U.S. Department
of Transportation
**National Highway
Traffic Safety
Administration**

NHTSA
www.nhtsa.gov

DOT HS 812 207 **November 2015**

Crash Warning Interface Metrics: Warning and Message Perception Under Ambient Noise Conditions Laboratory Experiments

Disclaimer

Suggested APA Format Citation:

Singer, J., Lerner, N., Kellman, D., & Robinson, E. (2015, November). *Crash warning interface metrics: Warning and message perception under ambient noise conditions laboratory experiments*. (Report No. DOT HS 812 207). Washington, DC: National Highway Traffic Safety Administration.

Technical Report Documentation

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT HS 812 207		

4. Title and Subtitle
Crash Warning Interface Metrics: Warning and Message Perception Under Ambient Noise Conditions Laboratory Experiments

5. Report Date
November 2015

6. Performing Organization Code

7. Author(s)
Jeremiah Singer, Neil Lerner, Daniel Kellman, Emanuel Robinson

8. Performing Organization Report

9. Performing Organization Name and Address
Westat, Inc.
1600 Research Boulevard
Rockville, Maryland 20850-3129

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
DTNH22-11-D-00237

12. Sponsoring Agency Name and Address
United States Department of Transportation
National Highway Traffic Safety Administration
1200 New Jersey Avenue SE.
Washington, DC 20590

13. Type of Report and Period Covered
Final Report/September, 2011 - August, 2014

14. Sponsoring Agency Code

15. Supplementary Notes
Eric Traube (COTR/TO)

16. Abstract
The Crash Warning Interface Metrics program addresses issues of the driver-vehicle interface for advanced crash warning systems. This report summarizes the methods and findings of three laboratory experiments that investigated acoustic signal detectability and perception under varied ambient noise conditions. These experiments replicated and expanded the findings of an on-road experiment that found that auditory signal detectability, perceived urgency, and perceived meaning can be substantially impaired under conditions of elevated ambient noise. Experiment 1 replicated the on-road experiment in a laboratory setting, where in-vehicle ambient noise and acoustic signals were presented to participants via headphones. This experiment closely replicated the findings of the on-road experiment. Experiment 2 included five new signals and seven ambient noise conditions, five of which had not been investigated in the previous on-road or lab experiments. This experiment found that while some ambient noise conditions had relatively little effect on signal perception (e.g., worn, bumpy asphalt or driving adjacent to a semi-trailer), others had a substantial effect (e.g., driver window open or driving in heavy rain). This experiment also found an interactive effect between signals and ambient noise conditions, meaning that different signals are affected in different ways by different ambient noise conditions. Experiment 3 directly investigated the effects of three different signal loudness levels, as well as adding annoyance and acceptability as dependent variables. As expected, increases in signal loudness led to higher ratings for noticeability, urgency, and annoyance. This effect was more pronounced in the louder ambient noise conditions (driver window half down and heavy rain) than in the relatively quiet baseline condition, which suggests that higher loudness levels are particularly beneficial for maintaining the noticeability and perceived urgency of warnings in loud ambient noise conditions. This experiment found that noticeability, perceived urgency, and annoyance were highly intercorrelated.

17. Key Words
human factors, driver vehicle interface (DVI), advanced crash warning system (ACWS), connected vehicle (CV), distraction, forward collision warning (FCW), lane departure warning (LDW), ambient noise

18. Distribution Statement
Document is available to the public from the National Technical Information Service www.ntis.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	65	N/A

Form DOT F 1700.7 Reproduction of completed page authorized.

Acknowledgements

The authors acknowledge the contributions of the Westat project team: Elisha Lubar, Shawn McCloskey, and Luis Romero (experimenters); Diane Snow (recruitment and scheduling); Michael Gill and Rick Huey (software development and programming). Dr. Carryl Baldwin (George Mason University) contributed experimental stimuli for evaluation. Dr. John Molino provided valuable insights regarding acoustic recording, playback, and psychoacoustic methods. As the NHTSA COTR, Eric Traube provided important guidance throughout the course of this work.

Table of Contents

List of Figures

List of Tables

1 Introduction

This report describes the methods and findings of a sequence of experiments on the effects of various vehicle interior ambient noise conditions on driver perception of warnings and messages. This task is part of a larger National Highway Traffic Safety Administration (NHTSA) project titled Crash Warning Interface Metrics (CWIM). The CWIM project deals broadly with the effectiveness of the driver interface for in-vehicle crash warnings. As part of this project, the work reported here addresses how acoustic interface effectiveness may be affected by various noise conditions that may realistically occur under different driving conditions. The laboratory studies presented here replicate and extend the findings of a previous on-road experiment, presented in an earlier report (Singer, Lerner, Walrath, & Gill, in press).

Important Note: Because this project focused specifically on the DVI rather than the ACWS as a whole, DVI effectiveness, as defined in this report, does not directly refer to crash reduction or mitigation. Rather, effectiveness refers to the ability of the DVI to influence attention, comprehension, or behaviors in a way that is appropriate for the hazard context.

To be considered effective, a DVI for crash avoidance situations must have a number of attributes. It must be detected reliably and rapidly by the driver and must be properly interpreted. They must convey the proper degree of urgency so that driver response is quick and appropriate. They should be distinguishable from less urgent alerts and messages, so that distraction, annoyance, and false alarm mistrust effects are minimized. Previous research on acoustic alerts and warnings for in-vehicle applications typically have been conducted under relatively benign in-vehicle ambient noise conditions. However, even with improved sound insolation in modern vehicles, the ambient noise within a vehicle may vary considerably under actual driving conditions and warnings must remain effective under the likely range of noise conditions that may be anticipated. Field measurements made in passenger vehicles as part of the CWIM project found that the ambient noise levels were in the mid-60s decibel, A-weighted (dBA) range at 60 mph on a smooth asphalt highway, but increased by about 10 dB for various conditions, such as normal music playback or windows down, and by even more during moderately heavy rain.

The Singer et al. (in preparation) experiment demonstrated important effects of ambient noise on acoustic signal perception. Two higher-level noise conditions – music playing and front windows down – were compared with a baseline condition while driving at 60 mph on a smooth asphalt highway. Various brief acoustic signals occurred during the course of these drives, at loudness levels equivalent to 65 or 75 dBA pink noise. Dependent measures included: response time to detect the sound; subjective rating of signal noticeability; subjective rating of signal urgency; and perceived category of signal meaning. Ambient noise conditions significantly affected all of these measures. Many of the 65 dBA sounds were frequently undetected under higher noise conditions, although the 75 dBA sounds remained well detected. However, even when the signal was heard, its interpretation was altered by ambient noise. Sounds that were perceived as highly urgent under the baseline condition lost urgency under noise and were less likely to be interpreted as an imminent crash warning. Not all sounds of equivalent loudness were equally affected by the ambient noise condition. The Singer et al. study thus identified an important issue in vehicle warning effectiveness and a need to understand the characteristics of acoustic signals that reduce their susceptibility to interference by in-vehicle noise. This study may be viewed as preliminary in that it defined the problem and indicated the need for better understanding of acoustic signal features and noise conditions.

Although the on-road method of Singer et al. (in preparation) was sensitive to ambient noise and acoustic signal properties, and offered good face validity, the method has some drawbacks. First, compared to laboratory methods, on-road driving experiments are not very efficient. Substantial time is spent training the participant, providing vehicle familiarity, driving to the data collection site, and sometimes suspending procedures when conditions are not appropriate (e.g., loud adjacent vehicles, change in road surface quality). Furthermore, data collection is subject to interruption by weather conditions, roadwork, or congestion. Another limitation is in the degree of experimental control. Conditions will not be identical from moment to moment or session to session. Vehicle speeds will vary, surrounding traffic will differ, transient events (e.g., bump in road) may occur, atmospheric conditions (e.g., wind, wet road) may vary, and so forth. Finally, certain ambient noise conditions cannot be reliably controlled (at least without considerable effort) on the road, such as rain or noise from adjacent vehicles. Therefore it would be desirable to be able to continue this line of research with a more efficient and controlled laboratory method.

There are significant challenges in collecting valid data on noise effects in a laboratory setting. The measurement, recording, calibration, and playback of in-vehicle noise are complex. Ambient noise is dynamic with a substantial low frequency component but with important higher frequency aspects, especially when there is an acoustic alert superimposed. The vehicle interior itself is a complex acoustic space that varies from vehicle to vehicle, with reflective and sound-absorbing surfaces that can significantly influence the signal at the driver's head position. There may be an important spatial component to elements of the noise (e.g., a passing large truck in the lane to the left of the driver) that must be captured. Measurement issues arise in specifying the loudness of acoustic signals that have diverse temporal characteristics, such as pulses, onset/offset ramps, and so forth. Developing and validating a practical laboratory method for realistically capturing the in-cab acoustic experience is therefore a critical step.

The set of experiments presented here adapted the on-road method of Singer et al. (in preparation) to a laboratory setting and extended the findings further. Experiment 1 implemented a laboratory method, using headphones, which paralleled the on-road procedures and used the same set of acoustic signals, sound levels (65 and 75 dBA), and ambient noise conditions. It provided a validation of the laboratory method against the on-road perceptual experience. Experiment 2 encompassed a broader range of ambient noise conditions, including some that could not reasonably be done using on-road methods (e.g., rain). This helped identify the ambient noise conditions that might be of particular concern and might best be used in subsequent work. In Experiment 2, all of the acoustic signals were presented at the same loudness (equivalent to 70 dBA pink noise). Experiment 3 used a smaller set of ambient noise conditions and acoustic signals and directly examined the effect of signal loudness (65, 70, or 75 dBA). It also included subjective response measures related to annoyance and consumer acceptability.

2 Experiment 1: Replication of on-road experiment

Experiment 1 closely replicated the design and methods of the on-road experiment, using the same set of acoustic signals and the same ambient noise conditions. The on-road procedure was adapted to a laboratory setting, with a simplified driving-like baseline task and all sounds (ambient noise and signals) were presented via headphones. The intent was to demonstrate essentially similar findings with the two methods, thus validating the more efficient and controlled laboratory method for use in subsequent work.

2.1 Method

2.1.1 Study design

The design of this experiment was nearly identical to the previous on-road study (Singer et al., in preparation). The key differences were:

- Whereas the on-road experiment included a between-subjects vehicle type factor, all recordings for the lab study were made in one vehicle (Toyota Camry). The on-road experiment found no main effect of vehicle type and no interactions of vehicle type (small car, sedan, SUV) with ambient noise conditions.
- Rather than actual on-road driving, participants performed a basic driving-like task involving lateral and longitudinal control over a simplified schematic road display presented on a computer screen.
- All ambient noise conditions and alerts were pre-recorded and played back to participants using headphones.
- In the laboratory experiment, the "music on" condition playback level was set at 75 dBA. In the on-road experiment, participants were allowed to adjust playback to their normal, preferred listening level, which ranged from 66 to 81 dBA across participants.

There were two within-group factors (interior noise condition, acoustic signal). Data were collected under three different interior noise conditions: (1) windows up, music off; (2) front windows down, music off; and (3) windows up, music on. The order in which each noise condition block was presented to participants was counterbalanced.

A set of 15 different acoustic signals was presented in each noise condition. These included three unique voice messages and eight unique non-voice sounds. All 11 of the unique sounds and voices were presented at a sound pressure level (SPL) of approximately 65 dBA as measured near the driver's right ear. One of the voice messages and three of the non-voice sounds were also presented at 75 dBA, with the resultant total of 15 signals. The lower 65 dBA level is representative of a number of acoustic alerts as measured in actual current practice (e.g., Lin & Green, 2013). The higher 75 dBA level is more consistent with human factors guidance (e.g., Campbell, Richard, Brown, & McCallum, 2007), assuming a moderate level of ambient vehicle cab noise.

Several different dependent measures were recorded to evaluate driver response. These included: 1) a measure of reaction time for the participant to detect the occurrence of a signal; 2) a rating of signal noticeability; 3) a rating of signal urgency; 4) a rating of speech intelligibility (for voice messages only); and 5) perceived meaning of the signal (chosen from a set of four alternatives).

2.1.2 Participants

There were 24 participants in this experiment. There were an equal number of males and females in each of three age categories (21-30, 31-40, and 41-50). Participant requirements were the same as for the on-road study. No participants reported having hearing decrements or using hearing assistive devices. All drove regularly, held valid U.S. driver's licenses and reported having no major moving violations on their record in the past three years. Participants were recruited through the Volunteers section of Craigslist and through a news item posted on Westat's Intranet homepage. Westat employees were not eligible, but could refer friends or family. Participants received $75 for completing the session. Prospective participants were screened via telephone. A recruitment ad and the telephone screener are shown in Appendix A and Appendix B, respectively.

2.1.3 Instrumentation and recording

2.1.3.1 Stimuli

Experiment stimuli were the same 15 sounds used in the previous on-road experiment (see Appendix C for more information about each sound). With the exception of the voice messages, these were adapted from actual acoustic signals used in production vehicles, encompassing various warning and alert applications. The set of sounds included:

1. FCW 1: One burst of 20 fast beeps with a relatively high frequency profile.
2. FCW 2: Four bursts of four fast beeps with a relatively low frequency profile.
3. Blind spot warning: Three bursts of four fast beeps, each with a smoothed onset and decay and a sustained low intensity sound between beeps.
4. Pedestrian warning: A constant tone with a duration of 2 seconds.
5. Seat belt alert 1: A single chime that decays to silence in the span of about two seconds, with intensity varying in a wavelike pattern.

6. Seat belt alert 2: Two chimes, each of which decays to silence in the span of about one second
7. Park assist 1: One burst of eight beeps.
8. Park assist 2: Two bursts of three beeps.
9. Female voice – not urgent: Female voice says "Attention."
10. Female voice – urgent: Female voice says "Warning, warning."
11. Male voice – urgent: Male voice says "Warning, warning."
12. FCW 1 (high): Same as FCW 1, but presented at 75 dB
13. Blind spot warning (high): Same as Blind spot warning, but presented at 75 dB
14. Park assist 1(high): Same as Park assist 1, but presented at 75 dB
15. Female voice – urgent (high): Same as Female voice – urgent, but presented at 75 dB

Each sound was recorded in a stationary 2014 Toyota Camry with the engine off. Sounds were played using the same playback equipment and methods that were used in the on-road study (i.e., Dell notebook computer connected to X-Mini II XAM4-B Portable Capsule Speaker located on top of the vehicle dashboard). For the on-road study, sounds were calibrated to a nominal 65 dB level by a jury panel procedure in which participants adjusted the loudness of each stimulus until it matched the perceived loudness of a calibrated 65 dB pink noise signal. The mean loudness level for each signal was selected as its playback volume. The 75 dB sounds were then created by adding 10 dB to each 65 dB signal using software editing.

Stimuli were recorded using a Bruel & Kjaer Type 2270 Sound Level Meter with Type 4101 binaural microphones worn by a researcher seated in the driver's seat. Each stimulus was recorded with all windows closed and again with both front windows down. The recordings with both front windows down were needed to match the ambient noise condition where both front windows were open because the open windows reduce interior acoustic reflections, and therefore alter stimulus characteristics. All recordings were 24 bit, 44,100 Hz WAV files using the sound level meter's low dynamic range setting to maximize recording gain. Recordings were made using the 65 dB-calibrated playback levels established

2.1.3.2 Ambient noise recordings

Recordings were made of the same ambient noise conditions that participants experienced in the on-road study: 1) windows up, music off; 2) windows up, music playing; and 3) both front windows down, music off. Recordings were made in a 2014 Toyota Camry traveling at 60 mph on smooth asphalt using a Bruel & Kjaer Type 2270 Sound Level Meter with Type 4101 binaural microphones worn by a researcher seated in the driver's seat. For the music recording, the song used was Café Amore by Spyro Gyra, which was also used in the on-road study. The on-road study allowed participants to set the music volume to their own preference, which resulted in a mean SPL of about 72 dB. For the recording, a nominal SPL of 75 dB was used. Recordings were made on the same road that was used for the on-road experiment (Maryland Route 200). Sound level meter recording settings were the same as were used when recording the stimuli.

From each recording, researchers selected a relatively uniform segment of 22 to 81 seconds that was free of extraneous noises such as passing vehicles or bumps in the road. Recordings were then set to loop indefinitely when played back to participants to simulate a continuous driving experience.

2.1.3.3 Playback apparatus

Sounds were played from a Dell notebook computer via USB to a Fiio E07K digital-to-analog converter, then out via a Fiio L7 line-out dock to a Schiit Magni amplifier. From the amplifier, sounds were played to participants using Beyerdynamic DT-880 Pro (250 ohm) headphones. All sound processing effects (e.g., auto-gain, equalization) were disabled. The playback equipment allowed sounds to be played back at the same bit depth and sampling frequency that was used for the original recordings. All sounds were played back at the same SPLs at which they were recorded.

2.1.3.4 Alert presentation

Alerts were presented using the same custom software that was used for the on-road experiment. Within each noise condition block, the experimental control software generated a random presentation order for the 15 auditory signals. The software provided a random time gap that ranged from 10 to 50 seconds and averaged 30 seconds from the completion of the previous sound's ratings to the presentation of the next sound. Once the random time had passed, the software indicated to the experimenter that the next signal could be activated. The actual triggering of the trial was done by the experimenter. When triggered, a trial began with a five-second pre-signal period. The signal was then automatically triggered at the end of the five seconds. When the participant detected the signal they pressed a microswitch button, worn on their finger or thumb, to provide a reaction time. The microswitch was attached to a Velcro strap that allowed the participant to locate the switch in a comfortable but easy-to-reach position, in a manner that was unlikely to result in unintentional switch activations. The precise location on the index finger or thumb was determined by the participant.

The data collection system recorded the reaction time and then cued the experimenter to verbally present a series of rating and choice questions. The questions were:

- "How noticeable was that that sound?" (1=not very; 7=extremely)
- "How urgent was that sound?" (1=not very; 7=extremely)

- "How intelligible was that sound"? (this question only asked for voice messages) (1=not very; 7=extremely)
- "Which of the following most closely matches the meaning conveyed to you by this sound?"
    - Urgent crash warning
    - Safety information
    - Information not related to safety
    - Incoming personal communication

The participant provided verbal responses which were manually entered by the experimenter. The definitions of key terms are shown in Table 1.

Table 1. Definition of rating factors and choice options

Term	Definition
Noticeability	The sound is easily noticeable among other sounds and noises in the vehicle
Urgency	The sound conveys a sense of importance, motivating you to make an immediate response
Intelligibility	The spoken words can be easily understood
Perceived Meaning	Choose the one that most closely matches the meaning conveyed to you by this sound
Urgent crash warning	… means that there is a situation in which you must react immediately to avoid a crash. For example, imagine you are about to hit a pedestrian or about to run off the road.
Safety information	… means that there is a safety issue that you need to pay attention to, but you are not in immediate danger of a crash. For example, imagine that you are approaching a work zone where two lanes are closed or there are reports of icy roads ahead.
Information not related to safety	… means exactly what it says – you are receiving information, but the information is not safety-related. This could include various types of information, such as traffic congestion several miles ahead, prices at nearby gas stations, or a navigation system telling you to make the next turn.
Incoming personal communication	… means that you are receiving an incoming call, text message, email, or other direct communication.

2.1.3.5 Secondary task (vehicle control)

In lieu of actual driving, participants engaged in a computer-simulated driving-like task (see Figure 1). The intent was to provide a simplified and uniform secondary task that required basic longitudinal and lateral vehicle control actions. The task required participants to use the steering wheel to keep their vehicle in a lane and the gas pedal to maintain a constant following distance behind a lead vehicle. The simulation included random fluctuations in lateral and longitudinal position to ensure that participants would maintain vigilance and be required to make input adjustments. Driving performance data were not recorded. The display was a 17-inch Dell monitor and the steering wheel and pedals were Logitech Driving Force GT. Steering wheel force feedback was disabled for this experiment.

Figure 1. Driving simulator task

2.1.4 Procedure

Upon arrival, the participant's driver's license was checked to confirm identity and the participant read and signed an informed consent form. They were then brought to the experiment room and seated in the participant chair. The participant sat facing the driving simulation apparatus. The experimenter was positioned to the left of the participant at a desk with two notebook computers. One computer controlled the auditory alerts, background noises, and program for running the session, while the other computer controlled the driving simulation task.

The complete set of instructions given to the participants is attached in Appendix D. The general purpose and procedure were first explained to the participant as an overview. Participants were asked to silence their cell phones to avoid disruptions. This was followed by a period of simulation familiarization, during which the participant used the steering wheel to practice the driving task. Participants were instructed to stay between the lane lines and to keep the blue lead car's rear wheels aligned with horizontal gray hash marks on the screen. The microswitch was then attached to the participant's finger or thumb and adjusted so that they could quickly and easily activate the switch without removing a hand from the steering wheel or altering their typical hand positions while driving. The experimenter confirmed that the switch mounting position was unlikely to result in unintended switch activations.

Next, the participant was introduced to the responses they were to make when they heard an auditory signal. An example sound (distinct from any in the set of test signals) was presented. The experimenter had the participant operate the microswitch to indicate that they heard the sound. The experimenter then walked the participant through the set of ratings and choice questions. The participant was provided with a definition of each of the factors to be rated and for each choice option for the meaning of the signal. The ratings for the three attributes of noticeability, urgency, and intelligibility were all made on a scale of one (not very) to seven (extremely).

Following this example, the participant was presented with a second practice trial. This time the signal was a voice message, distinct from other voice messages in the set of test signals. The participant clicked the microswitch after detecting the message and then made ratings about each attribute. During this trial, the experimenter introduced the intelligibility question, which was not asked for the previous practice question. Following this training, the data collection portion of the session began.

Data collection occurred in three blocks, each block under a different ambient noise condition. The sequence of the three noise conditions was counterbalanced among participants. The first block included only the core set of 15 auditory signals. The second and third blocks each began with two novel auditory signals (one voice, one non-voice). Different novel sounds were used for the second and third blocks. This was done to help preclude the participant from assuming that the same set of signals occurs for each block. The novel signals were then followed by the 15 signals of the primary set in a random order.

The computer program indicated to the experimenter when they were authorized to initiate the next trial. When the participant pressed the microswitch the response time was automatically recorded and the sequence of rating and meaning questions appeared on the experimenter's screen. The experimenter then read each question to the participant, who gave a verbal response. The experimenter then entered the response on the computer. Once the data for all questions were entered, the controlling software began timing the interval for the next trial. The time interval was a random time between 10 and 50 seconds. If the participant did not activate the microswitch within 8 seconds of activation of the auditory signal, the trial was recorded as a failure to detect the sound. The experimenter did not give the participant any feedback if they failed to hear a sound.

The entire session took approximately 65 minutes, with the data collection portion taking approximately 50 minutes.

2.2 Results

Analyses of variance (ANOVAs) for the dependent measures of response time, noticeability, and urgency all showed statistically significant effects of ambient noise, signal, and the ambient noise-by-signal interaction. The category-of-meaning data were analyzed by means of multiple logistic regression. These data also showed significant effects of ambient noise, signal, and the ambient noise-by- signal interaction. Bar charts showing the noticeability and urgency data are shown in Appendix E.

The primary objective of this experiment was to demonstrate that the laboratory method was able to closely replicate the findings of the on-road procedure. Table 2 shows the correlation between the group mean measures for each acoustic signal for the on-road experiment and Lab Experiment 1, for the dependent measures of percent detected, noticeability, urgency, and meaning. Note that the correlation for percent detected in the baseline condition is not applicable because all sounds were detected very close to 100 percent of the time. Figure 2 (percent detected), Figure 3 (noticeability), Figure 4 (urgency), and Figure 5 (meaning) show these data in scatterplot form. For all dependent variables other than percent detected, the correlations and scatterplots for the ratings and meaning data exclude cases where there were fewer than 11 observations in either the on-road or laboratory study due to frequent failures to detect the signal. This only occurred for four sounds, all in the windows-down condition. The scatterplots also show the value of R^2, indicating the proportion of variance in the ratings that is accounted for by the relationship. For purposes of this analysis, the categorical data on meaning was converted to a numeric rating, based on the ordering of driving-related importance for each category. Thus the least consequential category (personal communication) was coded as a 1, non-safety information as a 2, safety information as a 3, and urgent crash warning as a 4.

Table 2. Correlation coefficients between on-road and laboratory findings for response measures under each ambient noise condition

Dependent Measure	Baseline Noise	Music On	Front Windows Down
Percent detected	Not applicable	0.837	0.642
Noticeability	0.950	0.944	0.925
Urgency	0.951	0.937	0.942
Meaning	0.956	0.884	0.866

As Table 2 indicates, the correlations for noticeability, urgency, and meaning were very strong, all exceeding $r=0.90$ for all of the ratings and exceeding $r=0.86$ for the meaning data. All of these correlations were statistically significant at the $p<0.001$ level. The correlation for percent detected under the baseline noise condition is essentially meaningless, because the acoustic signal was almost always detected, as seen in Figure 2a, where most points overlap at 100 percent detection. For the music and windows-down conditions, where there were more cases of missed signals, the correlations for the percent detected were more substantial (0.837 and 0.642, respectively) and were statistically significant ($p<0.001$).

These correlations indicate strong agreement between the on-road and laboratory findings. The noticeability and urgency ratings were essentially the same in the two experiments, and no individual data point fell far (more than 0.5 rating unit) from the regression line. Likewise, the "meaning" data, as converted to a numeric score, had all data points close to the regression line, with one exception in the music condition.

The primary difference between the two experiments was in the percentage of signals detected in the open windows noise condition (see Figure 2c). In the on-road experiment, the group mean detection rate for the 15 acoustic signals was spread throughout the 0 to 100% range. In contrast, the laboratory experiment detection data were more binary. Most sounds (12 of 15) were well detected (at least 87.5%) and two sounds were rarely detected (4.2%). Only one signal was detected at an intermediate rate (45.8%). Although the reason for this difference is not known, it is probable that it is related to the fact that the ambient noise for the windows down condition was much more uniform in the sound clip used in the laboratory than it was under the more variable conditions of the on-road drive. Despite this difference in detection rate, given that the sound was detected, the perceptual aspects of noticeability, urgency, and meaning were highly similar across the experiments.

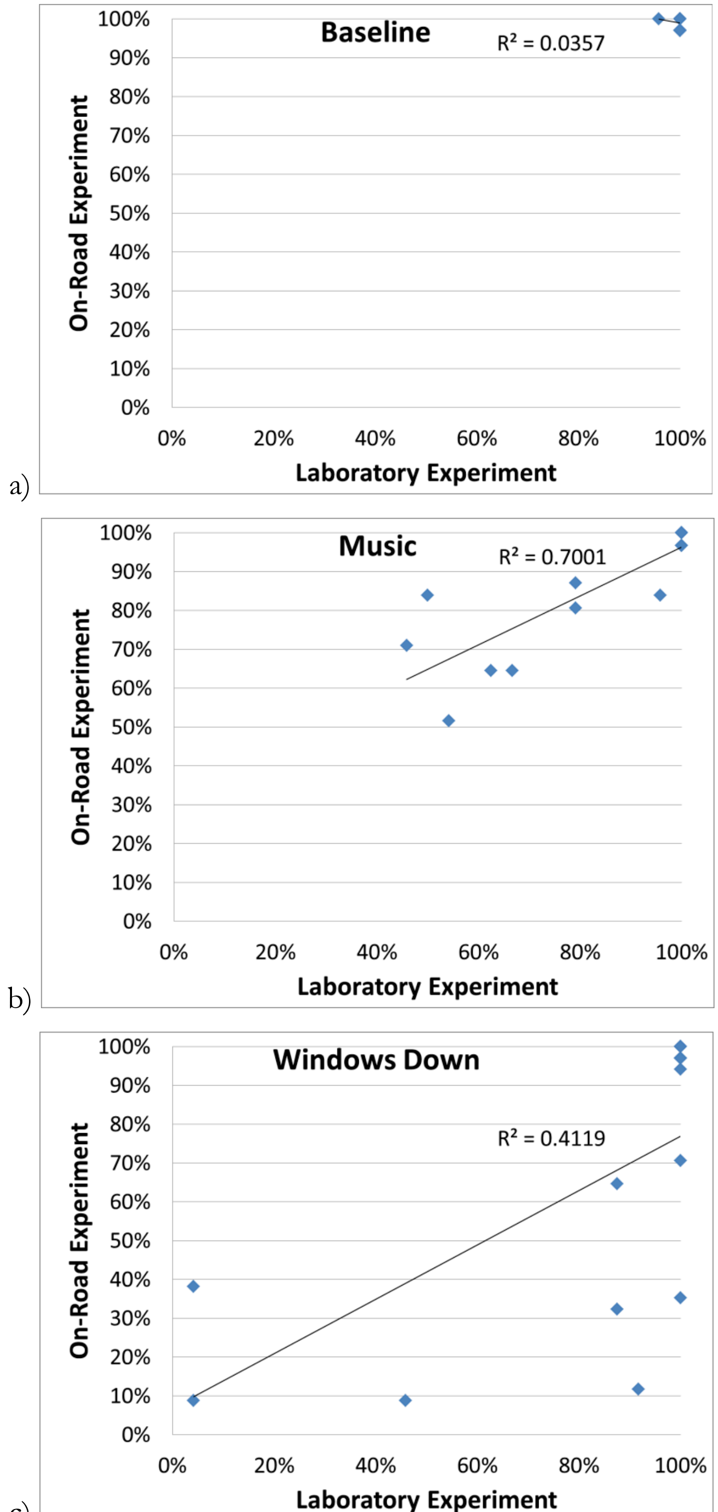
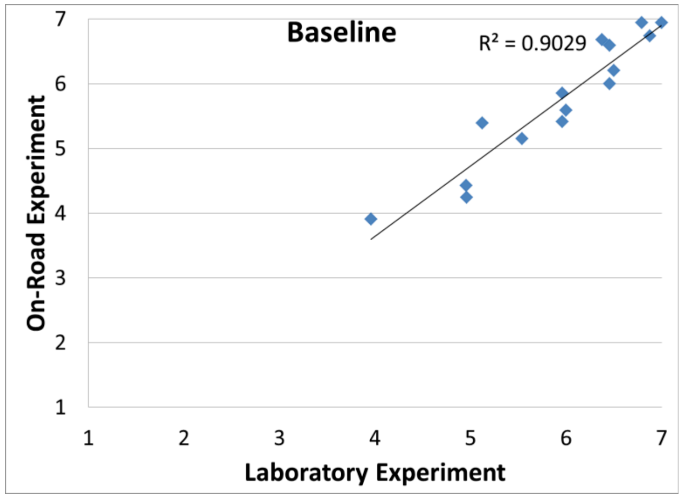
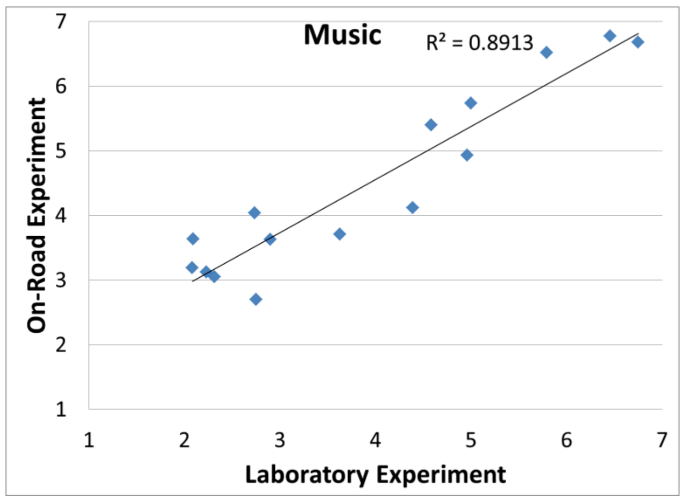


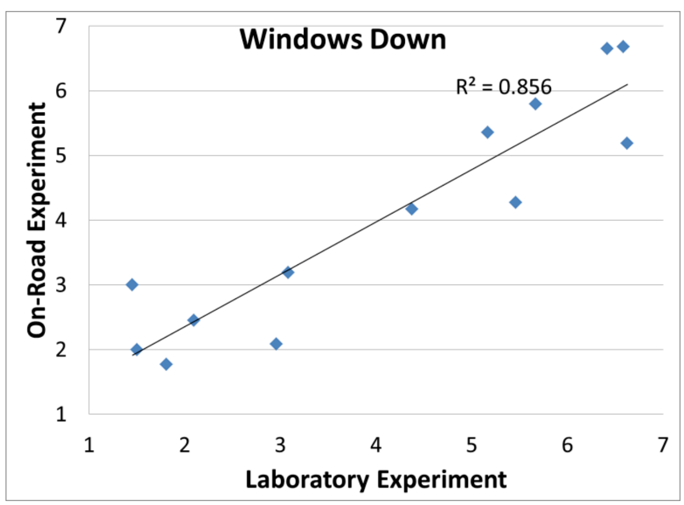
Figure 2. Relationship of percent detected findings for on-road and laboratory experiments, for each of three ambient noise conditions



a)



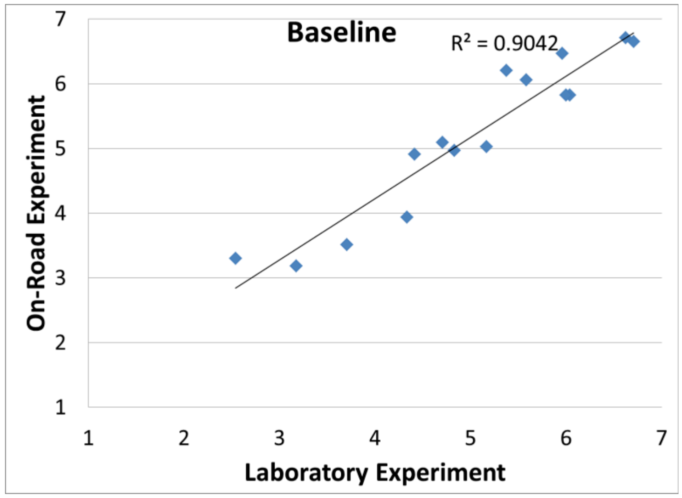
b)



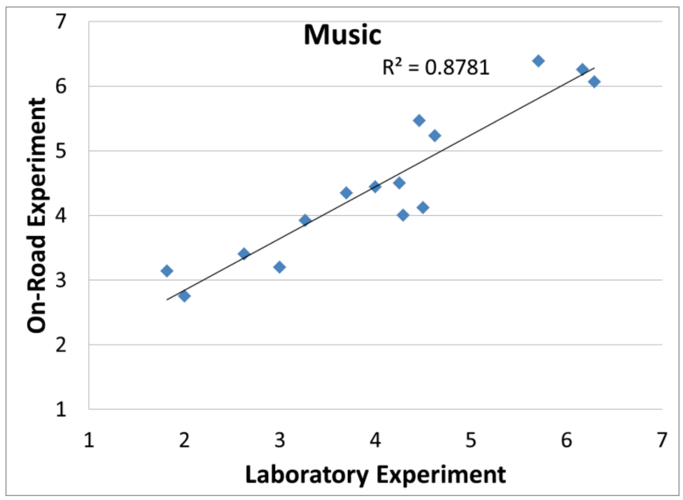
c)

Figure 3. Relationship of rated noticeability findings for on-road and laboratory experiments, for each of three ambient noise conditions

a)



b)



c)

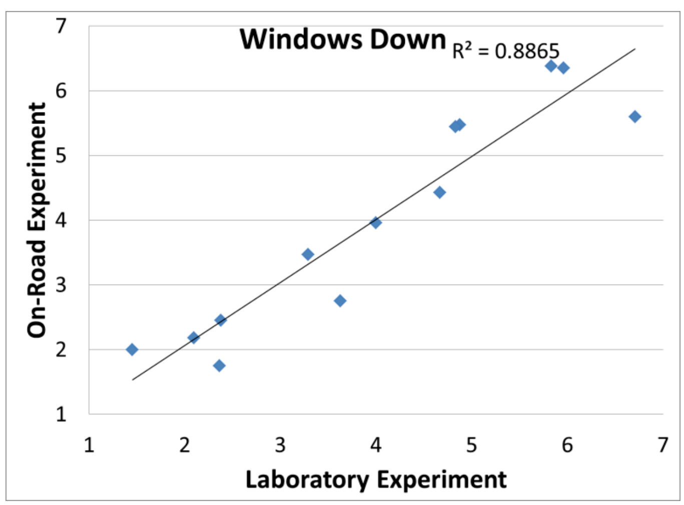
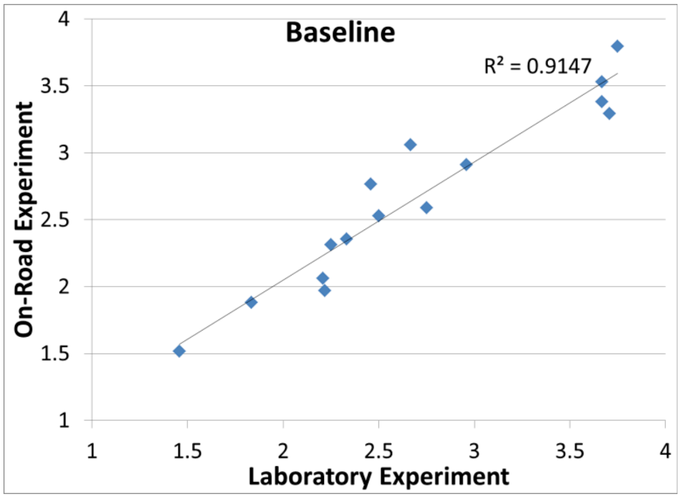
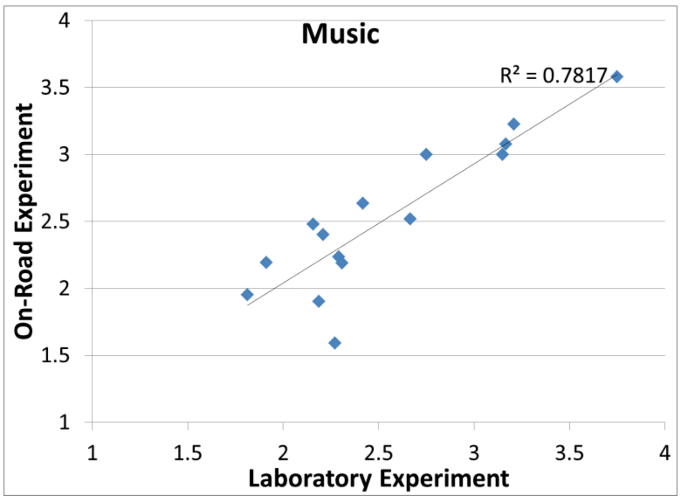


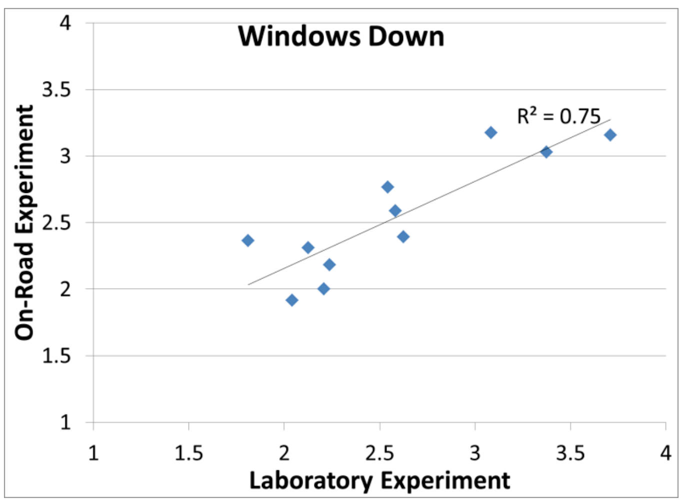
Figure 4. Relationship of rated urgency findings for on-road and laboratory experiments, for each of three ambient noise conditions



a)



b)



c)

Figure 5. Relationship of meaning category findings for on-road and laboratory experiments, for each of three ambient noise conditions

2.3 Discussion

2.3.1 Laboratory replication of on-road results

The laboratory experiment provided a close replication of the on-road findings. The relationship between the experiments accounted for 92% to 95% of the variance in the noticeability and urgency rating data under the various ambient noise conditions and from 75% to 91% of the variance in the perception of categorical meanings. This close agreement was demonstrated with only a moderate sample size of 24 participants in the laboratory experiment. Thus the laboratory procedure, using headphones and a simplified driving-like baseline task, appears to be a valid means of capturing accurate driver perceptual response to acoustic signals embedded in real-life vehicle cab noise environments. The validity of this method not only permits the use of a more efficient and controlled procedure but also promotes confidence in subsequent research findings that include noise conditions that cannot be reliably captured and controlled on the road, such as heavy rain or passing heavy vehicle noise.

2.3.2 Evaluation of laboratory procedure

Experience with the laboratory procedure points to some "lessons learned:"

- The driving simulation task appeared to be effective in maintaining participants' attention and proper seating position. The task was simple and was not intended to produce significant visual, motor, or cognitive demands, given that it was replicating steady-state driving on a low-demand limited-access highway.
- Some participants were very sensitive to any irregularities in the ambient noise sounds and occasionally clicked the finger switch in response to faint sounds in the recording such as wind gusts or passing vehicles. This was rare overall, and most often happened the first time a participant encountered a new ambient noise condition. This sensitivity may be a result of detachment of the sound environment from the visual environment. Given that there was no visual environment (other than the driving simulation task), participants may have been more intently focused on the ambient noises in the lab than they were on the road. In a general sense, the nature of the experimental task likely led participants to be closely attentive to any potential signal among the noise. Providing video recorded simultaneously with the audio might help to reduce this effect. Given an auditory-only environment, however, more uniform sound conditions without extraneous noises would

likely minimize false alarms. It is also important to avoid in-vehicle noises in recordings such as creaks, rattles, and sounds made by the driver.

- Although the on-road experiment included some experimenter control over ambient sound level (e.g., alerts were not presented while a loud truck was passing), there was still a notable amount of sound variability within and between experimental sessions. In comparison, the ambient noise recordings used in the lab study were intended to be as sonically "pure" as possible, had little variability within the signal, and eliminated all but subtle extraneous noises. While this was in many ways ideal for this experimental replication, it greatly reduced the natural variability of driving conditions. Similarly, the lack of variability in the ambient noise environment may have led to a tendency for any given sound to be heard nearly all of the time or nearly none of the time, whereas detection rates in the on-road experiment were somewhat more varied across signals.
- It is challenging to select playback equipment that can accurately play back on-road recordings at real-world intensity. The ambient noise recordings made for this experiment were extremely dynamic, with large differences between loud and quiet moments. To ensure realism, dynamic range compression and limiting cannot be applied, which means that significant amplification is required to reach adequate playback levels. The equipment used for this experiment was able to reproduce real-world levels, but only with volume levels set very close to maximum.
- Headphones differ from one another in many ways, and many headphones are unable to reproduce the complex, dynamic, and high-intensity sound environments experienced in the vehicle accurately (i.e., without distortion or deviations from neutral frequency response). The headphones that can do so with relatively high fidelity are often more expensive, high impedance, and less energy-efficient, requiring especially powerful amplifiers to drive them to adequate playback levels.

3 Experiment 2: Comparison of ambient noise conditions

Having established the appropriateness of the laboratory method, the next objective was to address a broader range of anticipated ambient noise conditions that are likely to arise in passenger vehicles. This will help identify the situations that are most problematic for good driver perception of acoustic signals and messages. The findings may also indicate whether there are important interactions between the particular ambient noise condition and specific auditory alerts. That is, one signal may suffer greater interference in one noise condition while another suffers more interference in a different noise condition. These findings will not only be of direct interest, but can also be used to inform subsequent research. Based on the results, a small number of carefully chosen ambient noise conditions can be carried forward for further research on optimal signal characteristics, driver response, consumer acceptability, and other research issues.

3.1 Method

3.1.1 Study design

The design of Experiment 2 was similar to the method and design of Experiment 1, with these key changes:

- All stimuli were played back at approximately 70 dBA. This change was made based on the results of the on-road experiment and Lab Experiment 1, which found that 65 dB sounds were often not well-detected under elevated ambient noise conditions, but that 75 dB sounds were almost always heard under all ambient noise conditions. The research team therefore hypothesized that 70 dB might be a reasonable nominal loudness for in-vehicle alerts.
- From the set of 15 acoustic signals used in Experiment 1, the "Female voice – not urgent" and the four higher loudness signals were replaced by five new tonal sounds. The new sounds were developed by researchers at George Mason University (GMU) as part of other work under the CWIM project. This work was intended to map acoustic signal characteristics to the perceived categorical meaning of the signal. In other words, it explored boundaries that result in an acoustic signal being classified as an urgent crash warning, a less urgent safety alert, or a non-safety communication. The research conducted by GMU did not include conditions of elevated ambient noise. A detailed description of this research may

be found in the CWIM project final report (Lerner et al., in press). These additional sounds are further described in Section 3.1.3.

- Five new ambient noise conditions were added: concrete road, rough asphalt road, driver window half open, driving between two heavy trucks, and driving in heavy rain. The baseline (windows closed, music off) and both-windows-open conditions were retained from Experiment 1, but the music condition was removed.
- Ambient noise conditions were made a partially between-subjects condition: All participants experienced the baseline condition plus three of the remaining six conditions. Half of the participants were exposed to the ambient conditions of "concrete," "front windows down," and "between trucks." The other half was exposed to conditions "driver window half down," "bumpy asphalt," and "heavy rain."

In summary, this was a two factor design, with 15 acoustic signals and seven ambient noise conditions. Dependent measures included acoustic signal detection, response time, rated noticeability, rated urgency, and category of meaning.

3.1.2 Participants

Thirty-six participants were recruited with the same requirements as in Experiment 1. There were equal numbers of males and females in each of the three age categories.

3.1.3 Instrumentation and recording

The 15 acoustic signals used in Experiment 2 are shown below. The first ten sounds on this list were all included in Experiment 1. The last five items in the list were sounds developed by GMU and reflect the inclusion or exclusion of various signal features found to be critical to the perception of a signal as an urgent crash warning (see Appendix F for additional details about each of these five signals). Previous GMU research found four key criteria for categorization of a signal as an urgent warning or alarm: (1) peak-to-total time ratio of at least 0.7; interburst interval of 125 ms or less; at least 3 harmonics; and a base frequency of 1000 Hz or higher.

1. FCW 1: One burst of 20 fast beeps with a relatively high frequency profile.
2. FCW 2: Four bursts of four fast beeps with a relatively low frequency profile.
3. Blind spot warning: Three bursts of four fast beeps, each with a smoothed onset and decay and a sustained low intensity sound between beeps.
4. Pedestrian warning: A constant tone with a duration of two seconds.
5. Seat belt alert 1: A single chime that decays to silence in the span of about two seconds, with intensity varying in a wavelike pattern.

6. Seat belt alert 2: Two chimes, each of which decays to silence in the span of about one second
7. Park assist 1: One burst of eight beeps.
8. Park assist 2: Two bursts of three beeps.
9. Female voice – urgent: Female voice says "Warning, warning."
10. Male voice – urgent: Male voice says "Warning, warning."
11. GMU 1 – meets all four warning signal criteria (
12. GMU 4 – meets three warning signal criteria (not interburst interval)
13. GMU 5 -- meets three warning signal criteria (not peak-to-total time ratio)
14. GMU A – meets only the ratio criterion
15. GMU B – meets only the ratio and interburst interval criteria

All 15 acoustic signals were presented at an intensity that resulted in loudness perceptually equivalent to a 70 dBA pink noise signal. The loudness for each signal was determined by a method-of-adjustment jury rating procedure. The method of adjustment procedure was used because objective, measurement-based methods do not reliably match the loudness of different signals in a way that matches human perception of loudness (Florentine, Popper, & Fay, 2010). Ten participants were individually seated in the stationary vehicle's driver's seat. Participants listened to each stimulus alternating with the pink noise every two seconds. Participants adjusted each sound's loudness by adjusting the volume slider in the digital audio player until it matched the loudness of the pink noise, which was calibrated to read 70 dBA near the participant's right ear. The group mean adjusted level for each signal was then used for presentation in the experiment.

The seven ambient noise conditions used in Experiment 2 were the following:

1. Baseline: 60 mph on smooth asphalt limited access highway
2. Concrete: 60 mph on smooth concrete limited access highway
3. Front windows down: Both front windows fully down at 60 mph on smooth asphalt limited access highway
4. Driver window half down: Driver window only down halfway at 60 mph on smooth asphalt limited access highway
5. Between trucks: Semi-trailer trucks on both sides of subject vehicle, windows closed, at 60 mph on smooth asphalt limited access highway
6. Bumpy asphalt: Driving over heavily potholed segment of road on an asphalt limited access highway
7. Heavy rain: 60 mph on smooth asphalt limited access highway during heavy rainfall, recorded during a summer thunderstorm (no meteorological data available)

All ambient noise conditions were recorded in a 2014 Toyota Camry. All recordings were newly made for this experiment (i.e., baseline and front windows down recordings were not the same used in Experiment 1) using the same recording procedures as used in Experiment 1.

3.1.4 Procedure

The procedure was the same as in Experiment 1. Each participant engaged in the driving-like secondary task and responded by activating the microswitch when they detected a signal, and then providing the ratings and meaning judgments. Each participant experienced four blocks of trials, each with a different ambient noise condition. The order of the ambient noise blocks was counterbalanced across participants.

3.2 Results

ANOVAs showed statistically significant effects ($p<0.001$) of ambient noise, signal, and the noise-by-signal interaction, for all dependent measures.

As anticipated, the 70 dBA acoustic signals in this experiment showed less dramatic effects of noise than were observed for the 65 dBA signals in Experiment 1. Nonetheless, these equally loud (under quiet conditions) sounds differed substantially in perceptual aspects of noticeability, urgency, and meaning, again demonstrating the importance of acoustic signal attributes for warning effectiveness. At the 70 dBA level, the effects of some of the ambient noise conditions on various acoustic signals were not substantial. However, some noise conditions were broadly disruptive even with 70 dBA signals, and some signals were more susceptible to noise effects than others.

Table 4, and Table 5 summarize the effects of the ambient noise condition on the ratings for each acoustic signal, for rated noticeability, urgency, and meaning (as quantified in the same manner as Experiment 1). For each signal (row), the table shows the group mean rating for the baseline condition and then the change in the group mean rating under each ambient noise condition. Thus a positive number indicates an increase in the rating, relative to baseline, and a negative number indicates a decrease in the rating. Changes of more than one-half unit on the seven-point rating scale are highlighted in Table 3 and Table 4: Green cells indicate an increase of at least 0.5 units and red cells indicate a decrease of at least 0.5 units. Changes of more than 0.3 units on the four-point meaning scale are highlighted in Table 5: Green cells indicate an increase of at least 0.3 units and red cells indicate a decrease of at least 0.3 units.

Table 3. Ambient noise effects on rated noticeability

Acoustic Signal	Baseline Rating	Change in Rating from Baseline: Noticeability					
		Concrete	Both windows down	Between trucks	Bumpy asphalt	Driver window down	Heavy rain
FCW 1	6.444	0.167	-0.856	-0.167	-0.167	-0.667	-1.111
FCW 2	5.278	-0.444	-1.333	-0.925	-1.500	-0.778	0.000
Blind Spot	6.361	0.286	-1.361	0.109	-0.139	-0.139	-1.639
Pedestrian Warning	5.914	-0.137	-2.503	-0.081	-0.137	-1.414	-1.359
Seat belt 1	5.028	0.417	-2.194	-0.322	-1.694	-1.440	-1.528
Seat belt 2	4.944	0.167	-2.217	-1.556	-1.500	-1.062	-2.356
Park asst 1	5.917	-0.028	-0.750	-0.361	-0.250	-0.417	-0.306
Park asst 2	5.611	0.389	-0.317	0.167	0.333	-0.222	-0.552
Female voice	5.972	0.139	-4.972	-0.266	-0.861	-2.855	-4.357
Male voice	5.472	0.528	-3.722	-0.766	-0.250	-2.590	-1.972
GMU 1	6.611	0.000	-0.905	0.212	0.056	-0.833	-3.222
GMU 4	6.694	0.083	-0.636	0.028	-0.306	-0.750	-2.250
GMU 5	6.611	0.000	-0.299	0.036	0.000	-0.278	-2.222
GMU A	4.829	0.060	-2.367	-1.182	-1.106	-1.593	-2.162
GMU B	4.472	-0.472	-2.847	-1.531	-1.750	-1.296	-2.694
Mean	**5.744**	**0.077**	**-1.819**	**-0.440**	**-0.618**	**-1.089**	**-1.849**

Table 4. Ambient noise effects on rated urgency

Acoustic Signal	Baseline Rating	Change in Rating from Baseline: Urgency					
		Concrete	Front windows down	Driver window half down	Between trucks	Bumpy asphalt	Heavy rain
FCW 1	5.889	-0.444	-0.242	-0.611	0.056	-0.222	-0.056
FCW 2	4.611	-0.611	-0.500	-0.552	-0.944	-0.278	0.000
Blind Spot	5.667	0.510	-0.196	0.098	0.111	0.556	-0.389
Pedestrian Warning	4.743	-0.021	-1.449	0.202	0.035	-0.132	0.035
Seat belt 1	2.806	0.361	-0.639	-0.100	-0.639	-0.806	-0.472
Seat belt 2	3.194	0.417	-0.013	-0.972	-0.250	-0.371	-0.194
Park asst 1	4.472	-0.028	0.028	0.639	0.306	0.639	0.806
Park asst 2	4.472	0.361	-0.413	0.472	0.583	-0.139	0.292
Female voice	6.444	-0.444	-3.844	-0.444	-0.444	-1.621	-4.368
Male voice	6.056	0.500	-3.306	-0.173	0.444	-1.232	-1.056
GMU 1	6.139	0.028	-0.257	-0.021	0.417	-0.250	-1.750
GMU 4	5.833	0.222	-0.363	0.000	0.500	-0.444	-0.556
GMU 5	6.139	-0.417	-0.139	0.332	-0.139	-0.194	-0.861
GMU A	3.800	-0.300	-1.108	-0.565	-0.078	0.141	-0.522
GMU B	4.028	-0.806	-0.903	-1.263	-0.250	0.031	-1.028
Mean	**4.953**	**-0.045**	**-0.889**	**-0.197**	**-0.020**	**-0.288**	**-0.675**

Table 5. Ambient noise effects on category of meaning

Acoustic Signal	Baseline Rating	Change in Rating from Baseline: Meaning					
		Concrete	Front windows down	Driver window half down	Between trucks	Bumpy asphalt	Heavy rain
FCW 1	2.972	-0.361	0.087	-0.250	0.472	0.194	0.417
FCW 2	2.278	-0.222	-0.111	-0.042	-0.389	-0.111	0.111
Blind Spot	2.889	0.052	-0.242	0.111	0.278	0.500	0.222
Pedestrian Warning	2.486	-0.097	-0.368	0.070	0.181	0.125	0.348
Seat belt 1	1.389	-0.056	0.111	-0.154	0.111	0.199	0.111
Seat belt 2	2.167	0.000	0.197	-0.056	-0.222	-0.049	0.245
Park asst 1	2.500	-0.111	-0.056	0.222	-0.111	0.222	0.278
Park asst 2	2.278	-0.167	-0.160	0.056	-0.056	-0.056	0.075
Female voice	3.778	-0.222	-1.578	-0.013	0.056	-0.601	-1.624
Male voice	3.806	0.028	-1.806	-0.217	-0.083	-0.570	-1.306
GMU 1	3.417	-0.250	-0.417	-0.181	0.083	-0.139	-0.417
GMU 4	3.139	0.028	-0.080	-0.083	0.194	0.028	0.139
GMU 5	3.500	-0.333	-0.375	-0.265	-0.056	-0.056	-0.222
GMU A	2.086	-0.197	-0.240	-0.027	-0.030	0.208	-0.030
GMU B	2.361	-0.028	0.014	-0.420	-0.083	-0.008	-0.028
Mean	**2.736**	**-0.129**	**-0.335**	**-0.083**	**0.023**	**-0.007**	**-0.112**

Table 3 summarizes the noise effects on noticeability. There was little difference between the concrete and asphalt road surfaces. However, noticeability was broadly affected by the other noise conditions. Using the 0.5 unit shift as a criterion, perceived noticeability was reduced in 48 of the 75 (15 signals under five noise conditions) cases, and for almost all sounds under the front windows down and the heavy rain conditions. There were fewer substantial shifts in the ratings of urgency (Table 4). However, rain and front windows down reduced urgency for seven of the 15 signals, and the driver window half down condition reduced urgency for five signals. The numeric conversion of the category meaning data was the least sensitive measure to ambient noise effects, but there were a cases of substantially reduced urgency. Front windows down accounted for five instances of sizable drops in meaning and heavy rain for three. The male and female voice messages were particularly susceptible to substantial drops in meaning.

As Table 3, Table 4, and Table 5 indicate, the ambient noise conditions of front windows down and heavy rain had broadly detrimental effects on noticeability, urgency, and meaning. However, not all signals were affected by these noise conditions in the same way, as the statistically significant noise-by-signal interaction terms suggest. The interaction of ambient noise with specific signals is illustrated in Figure 6. The figure shows the group mean urgency ratings for three signals (GMU 1, Pedestrian Warning, Male Voice Warning) under the baseline and "both windows down" and "heavy rain" noise conditions. The GMU 1 signal was little affected by "both windows down" but urgency ratings dropped under "heavy rain." In contrast, the Pedestrian Warning was little influenced by "heavy rain" but did drop under "both windows down." The Male Voice Warning urgency dropped under both "both windows down" and "heavy rain" noise conditions, but more so for "both windows down," and this drop was considerably larger than the others observed in the figure. Thus the figure shows that the effect of an ambient noise condition depends on the characteristics of the particular signal itself.

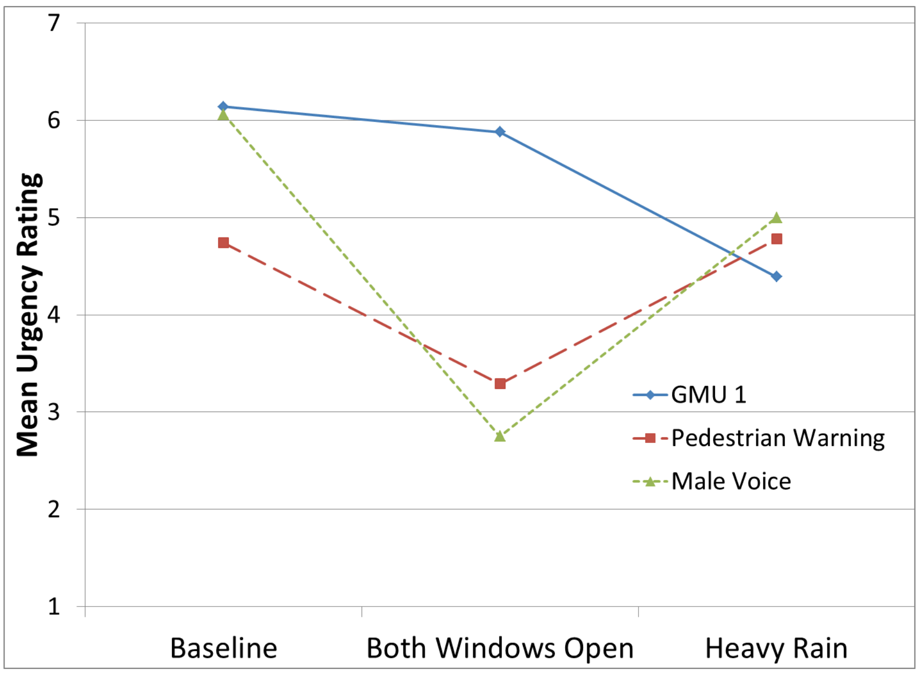


Figure 6. Comparison of the differing effects of rain and windows down noise conditions on the perceived urgency of three different warning signals

3.3 Discussion

As in the previous experiment, the significant main effect of signal demonstrates that sounds that are equally loud under quiet listening conditions nonetheless differ meaningfully in noticeability, urgency, and meaning under driving noise conditions. As anticipated, these effects were not as pronounced with the 70 dBA signals used in Experiment 2 as compared to the 65 dBA signals included in Experiment 1. However, this is an informal, between experiments observation. Experiment 3 addresses sound level more directly. The significant effect of signal confirms the importance of determining those signal characteristics that promote the perception of "crash warning" and distinguish such warnings from other acoustic signals.

Although the 70 dBA sound level resulted in generally good resistance to disruption of urgency and meaning in noise conditions, there were substantial effects in two noise conditions. Both the "front windows down" and the "heavy rain" noise conditions had broad negative impact on warning perception. Voice messages were particularly susceptible to noise effects, not only in the "front windows down" and the "heavy rain" noise conditions but also the "driver window half down" condition. This suggests that the urgency aspect is carried by the speech message, the intelligibility of which is vulnerable to noise effects. Therefore, evaluations of warnings that are limited to quieter conditions might overestimate the real world efficacy of voice signals.

The significant interaction of ambient noise condition and signal indicated that different sounds are differentially affected by different noise conditions. Acoustic signal characteristics affect the relative influence of different ambient noise conditions, as well as the magnitude of effects. Figure 6 shows this quite dramatically. Perceived urgency of the GMU 1 signal is little affected by the windows down condition, but drops by 1.75 rating scale units in the heavy rain condition. In contrast, the Pedestrian Warning was little affected by the rain condition, but dropped by 1.45 rating scale units in the windows down condition. The magnitude of theses drops, while large, is much less than the perceived urgency decrease observed for the Male Voice Warning in the windows down condition (3.31 rating scale units). These findings demonstrate the importance of developing in-vehicle signals that are resistant to a range of potentially disruptive noise conditions. This experiment found that windows down and rain conditions were particularly disruptive.

4 Experiment 3: Effects of signal loudness on perceived meaning, annoyance, and acceptance

Experiment 3 had two primary objectives. First, it provided a direct comparison of acoustic signal meaning degradation in ambient noise as a function of signal loudness. Experiment 1 (as in the on-road experiment upon which it was based) used loudness levels equivalent to 65 dBA and 75 dBA pink noise. The findings suggested that 70 dBA might be a reasonable level for in-vehicle acoustic signals. Experiment 2 then used the 70 dBA level for signals in its comparison of ambient noise conditions. In Experiment 3, the three sound levels of 65, 70, and 75 dBA were directly compared. The previous experiments used a variety of different sounds to investigate the effects of signal features other than loudness. Experiment 3 focused more directly on the loudness factor. The second objective was to bring driver annoyance and consumer acceptance dependent measures into the evaluation of the signals. One purported reason that automobile manufacturers sometimes use relatively low sound levels for warnings is that they are concerned about consumer acceptance if sounds are intrusive and annoying, particularly when there are nuisance alarms. Annoyance and acceptance must be considered in any comprehensive approach providing suggested criteria for warnings, alerts, and other acoustic signals. As characterized in this experiment, annoyance referred to the unpleasantness of the sound whereas acceptance judgments took into account the intended meaning or application of the acoustic signal (e.g., crash warning or personal communication). Thus, the rating indicated how willing the participant was to accept a particular sound for a particular purpose.

4.1 Method

4.1.1 Study design

Experiment 3 was comprised of two phases. The first phase was similar in procedure to the preceding experiments, collecting ratings and meaning category data after each presentation of a signal (assuming it was detected by the participant). This was a three-factor, within-subjects design, with the factors of acoustic signal, signal loudness, and ambient noise condition. The experiment used a subset of six signals and three ambient noise backgrounds from the preceding experiments. The signals were presented at three different sound levels (65, 70, and 75 dBA).

The second phase of the experiment collected participant judgments about the consumer acceptability of each of the 18 sounds used in the experiment (six signals, each at three loudness levels). These were presented in a random sequence.

4.1.2 Participants

Twenty-four participants were recruited with the same requirements as in Experiments 1 and 2. However, the boundaries of the age categories were different: 21-35 years old, 36-50 years old, and 51-65 years old. The age range was extended in this experiment (from a ceiling of 50 to 65) because a broader age representation might be important to accurately assess driver acceptance. Also, although all participants had no self-reported hearing problems, age-related losses in sensitivity, particularly at higher frequencies, might be of particular interest in the relationships among sound characteristics, sound level, perceived meaning/urgency, and annoyance/acceptance. There were equal numbers of males and females in each of the three age categories.

4.1.3 Instrumentation and recording

This experiment used a subset of the ambient noise recordings and acoustic signals from Experiment 2. The ambient noise conditions were baseline, driver window half open, and heavy rain. The acoustic signals were:

- FCW 2: Four bursts of four fast beeps with a relatively low frequency profile.
- Pedestrian warning: A constant tone with a duration of two seconds.
- Female voice – urgent: Female voice says "Warning, warning."
- GMU 1 – meets all four warning signal criteria
- GMU 4 – meets three warning signal criteria (not interburst interval)
- GMU A – meets only the ratio criterion

4.1.4 Procedure

The experimental session comprised two phases. The first phase was the same procedure as in Experiments 1 and 2, with the exception that an additional rating of "annoyance" (on a seven-point scale) was included for each trial. The instructions defined annoyance as meaning "that the sound is unpleasant or irritating. A 'one' means that the sound is not very annoying. A 'seven' means that the sound is extremely annoying." Each participant engaged in the driving-like secondary task and responded by activating the microswitch when they detected a signal, and then provided the ratings and meaning judgments. Each participant experienced three blocks of trials, each with a different ambient noise condition. The order of the ambient noise blocks was counterbalanced across participants, and the order of stimuli within each block was randomized.

The second phase of the session obtained participant judgments of the "acceptability" of each of the 18 sounds (six signals x three loudness levels) used in the experiment. Participants judged the acceptability of each sound for each of four categories of signal meaning: The specific instructions to the participant were:

"For the last part of this session, you will hear some more sounds and give your opinions about how acceptable those sounds would be to you as a driver. I will play you a sound and then ask you to rate how appropriate it is as an urgent crash warning, safety information, information not related to safety, and an incoming personal communication. This is important because some sounds might be good for letting you know that you're about to crash, but might not seem right for letting you know that you just received an email. On the other hand, some sounds might be good for letting you know that you just received an email, but might not seem right for letting you know that you're about to crash. You will rate how acceptable each sound is on a scale from 1 to 7, where 1 means absolutely unacceptable and 7 means completely acceptable. Please keep in mind that this question is not asking how effective the sound would be at getting your attention. It is specifically asking how acceptable the sound would be to you if you had it in your own car."

4.2 Results

ANOVAs were conducted for the five dependent measures of noticeability, urgency, meaning, annoyance, and acceptability. The outcomes of three-factor analyses for noticeability, urgency, meaning, and annoyance are summarized in Table 6 below. As the table indicates, the main effects of ambient noise condition, signal, and signal loudness were statistically significant for all measures (with the exception of being slightly above the p=0.05 level for meaning). Figure 7 shows the effect of loudness on mean ratings of noticeability, urgency, and annoyance for each of the three ambient noise conditions.

Table 6. Outcomes of Experiment 3 ANOVAs for noticeability, urgency, meaning, and annoyance

	Ambient Noise	Signal	Loud-ness	Ambient Noise* Signal	Ambient Noise* Loud-ness	Signal* Loud-ness	Ambient Noise* Signal* Loud-ness
Noticeability	p<0.0001	p<0.0001	p<0.0001	p<0.0001	p<0.0001	p<0.0001	p<0.0001
Urgency	p<0.0001	p<0.0001	p<0.0001	p<0.0001	p<0.0001	p=0.2843	p=0.0635
Meaning	p=0.0505	p<0.0001	p<0.0001	p<0.0001	p=0.4430	p=0.3084	p=0.5367
Annoyance	p<0.0001	p<0.0001	p<0.0001	p<0.0001	p=0.0029	p=0.9394	p=0.7760

*light shading means almost significant at p=.05; dark shading means not significant

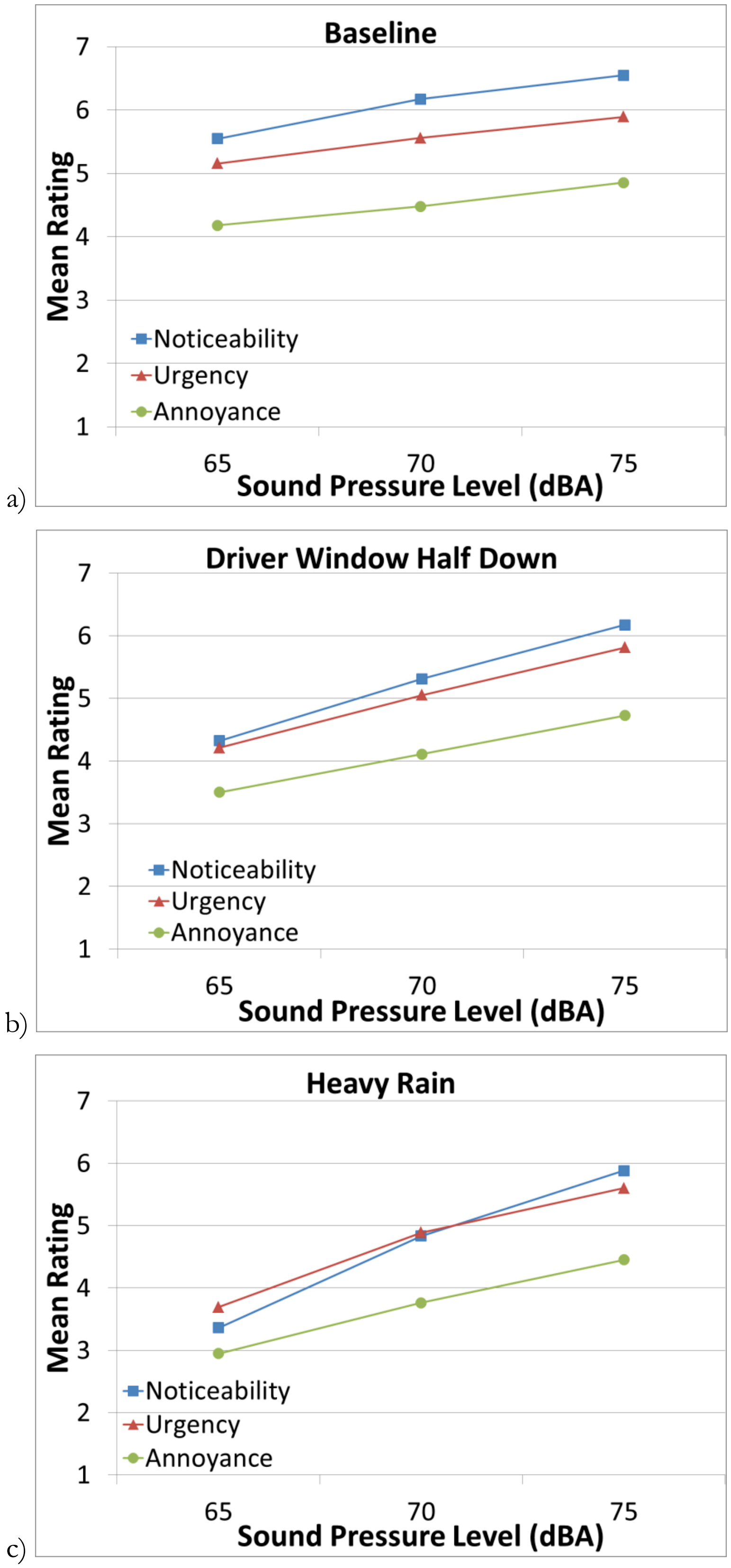


Figure 7. Main effects of loudness on mean ratings for each ambient noise condition

As in the other experiments, the ambient noise-by-signal interaction was significant. The significance of the other interaction terms varied somewhat among dependent measures. The ambient noise-by-sound level interaction was significant for three of the four dependent measures, reflecting the fact that the louder signals were less influenced by ambient noise. A complementary interpretation of this interaction is that increasing signal loudness had a more substantial positive effect on ratings in the louder ambient noise conditions (driver window half down and heavy rain) than in the quieter baseline condition. In contrast, the signal-by-sound level interaction was not significant except for the noticeability measure, indicating that the relative ordering of sounds on the various perceptual measures was similar at each signal loudness level.

Two-factor ANOVAs were conducted on the ratings of acceptability (ambient noise and loudness were not varied for this portion of the experiment; all sounds were played at 70 dBA against the baseline ambient noise). Participants rated the acceptability of each signal for each of four potential message meanings. Table 7 summarizes the analyses. The features of the signal significantly influenced acceptability for all four message meanings. The sound level of the signal only had a significant effect for the two safety-related meanings, wherein increasing signal loudness led to increased acceptance of the signal as an urgent crash warning or safety information. The interaction term was not statistically significant in any of these analyses.

Table 7. Outcomes of Experiment 3 ANOVAs for acceptability ratings

Alert type	Signal	Loudness	Signal*Loudness
Urgent crash warning	p<0.0001	p<0.0001	p=0.9769
Safety information	p<0.0001	p<0.0001	p=0.3408
Non-safety information	p<0.0001	p=0.7892	p=0.7871
Personal communication	p<0.0001	p=0.5249	p=0.8210

* dark shading means not significant

In contrast to the previous experiments, Experiment 3 included measures of annoyance and acceptability. The relationships of these measures to each other, and to the perceptual measures related to signal effectiveness as a warning, are important for consideration in designing warnings that are both effective and tolerated by consumers. Table 8 is a correlation matrix showing the relationship of each response measure to the other measures.

Table 8. Correlation matrix of dependent measures

	Urgency	Meaning	Annoyance	Acceptability as Urgent Crash Warning	Acceptability as Safety Information	Acceptability as Not Safety	Acceptability as Personal Communication
Noticeability	0.888	0.878	0.742	0.761	0.819	-0.569	-0.692
Urgency		0.892	0.922	0.899	0.864	-0.749	-0.820
Meaning			0.804	0.794	0.827	-0.543	-0.736
Annoyance				0.896	0.821	-0.808	-0.930
Acceptability as urgent crash warning					0.914	-0.807	-0.833
Acceptability as safety info						-0.592	-0.825
Acceptability as non-safety info							0.692

Several points may be pointed out in this matrix. As expected, the warning "performance" measures of noticeability, urgency, and meaning were strongly interrelated. Also as anticipated based on existing literature, annoyance was strongly related to perceived urgency. There was a strong positive relationship between rated annoyance and acceptability as an urgent crash warning and a strong negative relationship between annoyance and acceptability as a non-safety or personal communication message. In other words, participants felt that urgent and annoying signals were very acceptable as urgent crash warnings, and non-urgent/non-annoying signals were not acceptable as urgent crash warnings. Urgent and annoying signals were not at all acceptable as non-safety or personal communication messages. The intercorrelations among the four acceptability measures were generally strong. Of interest are the substantial negative correlations between the crash warning and safety categories on one hand and the non-safety and personal communication categories on the other. This demonstrates good agreement in categorizing sounds at least broadly as safety and non-safety messages.

The subsequent scatterplots illustrate some of these points. Figure 8 shows the strong positive relationship between perceived urgency and the annoyance of the signal. Although there is a strong linear relationship, highly urgent signals (in the 6 to 7 range) nonetheless vary in how annoying they are (ranging from 4.18 to 6.27). This indicates that some sounds that may be effective in conveying urgency are more annoying (potentially as false or unnecessary alarms) than others that are similarly effective. The most annoying signal was the GMU 4 signal at 75 dBA and the second and third most annoying were the GMU 4 signal at 65 dBA and 70 dBA. So although this signal was comparable in urgency to others, it may not be ideal as a warning signal because of its greater potential for annoyance. The three least annoying signals with urgency ratings greater than 6 were the three female voice signals. The voice signals conveyed the urgency of the message primarily through the word "warning" so they did not need to include annoying tonal characteristics to convey this message. As the correlations in Table 8 indicate, signals rated high in annoyance were also rated high in acceptability as urgent crash warnings. This indicates that participants found annoying signals appropriate if valid as crash warnings. However, consumer acceptance problems arise for false or nuisance warnings, which is where considerations of annoyance are most important.

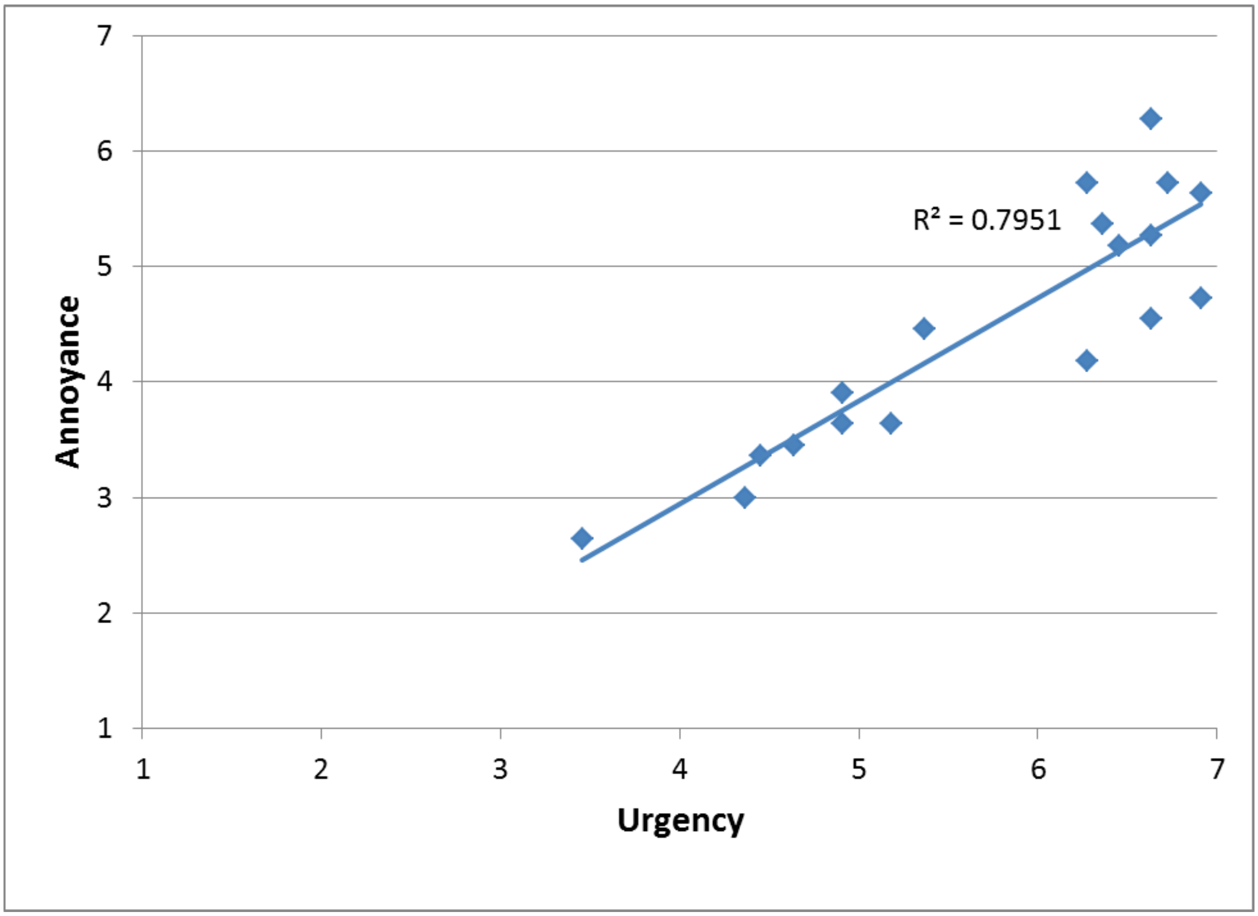


Figure 8. Relationship of urgency and annoyance ratings

Figure 9 shows the plot of urgency versus acceptability as an urgent crash warning. An interesting aspect of this chart is the range in acceptability among signals of similar urgency. The most highly rated signals in terms of acceptability were all at 75 dBA (GMU 1, female urgent voice, and GMU 4). This suggests that louder signals would be quite acceptable as crash warnings assuming they are not false or nuisance warnings. The contrast of appropriateness and annoyance factors points to consideration of nuisance alarm rates in determining what crash warning signals might be optimal.

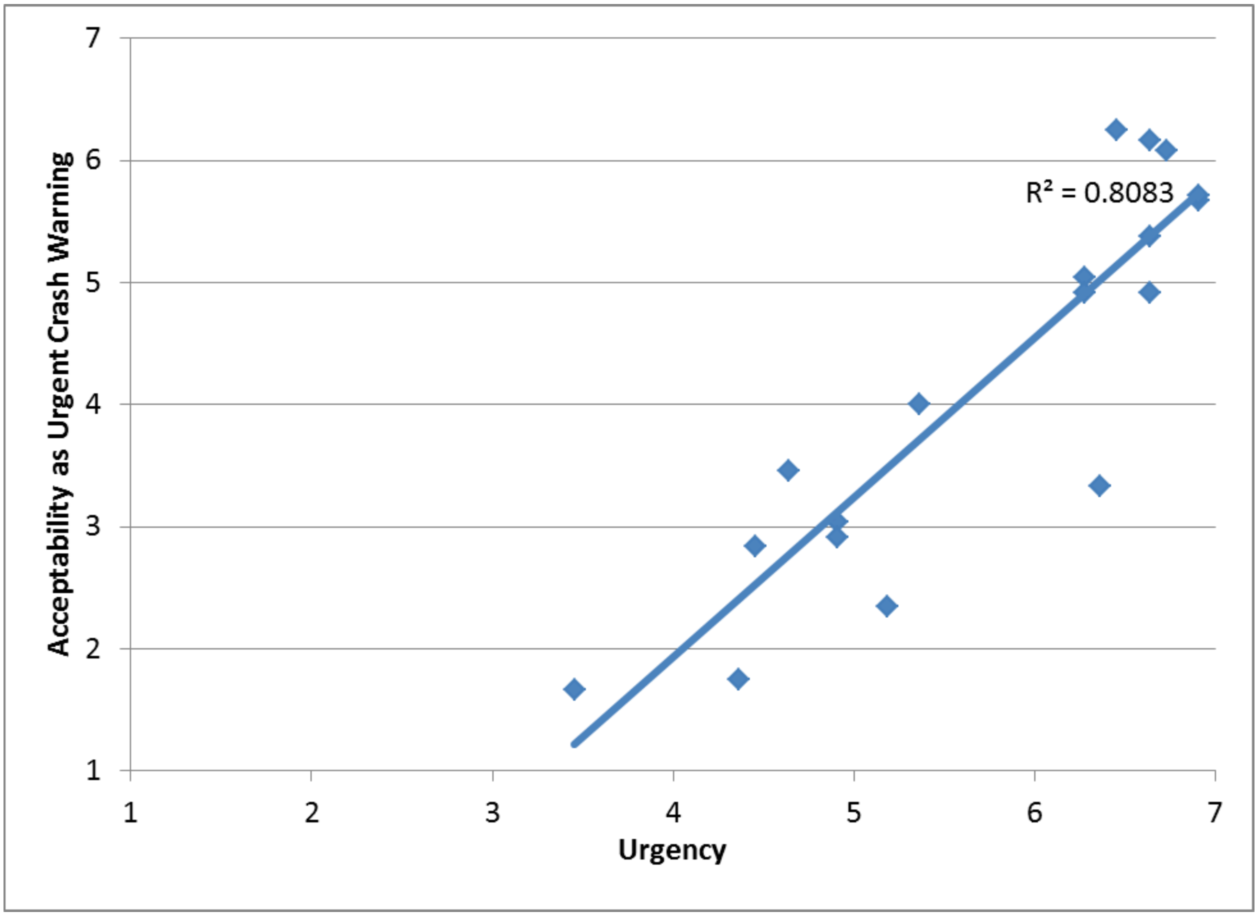


Figure 9. Relationship of urgency and acceptability as an urgent crash warning ratings

Figure 10 shows the relationship of perceived urgency and acceptability as a safety information alert. Although the relationship is positive, the function is much less steep than for the relationship between urgency and acceptability as an urgent crash warning (Figure 8). Even low urgency signals are still relatively acceptable as safety alerts, although higher urgency signals are preferred.

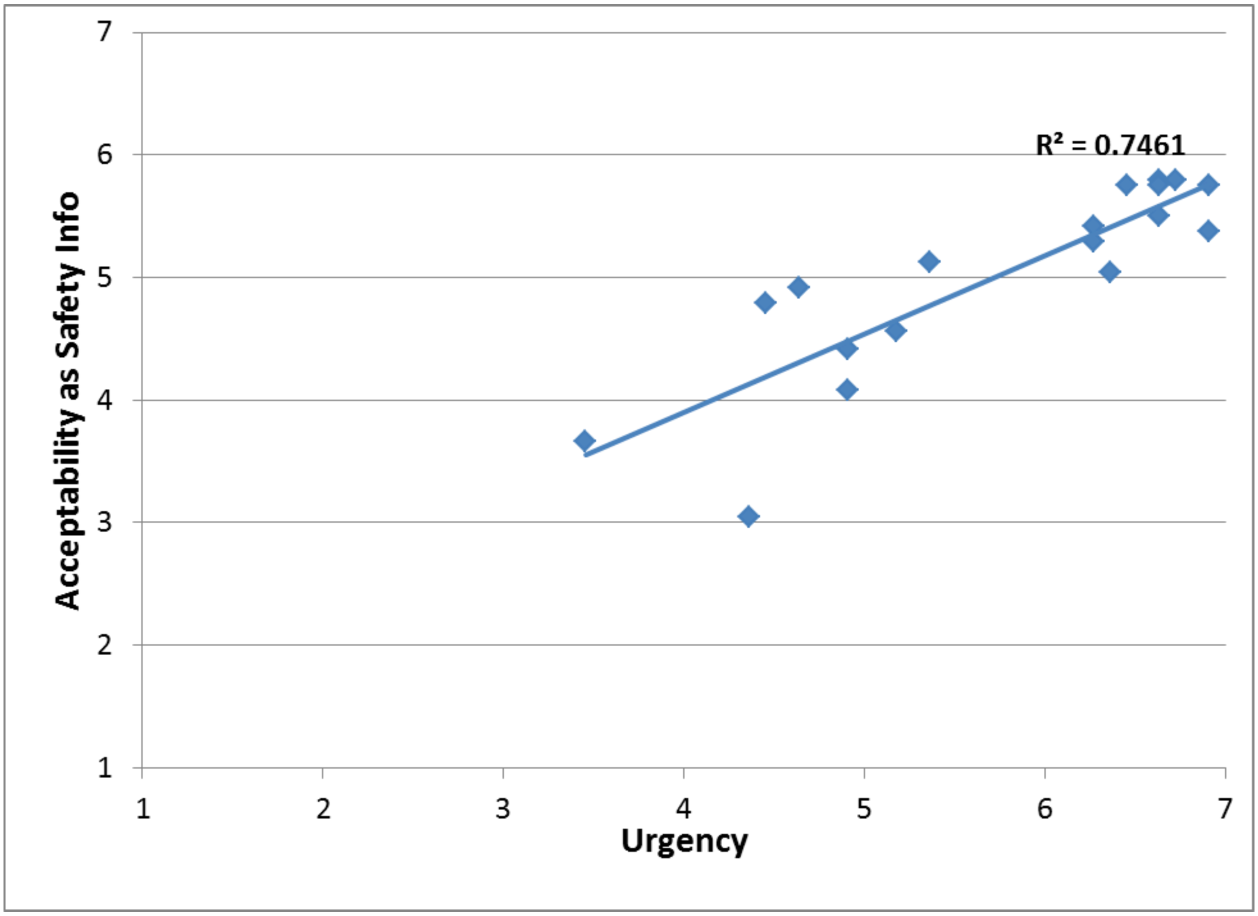


Figure 10. Relationship of urgency and acceptability as safety information ratings

Figure 11 plots acceptability as an urgent crash warning against acceptability as non-safety information. Ratings vary considerably for the crash warning message (range of about 4.5 units) but are much more restricted for the non-safety message (range of about 2.5 units). Despite the restricted range, there are some signals that fall clearly in the non-safety category without much acceptability as a crash warning, and vice versa. This provides some basis for separating signals based on clear distinctions among categories. Although not shown in any scatterplot here, it may also be noted that the FCW 2 signal stood out as a candidate for the personal communication category. It was rated highest of all signals for acceptability in this category and had very low ratings for annoyance and for acceptability as a crash warning.

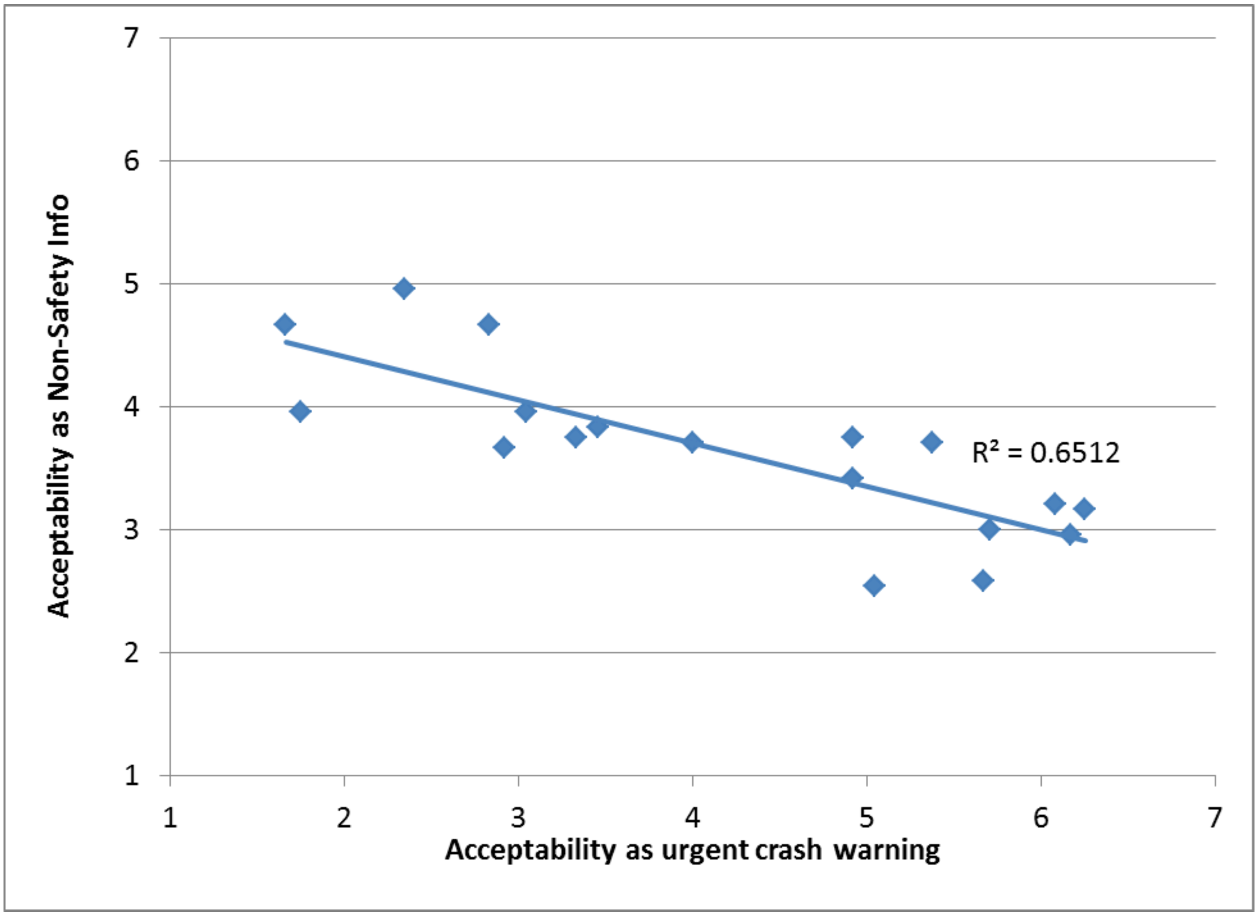


Figure 11. Relationship of acceptability as an urgent crash warning and acceptability as safety information ratings

4.3 Discussion

This experiment expanded upon the findings of Experiments 1 and 2 by specifically investigating the effects of signal loudness, and adding ratings for signal annoyance and acceptability for particular alerting purposes. As expected, increases in signal loudness led to higher ratings for noticeability, urgency, and annoyance. This effect was more pronounced in the louder ambient noise conditions (driver window half down and heavy rain) than in the relatively quiet baseline condition, which suggests that higher loudness levels are particularly beneficial for maintaining the noticeability and perceived urgency of warnings in loud ambient noise conditions.

This experiment found that noticeability, perceived urgency, and annoyance were highly intercorrelated. These variables were also highly correlated with appropriateness as an urgent crash warning, with perceived urgency and annoyance being particularly highly correlated. This suggests that signals that are perceived as urgent and annoying are also more likely to be perceived as crash warnings, and are largely deemed to be appropriate for use as urgent crash warnings. Drivers, however, are not likely to accept annoying signals for use as non-safety related alerts.

The findings of this experiment also show that for a given level of perceived urgency, some signals were rated as more annoying than others, so it might be possible to use signals that effectively convey the warning message without being excessively annoying to drivers. The female voice message effectively conveyed urgency to participants with relatively little annoyance; however, voice messages are generally not recommended for urgent crash warnings (Campbell, Richard, Brown, & McCallum, 2007).

When considering the acceptability of signals as urgent crash warnings, it is important to also consider the practical concerns of false and nuisance warning. In real-world experience, drivers might perceive many warnings as unnecessary, either due to limitations of the warning system or because drivers consider the warning to be too early, warning them of a situation they were already aware of, or otherwise unhelpful. Acceptability of annoying signals is likely to be negatively affected if false or nuisance warnings are relatively frequent.

5 General Discussion and Conclusions

5.1 Overview

The line of research described in this report began with the recognition that the intended meanings of acoustic sound and speech signals in vehicles need to be retained under the range of ambient noise conditions that may be expected when driving. There is no standard acoustic signal for urgent crash warnings or other in-vehicle messages for passenger vehicles, nor does it seem likely that there will be any time soon. Current practice varies widely among OEMs in terms of the acoustic characteristics, sound intensity, number and type of warnings and alerts, and so forth.

Furthermore, with the expectation of increased sensing capabilities and communications taking place in vehicles, as well as signals from portable and aftermarket devices, it will be critical that every auditory signal conveys the proper level of urgency to the driver. Crash warnings must be unambiguous with respect to the need for immediate driver action; non-crash warnings should not confuse the driver by conveying this warning message. Work conducted under the CWIM program, Connected Vehicles, and other NHTSA efforts has been addressing the issue of devising appropriate methods for insuring proper categorical perception of the intended meaning of acoustic warnings, alerts, and notifications in passenger vehicles. However, neither this work, nor the broader literature on in-vehicle warnings and alerts, has addressed the range of in-vehicle noise environments. Almost all of the research on acoustic crash warnings and other alerts has been conducted with only minimal or moderate ambient noise conditions in the vehicle. In real world applications, warnings must function well under a wide variety of potential noise conditions, such as when there is music, noise from lowered windows, noise from passing heavy vehicles, rain, rough road surfaces, and so forth. Existing literature does not indicate whether acoustic signals maintain their meaning under these conditions and what signal design attributes promote good communication under noise conditions.

The on-road experiment conducted as part of the CWIM project (Singer et al., in preparation) and the three subsequent laboratory experiments described in this report provide initial findings on these ambient noise issues. The initial on-road experiment demonstrated that ambient noise conditions could alter the perceived meaning and urgency of an acoustic message. The higher noise condition of having the front windows open led to many missed signals at the 65 dBA level, but very few at the 75 dBA level. But even when the signal was detected, noise could alter how the signal was interpreted by the driver. In particular, signals intended to convey an urgent warning lost some of their urgency; the degree to which this occurred varied substantially from signal to signal. Thus the on-road experiment established the practical significance of these perceptual changes for at least some signals under at least some realistic noise conditions.

5.2 Summary of Experiments

Experiment 1 in this set of laboratory experiments established the validity of a method based on the use of headphones and a simplified driving-like secondary task. The findings closely replicated the results obtained in the actual on-road setting. This replication was important to establish, since the complexities of recording, calibration, processing, and playback of the complex in-vehicle acoustic noise conditions and signals require a demonstration that the listener perceptual experience remains essentially intact. This validated laboratory approach thus can provide a more efficient and controlled method, relative to an on-road experiment. It also allows the inclusion of transient roadway and environmental conditions that cannot be easily controlled or sustained on the road, such as road surface irregularities, surrounding large truck traffic, or rain.

Experiment 2 expanded the findings of Experiment 1 by investigating the effects of a wider range of ambient noise conditions, including different pavement and environmental conditions. It also used a signal loudness of 70 dBA, which the research team hypothesized would be a practical loudness level between the 65 and 75 dBA levels used in Experiment 1. As in the previous experiment, the significant main effects of signal and ambient noise condition demonstrate that sounds that are equally loud under quiet listening conditions nonetheless differ meaningfully in noticeability, urgency, and meaning under driving noise conditions, and between various driving noise conditions. The windows down and heavy rain conditions were especially disruptive, providing evidence that these conditions should be considered when designing and evaluating warnings. The significant interaction of ambient noise condition and signal indicated that different sounds are differentially affected by different noise conditions. Acoustic signal characteristics affect the relative influence of different ambient noise conditions, as well as the magnitude of effects, meaning that a signal that is effective in one ambient noise condition, even one that is noisy, may not be as effective in a different ambient noise condition. As anticipated, these effects of ambient noise condition were not as pronounced with the 70 dBA signals used in Experiment 2 as compared to the 65 dBA signals included in Experiment 1. However, this was an informal, between experiments observation.

Experiment 3 directly investigated the effects of three different signal loudness levels, as well as adding annoyance and acceptability as dependent variables. As expected, increases in signal loudness led to higher ratings for noticeability, urgency, and annoyance. This effect was more pronounced in the louder ambient noise conditions (driver window half down and heavy rain) than in the relatively quiet baseline condition, which suggests that higher loudness levels are particularly beneficial for maintaining the noticeability and perceived urgency of warnings in loud ambient noise conditions.

This experiment found that noticeability, perceived urgency, and annoyance were highly intercorrelated. These variables, and especially perceived urgency and annoyance, were also highly correlated with perceived appropriateness of the signal as an urgent crash warning, which suggests that drivers are likely to accept annoying signals as urgent crash warnings, although this acceptance could potentially be moderated by false or nuisance warnings, which were not investigated in this experiment.

5.3 Limitations of this Research

The three laboratory experiments conducted in this series provide numerous insights into the interactions between ambient noise conditions and signal characteristics, and how they relate to driver perceptions of in-vehicle alerts. These experiments indicated that there is a significant concern with the effectiveness of warnings under realistic driving conditions and began to suggest some of the key ambient noise and acoustic signal attributes that are important. However, there are a number of limitations of this work.

While these experiments investigated many different signals, signal characteristics were not systematically manipulated in a way that would allow researchers to determine what specific characteristics of alerts influence drier perceptions. Research conducted by George Mason University within NHTSA's CWIM program (Lerner et al., in press) has systematically investigated the signal criteria that differentiate between alerts and less urgent signals, but this work did not include a range of ambient noise conditions.

Another limitation of the present study was that participants were focused on listening for signals, whereas in real driving, drivers are not likely to be so highly attuned to the occurrence of alerts due to lack of expectation or distraction. The alerts used in this study were also auditory-only, and did not provide participants with any supplementary information or context for the alerts (e.g., a dashboard indicator light or an on-road event such as an impending collision).

Finally, these experiments investigated how participants rated and categorized signals, but did not investigate actual responses to signals in a driving context. Ultimately, the most important test of a warning is whether it leads to appropriate driver responses (e.g., orienting and responding to a hazard). For less urgent signals, however, a rapid reorientation of attention might be undesirable.

6 References

Campbell, J. L., Richard, C. M., Brown, J. L., & McCallum, M. (2007). Crash warning system interfaces: Human factors insights and lessons learned (Report No. DOT HS 810 697). Washington, DC: National Highway Traffic Safety Administration.

Florentine, M., Popper, A. N., & Fay, R. R. (2010). Loudness. In: Richard R. Fay & Arthur Popper (Eds.), *Springer Handbook of Auditory Research*. New York: Springer.

Lerner, N., Singer, J., Huey, R., Brown, T., Marshall, D., Chrysler, S., … & Chiang, D.P. (in press). Crash Warning Interface Metrics: Phase 3 Final Report. Washington, DC: National Highway Traffic Safety Administration.

Lin, B., & Green, P. (2013). Measurements of driver-assistance warning-signal characteristics in 2013 cars. (Report N.o. UMTRI-2013-03). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Singer, J., Lerner, N., Walrath, J., & Gill, M. (in press). Warning and message perception under ambient noise conditions. Washington, DC: National Highway Traffic Safety Administration.

Appendix A: Recruitment Ad

Volunteers section, Craigslist Washington DC (Maryland suburbs)

Title: Participants needed for Driving Safety Study (receive $75)

Compensation: $75 for a 90-minute session

Location: Rockville, MD

Westat is seeking participants for a federally-funded research study on drivers' ability to detect and recognize sounds and voice messages in different driving conditions.

If you participate in the study, you will take part in a 90-minute session at Westat in Rockville. You will wear headphones to listen to background sound recorded in a moving vehicle. You will hear occasional sounds and messages while you are driving and you will be asked to answer questions about what you hear. At the same time you are listening to the recordings, you will perform a computer task that involves some of the same demands as driving.

Sessions will take place on weekday mornings, afternoons, and early evenings. Occasional weekend sessions may be available.

To be eligible to participate:

- You must have had a valid U.S. driver's license for at least 2 years and no major driving violations in the past few years.
- You must drive a car on a regular basis
- You must be between 21 and 50 years old
- You must have normal hearing; hearing aid users or those with functional hearing loss are not eligible.

If you are interested in participating or would like to learn more about this study, please call Diane at [Redacted].

Appendix B: Participant Screener

Thank you for your interest in the Ambient Vehicle Noise Study. If you participate in this study, you will listen to audio recordings of on-road driving while providing feedback about various sounds that will be played over those recordings. You will also perform a computer task that involves some of the same demands as driving.

I have a few questions I need to ask to verify your eligibility. Your ability to participate will depend on your eligibility and our need for participants with a variety of characteristics.

1. Did you take part in a Westat study in August or September of last year where you drove on the ICC while occasionally hearing sounds and alerts? Yes___ (*ineligible*) No___
2. In what year were you born? ___________ (*1993-1964 eligible; ask for exact age if close*)
3. For how many years have you had a valid U.S. driver's license? ___________ *(2 years minimum)*
4. Has your license ever been suspended or revoked within the past five years __Yes __No
5. How many days per week do you typically drive? _______ *(3 minimum)*
6. Have you ever been diagnosed with a hearing impairment? __Yes __No
7. Do you have any reason to believe you have a hearing impairment? __Yes __No
8. Do you use a hearing aid? __Yes __No
9. Which statement best describes your hearing (without a hearing aid)?

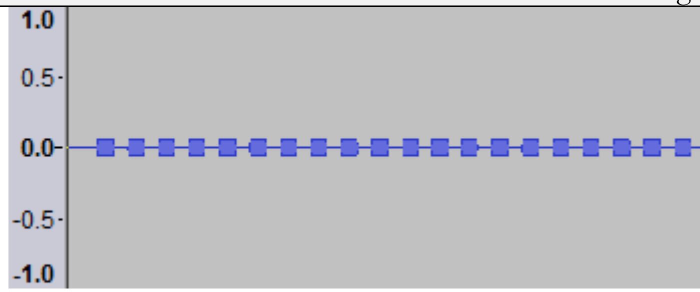
   ___ good ___a little trouble ___a lot of trouble
10. What times can you be available for a 90-minute session in Rockville?
    a. ___weekday mornings
    b. ___weekday afternoons
    c. ___weekday evenings
    d. ___weekend mornings
    e. ___weekend afternoons

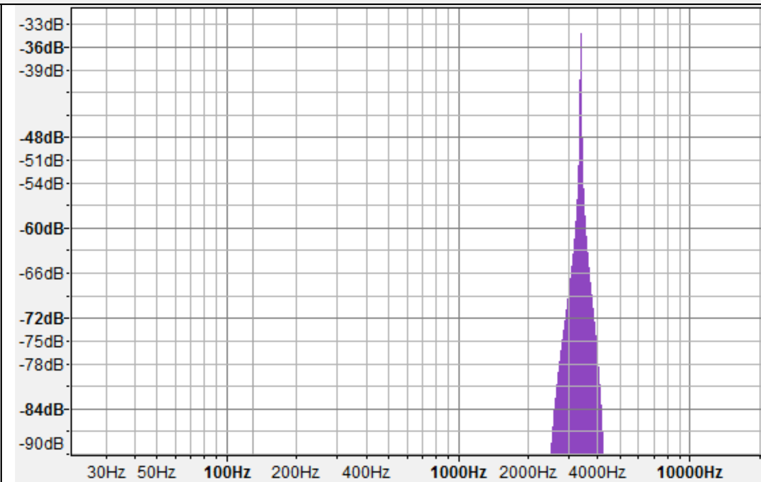
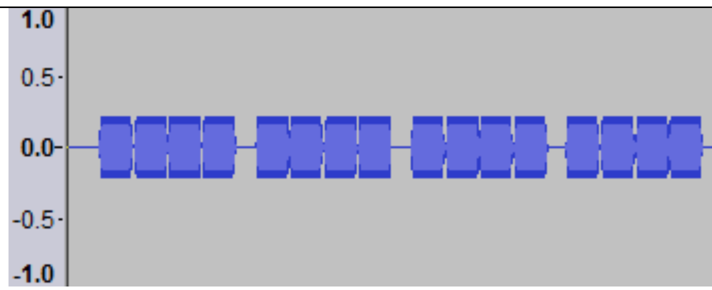
If eligible:

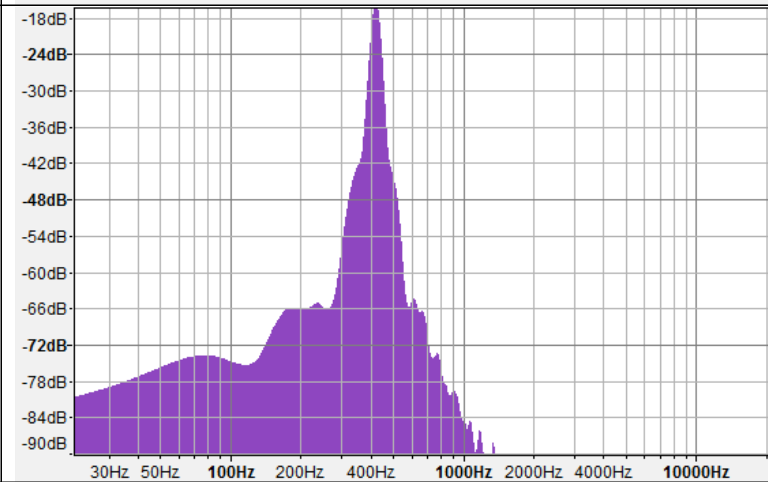
11. What is your full name? ___________________________
12. What is your daytime phone number? ________________________

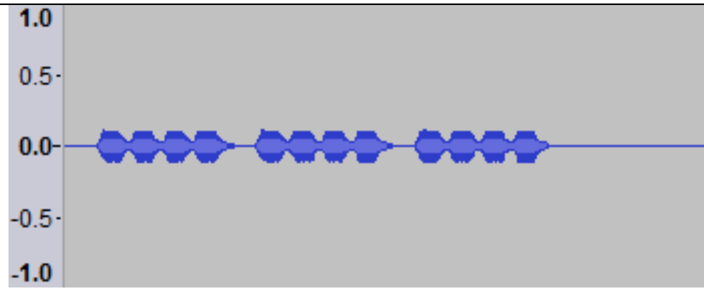
Is there an email address I can use to contact you about this study? _________________________

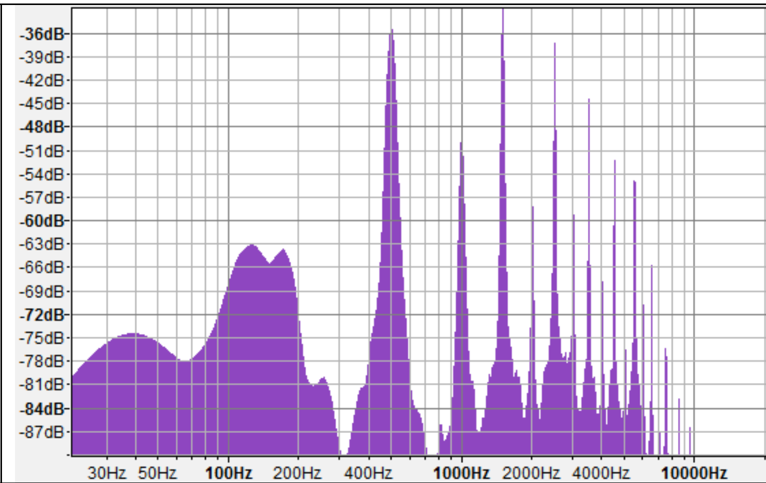
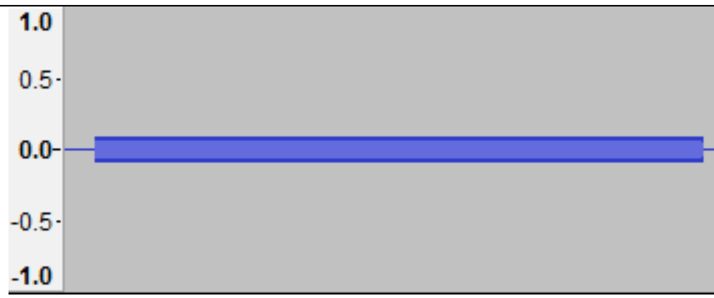
Appendix C: Sounds Used in Experiment 1

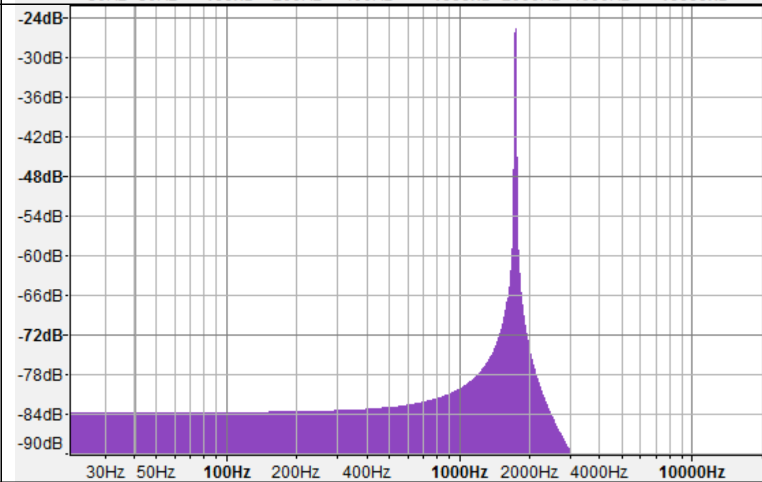
<table>
<tr><th>No.</th><th>Name</th><th>Amplitude waveform (2 s duration)</th><th>Frequency and intensity spectrograph (logarithmic)</th></tr>
<tr><td colspan="4">Tone Signals</td></tr>
<tr><td>1</td><td>FCW 1</td><td>

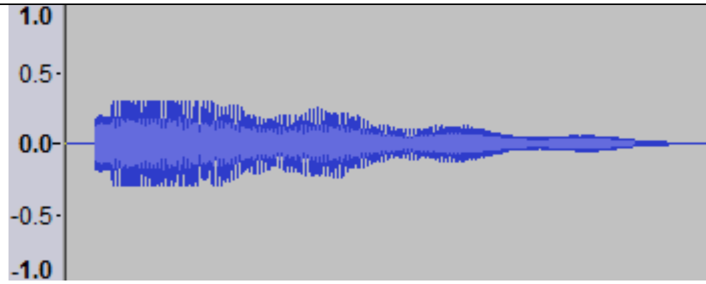
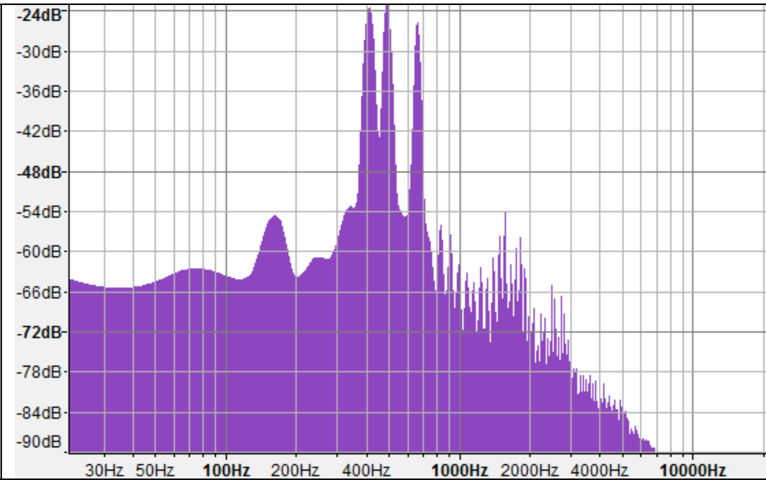
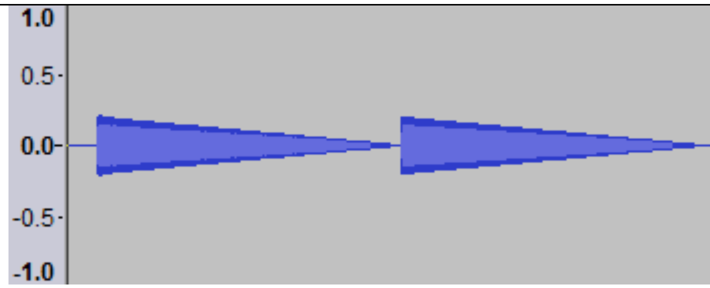
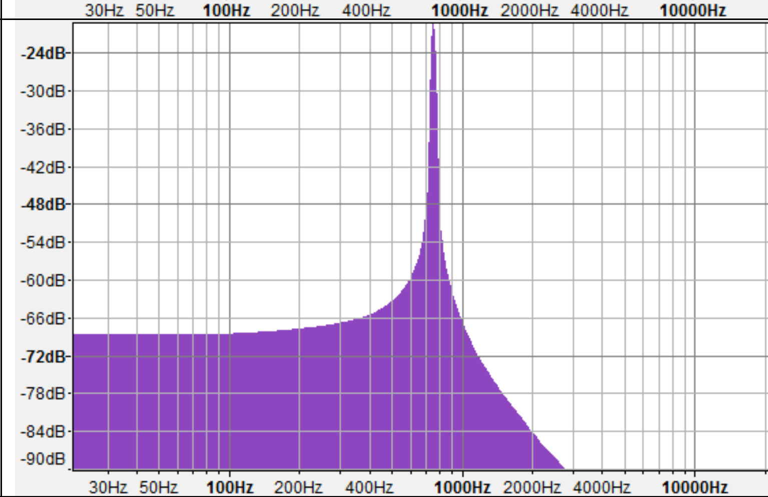
One burst of 20 fast beeps with a relatively high frequency profile.</td><td>
</td></tr>
<tr><td>2</td><td>FCW 2</td><td>

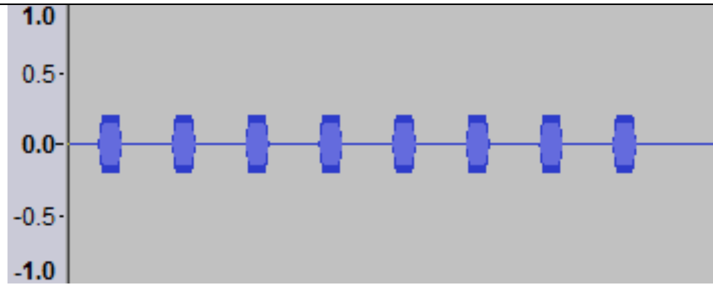
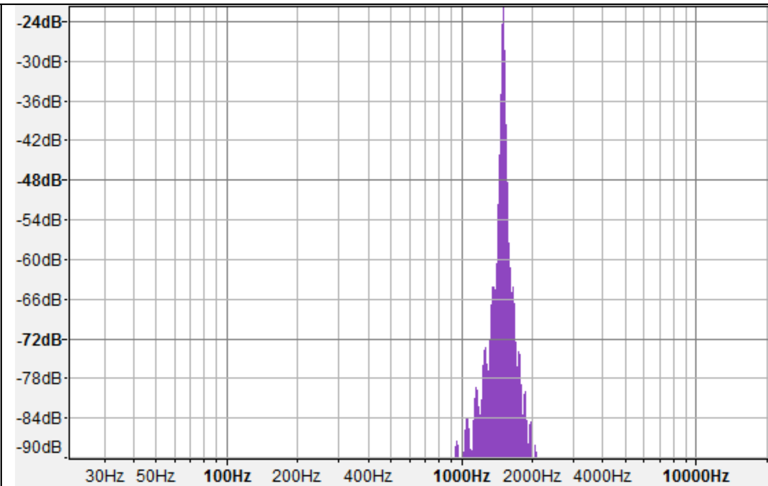
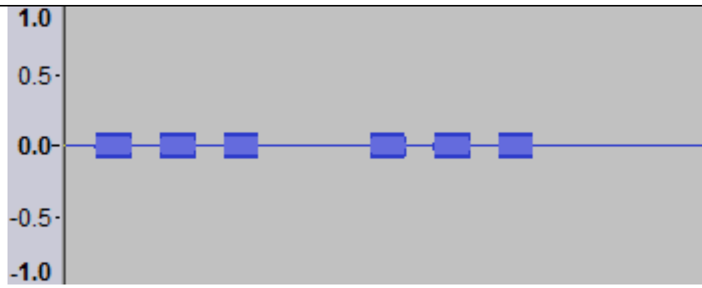
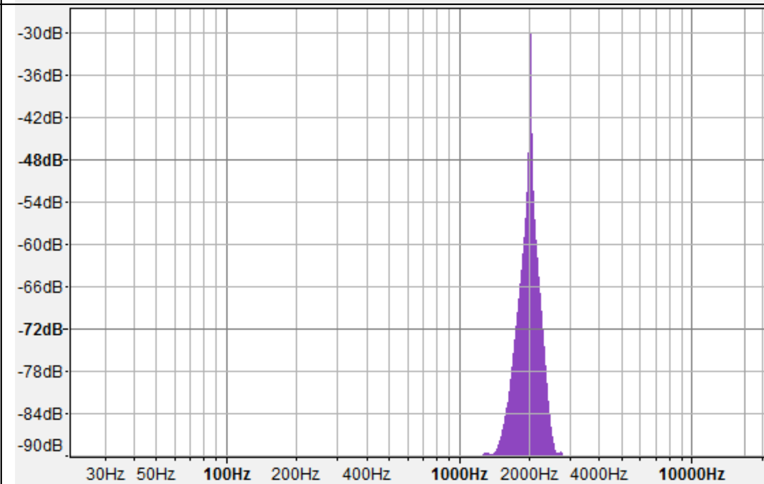
Four bursts of four fast beeps with a relatively low frequency profile.</td><td>
</td></tr>
</table>

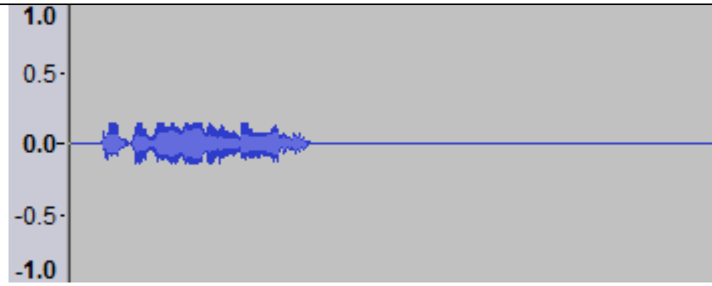
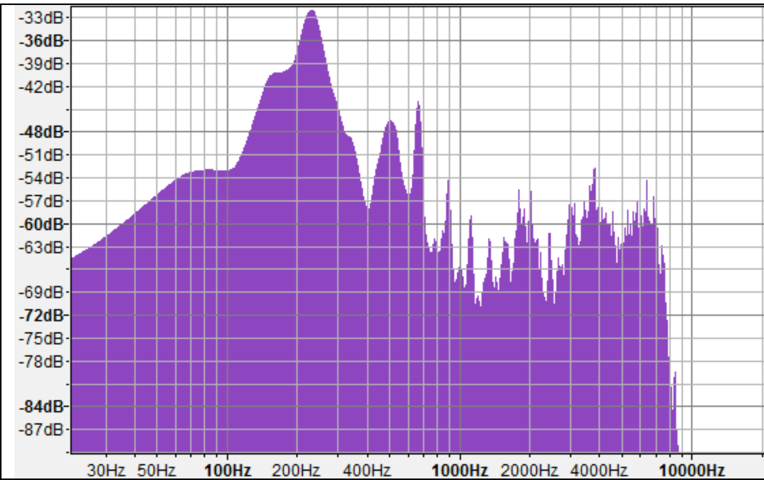
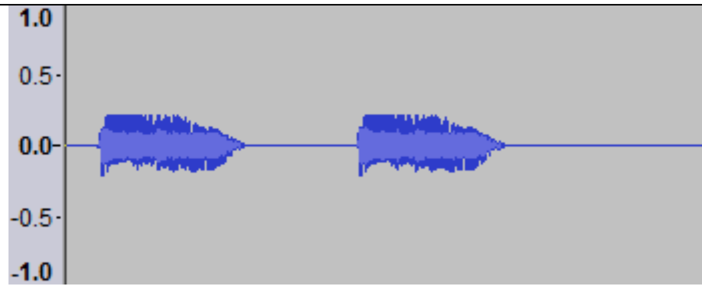
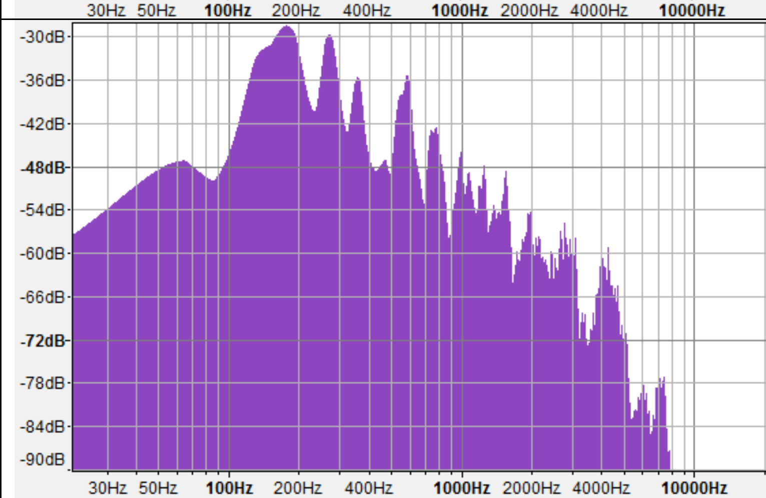
<table>
<tr><td>3</td><td>Blind spot warning</td><td>

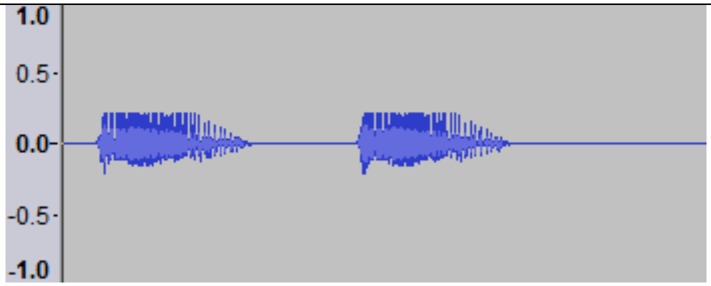
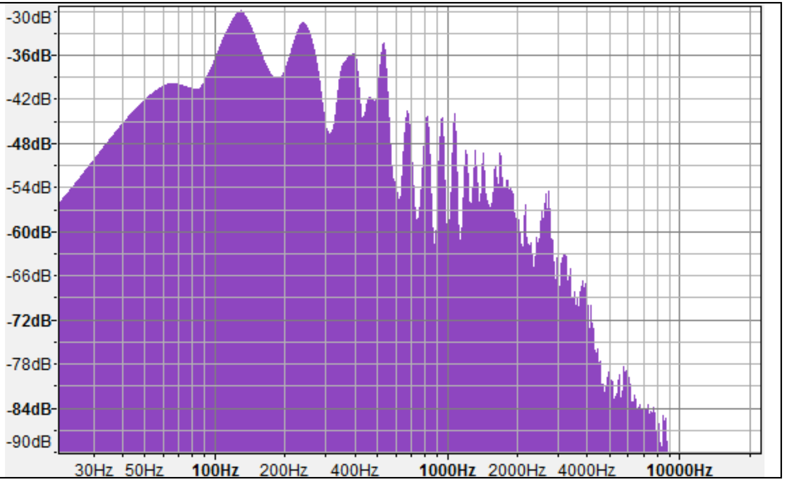
Three bursts of four fast beeps, each with a smoothed onset and decay and a sustained low intensity sound between beeps.</td><td>
</td></tr>
<tr><td>4</td><td>Pedestrian warning</td><td>

A constant beep with a duration of 2 seconds.</td><td>
</td></tr>
</table>

5	Seat belt alert 1	 A single chime that decays to silence in the span of about two seconds, with intensity varying in a wavelike pattern.	
6	Seat belt alert 2	 Two chimes, each of which decays to silence in the span of about one second.	

7	Park assist 1	 One burst of eight beeps.	
8	Park assist 2	 Two burst of three beeps.	
Voice signals			

9	Female voice – not urgent	 Female voice says "Attention."	
10	Female voice – urgent	 Female voice says "Warning, warning."	

11	Male voice – urgent	 Male voice says “Warning, warning.”	

Appendix D: Instructions to Participants

Instruction and Practice

Purpose and Procedure: This is a study about how people hear sounds and messages while they are driving. Some new vehicles can use sounds or voice messages to inform drivers about safety-related issues, the status of their vehicle, traffic conditions, incoming calls, and many other things. One important question is how well drivers can perceive these sounds in realistic driving conditions. In noisy conditions, it might be harder to hear and understand sounds and messages.

Today, I am going to ask you to listen to previously recorded sounds from a moving vehicle. The noise conditions under which the recordings were taken will vary. Every so often, I will present a sound over the recording of the moving car. Your job will be to let me know as soon as you hear the sound, and then make ratings about what you hear. As you are listening to these recordings and sounds, you will also be engaged in a computer task that is meant to simulate some of the basic aspects of driving a car.

Adjustments and calibration: Before we get started, please silence your cell phone. You can also adjust the seat to get comfortable. You should be seated so that it is comfortable to use the steering wheel and gas pedal. [*wait for participant to make adjustments*] Are you comfortable with your seat position?

Task familiarization: On the screen in front of you, there is a very basic driving simulation. You are driving the yellow car and you can see your yellow car hood in front of you. Your job is to follow the blue car at a constant distance, and to stay within your lane. As you press the gas pedal, you will speed up and get closer to the blue car. As you let off the gas pedal, you will slow down and get farther away from the blue car. Your job will be to keep the blue car's tires in line with the gray lines that you see just outside of the lane lines *[point/clarify as needed]*. Note that you will use only the gas pedal to control your speed; the brake pedal has no effect on your speed. You will also use the steering wheel to stay between the black lane lines. If you get too close to a lane line, a red square will appear to tell you to return to the center of the lane. Are you ready to practice the task? I'll let you know when to begin the driving task. *[Start the simulator task, and let the participant practice for one minute, or longer if necessary to get good at the task. Observe performance and provide feedback as necessary, make sure they are not trying to use brake pedal, then pause the driving simulator and ask if they have any questions.*

Now let's practice for just a bit longer, but this time you'll be listening to background noise that was recorded from the driver's position in a car with accurate stereo reproduction. Go ahead and put on the headphones now. Make sure that you put them on so the cord is on your left side. If they feel too big or too small I can show you how to adjust them. <*Make sure participant doesn't have hair under the ear pads.*> In a moment, you will hear the driving background noise, then the driving task will start.

Start the baseline ambient noise loop, start the driving task, and let the participant practice for 30 seconds. Observe performance and provide feedback as necessary, make sure they are not trying to use brake pedal, then ask if they have any questions.

You can take the headphones off now. Please put them down on the table in front of you with the left ear cup on the left and the right ear cup on the right. This will help to make sure that you always put the headphones back on the right way.

During today's session, you will do the driving task while hearing different background noise recordings. While you do this, you will occasionally hear other sounds or alerts. Let's go over what you will do when you hear a sound or voice message. When you hear a message, the first thing you have to do is click this little button [*give finger button to participant*]. That lets us know how quickly you recognized that there was a sound. You can attach it to your finger so you can click it easily without looking at it. *Attach the microswitch and have them operate it; have them adjust it so that they can quickly and comfortably operate the switch but where it will not likely be accidentally activated*] Once you push the button, I will ask you some questions about the sound. You can take your time with these answers. I'll play a practice sound for you, and then we will go through the ratings you will make about that sound. Go ahead and put the headphones on now. From this point forward you will be wearing the headphones, so please let me know if they get uncomfortable and you want a break. [*play kazoo practice sound*]

The first question I will ask you is "how NOTICEABLE was the sound?" Noticeability means that the sound is easily noticeable among other sounds and noises in the vehicle. You will rate the sound you just heard on a scale from one to seven. A "one" means that the sound is not very noticeable. A "seven" means that the sound is extremely noticeable. How would you rate this sound?

The next question I will ask you is "how URGENT was the sound?" Urgency means that the sound conveys a sense of importance, motivating you to make an immediate response. A "one" means that the sound is not very urgent. A "seven" means that the sound is extremely urgent. How would you rate the urgency of the sound you just heard?

Next, I will read you a list of four possible meanings for this sound. Choose the one the most closely matches the meaning conveyed by this sound. I'll read you the list of possible meanings, then I'll go back and explain what each one means. The options will be:

- <u>Urgent crash warning</u>… means that there is a situation in which you must react immediately to avoid a crash. For example, imagine you are about to hit a pedestrian or about to run off the road.
- <u>Safety information</u>… means that there is a safety issue that you need to pay attention to, but you are not in immediate danger of a crash. For example, imagine that you are approaching a work zone where two lanes are closed or there are reports of icy roads ahead.
- <u>Information not related to safety</u>… means exactly what it says – you are receiving information, but the information is not safety-related. This could include various types of information, such as traffic congestion several miles ahead, prices at nearby gas stations, or a navigation system telling you to make the next turn.
- <u>Incoming personal communication</u>… means that you are receiving an incoming call, text message, email, or other direct communication.

Any questions? Which meaning would you choose for the sound you just heard? [*record answer*] The list of options will be the same for all of the sounds you hear today. I'll read the list to you for each sound you hear. If you can't remember what a category means, let me know and I can try to clarify. Also, please remember that there isn't necessarily a correct or incorrect answer to this question – I want to know what the sound conveys to <u>you</u>.

Now let's try another sound for practice. [*play voice message; go through NOTICEABILITY and URGENCY; read full definitions again and indicate 1-7 scale*] Now the next rating that comes up is INTELLIGIBILITY. You did not make this rating before. That is because it will only come up when the sound is a voice message. "Intelligibility" means that the spoken words can be easily understood. A "one" means that the voice message was not very intelligible. In other words, you could not understand the words clearly. A "seven" means that the message was extremely intelligible. How would you rate this voice message for intelligibility? [*have participant say choice; go through meaning question; read full definitions again*] Do you have any questions about how to do the ratings and choices?

In a moment you will hear the background noise and start the driving task. I will pause the background noise to ask you questions about the sounds during the session so you can hear me clearly, but please continue to do the driving task while you answer questions. Would you like to make any more adjustments before we start? Any questions? Just one more thing – if I forget to turn the background noise back on after you finish answering questions, please let me know.

[*Select the correct recording on the experimenter computer and begin the secondary task simultaneously*]

Data Collection

- *Pause the ambient noise loop after participant clicks microswitch AND the alert finishes playing*
- *Click button quietly and avoid giving any subtle triggering cues*
- *If participant fails to hear a sound, you can trigger the next one without waiting for the countdown*
- *Try to be silent at all times*
- *Keep an eye on participant's driving behavior*
- *In between blocks, ask participants if they need a quick break*

Prior to Condition 1 (baseline): For the next set of sounds, you will be listening to a recording made with the car windows closed and no music playing.

Prior to Condition 2 (windows down): For the next set of sounds, you will be listening to a recording of the front two windows opened all the way during a drive.

Prior to Condition 3 (music on): For the next set of sounds, you will be listening to a song playing over the car speakers during a drive.

**[If the participant presses the button when there is no actual signal:]*

If this happens during a non-trial period, ask the participant what sound they heard, then record on paper as accidental or false alarm. If this happens during the 5 s pre-signal period of a trial, ask if they heard something or if it was an accidental button press. Then follow program prompts to redo the trial

Debrief

- When the participant has completed the session, pause the secondary task and the car recordings
- Any comments, questions?
- Pay participant, have them sign receipt and provide the yellow copy
- Guide the participant back to lobby

Appendix E: Experiment 1 Mean Noticeability and Urgency Ratings

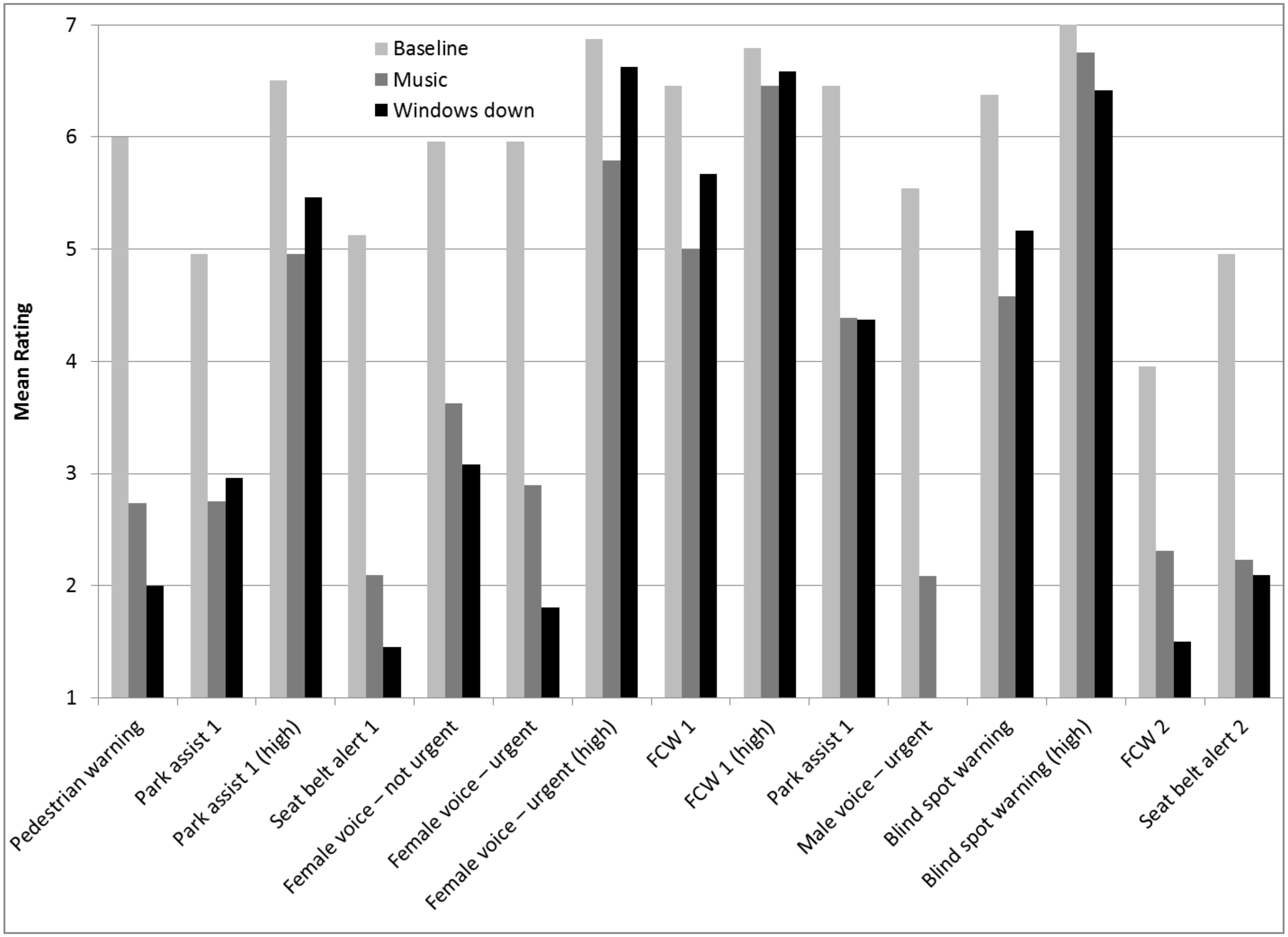


Figure E-1. Mean noticeability rating for each stimulus under each ambient noise condition

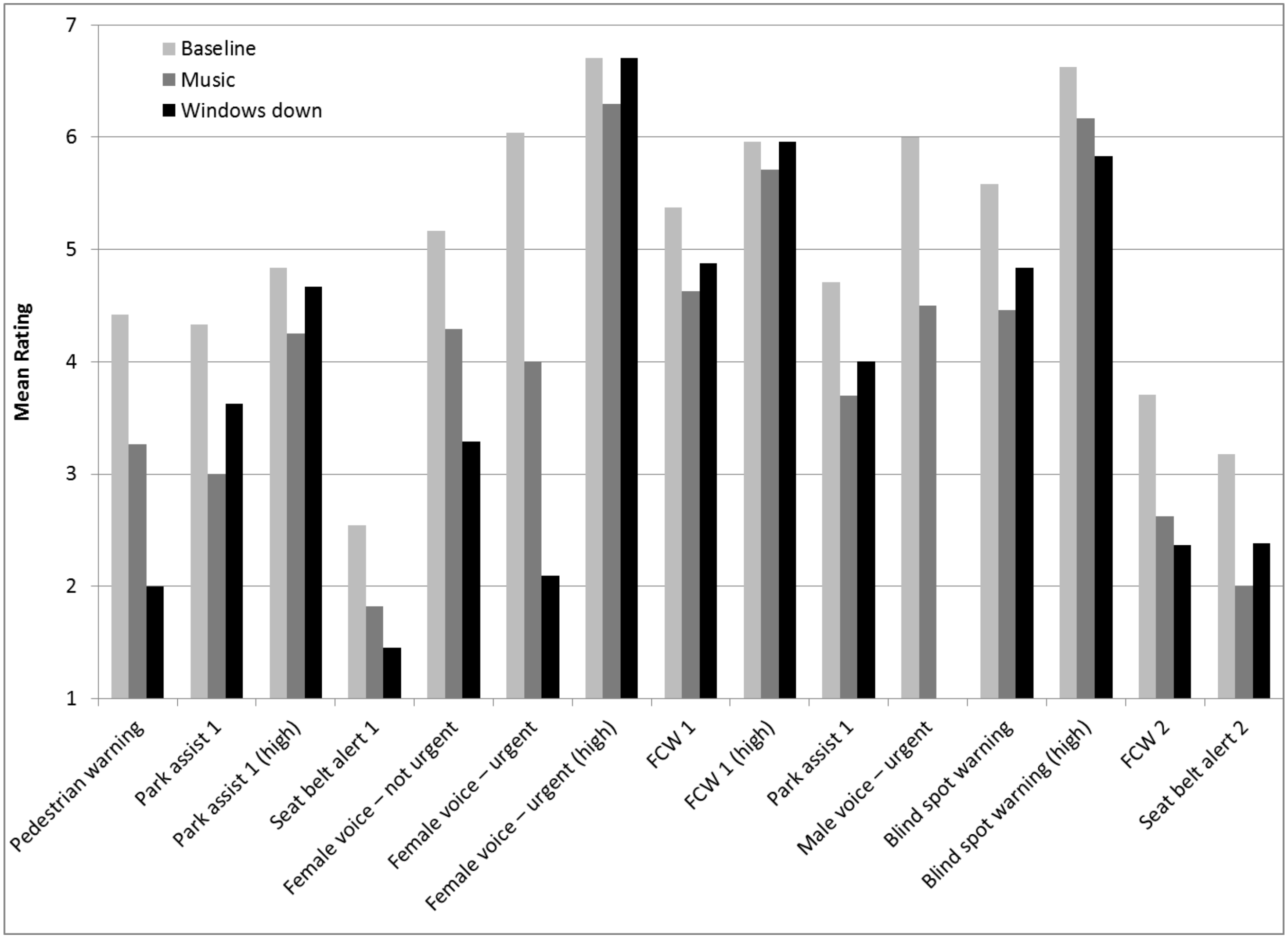
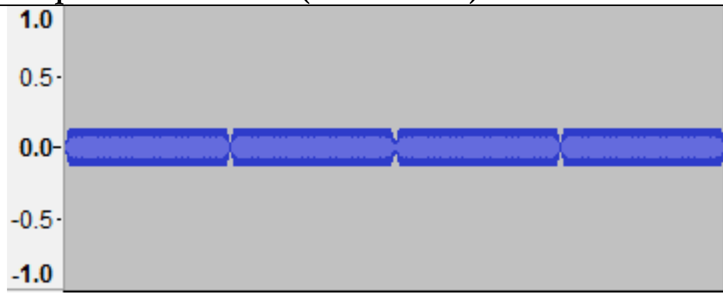
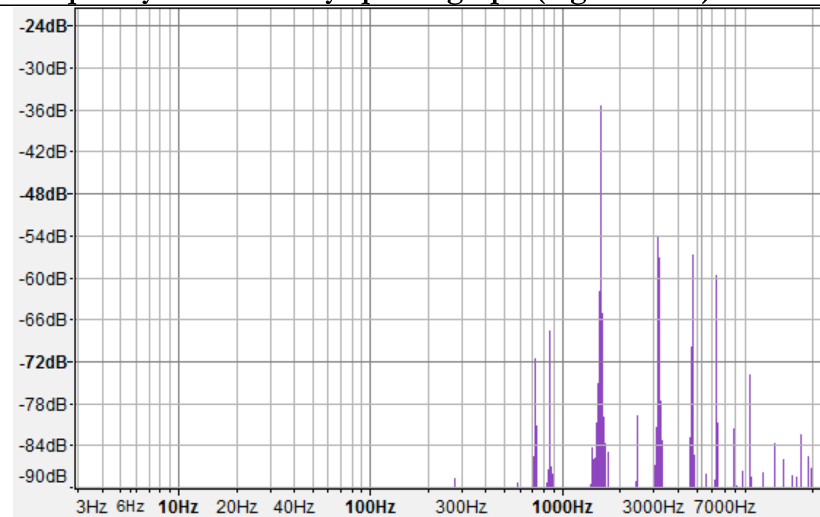
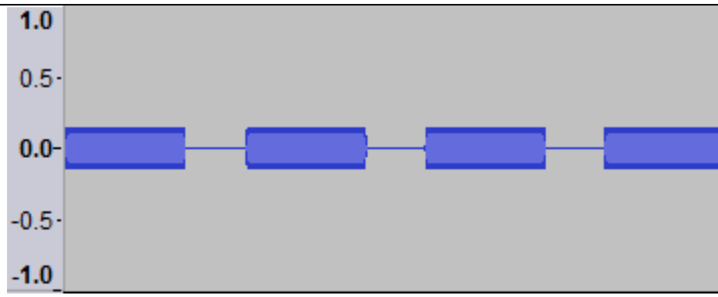
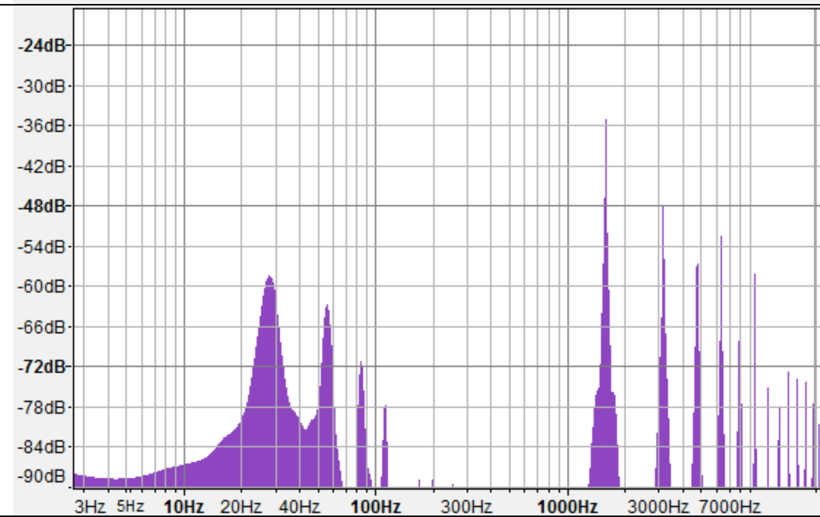
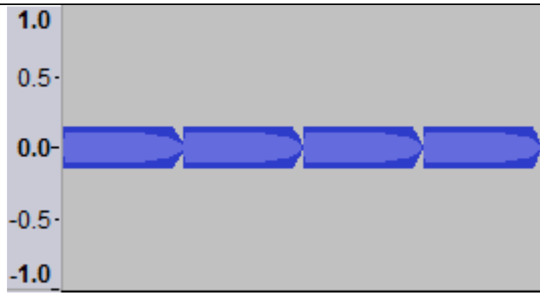
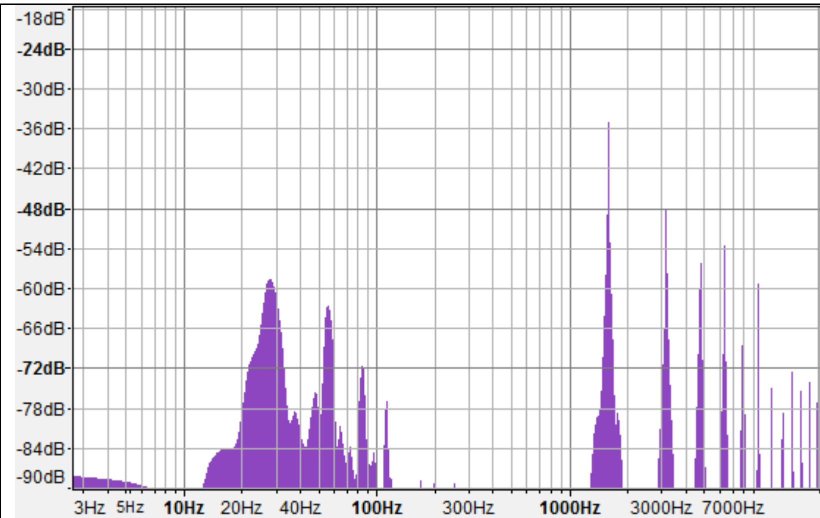
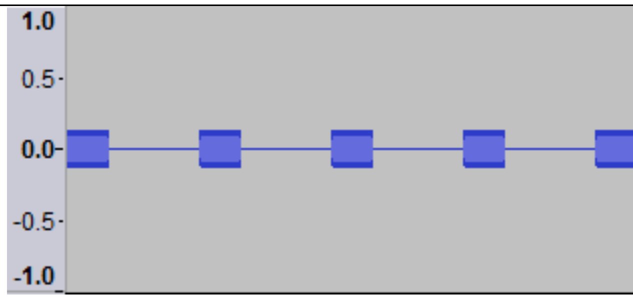
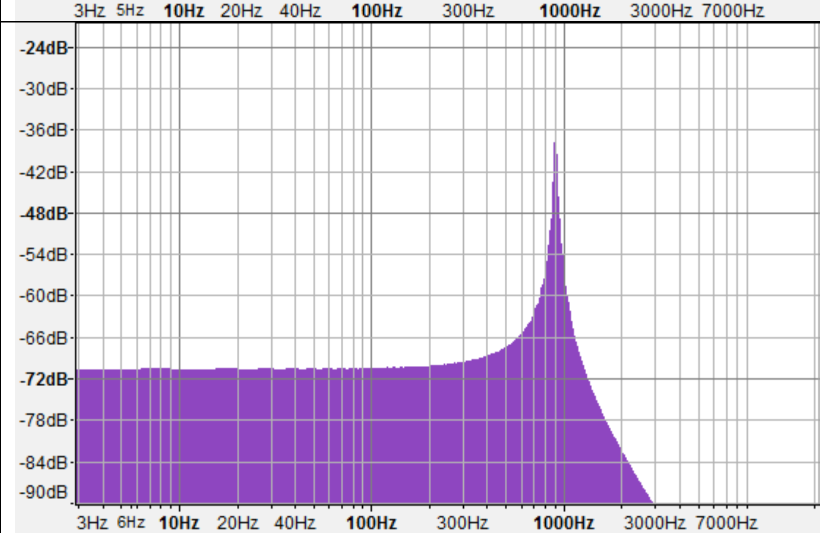
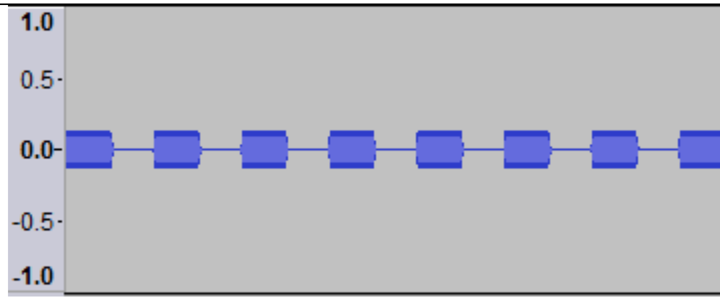
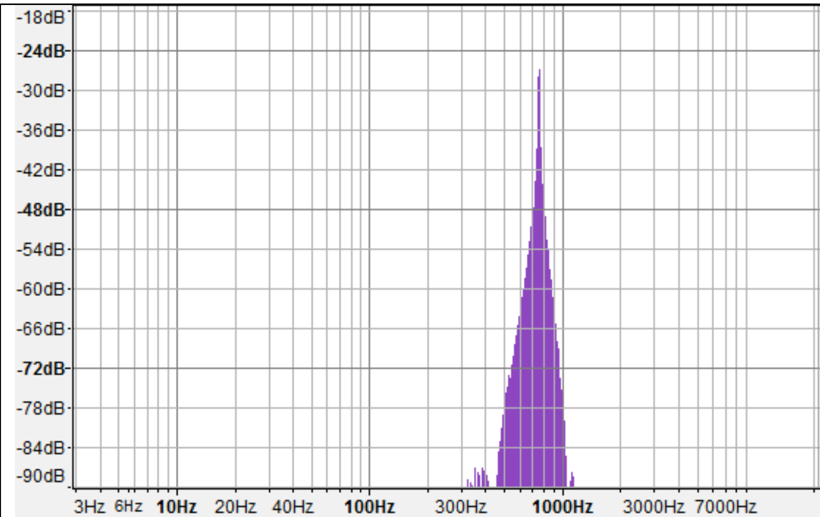


Figure E-2. Mean urgency rating for each stimulus under each ambient noise condition

Appendix F: Additional Sounds Used in Experiment 2

Name	Amplitude waveform (2 s duration)	Frequency and intensity spectrograph (logarithmic)
GMU 1	 Meets all four warning signal criteria.	
GMU 4	 Meets three warning signal criteria (not interburst interval).	

GMU 5	 Meets three warning signal criteria (not peak-to-total time ratio).	
GMU A	 Meets ratio criterion only.	

GMU B	 Meets ratio and IBI criteria.	

DOT HS 812 207
November 2015

U.S. Department
of Transportation
**National Highway
Traffic Safety
Administration**

www.nhtsa.gov

11708-110515-v2a