:V005T06A028. <https://doi.org/10.1115/DETC2013-12535>
 - [31] Yamauchi K, Sano T, Hasegawa S, Tamura F, Takeda Y. Detectability and hearing impression of additional warning sounds for electric or hybrid vehicles. *INTER-NOISE and NOISE-CON Congr Conf Proc* 2014;249:5279–85.
 - [32] Wall Emerson R, Naghshineh K, Hapeman J, Wiener W. A pilot study of pedestrians with visual impairments detecting traffic gaps and surges containing hybrid vehicles. *Transport Res F: Traffic Psychol Behav* 2011;14:117–27. <https://doi.org/10.1016/j.trf.2010.11.007>
 - [33] Dey D, Habibovic A, Löcken A, Wintersberger P, Pfleging B, Riener A, et al. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transport Res Interdiscip Perspect* 2020;7:100174. <https://doi.org/10.1016/j.trip.2020.100174>
 - [34] Mahadevan K, Somanath S, Sharlin E. Communicating awareness and intent in autonomous vehicle-pedestrian interaction. Paper presented at CHI Conference on Human Factors in

- Computing Systems, Montréal, Canada, 21-26 April 2018.
<https://doi.org/10.1145/3173574.3174003>
- [35] Deb S, Strawderman LJ, Carruth DW. Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment. *Transport Res F: Traffic Psychol Behav* 2018;59:135–49. <https://doi.org/10.1016/j.trf.2018.08.016>
- [36] Florentine E, Ang MA, Pendleton SD, Andersen H, Ang Jr MH. Pedestrian notification methods in autonomous vehicles for multi-class mobility-on-demand service. Paper presented at the Fourth International Conference on Human Agent Interaction, Biopolis, Singapore, 4-7 October 2016, 387–92. <https://doi.org/10.1145/2974804.2974833>
- [37] Moore D, Curran R, Sirkin D. Sound decisions: How synthetic motor sounds improve autonomous vehicle-pedestrian interactions. Paper presented at the 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Virtual Event, 21-22 September 2020, 94–103. <https://doi.org/10.1145/3409120.3410667>
- [38] De Winter J, Dodou D. External human-machine interfaces: Gimmick or necessity? *Transport Res Interdisc Perspect* 2022;15:100643. <https://doi.org/10.1016/j.trip.2022.100643>
- [39] Fidell S, Teffeteller S. Scaling the annoyance of intrusive sounds. *J Sound Vib* 1981;78:291–98. [https://doi.org/10.1016/s0022-460x\(81\)80039-9](https://doi.org/10.1016/s0022-460x(81)80039-9)
- [40] Fidell S, Teffeteller S, Horonjeff R, Green DM. Predicting annoyance from detectability of low-level sounds. *J Acoust Soc Am* 1979;66:1427–34. <https://doi.org/10.1121/1.383536>
- [41] Berglund B, Berglund U, Lindvall T. Scaling loudness, noisiness, and annoyance of community noises. *J Acoust Soc Am* 1976;60:1119–25. <https://doi.org/10.1121/1.381212>
- [42] Souaille T, Petiot JF, Misdariis N, Lagrange M. An interactive bi-objective optimisation process to guide the design of electric vehicle warning sounds. *Des Sci* 2022;8:e26. <https://doi.org/10.1017/dsj.2022.18>
- [43] Pallas MA, Chatagnon R, Lelong, J. Noise emission assessment of a hybrid electric mid-size truck. *Appl Acoust* 2014;76:280–90. <https://doi.org/10.1016/j.apacoust.2013.08.012>
- [44] Jonasson HG. Measurement and modelling of noise emission of road vehicles for use in prediction models. Nordtest Project 1452-99. Swedish National Testing and Research Institute; 2000
- [45] Talbot A, Lielens G. Pass-by noise synthesis from frequency domain exterior acoustic simulations. In: Siebenpfeiffer W, editor, *Proc Autom Acoust Conf*. WiesbadenL Springer Vieweg; 2019, p. 223–36. https://doi.org/10.1007/978-3-658-20251-4_13
- [46] Morse PM, Ingard KU. *Theoretical acoustics*. McGraw-Hill, New York; 1968, p 74
- [47] Ouyang K, Lu S, Zhang S, Zhang H, He Q, Kong F. Online Doppler effect elimination based on unequal time interval sampling for wayside acoustic bearing fault detecting system. *Sens* 2015;15:21075–98. <https://doi.org/10.3390/s150921075>
- [48] Randall's Rest & Relaxation. 12 hours of Cummins diesel engine sound idling = Dodge Ram truck engine noise for sleeping. <https://www.youtube.com/watch?v=2Y33bTIAA-E>; 2014
- [49] Zero Tailpipe. What sound does Hyundai iONIQ Electric make? Noisy from outside at speed? <https://www.youtube.com/watch?v=X0wpizkkHQ>; 2018
- [50] International Organization for Standardization. *Acoustics — Normal equal-loudness-level contours*. Switzerland; 2023. <https://www.iso.org/standard/83117.html>
- [51] Balastegui A, Romeu J, Clot A, Martín SR. New siren tones optimised for increased detectability distances of emergency vehicles. *Applied Acoust* 2013;74:803–11. <https://doi.org/10.1016/j.apacoust.2012.12.007>
- [52] Campbell T, Parida B, Ross JC, Zaouk AK, Harris M. Acoustical warning devices as emergency warning systems, Phase 1. No. DOT/FRA/ORD-19/40. United States. Department of Transportation. Federal Railroad Administration; 2019
- [53] Catchpole K, McKeown D. A framework for the design of ambulance sirens. *Ergon* 2007;50:1287–1301. <https://doi.org/10.1080/00140130701318780>
- [54] Subedi JK, Yamaguchi H, Matsumoto Y, Ishihara M. Annoyance of low frequency tones and objective evaluation methods. *J Low Freq Noise Vib Active Control* 2005;24:81–95. <https://doi.org/10.1260/0263092054531000>
- [55] Bazilinskyy P, Petermeijer SM, Petrovych V, Dodou D, De Winter JCF. Take-over requests in highly automated driving: A crowdsourcing survey on auditory, vibrotactile, and visual displays. *Transport Res F: Traffic Psychol Behav* 2018;56:82–98. <https://doi.org/10.1016/j.trf.2018.04.001>
- [56] Articulated Sounds. Quiet streets ambience sound library. https://www.youtube.com/watch?v=6CW_7BZBxQ; 2017
- [57] Bensound. Creative minds. Royalty Free Music - No Copyright Music. <https://www.youtube.com/watch?v=zFBwdFnezG0>; 2018

- [58] Hodgetts WE, Rieger JM, Szarko RA. The effects of listening environment and earphone style on preferred listening levels of normal hearing adults using an MP3 player. *Ear Hear* 2007;28:290–7. <https://doi.org/10.1097/AUD.0b013e3180479399>
- [59] Worthington DA, Siegel JH, Wilber LA, Faber BM, Dunckley KT, Garstecki DC, Dhar S. Comparing two methods to measure preferred listening levels of personal listening devices. *J Acoust Soc Amer* 2009;125:3733–41. <https://doi.org/10.1121/1.3125798>
- [60] Breinbauer HA, Anabalón JL, Gutierrez D, Cárcamo R, Olivares C, Caro J. Output capabilities of personal music players and assessment of preferred listening levels of test subjects: Outlining recommendations for preventing music-induced hearing loss. *Laryngosc* 2012;122:2549–56. <https://doi.org/10.1002/lary.23596>
- [61] De Leeuw J jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behav Res Methods* 2015;47:1–12. <https://doi.org/10.3758/s13428-014-0458-y>
- [62] Bazilinskyy P, De Winter JCF. Crowdsourced measurement of reaction times to audiovisual stimuli with various degrees of asynchrony. *Hum Factors* 2018;60:1192–206. <https://doi.org/10.1177/0018720818787126>
- [63] Bazilinskyy P, Kooijman L, Dodou D, De Winter JCF. How should external Human-Machine Interfaces behave? Examining the effects of colour, position, message, activation distance, vehicle yielding, and visual distraction among 1,434 participants. *Appl Ergon* 2021;95:103450. <https://doi.org/10.1016/j.apergo.2021.103450>
- [64] Sripada A, Bazilinskyy P, De Winter J. Automated vehicles that communicate implicitly: examining the use of lateral position within the lane. *Ergon* 2021;64:1416–28. <https://doi.org/10.1080/00140139.2021.1925353>
- [65] TTSM3P. Free text-to-speech and text-to-mp3 for US English. <https://ttsmp3.com>; 2022
- [66] Mok CS, Bazilinskyy P, De Winter JCF. Stopping by looking: A driver-pedestrian interaction study in a coupled simulator using head-mounted displays with eye-tracking. *Appl Ergon* 2022;105:103825. <https://doi.org/10.1016/j.apergo.2022.103825>
- [67] Onkhar V, Bazilinskyy P, Dodou D, De Winter JCF. The effect of drivers' eye contact on pedestrians' perceived safety. *Transport Res F: Traffic Psychol Behav* 2022;84:194–210. <https://doi.org/10.1016/j.trf.2021.10.017>
- [68] Oudshoorn M, De Winter JCF, Bazilinskyy P, Dodou D. Bio-inspired intent communication for automated vehicles. *Transport Res F: Traffic Psychol Behav* 2021;80:127–40. <https://doi.org/10.1016/j.trf.2021.03.021>
- [69] International Organization for Standardization. Acoustics — Methods for calculating loudness — Part 1: Zwicker method. Switzerland; 2017. <https://www.iso.org/standard/63077.html>
- [70] Zwicker E, Scharf B. A model of loudness summation. *Psychol Rev* 1965;72:3–26. <https://doi.org/10.1037/h0021703>
- [71] Zwicker E. Procedure for calculating loudness of temporally variable sounds. *J Acoust Soc Amer* 1977;62: 675–82. <https://doi.org/10.1121/1.381580>
- [72] International Organization for Standardization. Acoustics — Methods for calculating loudness — Part 2: Moore-Glasberg method. Switzerland; 2017. <https://www.iso.org/standard/63078.html>
- [73] Moore BC, Glasberg BR. A revision of Zwicker's loudness model. *Acta Acustica* 1996;82:335–45.
- [74] Swart DJ, Bekkers A. Interior and motorbay sound quality evaluation of full electric and hybrid-electric vehicles based on psychoacoustics. *INTER-NOISE NOISE-CON Congr Conf Proc* 2016;253:7029–7039.
- [75] International Telecommunication Union - Radiocommunication Sector. Algorithms to measure audio programme loudness and true-peak audio level. ITU-R BS.1770-4. Switzerland; 2015. https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1770-4-201510-I!!PDF-E.pdf
- [76] European Broadcasting Union. Loudness metering: 'EBU mode' metering to supplement EBU R 128 loudness normalization. Tech 3341. Switzerland; 2016. <https://tech.ebu.ch/docs/tech/tech3341.pdf>
- [77] Lund T. ITU-R BS.1770 revisited. In: Proc. NAB-2011 Convention; 2011
- [78] Bisping R. Emotional effect of car interior sounds: Pleasantness and power and their relation to acoustic key features. *SAE Technical Pap Ser* 1995;104:2207–13. <https://doi.org/10.4271/951284>
- [79] Yang W, Makita K, Nakao T, Kanayama N, Machizawa MG, Sasaoka T, et al. Affective auditory stimulus database: An expanded version of the International Affective Digitized Sounds (IADS-E). *Behav Res Methods* 2018;50:1415–29. <https://doi.org/10.3758/s13428-018-1027-6>
- [80] Haas EC, Casali JG. Perceived urgency of and response time to multi-tone and frequency-modulated warning signals in broadband noise. *Ergon* 1995;38:2313–26. <https://doi.org/10.1080/00140139508925270>

- [81] Hellier E, Edworthy J (1989) Quantifying the perceived urgency of auditory warnings. Canadian Acoust 1989;17:3–11
- [82] Hellier EJ, Edworthy J, Dennis IAN. Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. Hum Factors 1993;35:693–706. <https://doi.org/10.1177/001872089303500408>
- [83] Watts GR, Godfrey NS, Savill TA. Development of audible locating signals for use at pedestrian crossings. Appl Acoust 2001;62:15–27. [https://doi.org/10.1016/S0003-682X\(00\)00027-X](https://doi.org/10.1016/S0003-682X(00)00027-X)
- [84] Fastl H, Zwicker E. Stimuli and procedures. In: Fastl H, Zwicker E. Psychoacoustics: Facts and models. Berlin Heidelberg: Springer Science; 2007, p. 1–15.
- [85] Moore BC. Development and current status of the “Cambridge” loudness models. Trends Hearing 2014;18:1–29. <https://doi.org/10.1177/2331216514550620>
- [86] Skovenborg E, Lund T. Loudness descriptors to characterize programs and music tracks. In: Audio Eng Soc Convention 2008;125
- [87] Hove MJ, Vuust P, Stupacher J. Increased levels of bass in popular music recordings 1955–2016 and their relation to loudness. J Acoust Soc Amer 2019;145:2247–53. <https://doi.org/10.1121/1.5097587>
- [88] Skovenborg E, Nielsen SH. Evaluation of different loudness models with music and speech material. In: Audio Eng Soc Convention 2004;117
- [89] Bate AJ. Cockpit warning systems comparative study. Final Report AMRL-TR, 68–193; 2004. Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command; 1969
- [90] Bazilinskyy P, Dodou D, De Winter J. Survey on eHMI concepts: The effect of text, color, and perspective. Transport Res F: Traffic Psychol Behav 2019;67:175–94. <https://doi.org/10.1016/j.trf.2019.10.013>
- [91] Chi CF, Dewi RS, Surbakti YY, Hsieh DY (2017) The perceived quality of in-vehicle auditory signals: a structural equation modelling approach. Ergon 2017;60:1471–1484. <https://doi.org/10.1080/00140139.2017.1323121>
- [92] Yamauchi K, Choi JD, Maiguma R, Takada M, Iwamiya SI A basic study on universal design of auditory signals in automobiles. J Physiol Anthropol Appl Hum Sci 2004;23:295–298. <https://doi.org/10.2114/jpa.23.295>
- [93] Moore D, Currano R, Strack GE, Sirkin D The case for implicit external human-machine interfaces for autonomous vehicles. Proc 11 Int Conf Autom User Interf Interact Veh Appl 2019;295–307. <https://doi.org/10.1145/3342197>
- [94] Lee YM, Madigan R, Giles O, Garach-Morcillo L, Markkula G, Fox C, et al. Road users rarely use explicit communication when interacting in today’s traffic: implications for automated vehicles. Cogn Techn Work 2021;23:367–380. <https://doi.org/10.1007/s10111-020-00635-y>
- [95] Lee YM, Madigan R, Uzondu C, Garcia J, Romano R, Markkula G, Merat N. (2022). Learning to interpret novel eHMI: The effect of vehicle kinematics and eHMI familiarity on pedestrian crossing behavior. J Safety Res 2020;80:270–80. <https://doi.org/10.1016/j.jsr.2021.12.010>
- [96] De Clercq GK, Dietrich A, Núñez Velasco P, De Winter JCF, Happee R. External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. Hum Factors 2019;61: 1353–70. <https://doi.org/10.1177/0018720819836343>
- [97] Kaleefathullah AA, Merat N, Lee YM, Eisma YB, Madigan R, Garcia J, De Winter JCF. External Human-Machine Interfaces can be misleading: An examination of trust development and misuse in a CAVE-based pedestrian simulation environment. Hum Factors 2022;64:1070–85. <https://doi.org/10.1177%2F0018720820970751>
- [98] Maunder M. Experiences tuning an augmented power unit sound system for both interior and exterior of an electric car. SAE Tech Paper 2018:2018-01-1489. <https://doi.org/10.4271/2018-01-1489>

Appendix A. Sound signals and discrete Fourier transforms of the auditory stimuli

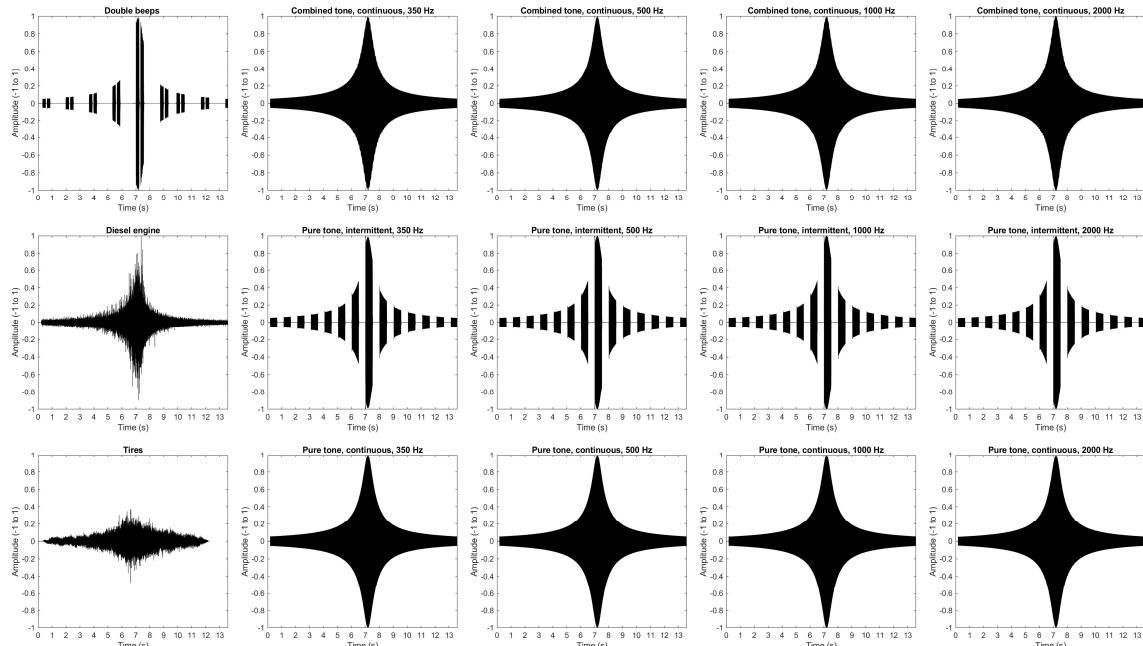


Figure A1. Sound signal as a function of the elapsed time for the auditory stimuli (without background noise)

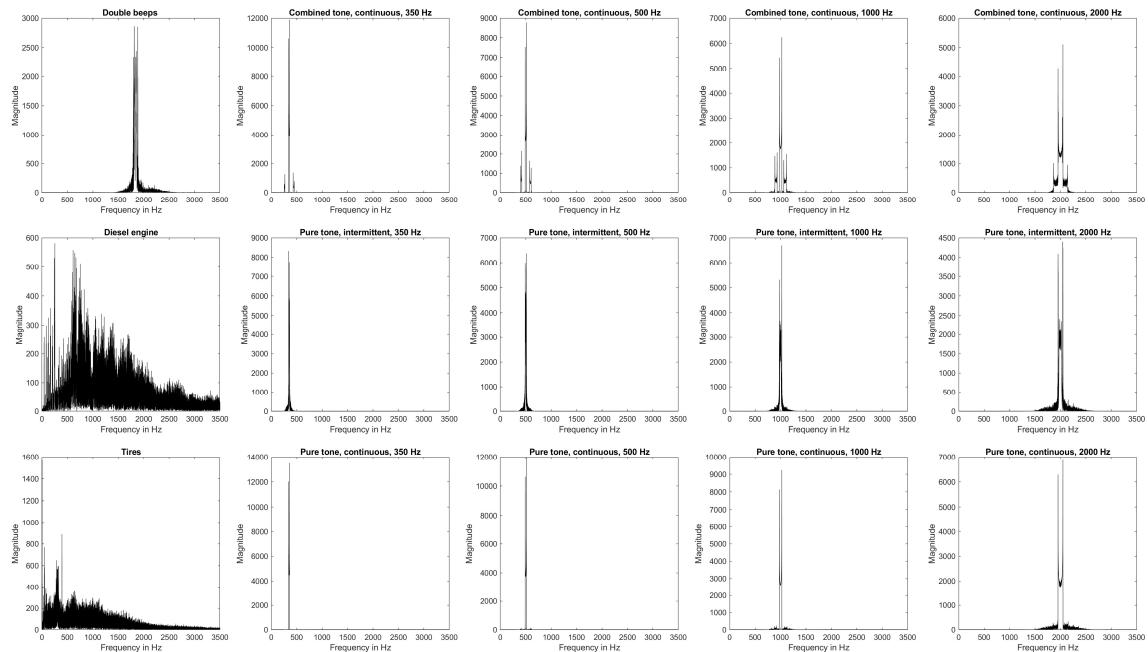


Figure A2. Fourier transforms of the auditory stimuli (without background noise)

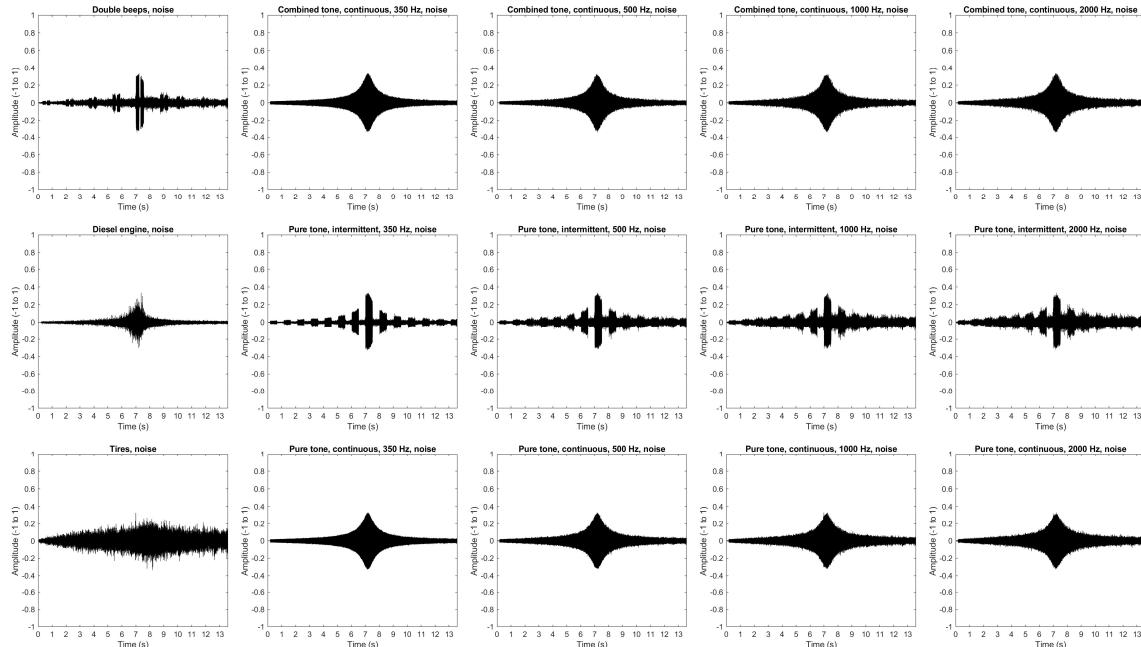


Figure A3. Sound signal as a function of the elapsed time for the auditory stimuli (with background noise)

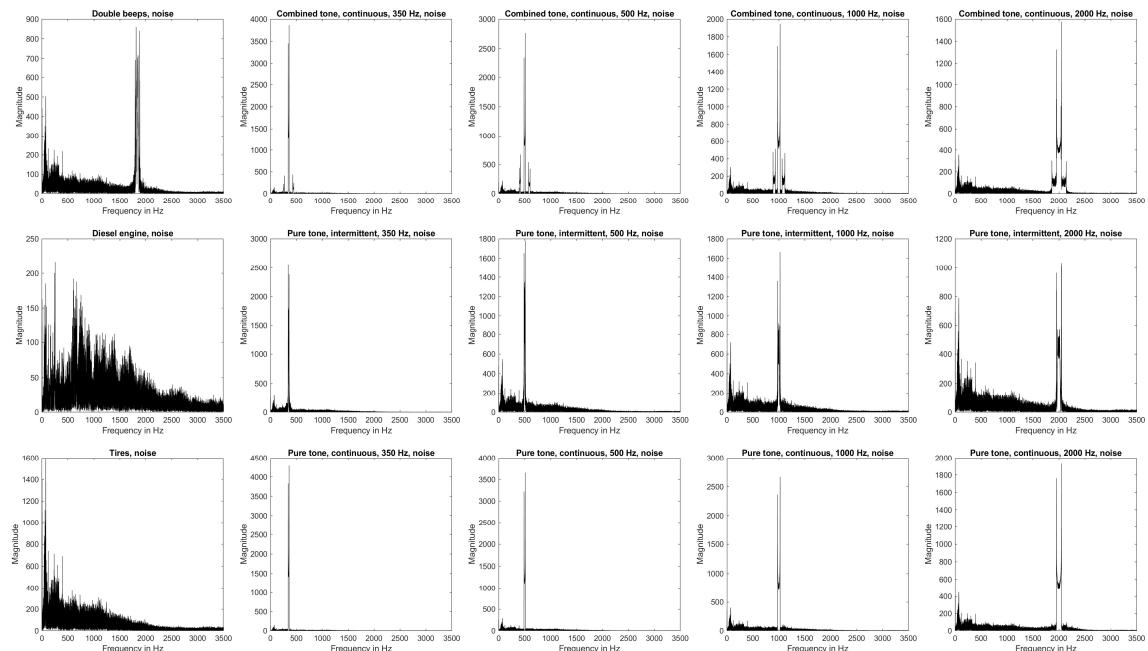


Figure A4. Fourier transforms of the auditory stimuli (with background noise)