

Article

# Effect of Emotionalizing Sounds on the Estimation and Evaluation of Displayed Safety Distances

Manuel Petersen <sup>\*</sup>, Deniz Yüksel and Albert Albers 

IPEK—Institute of Product Engineering, Department of Mechanical Engineering, Karlsruhe Institute for Technology (KIT), Kaiserstraße 10, 76131 Karlsruhe, Germany; albert.albers@kit.edu (A.A.)

\* Correspondence: manuel.petersen@kit.edu

**Abstract:** Musicological and traffic psychology research shows that emotions can be changed by certain tone combinations or sound characteristics and that emotions, in turn, influence our driving behavior. Nevertheless, there are no studies on how a dynamic active sound design could influence driving behavior via changing the emotional state of drivers in certain driving situations. Based on a previous study, emotionalizing sounds, characterized by their capacity to evoke specific emotional responses in individuals, were created and used to investigate their effect on the perception of safety distances in an online study. To test this, participants made statements on the safety distance shown in videos of cars following scenarios combined with emotionalizing sounds. The results show a significant difference in the estimated safety distance for videos combined with sounds invoking positive emotions like light-heartedness vs. sounds invoking negative emotions like feeling threatened. The odds of the safety distance being evaluated as too small compared with appropriate were two to three times higher for some threatening sounds vs. the positive sounds. The results further suggest that threatening sounds influenced participants' wishes to increase the depicted safety distances. The results show that emotionalizing sounds had effects on the participants, though not all were statistically significant.

**Keywords:** vehicle interior noise; active sound design; electric vehicle; psychoacoustics; sound evaluation; emotions; safety distance; participant study



**Citation:** Petersen, M.; Yüksel, D.; Albers, A. Effect of Emotionalizing Sounds on the Estimation and Evaluation of Displayed Safety Distances. *Acoustics* **2024**, *6*, 386–407. <https://doi.org/10.3390/acoustics6020021>

Academic Editors: Simone Torresin and Jian Kang

Received: 12 March 2024

Revised: 23 April 2024

Accepted: 26 April 2024

Published: 30 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The field of sound design is becoming increasingly relevant in the development of automobiles, especially for electric vehicles, which usually lack a discernible engine sound. Approximately 5% of the development budget for a new car is allocated to the different aspects of sound design in the context of electrical vehicles [1]. This includes various auditory warning systems in vehicles for different hazards and manufacturers creating a brand-specific sound through artificially generated engine noise [2]. Streicher et al. note the following regarding vehicle sounds: “The sound inside the vehicle is a significant sensory impression for the occupants, conveying numerous pieces of information”, meaning on the one hand important feedback on the driving condition, and on the other hand, the representation of the character and value of a vehicle [3]. A study about the acceptance of synthetic vehicle sounds showed that the subjects generally preferred a quiet electric vehicle, but at the same time, demanded an adequate acoustic load feedback [4]. The results of Pilgerstorfer et al.'s study showed that electric vehicles are much quieter than combustion engine vehicles at speeds below 50 km/h. At the same time, in this speed range, people drove faster in electric vehicles than in combustion engine vehicles. This indicates that the absence of acoustical feedback in electrical vehicles could lead to increased vehicle speeds in low-speed situations [5].

The assessment of the current driving situation by the driver involves processing various sensory inputs. Roach et al. highlight that the perceptual system integrates

sensory cues, such as visual and auditory cues, to derive the most reliable estimation of the present situation [6]. Sounds, instrumental tones, and noises possess various characteristics, including pitch, tone/timbre, tonality, loudness, and rhythm. These characteristics can be related to acoustical and psychoacoustic parameters, with examples such as harmonics, sharpness, roughness, and fluctuation strength describing the tone/timbre of a sound, and periodicity or impulsiveness characterizing rhythmic qualities [7,8]. Studies have demonstrated that alterations in auditory stimuli, including factors like loudness [9,10], frequency composition [11], and unrelated auditory stimuli like music [12], can impact the perception of driving speeds. Brodsky conducted an experiment showing that music had an impact on physiological and behavioral aspects during simulated driving. The study consistently demonstrated that an increasing tempo of background music increased both simulated driving speed and speed estimates. Additionally, a higher tempo of background music consistently increased the frequency of virtual traffic violations, such as disregarding red traffic lights [13].

When analyzing the influence of auditory stimuli on a driver, one should not only consider cognition but also the effect on emotions. Generally, individuals are inclined to seek situations with high valence ratings (pleasant) and avoid those with low valence ratings (unpleasant). In situations with high/low valence ratings, the desire to approach/avoid said situations intensifies with the level of arousal [14]. Specific intervals between two simultaneously played tones evoke different emotions based on the frequency difference between the two tones [15]. Three or more simultaneously played tones in a musical scale are called chords. Chords containing dissonant intervals like diminished chords are generally rated as more negative sounding compared with low-dissonance major chords [16]. Higher tones or chords are generally associated with more positive emotions, while lower tones are linked to more negative emotions [17]. These aspects and characteristics of music can influence the overall valence and arousal experienced by individuals, eliciting emotions such as fear, sadness, or happiness [7,18,19]. A key focus of research in the domain of active sound design is increasing the positivity of artificial sounds or inducing positive emotions through the sounds produced by electric vehicles [20,21]. However, emotionally induced states of anxiety and fright have been shown to influence driving behavior, such as velocity and acceleration behavior [22]. Further, negative affective auditory stimuli have been shown to reduce reaction times for braking in both risky and nonrisky situations [23], and they are also associated with lower tendencies of risk-taking [24]. These results indicate that the utilization of an active sound design (ASD) has the potential to improve feedback about the current driving situation and influence safety-related factors, such as risk-taking and speed estimation.

Falling below a safe following distance while driving represents one of the most common causes of car accidents and is a leading factor in casualties [25]. To address this issue, numerous emergency braking systems exist that intervene in critical situations [26]. However, these systems do not ensure adherence to safe distances. While systems aiding in maintaining proper following distances exist, they are not widely adopted, and there is currently no mandatory implementation as with emergency braking systems [27]. The challenge lies in the fact that their functionality closely resembles conventional warning systems, causing drivers to overlook critical cues amid the multitude of stimuli [28].

An alternative approach could be offered through the implementation of an emotion-alizing sound design. This entails an active approach to sound design aimed at altering the emotional state of a driver in specific driving situations, thereby influencing their driving behavior. This shift in the driver's emotional state is achieved by modifying the valence of the active sound design itself, transitioning, for instance, from a positive to a negative mood in the sound. Numerous contemporary vehicles come equipped with sensors that continuously measure distances from other road users [27]. By utilizing vehicle parameters and data from these sensors, an active sound design module could adjust the mood of the sound design (from positive to negative) based on deviations from the necessary safety

distance. This adjustment would consequently impact the driver and likely influence their driving behavior [29].

In order to influence the emotional state of a driver effectively via an active sound design, in a previous study, Petersen et al. created and identified fitting affective sounds as stimuli. For this, 16 sounds were created using professional tools and then validated regarding their inherent emotional character and their emotionalizing impact in a study. The participants' evaluations were then correlated with the psychoacoustic features of the sounds, showing that experiencing a sense of threat while hearing the sound was associated with fluctuation strength and impulsiveness. Conversely, feeling calm or light-hearted showed a strong positive correlation with low fluctuation strength, low impulsiveness, and a high level of tonality [30]. These results could be useful to enhance targeted emotional effects and facilitate sound optimization for the validation phase of an emotionalizing sound design.

This paper aims to investigate whether different emotionalizing sounds could have an impact on relevant factors for driving behavior in regard to safety distance. For the relevant factors, we decided on the estimation, desire to change, and evaluation of the safety distance and propose the following hypotheses:

**Hypothesis 1.** *Sounds that invoke/are associated with (a) negative/(b) positive emotions lead to a (a) decrease/(b) increase in the estimated safety distance depicted in videos.*

**Hypothesis 2.** *Sounds that invoke/are associated with negative or positive emotions have an influence on the desired increase in the safety distance depicted in videos.*

**Hypothesis 3.** *Sounds that invoke/are associated with (a) negative/(b) positive emotions lead to increased odds of a depicted safety distance being evaluated as (a) too small/(b) appropriate.*

Furthermore, this paper investigates whether the previously identified psychoacoustic parameters [30] can be used to optimize the emotionalization of participants through sound. Additionally, we want to investigate if the sounds that we create with our own sound generator are as effective as the sounds created with professional tools. For the investigation, a video-based online participant study was conducted.

## 2. Materials and Methods

### 2.1. Acoustic Stimuli

The acoustical stimuli for the study are 5 different emotionalizing sounds. The sounds were selected and created with the goal of effectively emotionalizing the participants positively or negatively while listening to the sounds. The 5 emotionalizing sounds used in this study consist of two sounds from a previous study (the most positive and the most negative/threatening) created with professional tools [30], and three new sounds created with our own sound generator. The first new sound is a recreation of the most threatening sound of the previous study created in our own sound generator. It was recreated to evaluate if we were able to create equally affective sounds with our sound generator as with the professional tools used for the sounds in the first study. The other two new sounds were created based on the found correlations between psychoacoustic parameters (tonality, fluctuation strength, impulsiveness) and a positive and negative/threatening perception of the sounds [30]. All of the created sounds are stereo. To consider the temporal and logical structure of the underlying choice and creation process of the sounds, this paragraph first discusses the reused sounds and then the newly created sounds. The sounds are numbered in ascending order from positive to negative emotional effectiveness throughout this paper. An overview of the parameters and characteristics of the 5 different sounds can be seen in Table 1. The sounds can be accessed via the link in the Supplementary Materials section.

**Table 1.** Overview of the characteristics of the 5 different acoustic stimuli.

Sound Name	Chord/Harmonic Structure	Harmoniousness	Primary Sound Characteristics	Targeted Emotion
Sound 1	Major Chord	Very Harmonic	Stacked sine waves in 100–1000 Hz range, very pure sound. Very slight movement in amplitude.	very positive, calming
Sound 2	Harmonious Note Cluster	Harmonic	Church-organ-like sound in 100–2000 Hz range, very tonal, very slight movement in amplitude.	positive
Sound 3	Disharmonious Note Cluster	Disharmonic	Hollow-sounding tonal component (100–900 Hz), unsettling sound. Metallic hissing sound component (2000–10,000 Hz) with sinusoidal amplitude modulation at 8–10 Hz for rhythmic character.	negative, unsettling, threatening
Sound 4	Diminished	Disharmonic	Hollow-sounding tonal component (100–1000 Hz), unsettling, slightly aggressive sound. Metallic hissing sound component (2000–10,000 Hz) with sinusoidal amplitude modulation at 8–10 Hz for rhythmic character.	negative, unsettling, threatening
Sound 5	Major 7 suspended 2	Very Disharmonic	Hollow-sounding tonal component (100–1200 Hz), very out-of-tune-sounding. Metallic hissing sound (2000–10,000 Hz), strong square wave amplitude modulation at 2 Hz for hammering aggressive character.	very negative, disturbing, very threatening

The two sounds taken from the previous study were reused, since their emotionalizing effect had already been established [30]: one sound that was perceived as most light-hearted and calming (Sound 2) and one sound that was perceived as the most threatening in the previous study (Sound 3). Sound 2 has a harmonic church-organ-like sound, is very tonal, and has some slight sinusoidal amplitude modulation to give it a bit more liveliness. The sound ranges from 100 Hz to 2000 Hz. The foundation for the sound was not a musical chord but rather a harmonious note cluster, consisting of multiple notes, generated with a neural network trained on positive music [30]. To create the sound, these notes were then used as input for the Cinematic Synth from Acid Pro using the “Jingling Waterpot” preset (Acid pro, version 11, MAGIX Software GmbH, Berlin, Germany). Sound 3 consists of two primary sound characteristics: A hollow more disharmonic-sounding tonal component in the frequency range of 100 Hz to 900 Hz, which targets an unsettling feeling. The foundation for the sound was a disharmonious note cluster generated with a neural network trained on negative and threatening music [30]. To create the sound, these notes were then again used as input for the Cinematic Synth using the “Black Shadow” preset (Acid pro, version 11). Further, it contains a metallic hissing sound in the frequency range of 100 Hz to 900 Hz, which is amplitude-modulated at 8–10 Hz with a sinusoidal modulator for a rhythmic character, as these were also found to be related to a threatening perception of sounds [7].

For the creation of the new sounds, we used our sound generator [29], which is aimed at enabling us to create emotionalizing active sound designs that can be dynamically controlled in real time to use them in a driving simulator in future studies. For the creation of the new sounds, the existing sound generator [29] was extended by the possibility for variable amplitude modulation, to increase the psychoacoustic impulsiveness and fluctuation strength. It was further extended by a modulated high-pass-filtered white noise module in conjunction with a delay effect set to a very low feedback time (10.2 ms) to recreate the metallic hissing prevalent in the threatening Sound 3 from the evaluation study that needed to be recreated.

The first newly created Sound 4 is an approximated recreation of the threatening sound of the previous study (Sound 3). We recreated it to test whether our sound generator is capable of creating equally affective and effective sounds. Sound 4 is similarly hollow-sounding compared to Sound 3, but it has a slightly more forefront and aggressive quality to it. It consists of 3 stacked disharmonic tones per each of the chord notes to approximate the tonal characteristics of Sound 3. Its underlying chord structure is a disharmonic diminished chord, which is associated with negative emotions [17]. The hollow sound characteristic was achieved with a chorus effect that combines multiple instances of a sound

with continuously shifting phase information between the combined sounds. The sound also incorporates the amplitude-modulated rhythmic metallic hissing sound apparent in Sound 3.

Two additional sounds were created based on the correlations found between the psychoacoustic parameters of the sounds and the emotional reception of the sounds by the participants in the previous study [30]: one positive sound (Sound 1) and one threatening sound (Sound 5). Positive Sound 1 is a very harmonic sound targeting a positive emotion. It consists of 3 stacked sinus tones for each of the chord notes (fundamental tone, 1 octave above, and 2 octaves above fundamental tone). It is based on a major chord, which has a positive, very harmonious sound. It further targeted a high psychoacoustic tonality, which correlated positively with positive emotions, and low impulsiveness and fluctuation strength, which correlated negatively with positive emotions in the previous study. The newly created negative/threatening sound (Sound 5) used the recreated threatening Sound 4 as a foundation to create an even more affective and effective sound. A strong square-wave-based amplitude modulation over the whole frequency spectrum at around 2 Hz was added to Sound 4 to create Sound 5. This was performed to increase the psychoacoustic impulsiveness and fluctuation strength of Sound 5 which were previously found to correlate strongly with a sound being perceived as more threatening.

To increase the level of realism, and make the acoustic stimuli more immersive, the created sounds were combined with a binaural soundscape recorded in the driving cabin of a small-car electrical vehicle without an active sound design. During the recording, all auxiliary systems that could induce additional noise into the vehicle cabin, like the heating, ventilation, or air conditioning systems, were turned off. The binaural recordings were conducted during the recording of the video material at constant velocities between 50, 60, and 80 km/h. An artificial head (Manakin Mk1 Cortex, dBsonic, Eseneler, Istanbul) placed on the passenger seat was used for the recordings. To exclude the possibility of an influence from the binaural recordings at different speeds, the same binaural measurement at 60 km/h was used for all of the sounds in the study. The combination of the created sounds and the binaural measurement was performed in Ableton Live, a digital audio workstation (Ableton Live, version 11.3, Ableton AG, Berlin, Germany). The level of the binaural soundscape layer was the same for all sounds and was set to a level at least 10 dB below the quietest sound, so the soundscape was noticeable but did not mask any of the actual stimuli sounds. The combined sounds (artificial sounds + vehicle cabin noise) were then adjusted in Ableton Live to have the same Loudness Units Full Scale [31]. With the exported sounds, an iterative adjustment of the individual levels of the sound was conducted to reduce the differences in perceived loudness between the combined sounds. The adjustments were carried out based on the feedback from a subjective evaluation of the loudness differences in the sounds by a consortium of five scientific employees of the vehicle acoustics department of the institute. The initially perceived loudness by the consortium of positive sounds (Sounds 1 and 2) was lower than the threatening sounds (Sounds 3, 4, and 5). Based on iterative feedback, the level of Sound 1 was increased by 5 dB, and Sound 2 was increased by 4 dB. After this adjustment, the consortium agreed that all the sounds were perceived as equally loud. The level of the binaural background layer was not adjusted and stayed equal for all sounds. All sounds had a level fade in and fade out of 0.1 s to fit the fade in of the videos and to reduce the sensation of sudden impulses like loud noises in participants when playing back the sounds.

## 2.2. Visual Stimuli

For the visual stimuli, 5 different real-world car-following scenarios were recorded with a GoPro 5 camera (GoPro, version 5, San Mateo, CA, USA) from a position close to the driver's head position with a constant velocity and a constant safety distance during the recording, as suggested by [32]. The drives were recorded on country roads around Karlsruhe, Germany, and on the South Campus of the KIT. The scenarios presented to participants aligned with probable scenarios encountered in real-world car-following

situations to increase the validity of the experiment [33]. The recorded drives differed in location and driving speeds, and they depicted safety distances to avoid a high similarity between videos and significant learning effects in participants [34]. The videos for this study were recorded at a resolution of 1080 p and 60 fps in fine weather conditions based on the approach of Horswill et al. on video-based measures of drivers' following distance and gap acceptance behaviors [32]. The different locations, vehicle speeds, and actual safety distances depicted in the videos can be seen in Table 2. The driven vehicle was a small-car electrical vehicle, and the car driving ahead was a gray upper-middle-class fuel-cell sedan. The videos can be accessed via the link in the Supplementary Materials section.

**Table 2.** Boundary conditions of the 5 different depicted videos of car following scenarios.

Video Name	Distance (m)	Velocity (km/h)	Location	
Video 1	15	60	City road (alley)	
Video 2	12	50	KIT internal road	
Video 3	12	80	Country road (alley)	
Video 4	15	80	Country road	
Video 5	15	80	Country road	

The depicted safety distances were chosen by a consortium of 5 scientific employees of the institute based on the target to look at the threshold of being too close with respect to the size of the video playback window inside the online experiment software (Unipark, version 21.1, Tivian XI GmbH, Cologne, Germany). The instruments depicting the velocity of the vehicle used in the videos were blackened to increase the necessity to make an intuitive estimation of the shown safety distance and minimize the influence of the evaluation regarding the appropriateness of the shown safety distance.

To avoid videos being shown in a determined sequence and thus risk distorted results [35], the videos were presented to the participants in a random manner, though each

video was only evaluated once per participant. To further necessitate the re-evaluation for each shown video, the videos randomly alternated between the 12- and 15-m videos in regards to the depicted safety distances.

### 2.3. Participants

A total of 169 people participated in this study. The sample was of rather young age (mean = 33.3, SD = 12.1, median = 28.5, range: (17, 80)), predominantly identifying as male (female: 59, male: 107, other: 1). A total of 159 participants had owned a driving license (7 had no driving license) for approximately 11 years (mean = 10.9, SD = 23.8, median = 10.0, range: (1, 74)). Only 9 participants stated they had had an accident in the last 2 years (9 made no statement). Most participants reported driving up to 6000 km per year (up to 6000 km: 68 participants; 6001 km–9000 km: 33 participants; 9001 km–12,000 km: 25 participants; 12,001 km–15,000 km: 16 participants; more than 15,000 km: 17 participants). A similar number of participants reported usually driving in cities (104 participants), in rural areas (84 participants), and on motorways (85 participants).

### 2.4. Methods

The experiment was a video-based online study in the German language, which the participants carried out on their own computer and audio hardware.

When beginning this study, the participants were briefed to use a PC/laptop/tablet to keep the screen size comparable between participants. They were further instructed to use good headphones throughout the experiment and no external speakers like laptop speakers or tablet speakers. The participants predominantly followed these instructions. Participants who used speakers were exempt from the study. The choice of headphones was diverse, from high-quality neutral studio headphones to small earbuds. The participants were further briefed to estimate and evaluate the safety distance intuitively and not use any methods for estimation of safety distances. With this, we wanted to reduce the probability of people using estimation methods for the safety distance. Common methods are counting the seconds between vehicles to derive a time headway or using the guide posts that are omnipresent in Germany and are often known to be 50 m apart. The participants were not told these examples of possible methods. The participants were told that this study is only about subjective estimation/evaluation of safety distances, and they did not know prior to the study that there were emotional aspects considered in the study.

This study consisted of two main parts. The first part was about the estimation and evaluation of the safety distances in the shown video–sound combinations. A total of 5 video–sound combinations from a pool of 25 videos total (each a unique combination of 1 of the 5 video stimuli and 5 different acoustic stimuli) were randomly assigned to be shown to each participant. A length of 7 s per video was deemed suitable for the following reasons: It is short enough so that the test participants would ideally not have time to resort to methodologies for the estimation of the safety distance instead of making intuitive decisions. On the other hand, the videos are long enough so that the test participants have enough time to evaluate the safety distance and so that the acoustical stimulus can take effect. If the videos or stimuli were much longer, this could lead to a carry-over effect, where the participants would take the effect of one stimulus into the next stimulus [36]. The selection of the 5 sound and video combinations and the sequence of the combinations was randomized for each participant, and each underlying video and sound stimulus was presented to each participant only once. To further necessitate the re-evaluation for each shown video, the videos randomly alternate between the 12- and 15-m videos in regard to the depicted safety distances. The participants were told to not maximize the videos to reduce the impact of different screen sizes used by the participants during the study. After watching a video, the participants made their statements about the estimation and evaluation of the safety distances. This was performed by directly entering numbers for the estimated safety distance in meters and choosing an answer about the sufficiency of the depicted safety distance via radio buttons. If the safety distance was evaluated as too small

or too large, the participants had the possibility to state by how much they would want to change the safety distance. Entering numbers and radio buttons were used to keep the input process as simple as possible, as suggested in [37].

The second part of this study was about the evaluation of the sounds in regard to how participants felt while listening to the sounds, as well as the emotions the participants associated with the sounds. The descriptors for stating the emotions felt, as well as the descriptor pairs for the semantic differentials for the associated emotions (7-level bipolar Likert scales) were taken from existing scales concerning the emotional reception of sounds and noises, and in particular, vehicular interior noises [38–40]. Descriptors and descriptor pairs that were not relevant to the research question were filtered out to keep the processing time of the study at a reasonable level. The remaining descriptors for the emotions felt were serious, sad, afraid, light-hearted, threatened, calm, and excited. The descriptor pairs for the Likert scale were threatening–harmless, unpleasant–pleasant, repulsive–attractive, obtrusive–unobtrusive, exhausting–relaxing, calming–exciting, boring–stimulating. The order of all the descriptors was randomized to prevent order effects. The descriptors from the existing scales were chosen because they were readily available in the German language, were used in the context of vehicle sounds, and described the different emotional states sufficiently in a concise way. Further, the descriptors were used in the previous study about the emotional effect and associated emotions with sounds [30]. It was necessary to use the same descriptors to evaluate whether the newly created sounds would be as effective as the sounds of the previous study. Further, we wanted to compare if the sounds of the previous study would be evaluated differently if they were evaluated between the newly created sounds. There was an introductory example at the beginning of both parts explaining the questionnaire. The introduction for the first part was also used to properly set up the listening conditions by the participants to hear the sounds loudly and clearly. Participant were told to not change their sound settings after this initial calibration step.

For the analysis of the acoustical characteristics of the sounds, ArtemiS SUITE was used (ArtemiS SUITE, version 14.2, HEAD Acoustics GmbH, Herzogenrath, Germany). The following psychoacoustical analyses were used: Loudness vs. Time (ISO 532-1) [41], Fluctuation Strength vs. Time (ArtemiS SUITE), Tonality (Hearing Model) vs. Time (ECMA-418-2) [42], Impulsiveness (Hearing Model) vs. Time (ArtemiS SUITE). We focused on these psychoacoustic analyses because they showed a good correlation with positive (light-hearted, calm) or negative (threatened) emotions in the previous study [30]. These psychoacoustic parameters were also used to optimize the newly created sounds in regard to their emotional effectiveness.

For the statistical analysis, a one-tailed *t*-test for repeated measures was used to find significant directed differences between the interval-scaled estimated safety distances (less for negative sounds), as well as the quantified desired change in the safety distance between the different sounds (greater for negative sounds), within participants. Cohen's *d* was used for the calculation of the statistical effect size of the sounds on the estimation and desire to change [43].

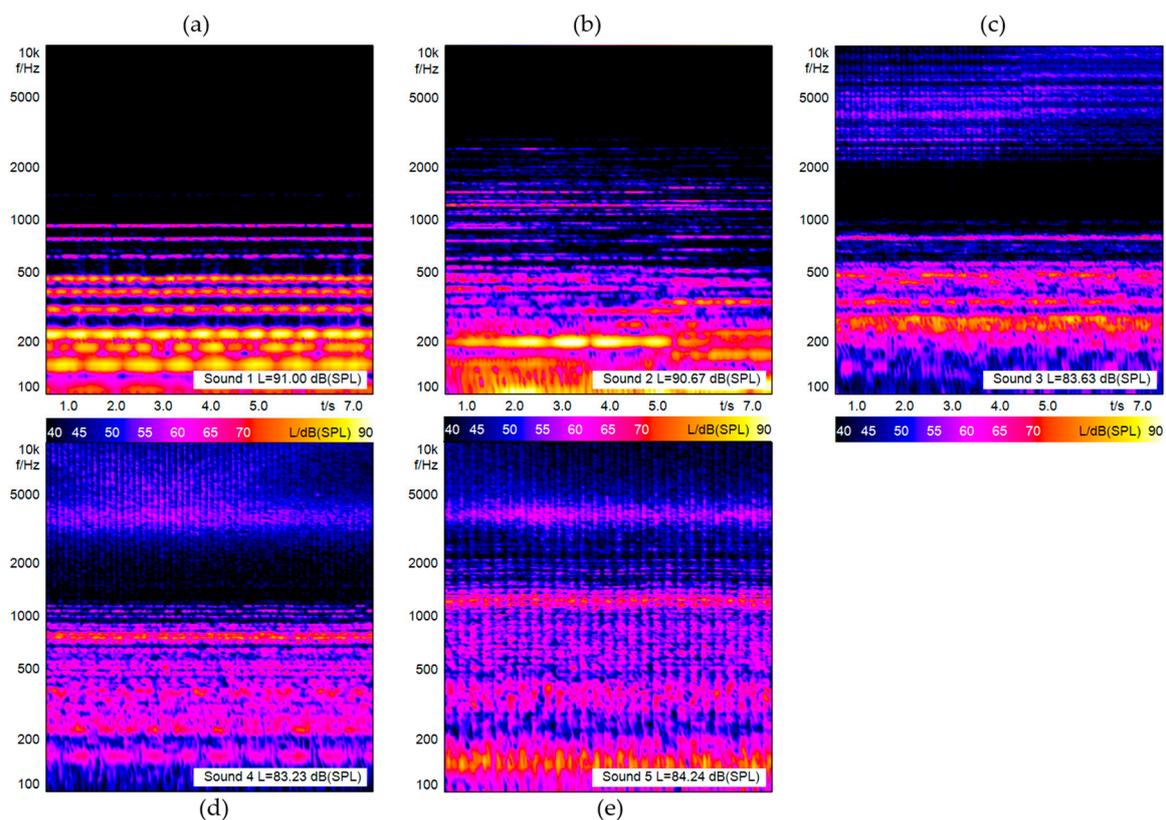
For the analysis of the effect of the sounds on the evaluation of the safety distance in regard to appropriateness, a binomial logistic regression for repeated measures was used. These models establish a connection between a categorical response variable and various independent variables (factors or covariates). In this context, the model involves a linear combination of independent variables, resulting in odds ratios representing the probability of a specific characteristic in relation to a binary variable (such as evaluating the adequacy of safety distance). This estimation is based on the values of independent variables like sounds, displayed safety distance, and velocity. As model selection criteria, we used Akaike's information criteria (AIC), Schwarz's Bayesian information criteria (BIC), and the deviance, pseudoconditional, and marginal *R*<sup>2</sup> values to select the best predictive model. AIC estimates the relative distance between the true and fitted likelihood functions of the data and model plus a constant. The AIC is to choose the model which yields the smallest value of AIC. The BIC gives a function of the posterior probability of a true model

under a certain Bayesian setup. The BIC criteria are to choose the model which yields the smallest value of BIC. Marginal  $R^2$  ( $R^2$  (fixed effects)) represents the proportion of variance explained by the fixed effects relative to the overall variance. Conditional  $R^2$  ( $R^2$  (total)) represents the proportion of variance explained by both fixed and random effects relative to the overall variance. Further evaluation of the model was conducted with the “DHARMA” package (Version 0.4.6) for R. It employs a simulation-based strategy to generate scaled (quantile) residuals that are easily interpretable for fitted (generalized) linear mixed models [44]

### 3. Results

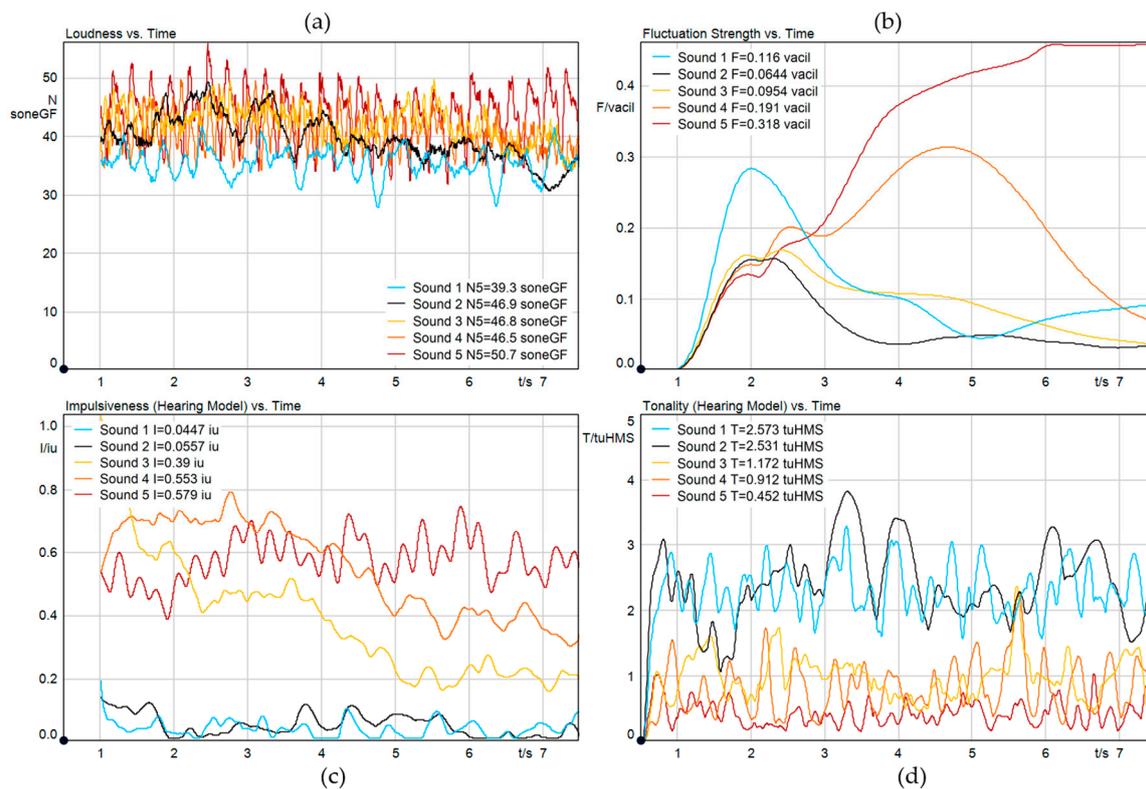
In this section, we first look at the sounds’ characteristics and what emotions they invoked in the participants, as well as how the sounds were perceived by the participants. Afterward, the influence on the estimation/evaluation of the safety distance is presented.

In Figure 1, the spectrograms, as well as the sound pressure levels, of the five different sounds used in the experiment are shown. The sounds that targeted positive emotionalization (Sound 1 (a), Sound 2 (b)) contain no contents in the frequency area between 2000 and 10,000 Hz compared with the targeted threatening sounds (Sound 3 (c), Sound 4 (d) and Sound 5 (e)). Furthermore, the strong amplitude modulation in the threatening sounds due to the targeted impulsiveness and fluctuation strength can be seen in the spectrograms.



**Figure 1.** Spectrograms and sound pressure levels of the used acoustical stimuli: (a) Sound 1; (b) Sound 2; (c) Sound 3; (d) Sound 4; (e) Sound 5.

To evaluate whether the acoustical stimuli contain the psychoacoustic properties that were found to correlate with a change in emotion or a sound being perceived as threatening or positive, the psychoacoustical loudness, fluctuation strength, impulsiveness, and tonality were analyzed. The results were averaged over both stereo channels. The analysis vs. time of the loudness, the fluctuation strength, the impulsiveness, and the tonality, as well as the respective averaged single values, can be seen in Figure 2.

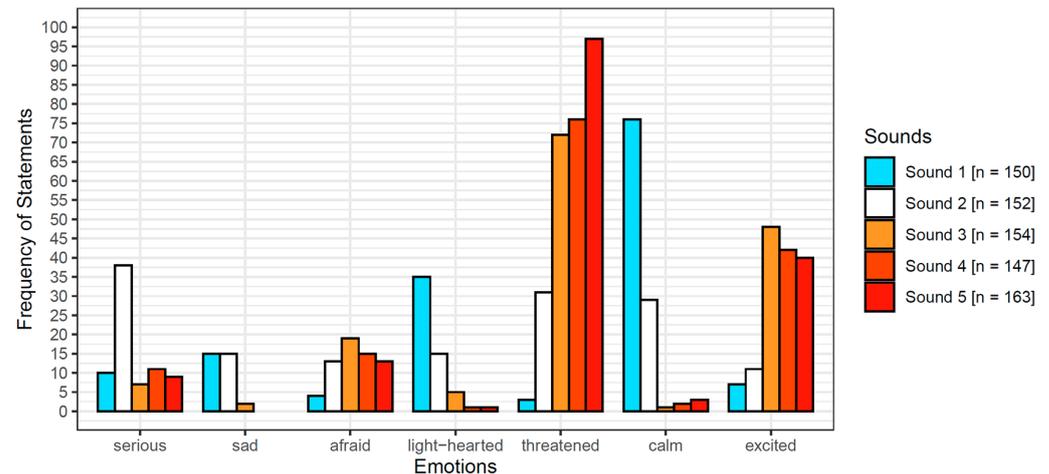


**Figure 2.** Psychoacoustic analysis vs. time and averaged single values of the used acoustical stimuli: (a) loudness; (b) fluctuation strength; (c) impulsiveness; (d) tonality.

In Figure 2a, it can be seen that despite the iterative adjustment of the volume levels, Sound 1 is on average still around 7 Sone less loud than Sounds 2, 3, and 4. Further, Sound 5 is 4 Sone louder than Sounds 2, 3, and 4. This was accepted by the researchers, because the arousal was found to correlate with loudness [30], and thus, the threatening effect of the acoustical stimulus should be increased, which fits the desired effects for the sounds. In Figure 2b, we see the fluctuation strength vs. time, as well as the averaged values for the five sounds. As seen in the spectrograms (Figure 1), the sounds are fairly stationary in regard to their characteristics in the beginning, so the fluctuation strength method of HEAD Artemis seems to show some settling time due to the quick fade in of the sounds. Despite this initial artifact, we can see that the average values for newly created positive Sound 1 (0.12 vacil) and the threatening Sound 4 (0.19 vacil) and Sound 5 (0.32 vacil) differ substantially in regard to fluctuation strength. As targeted, the optimized Sound 5 has the highest fluctuation strength. Although Sound 4 was a recreation of Sound 3, they differ in fluctuation strength (0.09 vacil for Sound 3 vs. 0.19 vacil for Sound 4). With the target of keeping the perceived sound characteristics similar, and a higher fluctuation strength correlating with a threatened feeling, this was an accepted deviation between the two sounds. Similar circumstances can be seen in Figure 2c. The positive sounds (Sound 1 and Sound 2) have a very low impulsiveness at 0.05 iu and 0.06 iu compared with the targeted threatening Sound 3 (0.39 iu), Sound 4 (0.56 iu), and Sound 5 (0.58 iu), with a slight offset between Sound 3 and 4. Again, we accept this offset in the context of the target of an increased threatening effect. In Figure 2d, we see the tonality vs. time, as well as the averaged tonality. It can be seen that the target to have high tonality for the positive Sound 1 (2.57 tuHMS) and low tonality for the threatening Sound 4 (0.91 tuHMS) and Sound 5 (0.45 tuHMS) was accomplished in the newly created sounds.

### 3.1. Felt and Associated Emotions vs. Sounds

To investigate the emotionalizing effect of the sounds, the first part of the questionnaire asked about the emotions the participants felt while listening to the different sounds. The emotional descriptors of the single-choice answers were serious, sad, afraid, light-hearted, threatened, calm, and excited. Figure 3 illustrates the frequencies of responses given by participants regarding the emotional descriptors while listening to the sounds.

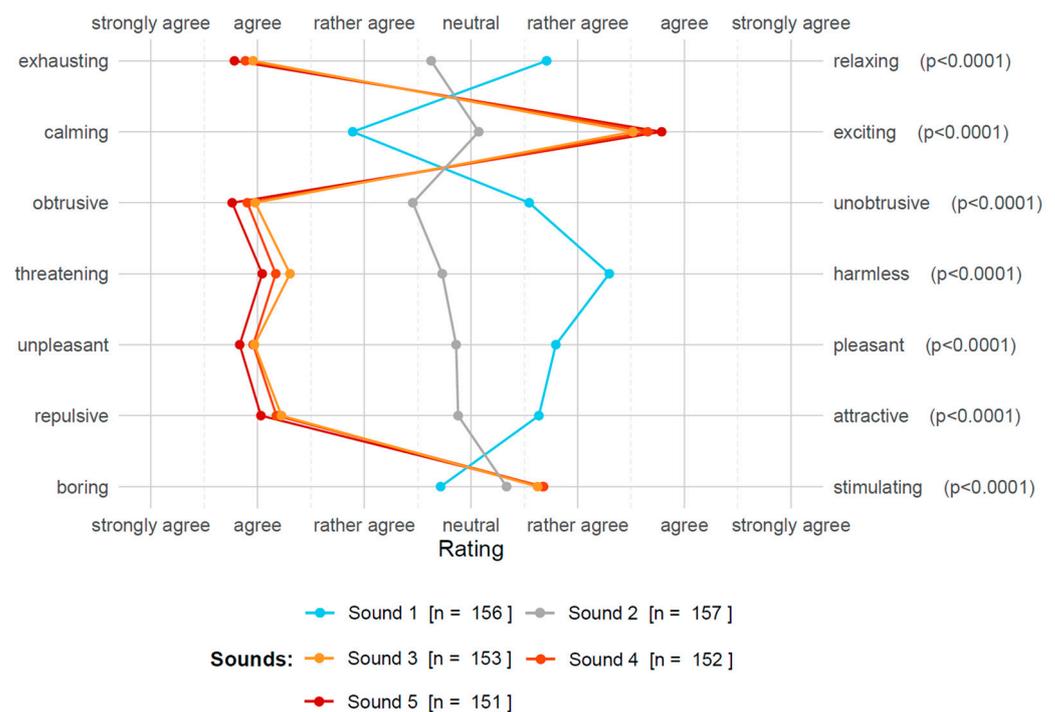


**Figure 3.** Frequency of responses by participants about felt emotions for each of the 5 presented sounds.

For the evaluation of the emotionalizing sounds, the interest is primarily on activating positive and negative emotions that could lead to a change in behavior or evaluation of the situation. For the positive emotions, these are light-hearted and calm, and for the negative emotions, threatened or afraid. Further of interest is the level of arousal in participants, which especially increases the effect of negative emotions on the urge to avoid or leave a situation. A high arousal level is represented by the descriptor excited.

As shown in Figure 3, Sound 1 was evaluated by 150 participants, of which only 5% felt threatened or afraid. A total of 75% of participants had positive emotions (light-hearted or calm) while listening. Only 5% felt excited by listening to the sound. Sound 2 was evaluated by 152 participants, of which 30% felt threatened or afraid, 30% felt light-hearted or calm, and 26% felt neutral (serious) while listening. A total of 8% felt excited by listening to the sound. Sound 3 was evaluated by 154 participants, of which 68% felt threatened or afraid. Only 4% had positive emotions (light-hearted or calm) while listening. A total of 25% of participants felt excited by listening to the sound. Sound 4 was evaluated by 147 participants, of which 62% felt threatened or afraid. Only 2% had positive emotions (light-hearted or calm) while listening. A total of 25% of participants felt excited by listening to the sound. Sound 5 was evaluated by 163 participants, of which 68% felt threatened or afraid while listening. Only 3% had positive emotions (light-hearted or calm). A total of 25% of participants felt excited by listening to the sound.

The participants were further asked to evaluate the sound based on descriptor pairs in a 7-level Likert scale. In Figure 4, the averaged agreements for the bipolar descriptor-pairs are displayed. The x-axis contains the averaged level of agreement, while the y-axis contains the verbal descriptor pairs.



**Figure 4.** Average agreement of participants on the descriptors for associated emotions of the presented sounds.

In general, the average level of agreement was stronger for the negative sounds compared with the positive sounds. The results for the threatening descriptor very much fit the statements of feeling threatened (Figure 3). Sounds perceived as threatening were also assessed as unpleasant and repulsive, and vice versa. Furthermore, all the threatening sounds were perceived as more exhausting and obtrusive, whereas positive sounds were perceived as more relaxing and unobtrusive. Threatening perceived sounds were rated more exciting, whereas harmless sounds were perceived more calming. In regard to the descriptor pair boring/stimulating, all of the sounds were rated rather similarly on average, independent of a perceived fearful or positive sound. A significance test with a Friedman rank sum test for unreplicated block data showed significances ( $p < 0.0001$ ) for all descriptors when calculated over all five sounds. To further evaluate the individual significances between two different sound pairs, a Friedman post hoc test was conducted. It showed statistically significant differences ( $p < 0.0001$ ) between all possible pairs between either Sound 1 or 2 and Sounds 3, 4, or 5 in all descriptors except for boring–stimulating.

In summary, Sound 1 invoked positive emotions (light-hearted/calm) in 75% of the participants, and only 5% felt negative emotions (threatened/afraid). Although the results from the Likert scale for Sound 1 are less pronounced compared with Sounds 5, 4, and 3, they still resemble the invoked emotions (Figure 3).

The emotions invoked by Sound 2 were very much divided over all of the possible answers, with 30% of the participants feeling negative emotions, 30% feeling positive emotions, and 25% feeling neutral while listening. In the results of the Likert scale, however, we see a slight shift towards a negative evaluation similar to Sounds 3, 4, and 5.

Regarding the positive and negative valence of the invoked emotions, Sounds 3, 4, and 5 invoked negative emotions (threatened/afraid) in approximately 65% of the participants, and only rarely positive emotions, with only between 4 and 2% feeling light-hearted or calm while listening to either of these sounds. These evaluations resemble the evaluation of the sounds in the Likert scale (Figure 4).

Regarding the arousal (increasing the urge to act on emotions) invoked by the sounds, the percentages of participants feeling excited while listening to the sounds were lower for

Sound 1 (5%) and Sound 2 (8%) compared with Sound 3 (33%), Sound 4 (25%), and Sound 5 (25%).

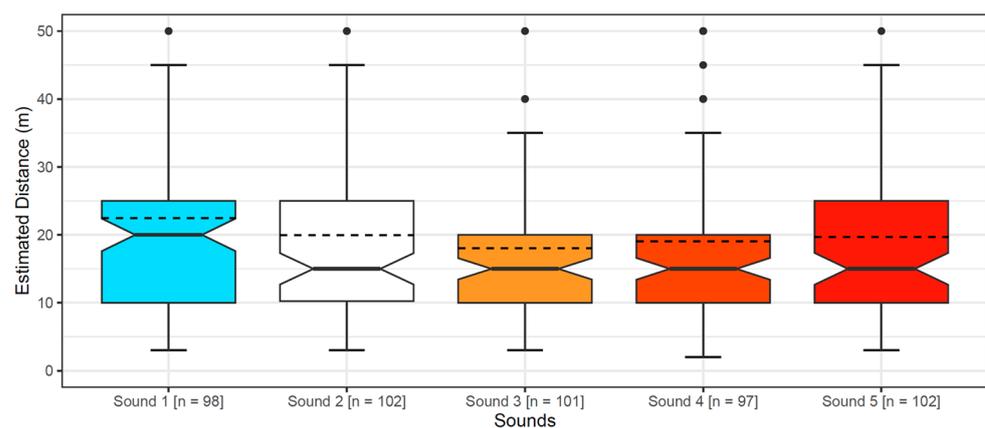
With these results, we can say that the sounds created with our own sound generator (Sounds 5, 4, and 1) can indeed invoke the targeted emotions and are positively associated with the desired emotions. For the question regarding the increase in the effectiveness of the sounds due to the optimization parameters found in the previous study [30], we saw an increase in the invocation of feeling threatened with Sound 5 with its higher levels of impulsiveness and fluctuations strength compared with Sound 3 and 4. But we saw a slight decrease in feeling afraid. For Sound 1, we saw a substantial improvement in regard to invoking a positive feeling, in contrast to the most positive sound of the previous study (Sound 2), even though they were similar in regard to their psychoacoustic values. However, Sound 1 has a lot less harmonic content and utilizes a clean harmonic major chord.

In the next section, we investigate if there is an effect of Sounds 5, 4, 3, and 2 compared with Sound 1 in regard to the estimation of safety distance, desired change, and evaluation of safety distance. The focus is on comparing the positive Sound 1 to the negative Sounds 3, 4, and 5, because Sound 2 had a rather neutral emotional effect.

### 3.2. Estimated Safety Distance, Desired Change in Safety Distance, and Evaluation of Safety Distance vs. Sounds

For the first part of the experiment, participants were presented with five different video–sound combinations. They had to estimate the shown safety distance and were asked if they evaluated the shown safety distance as appropriate. Participants who deemed the safety distance as too small or too large could state by how much they would want to change the safety distance. The results for these three different questions are displayed in aggregated form over all the videos. Only participants who stated that they did not use a method to determine the safety distance, but estimated them intuitively, are included in the analysis.

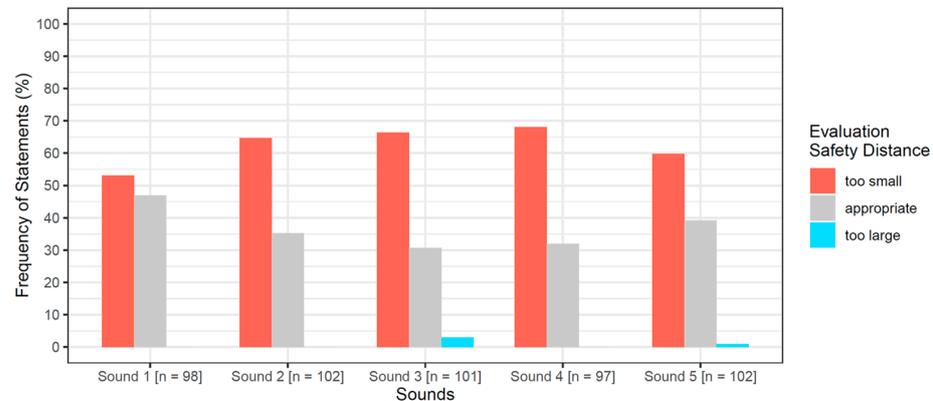
Figure 5 shows the boxplots for the estimated safety distance by the participants for each of the sounds aggregated over all of the videos. The thick line represents the median and the dotted line the mean. For Sound 1, the median is 20 m. The means are Sound 1: 22.47 m (SD = 20.68), Sound 2: 19.96 m (SD = 14.11), Sound 3: 18.8 m (SD = 20.68), Sound 4: 19.05 m (SD = 13.70), and Sound 5: 19.7 m (SD = 16.64).



**Figure 5.** Box plots for the estimation of the safety distances from the shown videos: aggregated and displayed vs. the 5 sounds. The thick line represents the median and the dotted line the mean. The lower and upper borders of the box represent the 1st and 3rd quartiles, respectively. The whiskers represent 0.25–1.5 IQR/0.75 + IQR (interquartile range), and outliers are drawn interdependently.

Based on the mean and median, it can be assumed that the positive Sound 1 seems to lead to higher values of the estimation of the safety distance compared with the rest of the sounds.

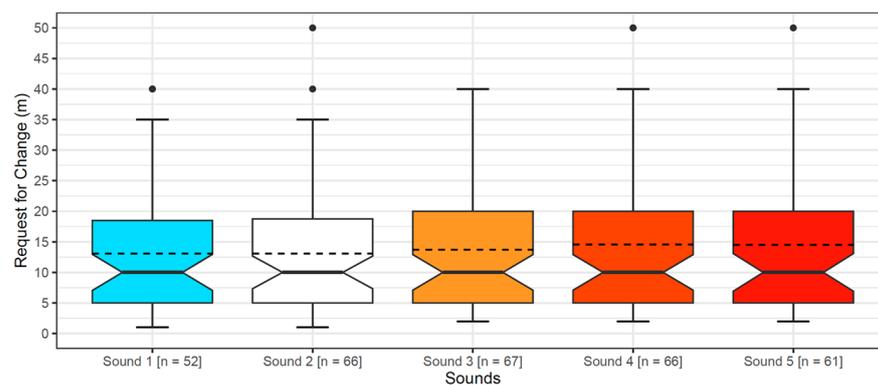
In Figure 6, the percentage values of participants' frequency of response when rating the safety distance as appropriate, too small, or too large are shown for each of the five videos presented to the participants aggregated by the five sounds.



**Figure 6.** Percentages of participants' frequency of response for the evaluation of the shown distances in the 5 presented videos, aggregated and displayed by the 5 sounds.

It can be seen that almost no participants evaluated the shown safety distances as too large, and further, at least 30% of participants still evaluated the shown safety distances as appropriate, which suggests the used videos are within a good range of being open to interpretation. It can also be seen that the difference between the statements appropriate and too small is the smallest for positive Sound 1. Interestingly, the next smallest is the optimized Sound 5.

Figure 7 shows the boxplots for the desired change in meters of the safety distance by the participants for each of the sounds aggregated over all of the videos. The means are Sound 1: 13.10 m (SD = 10.62), Sound 2: 13.06 m (SD = 9.91), Sound 3: 14.70 m (SD = 12.24), Sound 4: 14.55 m (SD = 12.36), and Sound 5: 15.92 m (SD = 13.41).



**Figure 7.** Box plots for the desired change in safety distance from the shown videos in meters: aggregated and displayed vs. the 5 sounds. The thick line represents the median and the dotted line the mean.

Based on the mean, it can be assumed that the positive Sound 1 and neutral Sound 2 seem to lead to lower values in meters of the desired change in the safety distance compared with Sounds 5, 4, and 3.

### 3.3. Statistical Analysis of the Effect of the Sounds on the Estimated Safety Distance and Requested Change in Safety Distance

To test whether the different sound–video combinations had an effect on participants' estimation of the safety distance displayed in the videos (Hypothesis 1), as well as the

desired increase in safety distance (Hypothesis 2), a statistical analysis for repeated measures was conducted. The method chosen is the one-tailed pairwise *t*-test for repeated measures. It tests for a decrease in the estimation of the safety distance between the sounds. It was implemented in R (Version 4.3.1) using the R-function “t.test” of R’s native “stats” package. The method was chosen because it allows for pairwise (Sound A vs. Sound B) comparison of the increase or decrease in means in the differences in at least ordinal dependent variables (estimation of safety distance (m) or desired change in safety distance (m)) based on a categorical independent variable (Sounds 1–5) for repeated measures (one participant evaluated multiple sound–video combinations). Participants who stated that they used a method to determine the safety distance, and did not estimate them intuitively, as well as participants who did not answer all of the questions relevant to the analysis, were excluded from the analysis. Potential outliers were checked and deleted if an erroneous input seemed plausible. However, some people who reported very high values for the estimation and desired change in safety distance were kept in the sample, since these values seemed consistent over the participants’ respective estimations/desires. This left 89 participants for the analysis. Estimations and desires for change in safety distance were not normally distributed, as assessed by Shapiro–Wilk’s test ( $p < 0.05$ ), though at sample sizes  $n > 30$ , the *t*-test also can be used for non-normally distributed data [45]. Furthermore, there was homogeneity of variances, as assessed by Levene’s test for equality of variances. The analyzed pairs were selected as the comparison between the most positive sound (Sound 1) vs. all of the other sounds (Sounds 2–5). The effect size was calculated in the form of Cohen’s *d* using the “cohens.d” function of the “misty” package in version 0.6.2. For the desired change in safety distance, a statement in meters was only possible when participants evaluated the shown safety distance as inappropriate or too large. For statements of appropriate safety distance, a missing desired change in meters was replaced with 0.

The results of the one-tailed paired *t*-test testing for a decrease in the means in differences in estimated safety distance in meters between Sounds 2–5 compared with Sound 1 are shown in Table 3. It contains the mean differences, Cohen’s *d* values, and *t*-values, as well as the significances of the estimated safety distances for all of the compared pairs of sounds.

**Table 3.** Results of one-tailed pairwise *t*-test testing for a decrease in estimated safety distance for Sounds 2–5 compared with Sound 1.: mean differences, effect size (Cohen’s *d*), *t*-values, and significances. Significance levels: \* =  $p \leq 0.05$ .

One-Tailed <i>t</i> -Test	Estimated Safety Distance				
	Sound Pair	Mean Differences (m)	Cohen’s <i>d</i>	<i>t</i> (88)	<i>p</i> -Value
	Sound 1–Sound 2	−1.71	−0.09	1.09	0.14
	Sound 1–Sound 3	−3.79	−0.21	2.13	0.02 *
	Sound 1–Sound 4	−3.34	−0.19	2.08	0.02 *
	Sound 1–Sound 5	−2.78	−0.15	1.97	0.03 *

From Table 3, it can be derived that there is a significant ( $p < 0.05$ ), small effect size ( $|\text{Cohen’s } d| > 0.1, < 0.3$ ) and negative mean of differences for the estimation of the shown safety distance between Sound 1 compared with Sound 3, 4 and 5. For Sound 1 compared with Sound 2, the effect and the mean in differences are negative, statistically not significant, and negligible ( $|\text{Cohen’s } d|, < 0.1$ ). All means in differences in safety distance for Sound 1 compared with either Sounds 2–5 are negative, showing a systematic decrease in estimation under more negative sounds. Hypothesis 1 states that sounds that invoke/are associated with negative/positive emotions lead to a decrease/increase in the estimated safety distance depicted in videos. Based on the results, we can reject the null hypothesis (negative/positive emotions lead to no decrease/increase in the estimated safety distance)

with a certainty of 95%, with the  $t$ -values for the positive–negative comparison being greater than the critical  $t$ -value ( $t(88) = [1.99, 2.08, 2.13] > t_{critical}(88) = 1.662$ ), and thus accept the stated alternative hypothesis (H1).

The results of the one-tailed paired  $t$ -test testing for the increase in means in differences in desired change in the safety distance in meters between Sounds 2–5 compared with Sound 1 are shown in Table 4. It contains the mean differences, Cohen’s  $d$  values, and  $t$ -values, as well as the significances of the desired change in the safety distance for all of the compared pairs of sounds.

**Table 4.** Results of one-tailed pairwise  $t$ -test testing for an increase in desired change in the safety distance for Sounds 2–5 compared with Sound 1.: mean differences, effect size (Cohen’s  $d$ ),  $t$ -values, and significances. Significance levels: \* =  $p \leq 0.05$ .

One-Tailed $t$ -Test Sound Pair	Desired Safety Distance			
	Mean Differences (m)	Cohen’s $d$	$t(88)$	$p$ -Value
Sound 1–Sound 2	0.22	0.04	−0.27	0.39
Sound 1–Sound 3	1.58	0.25	−1.78	0.04 *
Sound 1–Sound 4	1.66	0.26	−2.00	0.02 *
Sound 1–Sound 5	0.63	0.01	−0.74	0.23

From Table 4, it can be derived that there is a significant ( $p < 0.05$ ) and small effect size ( $| \text{Cohen’s } d | > 0.1, < 0.3$ ) in the mean of differences in the estimation of the shown safety distance between Sound 1 compared with Sounds 3 and 4. For Sound 1 compared with Sound 2, the effect and difference in mean are positive, statistically not significant, and negligible ( $| \text{Cohen’s } d |, < 0.1$ ). All means in differences in desire to change safety distance for Sound 1 compared with either Sounds 2–5 are positive, showing a systematic increase in the desired increase in the safety distance under more negative sounds.

Hypothesis 2 states that sounds invoking positive or negative emotions have an influence on the desired increase in the safety distance depicted in videos. Based on the results, we cannot fully reject the null hypothesis that positive/negative sounds have no influence on the desired increase in safety distance in videos. Even though there was a positive effect on the desired change in safety distance, the  $t$ -values for the positive–negative comparison are not all greater than the critical  $t$ -value ( $t(88) = [0.74, 2.00, 1.78] > t_{critical}(88) = 1.662$ ) for the negative sounds.

### 3.4. Statistical Analysis of the Effect of the Sounds on the Evaluation of Safety Distance

To test whether the different sound–video combinations had an effect on the evaluation of the safety distance displayed in the videos, a statistical analysis of the data based on the mixed-effect logistic regression was conducted. This method estimates how well the fixed effects (terms of the regression model) predict a change in the dependent variable (evaluation of safety distance as appropriate or too small) within participants. This method was chosen because it allows for a comparison between the binomial dependent variable—evaluation of shown safety distance in regards to being evaluated appropriate or too small—and the different sounds. Participants who stated that they used a method to determine the safety distance were excluded from the analysis. Furthermore, the four participants’ statements that the shown safety distances were too large were excluded to fulfill the binomial condition required by the method. This leaves a sample size of  $n = 99$  for the calculation of the binomial logistic regression. The participant number was modeled as a random effect to incorporate the repeated measurements design of the study. We begin by looking at the predictors of the evaluation of the safety distance. We tried all possible combinations of predictors and then compared each model using the goodness of fit test and model selection criteria. We created four candidate models while performing the goodness of fit test. Table 5 shows results from fitting the logistic regression models, in which we

compared results across the four different candidate regression specifications. Since the logistic model is susceptible to high correlations between the independent variables, a Spearman correlation was calculated between the actual safety distance and the velocity in the videos. There was only a low correlation ( $r = 0.32, p < 0.05$ ) between the two. Since the sounds and videos are matched randomly, there is no significant correlation between the sounds and either the actual distance ( $r = -0.01, p = 0.77$ ) or speed ( $r = 0.04, p = 0.32$ ). To lower the potential influence of the correlation between the distance and speed toward the model, the terms will only be used additively in the formula. The models were implemented in R (Version 4.3.1) using the “glmer” function in the binomial form of the “lme4” package in version 1.1–35.1. Model 1 represents estimates from only the sounds, while Model 2 estimates the model with sounds and actual shown distance. Model 3 represents estimates from the sounds and the velocity, while Model 4 estimates the model with all predictors. We fitted the logistic mixed model (estimated using ML and BOBYQA optimizer) to predict the statements of evaluation of the distance with the different Sounds 1–5, the actual distance, and the actual speed (formula: Evaluation of the distance~Sound Number (All Models) + Actual Distance (Model 2 and 4) + Actual Speed (Model 3 and 4). The models included the participant numbers as random effect (formula: ~1 | Participant Number).

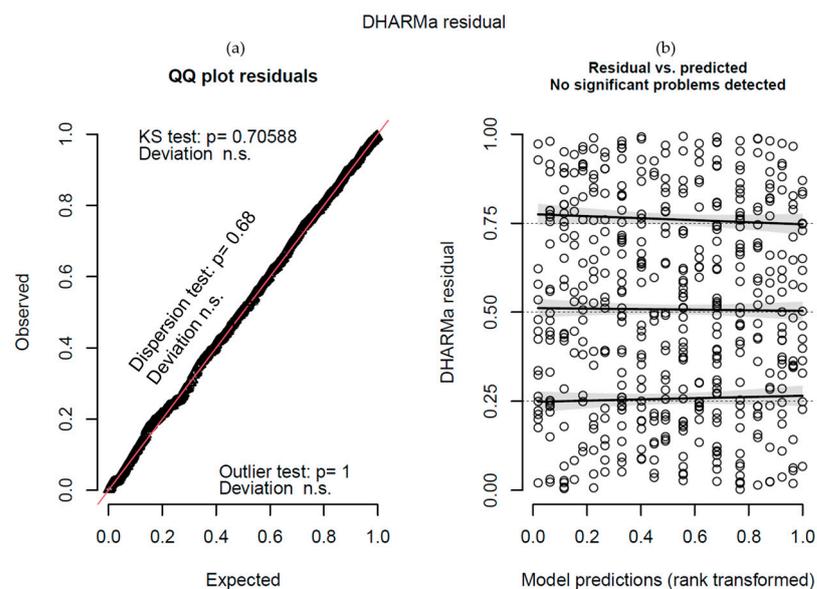
**Table 5.** Predictive models for evaluation of safety distance.

	Model 1	Model 2	Model 3	Model 4
Intercept				
(Sound 1				
Distance 12 m	1.18 (0.28)	2.82 ** (0.35)	1.25 (0.41)	1.21 (0.46)
Vel. 50 km/h)				
Sound 2	1.95 * (0.34)	2.28 * (0.38)	1.95 (0.40)	2.35 (0.45)
Sound 3	2.41 ** (0.35)	2.68 ** (0.38)	2.01 (0.40)	2.66 * (0.45)
Sound 4	2.44 ** (0.89)	2.89 ** (0.39)	2.30 (0.41)	2.90 * (0.46)
Sound 5	1.47 (0.34)	1.46 (0.37)	1.09 (0.40)	1.02 (0.46)
Actual Distance 15 m		0.23 *** (0.27)		0.05 *** (0.57)
Velocity 60 km/h			0.16 *** (0.41)	2.58 (0.15)
Velocity 80 km/h			2.58 ** (0.32)	26.70 *** (0.59)
AIC	606	573	540	479
BIC	632	603	574	535
Log Likelihood	−297	−279	−262	−239
Deviance	594	559	524	479
R <sup>2</sup> (total)	0.41	0.52	0.61	0.74
R <sup>2</sup> (fixed effects)	0.02	0.10	0.15	0.26
Participants	99	99	99	99
Observations	496	496	496	496

Note: Figures represent the odds ratios. Figures in parentheses indicate standard errors. \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ .

We proceed to obtain the best possible model specified by the best subset selection method. For this purpose, we used AIC, BIC, Deviance, log-likelihood,  $R^2$  (total), and  $R^2$  (fixed-effects) criteria. Since the residual deviance, as well as the AIC and BIC values, decrease with the successive inclusion of predictor variables, while the explanatory power of the model  $R^2$  (total) and explanatory power by the predictors  $R^2$  increase, we can assume that Model 4 with the inclusion of all predictor variables is a better fit than the other models. Finally, further tests were conducted to identify potential issues of the model using the “DHARMA” package. Figure 8 shows the aggregated results of the analysis. The fitness of the model was further evaluated, simulating new data from the model and comparing them with the observed data. The QQ plot (a) of the residuals shows a rather straight line, indicating that both sets of quantiles (observed vs. predicted) come from the same distribution. This is further supported by the not-significant Kolmogorov–Smirnov (KS) test. It was tested for over- or underdispersion, meaning the residual variance is larger/smaller than expected under the fitted model. The dispersion test was not significant, indicating no

over- or underdispersion between observed and simulated data. No outliers are simulated from the fitted model compared with the observed data. Plot (b) displays the residuals against the predicted value calculated via a quantile regression, which compares the empirical 0.25, 0.5, and 0.75 quantiles in the y direction (solid lines) with the theoretical 0.25, 0.5, and 0.75 quantiles (dashed black line). The test for deviation between observed and predicted quartiles is not significant ( $p = 0.86$ ). The analysis indicated no issues with the fitted Model 4.



**Figure 8.** Model fitness analysis of comparison between observed data and predicted data by fitted Model 4: (a) shows the QQ plot of residuals with results for KS test for deviation, dispersion test, and outlier test. The Plot displays the observed quantiles of the data (depicted as triangles) compared to quantiles that we would expect to see if the data were normally distributed (red line) (b) shows the residuals against the predicted value calculated via a quantile regression, which compares the empirical 0.25, 0.5, and 0.75 quantiles in y direction (solid lines) with the theoretical 0.25, 0.5, and 0.75 quantiles (dashed black line).

Model 4's total explanatory power is substantial ( $R^2$  (total) = 0.74), and the part related to the fixed effects alone ( $R^2$  (fixed effects)) is 0.26. The model's intercept, corresponding to sound number = Sound 1, actual distance = 12 m, and actual speed = 50 km/h, is at 0.19 (95% CI [−0.72, 1.11],  $p = 0.681$ ), with an odds ratio of 1.21 (Standard Error = 0.46). Within this model, the following was found:

- The effect of Sound 2 is statistically nonsignificant and positive (beta = 0.86, 95% CI [−0.03, 1.74],  $p = 0.058$ ), with an odds ratio of 2.35 (SE = 0.45);
- The effect of Sound 3 is statistically significant and positive (beta = 0.98, 95% CI [0.09, 1.87],  $p = 0.030$ ), with an odds ratio of 2.66 (SE = 0.45);
- The effect of Sound 4 is statistically significant and positive (beta = 1.06, 95% CI [0.16, 1.97],  $p = 0.021$ ), with an odds ratio of 2.90 (SE = 0.46);
- The effect of Sound 5 is statistically nonsignificant and positive (beta = 0.02, 95% CI [−0.87, 0.91],  $p = 0.968$ ), with an odds ratio of 1.02 (SE = 0.46).

Standardized parameters were obtained by fitting the model on a standardized version of the dataset. Using a Wald z-distribution approximation, 95% Confidence Intervals (CIs) and  $p$ -values were computed.

Based on the odds ratios of the best-fitting model to predict a change in the dependent variable (evaluation of safety distance as appropriate or too small) based on the different sounds, it can be seen that only Sounds 3 and 4 have a significant odds ratio (2.66 and 2.90,  $p < 0.05$ ). Sound 5 has a nonsignificant odds ratio of 1.02. Even though the sounds

invoked fairly similar emotions in the participants, they seem to differ in the influence on the evaluation of the appropriateness of the safety distances shown in the videos.

Sounds that invoke positive or negative emotions or are associated with such emotions have an influence on the evaluation of the appropriateness of depicted safety distances in videos displaying tailgating. Hypothesis 3 states that sounds that invoke/are associated with negative/positive emotions lead to increased odds of a depicted safety distance being evaluated as too small/appropriate. Based on the results, we cannot fully reject the null hypothesis (negative/positive emotions have no influence on a depicted safety distance being evaluated as too small/appropriate). Even though the odds were 2.66- or 2.9 higher than the shown safety distance being evaluated as too close under Sounds 3 and 4 compared with Sound 1, the probability was basically 50/50 when the videos were combined with Sound 5. Although based on the raw percentages for the statements of depicted safety distances (Figure 6), there was a significant difference in statements for Sound 5 (60% stating too small) compared with Sound 1 (52% stating too small).

#### 4. Discussion

In this paper, we investigated whether the sounds created with our self-implemented sound generator can invoke similar emotions as the sounds we created in a previous study [30], and further, if the optimized sounds we created increase the effectiveness of sounds in regards to the invocation of emotions. Further, it was investigated whether sounds invoking positive or negative emotions or are associated with such emotions have an influence on the estimation of the safety distance, on the desire to increase the safety distance, or influence the evaluation of the appropriateness of depicted safety distances in videos displaying tailgating. Regarding the effectiveness of the sounds created in the sound generator (Sounds 1, 4, and 5) compared with the sounds from the previous study (Sounds 2 and 3) [30], it can be said that for the sounds targeting feeling afraid/threatened, we reached very similar results for the recreated sound (Sound 4) compared with the original one (Sound 3). For the optimized sound (Sound 5), we were indeed able to increase the invocation of feeling threatened. However, this also leads to decreasing the invocation of feeling afraid, which was not the intent. For the optimized sound targeting positive emotions, we drastically increased the invocation of positive feelings (light-hearted/calm) compared with the most positive sound of the previous study (Sound 2). The newly created sounds were designed utilizing the findings about relations between emotionalizing positive and negative effects and harmonic structures [15] and underlying chords [16], as well as aspects like sharpness, roughness, and fluctuation strength describing the timbre or impulsiveness characterizing rhythmic qualities [7,8] and thus supporting the findings in these studies conversely. For Sound 2 from the previous study, it is interesting to note that it invoked very diverging emotions (threatened/afraid and light-hearted/calm) in equal numbers of participants (30% of participants stated feeling positive and negative emotions, respectively) in this study compared with the previous study (64% of participants stated feeling positive emotions and 21% participants feeling negative emotions). This potentially shows how different individual sounds are evaluated depending on the overall pool of sounds that are evaluated. The previous study definitely had an emphasis on threatening sounds based on the evaluation by the participants.

The investigation of the effect of the different emotionalizing sounds on the perceived safety distances in the videos generated mixed results. Hypothesis 1, stating that sounds have an effect on the estimation of the depicted safety distances, was accepted based on the statistical results. There was a significant difference in estimated safety distances between the positive sound (higher estimates) and the negative sounds (lower estimates).

Hypothesis 2, stating that sounds affect the desired increase in the depicted safety distance, was partially rejected based on the statistical results. There was a significant difference in the desired change in safety distance between the positive Sound 1 and two of the negative sounds (Sounds 3 and 4), but not for negative Sound 5. Looking at the difference between the desired change for the positive sound and all of the negative sounds,

the desired change regarding the negative sounds was higher compared with the positive sound, suggesting a stronger desire for change when stimulated by the threatening sounds.

Hypothesis 3, stating that the sounds have an effect on the evaluation of the appropriateness of the depicted safety distance (too small/appropriate), was partially rejected based on the statistical results. There were significant positive odds ratios in the predictive logistic models for two of the negative sounds (Sounds 3 and 4), suggesting two to three times higher odds of the safety distance being evaluated as too small compared with positive Sound 1. Sound 5, on the other hand, showed a nonsignificant odds ratio close to 1, suggesting no influence on the evaluation.

Even though Sound 5 was rated very similarly to Sound 3 and 4 in regards to emotions, it had a much lower effect on the desire to change and the evaluation of the appropriateness of the shown safety distances. One explanation could be that because of its high level of impulsiveness and fluctuation strength at around 2 Hz (which sounds potentially similar to more aggressive techno or similar electronic music), there could have been an effect similar to the influence of music tempo on driving in a simulation [13]: a higher tempo in music was observed to increase speeds and higher occurrence of traffic violations.

In general, individuals have different preferences regarding the sufficiency of safety distances. The shown safety distances ideally should be close to the individual tipping point for each and every participant to increase the possibility of an emotion-based shift in evaluating whether a safety distance is appropriate or not. This would of course imply a tremendous effort in regard to the experimental setup but can be regarded as a limitation to this experiment nonetheless.

## 5. Conclusions and Outlook

The results show that we can create sound stimuli that invoke targeted emotions, such as feeling threatened or light-hearted, with our sound generator. They further show that the direction of the effects of the different emotionalizing sounds on the participants' statements concerning their perception of safety distances aligned with the stimulus direction, even though there was no real danger for the participants and they only saw the videos on a small screen. In the case of the estimation of the safety distance, the effects were statistically significant between the positive and the threatening sounds. In the case of the effect of the sounds on the evaluation of the safety distance, as well as the desired change in safety distance, the effects were statistically significant for two out of the three pairs between the positive and the threatening sounds. With these findings, and the technical means to generate the used sounds in real time based on a variety of vehicle parameters, the next step will be a simulator study. This will offer the possibility to investigate the effects of the emotionalizing sounds on driving behavior in realistic driving scenarios under much more immersive conditions.

**Supplementary Materials:** The videos 1–5 as well as the sounds 1–5 can be downloaded at <https://publikationen.bibliothek.kit.edu/serie/750> (accessed on 29 April 2024).

**Author Contributions:** Conceptualization, M.P., D.Y. and A.A.; methodology, M.P., D.Y. and A.A.; software, M.P.; validation, M.P. and A.A.; formal analysis, M.P.; investigation, M.P. and D.Y.; resources, M.P. and A.A.; data M.P.; writing—original draft preparation, M.P. and D.Y.; writing—review and editing, A.A.; visualization, M.P.; supervision, M.P. and A.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Your Request to the Ethics Committee of KIT of 25 May 2023—Research Project: “Studying the Effects of a Suggestive Sound Design on Estimating Safe Spacing Between Driving Vehicles”: We have no ethical concerns about the permissibility of the research project. Document provided.

**Data Availability Statement:** The codebook and dataset are available online at <https://publikationen.bibliothek.kit.edu/1000168515> (accessed on 29 April 2024).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Mortsiefer, H. Der neue Elektro-Sound. Tagesspiegel Background. 2019. Available online: <https://background.tagesspiegel.de/mobilitaet/der-neue-elektro-sound> (accessed on 15 February 2024).
2. Czolbe, F.; Pfeleiderer, M. *Klangwelten Gestalten. Zur Aktualität des Bauhauses in Sound Design und Auditiver Stadtplanung*, 1st ed.; Mensch & Buch: Berlin, Germany, 2021; pp. 104–112, ISBN 978-3-96729-089-9.
3. Streicher, B.; Lange, S.; Eisele, G.; Steffens, C. Effiziente Methoden im Active Sound Design. *ATZ Automob. Z* **2021**, *123*, 58–61. [[CrossRef](#)]
4. Fiebig, A.; Schulte-Fortkamp, B. Acceptance of synthetic driving sounds in the interior of electric vehicles. In *3. Internationale ATZ-Fachtagung. Automotive Acoustics Conference 2015*; Proceedings; Siebenpfeiffer, W., Ed.; Springer: Wiesbaden, Germany, 2015; pp. 28–44, ISBN 978-3-658-27648-5.
5. Pilgerstorfer, M.; Runda, K.; Conter, M.; Gatscha, M. drivEkustik—Fahrverhalten in und akustische Wahrnehmung von Elektrofahrzeugen. In *Endbericht—Forschungsarbeiten des Österreichischen Verkehrssicherheitsfonds*; Elsevier: Amsterdam, The Netherlands, 2013.
6. Roach, N.W.; Heron, J.; McGraw, P.V. Resolving multisensory conflict: A strategy for balancing the costs and benefits of audio-visual integration. *Proc. Biol. Sci.* **2006**, *273*, 2159–2168. [[CrossRef](#)] [[PubMed](#)]
7. Koelsch, S.; Skouras, S.; Fritz, T.; Herrera, P.; Bonhage, C.; Küssner, M.B.; Jacobs, A.M. The roles of superficial amygdala and auditory cortex in music-evoked fear and joy. *Neuroimage* **2013**, *81*, 49–60. [[CrossRef](#)]
8. Alt, N.W.; Jochum, S. Sound Design under the Aspects of Musical Harmonic Theory. In Proceedings of the SAE 2003 Noise & Vibration Conference and Exhibition, Grand Traverse, MI, USA, 5–8 May 2003; SAE Technical Paper Series. SAE International: Warrendale, PA, USA, 2003.
9. Horswill, M.S.; Plooy, A.M. Auditory feedback influences perceived driving speeds. *Perception* **2008**, *37*, 1037–1043. [[CrossRef](#)] [[PubMed](#)]
10. Hellier, E.; Naweed, A.; Walker, G.; Husband, P.; Edworthy, J. The influence of auditory feedback on speed choice, violations and comfort in a driving simulation game. *Transp. Res. Part F Traffic Psychol. Behav.* **2011**, *14*, 591–599. [[CrossRef](#)]
11. Wang, E.Y.; Wang, E.M. In-car sound analysis and driving speed estimation using sounds with different frequencies as cues. *Int. J. Ind. Ergon.* **2012**, *42*, 34–40. [[CrossRef](#)]
12. Dekker, C. The Impact of Movie Soundtracks on the Estimation of Driving Speeds. Graduation Thesis, Leiden University, Leiden, The Netherlands, 2015. Available online: <https://coendekker.nl/media/pdf/Coen-Dekker-The-impact-of-movie-soundtracks-on-the-estimation-of-driving-speeds-Version-3-1.pdf> (accessed on 4 April 2024).
13. Brodsky, W. The effects of music tempo on simulated driving performance and vehicular control. *Transp. Res. Part F Traffic Psychol. Behav.* **2001**, *4*, 219–241. [[CrossRef](#)]
14. Russell, J.A.; Mehrabian, A. Approach-Avoidance and Affiliation as Functions of the Emotion-Eliciting Quality of an Environment. *Environ. Behav.* **1978**, *10*, 355–387. [[CrossRef](#)]
15. Costa, M.; Ricci Bitti, P.E.; Bonfiglioli, L. Psychological Connotations of Harmonic Musical Intervals. *Psychol. Music* **2000**, *28*, 4–22. [[CrossRef](#)]
16. Pallesen, K.J.; Brattico, E.; Bailey, C.; Korvenoja, A.; Koivisto, J.; Gjedde, A.; Carlson, S. Emotion processing of major, minor, and dissonant chords: A functional magnetic resonance imaging study. *Ann. N. Y. Acad. Sci.* **2005**, *1060*, 450–453. [[CrossRef](#)]
17. Coutinho, E.; Cangelosi, A. Musical emotions: Predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. *Emotion* **2011**, *11*, 921–937. [[CrossRef](#)] [[PubMed](#)]
18. Coutinho, E.; Dikken, N. Psychoacoustic cues to emotion in speech prosody and music. *Cogn. Emot.* **2013**, *27*, 658–684. [[CrossRef](#)] [[PubMed](#)]
19. Västfjäll, D. Emotion induction through music: A review of the musical mood induction procedure. *Music. Sci.* **2001**, *5*, 173–211. [[CrossRef](#)]
20. Chang, K.-J.; Park, D.C. Technology of an Emotional Engine Sound Designing for Active Sound Control Using Order Balance and Musical Instrument Sound. In Proceedings of the 9th International Styrian Noise, Vibration & Harshness Congress: The European Automotive Noise Conference, Graz, Austria, 22–24 June 2016; SAE Technical Paper Series. SAE International: Warrendale, PA, USA, 2016.
21. Gwak, D.Y.; Yoon, K.; Seong, Y.; Lee, S. Subjective evaluation of additive sound designed to reinforce acoustic feedback of electric vehicle. In Proceedings of the INTER-NOISE and NOISE-CON Congress and Conference Proceedings, Melbourne, Australia, 16–19 November 2014; pp. 2157–2161.
22. Roidl, E.; Frehse, B.; Höger, R. Emotional states of drivers and the impact on speed, acceleration and traffic violations—A simulator study. *Accid. Anal. Prev.* **2014**, *70*, 282–292. [[CrossRef](#)] [[PubMed](#)]
23. Serrano, J.; Di Stasi, L.L.; Megías, A.; Catena, A. Affective-sound effects on driving behaviour. *Transport* **2013**, *29*, 100–106. [[CrossRef](#)]
24. Megías, A.; Di Stasi, L.L.; Maldonado, A.; Catena, A.; Cándido, A. Emotion-laden stimuli influence our reactions to traffic lights. *Transp. Res. Part F Traffic Psychol. Behav.* **2014**, *22*, 96–103. [[CrossRef](#)]

25. Kühn, M.; Hannawald, L. Verkehrssicherheit und Potenziale von Fahrerassistenzsystemen. In *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für Aktive Sicherheit und Komfort*; Winner, H., Hakuli, S., Lotz, F., Singer, C., Eds.; Springer: Wiesbaden, Germany, 2015; pp. 55–70, ISBN 978-3-658-05733-6.
26. Jentsch, M.; Lindner, P.; Spanner-Ulmer, B.; Wanielik, G.; Krems, J. Nutzerakzeptanz von Systemausprägungen zur Aktiven Gefahrenbremsung: Mensch, Technik, Organisation-Vernetzung im Produktentstehungs-und-herstellungsprozess. In *Tagungsband GfA Frühjahrskongress Chemnitz*; GfA-Press: Dortmund, Germany, 2011.
27. Kroher, T. *Fahrerassistenzsysteme in der Übersicht: So können sie Autofahrer Unterstützen*; ADAC: Allgemeiner Deutscher Automobil-Club: Munich, Germany, 2022.
28. Birbaumer, N.; Schmidt, R.F. *Biologische Psychologie*, 6th ed.; Birbaumer, N., Schmidt, R.F., Eds.; Springer: Wiesbaden, Germany, 2006; p. 508, ISBN 10 3-540-25460-9.
29. Petersen, M.; Etri, M.; Behrendt, M.; Albers, A.; Spekker, M.; Lefringhausen, T.J. Suggestive Sound Design—How to use Active Interior Sound Design to improve traffic safety. *INTER-NOISE NOISE-CON Congr. Conf. Proc.* **2021**, *263*, 1561–1573. [[CrossRef](#)]
30. Petersen, M.; Zaimovic, M.; Albers, A. Evaluating emotionalizing effects of active sound designs. *INTER-NOISE NOISE-CON Congr. Conf. Proc.* **2023**, *268*, 4689–4700. [[CrossRef](#)]
31. EBU. *Recommendation R 128: Loudness Normalisation and Permitted Maximum Level of Audio Signals*; European Broadcasting Union: Geneva, Switzerland, 2011; p. 3. Available online: <https://tech.ebu.ch/docs/r/r128.pdf> (accessed on 4 April 2024).
32. Horswill, M.S.; Hill, A.; Silapurem, L. The development and validation of video-based measures of drivers' following distance and gap acceptance behaviours. *Accid. Anal. Prev.* **2020**, *146*, 105626. [[CrossRef](#)]
33. Breuer, J.; von Hugo, C.; Mücke, S.; Tattersall, S. Nutzerorientierte Bewertungsverfahren von Fahrerassistenzsystemen. In *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für Aktive Sicherheit und Komfort*; Winner, H., Hakuli, S., Lotz, F., Singer, C., Eds.; Springer: Wiesbaden, Germany, 2015; pp. 183–196, ISBN 978-3-658-05733-6.
34. Babić, D.; Babić, D.; Sucha, M.; Stanić, V.; Toman, M. The influence of music genres on the driving behaviour of young drivers and their visual scanning of the environment. *Transp. Res. Part F Traffic Psychol. Behav.* **2021**, *81*, 396–407. [[CrossRef](#)]
35. Groeger, J.A.; Chapman, P. Judgement of Traffic Scenes: The Role of Danger and Difficulty. *Appl. Cogn. Psychol.* **1996**, *10*, 349–364. [[CrossRef](#)]
36. Anund, A.; Lahti, E.; Fors, C.; Genell, A. The effect of low-frequency road noise on driver sleepiness and performance. *PLoS ONE* **2015**, *10*, e0123835. [[CrossRef](#)] [[PubMed](#)]
37. Funke, F.; Reips, U.-D.; Thomas, R.K. Sliders for the Smart: Type of Rating Scale on the Web Interacts with Educational Level. *Soc. Sci. Comput. Rev.* **2011**, *29*, 221–231. [[CrossRef](#)]
38. von Bismarck, G. Timbre of Steady Sounds: A Factorial Investigation of its Verbal Attributes. *Acta Acust. United Acust.* **1974**, *30*, 146–159.
39. Solomon, L.N. Search for Physical Correlates to Psychological Dimensions of Sounds. *J. Acoust. Soc. Am.* **1959**, *31*, 492–497. [[CrossRef](#)]
40. International Organization for Standardization. *International Standard, Methods for Calculating Loudness—Part 1: Zwicker Method*; ISO: Geneva, Switzerland, 2017.
41. ECMA International. *ECMA-418-2: Psychoacoustic Metrics for ITT Equipment—Part 2 (Models Based on Human Perception)*; ECMA: Geneva, Switzerland, 2022.
42. Hempel, T. *Untersuchungen zur Korrelation Auditorischer und Instrumenteller Messergebnisse für die Bewertung von Fahrzeuginnenraumgeräuschen als Grundlage eines Beitrags zur Klassifikation von Hörereignissen*; Herbert Utz Verlag: München, Germany, 2001; ISBN 3831600694.
43. Cohen, J. A power primer. *Psychol. Bull.* **1992**, *112*, 155–159. [[CrossRef](#)]
44. Hartig, F. DHARMA: Residual Diagnostics for Hierarchical (Multi-Level/Mixed) Regression Models: R Package Version 0.4.6. Available online: <https://cran.r-project.org/web/packages/DHARMA/vignettes/DHARMA.html> (accessed on 15 February 2024).
45. Kwak, S.G.; Kim, J.H. Central limit theorem: The cornerstone of modern statistics. *Korean J. Anesthesiol.* **2017**, *70*, 144–156. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.