EVALUATION OF SOUNDS FOR HYBRID AND ELECTRIC VEHICLES OPERATING AT LOW SPEED

John K. Pollard[1]; Catherine Guthy[1]; Aaron Hastings[1]; Mary D. Stearns[1]; Lisandra Garay-Vega[2]

U.S. Department of Transportation

[1]John A. Volpe National Transportation Systems Center

[2]National Highway Traffic Safety Administration

Electric vehicles (EV) and hybrid electric vehicles (HEVs), operated at low speeds may reduce auditory cues used by pedestrians to assess the state of nearby traffic creating a safety issue. This field study compares the auditory detectability of numerous synthetic sounds for hybrid and electric vehicles operating at a low speed. The sample includes pedestrians who are sighted and legally blind, independent travelers, with self-reported normal hearing. The test site has the acoustic characteristic of an urban area with a typical ambient noise level of approximately 58-61 dB (A). Dependent variables include proportion of detection and detection distance. Synthetic sounds tested, that resemble those of an internal combustion engine (ICE) vehicle, produce similar detection distances as the actual ICE vehicle tested for a 6 mph constant speed operation. In some instances, synthetic sounds designed according to psychoacoustic principles were detected much sooner than the reference ICE vehicle tested. Synthetic sounds that contain only the fundamental component of combustion noise, but lack the harmonics and other high-frequency characteristics of an actual ICE were relatively ineffective, with their detection distances being only about half of those of an ICE vehicle.

INTRODUCTION

The National Highway Traffic Safety Administration (NHTSA) conducted an analysis on a total of 24,297 HEVs and 1,001,000 ICE Honda and Toyota vehicles. A total of 186 HEVs and 5,699 ICE vehicles were involved in pedestrian crashes. The incidence and the odds of HEVs being involved in a pedestrian crash are 1.35 times (35% higher) as much as the corresponding incidence and odds of an ICE vehicle being involved in a similar crash (Wu, Austin, and Chen, 2011).

The vehicles involved in such crashes are likely to be moving at low speeds where the difference between the sounds emitted by HEVs and some ICEs can be substantial. A reduction in the sound emitted by vehicles operating at low speeds and in electric mode may have implications for all pedestrians; pedestrians who are blind may be particularly affected because they depend almost entirely on auditory cues to navigate. Groups representing people who are blind have expressed concern about the lack of sound emitted by HEVs (Maurer, 2008).

In April 2010, NHTSA issued a report presenting results of the agency's research. The report documented the overall sound levels and spectral content for a selection of HEVs and ICE vehicles in different operating conditions. The auditory detectability of HEVs and ICE vehicles was evaluated for two background noise levels. In addition, countermeasure concepts (e.g., vehicle-based, infrastructure-based, and systems requiring vehicle-pedestrian communications) were reviewed.

The results showed that the overall sound levels for the HEVs tested are noticeably lower at low speeds than for the ICE vehicles tested. Overall, study participants were able to detect any vehicle sooner in the low ambient condition. ICE vehicles tested were detected sooner than their HEV pairs in two of the three operating conditions tested. Detection time varied by vehicle operating condition, ambient sound level, and, vehicle type (i.e., ICE versus HEV in EV mode) (Garay-Vega, Hastings, Pollard, Zuschlag, and Stearns, 2010).

PURPOSE

Human-subject testing was conducted to compare alternative sounds, in terms of the time and distance a subject detected these sound sources and recognized them as a vehicle in representative urban-ambient noise conditions. Sounds were evaluated at two sound pressure levels typical of ICE vehicles at low speeds (i.e., 59.5 and 63.5 dB (A)). An ICE vehicle that produced 60 dB (A) in the 6 mph pass-by test was used as a reference in this evaluation.

METHODS

Stimuli

The sounds evaluated included sounds emitted by HEVs with integrated sound systems; sounds emitted by prototype sound systems played back by loudspeakers temporarily mounted on HEVs; and sound emitted by an ICE vehicle, included in the study as a baseline for comparison purposes.

Five vendors, motor vehicle manufacturers or suppliers of automotive electronics, provided prototypes of synthetic sound generators for EVs or HEVs. These five systems were labeled "A" to "E". A total of nine sounds were evaluated: A1, A2, A5, B, C, D, E1, E3, and E4. The ICE vehicle was labeled 'R'.

The outputs of sound systems "A", "B", and "E" were recorded during 6 mph forward driving to present sounds to participants at the same sound pressure levels. The sound pressure levels of the recorded '.wav files' were normalized to 63.5 dB (A) (the higher level used in the human-subject study). Additional versions of each recording were normalized to 59.5 dB (A) (the lower level used in the human-subject study) and to 70 dB (A) (the level used for playback level calibration). Prior to each experimental session, the volume controls of the amplifiers that reproduced these sounds were calibrated so that the 70 dB (A) version produced 70 dB (A) as measured by the SAE J-2889 draft test procedure (SAE J-2889, draft, 2009). The speaker systems for these sounds were mounted on the hood, just above the front grill of the vehicle. Two of the systems ("C" and "D") were fully integrated into vehicles and had no accessible electrical connection from which a direct recording could be made. For system "C", the provided volume control was manually adjusted to the target values of 59.5 and 63.5 dB (A) used in the experiments. For system "D", the maximum level was used, which could only achieve the lower level desired for the human-subject testing, 59.5 dB (A).

Study Design

Human-subject testing collected data outdoors during three independent sessions conducted on three days in July and August 2010. This paper report data for the two sessions which focused on one vehicle operation: 6 mph forward pass-by.

Each participant was presented with the same experimental conditions (within-subject design). Trials were presented in four blocks. Each block contained 24 trials. There were four trials without a vehicle or sound per block. Participants experienced four repetitions of each experimental condition. Half of the 6 mph forward pass-by trials were presented from the left side and half from the right side. Varying the direction in which the stimuli (i.e., vehicle approaching at constant speed) is presented reduces subjects predisposed to focus on one direction (e.g., left or right) more than the other thus providing a more realistic listening situation. The presentation order was randomized within a block.

The independent variables include:

- Vehicle sounds
- Level of the sound stimuli (Low = 59.5 dB(A); High = 63.5 dB(A))
- Masking effect due to vehicle direction of travel (from left and from right)

- Study session (to examine any effect due to the difference in the overall conditions at the test site from session to session).

The dependent variables include:

- Raw Detection Distance: the number of feet the vehicle was from the participant when the participant indicated she or he heard the sound. A failure to detect the sound before the vehicle passed was treated as missing data.
- Proportion of Detection: the proportion of trials of a given condition in which the participant detected the sound anytime before the vehicle passed the participant.
- Time-to-vehicle-arrival: time, in seconds, from detection of a target vehicle sound to the instant the vehicle passes the pedestrian location.
- Detection Distance: calculated distance, feet, to the target vehicle at the moment each subject responded.

Participants

The sample included 79 sighted and blind adults. The subjects met the eligibility criteria listed below:

- 18 years or older
- self -reported as an independent traveler
- travel regularly (i.e., cross streets at least 10 times per week, on a regular week)
- blind participants: legally blind regardless of whether they are totally blind, blind with light perception or blind with partial vision
- sighted participants must be able to guide a blind participant to their seats if needed;
- self-report to have normal hearing on both ears without hearing aids
- willing to be screened for hearing threshold shift
- willing to wear blindfolds
- have normal manual dexterity in both hands for prompt button pressing

Table 1 Distribution of Participants by Sex and Vision Category

Vision Category	Number of Males	Number of Females	Total
Blind	25	20	45
Sighted	17	17	34
Total	42	37	79

Test Site

The study took place in a parking lot located on the USDOT/Volpe Center campus in Cambridge, Massachusetts. The test site has the acoustic characteristic of an urban area with a typical ambient noise level of approximately 58-61 dB (A). Most of the ambient noise at the site was generated by fans on the roofs of buildings about 400 feet west of the test site, which resulted in a greater masking effect for vehicles approaching from the west (the "a-to-b" direction) than for the reverse.

The test site has several important characteristics essential for the conduct of the study. Vehicles that were not part of the study did not pass through the test site; all vehicular traffic was controlled by Volpe Center security guards. There was adequate approach distance for vehicles to get up to 6 mph speed before they became audible. There was an alternate route available for test vehicles to go to the designated starting points while recharging their batteries. The participants were seated next to each other in a line and shaded from the sun. The road surface consisted of bituminous asphalt and was swept and washed down prior to each test session.

Procedure

All participants were briefed by a Volpe Center investigator and completed an informed consent form. Each session began with a practice session during which subjects heard samples of the trials included in the evaluation. The practice session allowed participants to experience the range of sounds before the testing began, familiarize themselves with the task and use of the response button, and ask questions to the investigator. The investigator described the instructions and traveling situation before each experimental block. The experimenter described the task (in Sessions 2 and 3) as follows:

"Please imagine you are standing on the curb waiting to cross a street where there may be vehicles approaching from both your right and left. You will hear distant vehicles in the background in all trials. If and when you detect and recognize a vehicle that would affect your decision about when to start crossing the street, press and release your push button firmly."

The investigator provided general feedback during the practice trials about whether or not there was a target sound present that they may or may not associate with a vehicle. Feedback was not offered during the experimental trials.

DATA COLLECTION

The Volpe Center staff constructed a data acquisition system to capture the independent variables and record the occurrence of each subject response to each event. The input devices consisted of a pair of photo-electric sensors, positioned at measured locations at each end of the line of participants to determine the precise moment that each test vehicle passed the fixed point with the photo detector at each end. Each subject had a push button device which they used to indicate when they detected a nearby vehicle.

The input devices were connected to a digital input card in a desktop computer running LabVIEW software. If a subject responded by pressing his/her button, the data acquisition system recorded the exact moment (plus/minus one millisecond) in relation to the time the trial started and the times that the vehicle passed through the beams of each photo-detector.

The LabVIEW software was programmed to generate a spreadsheet for each block of trials containing the following data items for each subject response:

- Trial number
- Direction of vehicle passage
- Time/date
- Vehicle speed between the photo detectors
- Number of milliseconds after the trial started until each subject responded
- Calculated distance to the target vehicle (feet) at the moment each subject responded
- Calculated time (milliseconds) between the moment each subject responded and the vehicle arrived at the subject's position

The calculated values are based on the assumption that the vehicle's speed is the same at all points as its average speed while traveling between detectors.

DATA ANALYSIS

The near real-world conditions of these experiments expose the participants to many more possible unintended stimuli than a normal laboratory experiment. Aircraft, distant noisy road vehicles, lawn mowers, and leaf blowers introduced sounds that could have been confused with approaching ICE vehicles on numerous occasions. Other trials were disrupted by a malfunction in the steam-pressure- relief system on an adjacent building. Furthermore, because the presentation of stimuli was not controlled directly by computer, but rather depended on human actuation in response to a radioed instruction, there was a possibility for human error. To monitor all of these possibilities to determine which trials were invalid, the study used four systems:

- High-definition video recordings of subject behavior and vehicle passage including a "shotgun" microphone recording of sounds from the area in the immediate vicinity of the participants.
- Recordings from a binaural head placed just behind the participants to approximate what the participants heard.
- Continuous recording of the noise level near the participants with the same acoustic measurement system used to characterize the vehicles.
- A-weighted, fast-response, sound pressure level values at 100 msec intervals during each trial.

Sessions 2 and 3 were planned to each generate 96 trials in total (4 blocks of 24 trials each). During the course of each block, it became obvious that some trials were tainted by extraneous noises, such as, aircraft or malfunctioning steam valves. These trials were redone at the end of each block when time allowed. During data analysis, a few additional trials were discovered to have been invalidated by more subtle extraneous noises (e.g., intermittent, distant traffic on nearby street). The final data set contained 79 valid trials in Session 2 and 95 valid trials in Session 3. No response and detection after the vehicle passed the pedestrian location are considered missed detection.

RESULTS

Detection Distance and Proportion of Detection

Trials with no detection or detection after a vehicle passed the participant are considered missed detections instead of missing data. This is done by assigning 0 feet to all cases where a sound was not detected. This procedure allows analyses of detection distance to include more participants for greater statistical power. In contrast, if no-detection trials were treated as missing data, the listwise deletion of data required for repeated measures analysis would delete any participant who does not have valid detection distances for all two or four trials of a given sound-amplitude direction condition.

The analyses indicated relationships between raw detection distances and proportion of detection. Across all conditions, participants who detected relatively more sounds also tended to detect the sounds at greater distances than participant who detected relatively fewer sounds ($r = 0.326$, $n = 53$, $p = 0.0168$). Furthermore, high amplitude sounds were detected more often and at greater distances than low amplitude sounds (for raw detection distance, Ms = 55 and 44 feet, t (52) =11.45, p < 0.0001; for proportion of detections, Ms = 0.95 and 0.92, t (52) = 3.51, p =0.0009).

These positive relations between raw detection distance and proportion of detection suggest they represent the same underlying construct. Thus, they are combined into a single "detectability" variable by assigning 0 feet to all cases where a sound was not detected. Repeated trials for a given sound, amplitude and direction were averaged to create a single value per condition per participant for subsequent analyses.

Table 2 shows the mean detection distances for the sounds evaluated in the human-subject studies; sounds at the top of the list can be described as alternative sounds designed according to psychoacoustic principles and sounds at the end of the list can be described as ICE-like sounds with only the fundamental combustion noise or otherwise lacking in the qualities that support detectability.

Table 2 Mean Detection Distance (ft) for all Sounds at two Amplitudes and for the Reference ICE Vehicle

Sound Number	Average Detection Distance (feet) for amplitude equal 59.5 dB(A)	Average Detection Distance (feet) for amplitude equal 63.5 dB(A)
E4	72	85
A2	57	77
E3	52	70
A5	50	47
ICE vehicle, 60 dB(A)	41	NA
A1	35	44
C	32	41
E1	30	32
B	20	25
D	19	NA

Detection Relative to Reference Vehicle

To compare the sounds to the detectability of the reference sound from an ICE powered vehicle ('R'), a mixed design, analysis of variance (ANOVA) was performed on detectability with session and vision as between-subjects independent variables, and sound and direction as within-subject independent variables. Session refers to the date of the study; vision refers to whether the participant is blind or sighted; sound refers to the test sound or ICE vehicle sound; and direction refers to vehicle direction of travel ('a' to 'b' = left to right; 'b' to 'a' = right to left).

The analysis included low-amplitude sounds since their amplitude, 59.5 dB (A), correspond closely to the sound pressure level measured for the reference ICE vehicle, 60.0 dB (A). Using only low-amplitude sounds also allowed the inclusion of sound 'D', which produced only a low amplitude sound.

The analysis indicated significant main effects of sound (F (4.3, 185.8) = 75.5, $p < 0.001$, Greenhouse-Geisser corrected for sphericity), but, more importantly, a significant three-way interaction of session, sound, and direction (F (5.6, 241.3) = 4.3, $p = 0.001$, Greenhouse-Geisser corrected for sphericity). This implies that the relative performance of each sound, including the reference sound, is jointly contingent on the direction it comes from and the session it was presented in. The directional effect results primarily from the fact that the roof-top fans on buildings to the west were the predominant source of ambient noise, which can mask vehicles approaching from the west (the "a-to-b" direction), compared with the reverse.

Thus, each direction-by-session condition may be regarded as constituting an independent test of the performance of each sound relative to the reference sound. The frequency of Type I errors should thus follow a binomial distribution. Based on this, the detectability of each sound relative to the reference was evaluated by *t*-tests comparing each sound to the reference vehicle for the corresponding session and direction condition of each. With nine sounds to compare to the reference, a Bonferroni post-hoc adjustment for family-wise error rate yielded a critical *p*-value of 0.0056 for deciding if a particular artificial sound is overall significantly different than the reference. Treating the four session-by-direction test conditions as independent samples of ambient sound, this critical p-value is achieved if at least two conditions show a significant difference at the 0.0311 level, according to binomial calculations. If this is observed then one can conclude that the sound has significantly different detectability than the reference sound. Results show that A2, A5, E3, and E4 have significantly better detectability than the reference sound in at least two of the four conditions. These sounds never have significantly worse detectability in any condition.

Thus, these sounds overall have better detectability than the reference sound. In contrast, sounds A1, B, C, D, and E1 all have significantly worse detectability than the reference sound in at least two of the four conditions. These sounds never have significantly better detectability in any condition. Thus, these sounds overall have worse detectability than the reference sound.

Ranking of Alert Sounds

To compare the detectability of the sounds to each other, a mixed design ANOVA was performed on detectability with session and vision as between-subjects independent variables, and sound, direction, and amplitude as within-subject independent variables. The reference sound 'R' and sound 'D' were excluded from this analysis since they did not differ in amplitude.

Sound 'D' was significantly worse than the reference sound, so it is of minor interest for subsequent analyses. The analysis indicated significant main effects of sound (F (3.6, 132.0) = 78.3, $p < 0.001$, Greenhouse-Geisser corrected for sphericity). There also was, more importantly, a significant four-way interaction of session, sound, direction, and amplitude (F (5.2, 192.8) = 4.5, $p = 0.001$, Greenhouse-Geisser corrected for sphericity). This interaction where 'Lo' refers to the low amplitude tested (59.5 dB (A)) and 'Hi' refers to the high amplitude tested (63.5 dB (A).This implies that the relative performance of each artificial sound is jointly contingent on the direction it comes from, the session it was presented in, and the amplitude that was used. With this in mind, sounds are ranked by comparing each to the other with t-tests for each session by-direction-by-amplitude condition. To assist in the control for family-wise error rate, the analyses only included the four sounds shown in the previous section to be superior to the reference sound. With four sounds to compare to each other, a Bonferroni post-hoc adjustment for family-wise error rate yielded a critical p-value of 0.00833 for deciding if a particular sound is overall significantly different than another sound. Treating the eight session-by-direction-by amplitude test conditions as independent samples, this critical p-value is achieved if at least two conditions show a significant difference at the 0.0178 level, according to binomial calculations. If this is observed then one can conclude that the sound has significantly different detectability than the reference sound. Results show that E4 has overall significantly better detectability than the other sounds, and within each condition it is never worse than any other sound, except for one condition when compared to A2. Sounds A2 and E3 are overall not significantly different than each other, showing only a difference in a single condition. Sound A5 has overall significantly worse detectability than the other sounds, and within each condition is it never better, except for one condition when compared to E3. The overall ranking of the sounds from most to least detectable is therefore: E4, A2 and E3, and A5.

None of the analyses found a significant effect of vision ability. Blind participants, on average, were no better or worse

than sighted participants in detecting the approach sounds (smallest p = 0.636).

SUMMARY

- The relative performance of each sound, including the reference ICE sound, was jointly contingent on the direction it came from and the session it was presented in.
- Synthetic sounds tested, that resemble those of an internal combustion engine (ICE) vehicle, produce similar detection distances as the actual ICE vehicle tested for a 6 mph constant speed operation.
- In some instances, synthetic sounds designed according to psychoacoustic principles were detected sooner than the reference ICE vehicle tested.
- Synthetic sounds that contain only the fundamental combustion noise are relatively ineffective.

ACKNOWLEDGMENT

REFERENCES

Garay-Vega, L., A. Hastings, J. K. Pollard, M. Zuschlag, and M. Stearns. (2010) Quieter Cars and the Safety of Blind Pedestrians: Phase 1. (Report DOT HS 811 304). Washington, DC: National Highway Traffic Safety Administration. http://www.nhtsa.gov/staticfiles/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2010/811304.pdf

Hanna, R. (2009) Incidence of Pedestrians and Bicyclist Crashes by Hybrid Electric Passenger Vehicles. National Center for Statistical Analysis (Technical Report DOT HS 811 204). Washington, DC: National Highway Traffic Safety Administration. http://www.nrd.nhtsa.dot.gov/Pubs/811204.PDF.

Hastings, A., Pollard, J. K., Garay-Vega, L., Stearns, M. D., & Guthy, C. (2011). Quieter Cars and the Safety of Blind Pedestrians, Phase 2: Development of Potential Specifications for Vehicle Countermeasure Sounds. (Report No. DOT HS 811 496). Washington, DC: National Highway Traffic Safety Administration. http://www.nhtsa.gov/DOT/NHTSA/NVS/Crash%20Avoidance/Technical%20Publications/2011/811496.pdf.

Maurer, M. The Danger Posed by Silent Vehicles. Presented at Work Forum for Harmonization of Vehicle Regulations (United Nations Working Party 29), Geneva, Switzerland, Feb. 20, 2008. National Federation of the Blind, Baltimore, Md., 2008.

SAE. Measurement of Minimum Noise Emitted by Road Vehicles. SAE J2889-1 Draft 2009. Society of Automotive Engineers.

Wu, J., Austin, R., & Chen, C-L. (2011). Incidence Rates of Pedestrian and Bicyclist Crashes by Hybrid Electric Passenger Vehicles: An Update. (DOT HS 811 526). Washington, DC: National Highway Traffic Safety Administration.