A19 / PROCESS FOR SELECTING THE SOUNDS USED FOR CRASH ALERTS IN THE THREE DRIVER INTERFACE STUDIES

Auditory Alert Development

A total of 18 sounds were tested as candidates for an auditory crash alert, which are listed and briefly described in Table 8. The 18 sounds were from five categories:

(1) Standard production vehicle chime
(2) The five top-rated sounds (all Non-Speech) evaluated by Tan and Lerner (1995)
(3) Production-oriented non-speech sounds
(4) Speech message alerts
(5) Non-speech sounds developed by the General Motors Noise and Vibration laboratory

The various sounds within each of these last 4 categories will now be discussed in turn. The five Category (2) sounds evaluated were composed of the top 5 (of the 26) rated sounds evaluated in the Tan and Lerner (1995) laboratory study. In this previous study, participants were asked to rate sounds on various attributes including annoyance, appropriateness, discretion, startle, and urgency. (A modified version of this procedure was employed here.) The mean rating on each attribute was then weighted according to “expert” rankings of the importance of each attribute to an auditory crash alert. The five sounds included in the present study received the five highest total weighted scores. These top sounds were all non-speech sounds, which received higher total weighted ratings than any of the ear con (car horn, tire skid) and speech sounds examined. Unlike the current study, the sounds evaluated in the Tan and Lerner study were examined for their merit as a “master” auditory crash alert, which was intended to precede a subsequent alert indicating direction of threat (e.g., forward).

The seven Category (3) sounds evaluated were modified standard production chimes. These modified chimes had frequencies of 750 Hz, 2000 Hz, or both. In general, the attack/decay characteristics and the cadence of the production chimes were modified to create warning-like sounds (e.g., ambulance, and alarm clock).

The three Category (4) sounds evaluated were the speech alerts “danger”, “warning”, and “look out”. To create these candidates, a male professional broadcaster repeated these warnings three times in sequence. Reverberation was added to the recording of each alert.

The two Category (5) sounds evaluated developed by the General Motors Noise and Vibration laboratory specifically for this test. These sound candidates were created by mixing pulses at frequencies of 2000 Hz and 2500 Hz. The two sounds were identical except that one had a faster cadence.

All 18 sounds were digitized with the assistance of the General Motors Noise and Vibration Center. With the exception of the Category (4) sounds, each of the sounds were 2.10 seconds in length. The category (4) speech alert sounds “danger”, “warning”, and “look out”, were 2.60, 2.49, and 2.42 seconds in length, respectively.

Loudness Adjustment Procedure

A staircase threshold procedure was conducted to attempt to equate the sounds for subjective loudness, so that sounds could be subsequently evaluated for their “pure” crash alert properties independent of subjective loudness. Previous work has indicated that subjective loudness is highly correlated with crash alert properties (e.g., a louder sound is perceived as more urgent) (Tan and Lerner, 1995). The loudness adjustment procedure involved comparing each candidate sound to the standard production chime and judging whether the candidate sound was louder or softer than the standard chime. On each presentation of a sound pair, the loudness of the candidate sound was adjusted one decibel until the rater’s response changed. The initial direction of the decibel change, increasing or decreasing, was randomly varied across the candidate sounds. Once the rater’s response changed, the direction of the loudness adjustment was reversed. This adjustment sequence continued until five response changes occurred. The decibel level of the last four response changes was averaged for each candidate sound. This average represented the decibel level at which the rater judged the loudness of the candidate sound to be equal to that of the standard chime. The loudness adjustment procedure was used with four raters (2 females, average age 30; 2 males, average age 40). The mean of the four raters’ average decibel levels for each sound was then used to compute the decibel adjustment. The decibel adjustments for the candidate sounds were as follows: #4, -6 dBa; #5, -7 dBa; #6, -4 dBa; #7, -6 dBa; #8, -4 dBa; #12, -1 dBa; #19, 0 dBa; #20, -1 dBa; #21, -1 dBa; #22, -4 dBa; #24, -8 dBa; #25, -10 dBa; #26, -10 dBa; #27, -7 dBa; #28, -10 dBa; #29, -2 dBa; and #30,-2 dBa.

Table 8 Brief Description of Collision Alert Sound Candidates

Sound #	Description
	Standard Production Chime
1	2000 Hz production chime, cadence 3.3 per second
	The five top-rated sounds (all Non-Speech) evaluated by Tan and Lerner (1995)
4	Stimuli 1 (low fuel warning)
5	Stimuli 4 (high-pitched, ambulance-like siren)
6	Stimuli 5 (low-pitched, ambulance-like siren)
7	Stimuli 8 (2500 & 7500 Hz 100 ms broad pulse of 110 ms each, repeated at 8 ms intervals, pause of 110 ms)
8	Stimuli 10 (2500 & 2650 Hz peaks, temporally similar to preceding sound)
	Production Oriented Non-Speech Sounds
12	2000 Hz production chime, pulse=7.5 ms attack followed by 142.5 ms decay, cadence 3.3 per second
19	Same as sound 12, using 750 Hz zone
20	Sounds 12 & 19 together (2000 & 750 Hz)
21	Beep 4H33, 2000 Hz, cadence 100 (3.3 sec)
22	2000 Hz production chime, pulse=7.5 ms attack followed by 142.5 ms decay, 4 pulse sequence separated by 110 ms silent pause
27	2000 & 750 Hz production chime overlaid, cadence 3.3 per sec
28	750 & 2000 Hz chimes, alternating (ambulance-like siren)
	Speech Message Alerts
24	"Danger, danger, danger"
25	"Look out, look out, look out"
26	"Warning, warning, warning"
	GM Noise and Vibration Laboratory
29	2000 & 2500 Hz triangular wave tones overlaid
30	Same as sound 29, faster cadence

Sound Evaluation Ratings

Ten DAT recordings of the 18 candidate sounds were created for the sound evaluations. A different random order of the candidate sounds was used for each recording. The interior sound of a 1997 Ford Taurus SHO traveling on dry, smooth pavement at 70 mph was used as background noise for the recordings. The background noise was presented continuously on each recording. The candidate sounds were presented at 12-second intervals "on top of" (or overlaid upon) the background noise.

After listening to verbal instructions (which are described below), participants were asked to rate each sound on the 13 statements shown in Table 9. The order of the statements shown in this table corresponds to the order in which the participants experienced the statements. Participants provided their general opinion of each sound by rating the sounds on the statement, "this sound is a good choice for a collision warning sound." The participants rated each sound on this general opinion statement twice, initially on the first trial (Statement 1) and then again on the second from last trial (Statement 12). The practice statement, "this sound is very musical", was used to acquaint participants to the sounds and the sound rating procedure.

Eleven of the 12 remaining statements were related to attributes considered critical for an effective warning sound. These attributes were notability, confusability, attention-getting, startle, interference, annoyance, appropriateness, emergency, and loudness. With the exception of the annoyance and interference attributes, each attribute was addressed by one corresponding statement. For the interference attribute, one statement asked whether the sound would interfere with the driver's ability to decide on an emergency driving action (Statement 6). Another related statement asked whether the sound would interfere with the driver's ability to perform an emergency driving action (Statement 7). For the annoyance attribute, one statement asked whether the sound would annoy the driver if the alert came on when no driving action was required once a day (Statement 8). Another related statement asked whether the sound would annoy the driver if the alert came on when no driving action was required once a week (Statement 9). One critical difference between the current study and the Tan and Lerner (1995) study which should be stressed is that drivers in the latter study were told to assume "minimal" false alarms, where minimal was left undefined. It is quite possible that the Tan and Lerner participants idea of "minimal" corresponded to a false alarm (or nuisance alert) frequency of substantially less than once a week.

Table 9 Rating Scale and Statements Used for Sound Ratings

-3	-2	-1	0	1	2	3
Strongly Disagree	Moderately Disagree	Perhaps Disagree	Neutral	Perhaps Agree	Moderately Agree	Strongly Agree

Practice: This sound is very musical.

1. This sound is a good choice for a collision warning sound.
2. This sound would clearly stand out and be noticeable among the other noises inside and outside the vehicle such as engine noise, the fan blowing, talking and music on the radio, horns, and sirens. (Notability)
3. This sound would be confused with other sounds inside and outside the vehicle such as engine noise, talking and music on the radio, horns, sirens, car phones, or other electronic devices. (Confusability)
4. This sound would get my attention immediately. (Attention-getting)
5. This sound would startle me, that is, cause me to blink, jump, or make a rapid reflex-like movement. (Startle)
6. This sound would NOT interfere with my ability to make a quick and accurate decision about the safest driving action to take. (Interference)
7. This sound would NOT interfere with my ability to quickly and accurately perform an emergency driving action. (Interference)
8. This sound would annoy me if it came on once a day in a situation where NO driving action was required. (Annoyance)
9. This sound would annoy me if it came on once a week in a situation where NO driving action was required. (Annoyance)
10. This sound would appear out of place as a warning in a car or truck. (Appropriateness)
11. This sound would clearly tell me that I'm in danger and I need to react immediately. (Emergency)
12. This sound is a good choice for a collision warning sound.
13. This sound seemed louder than the other sounds in the test. (Loudness)

At the start of the session, the experimenter told participants that the evaluation was part of the selection process for a collision warning sound. The text of the verbal instructions are shown on the last page of this Appendix. As a means of explaining the context and requirements of the warning sound, participants were asked to recall their experiences from CAMP Study 1 in which they had to brake hard at the last second possible to avoid colliding with a lead (surrogate) vehicle. They then were told to imagine that they were driving on a real road and to suppose that they were distracted or not paying attention to their driving. Further, when their vehicle rapidly approached a slower or stopped vehicle, the collision warning would sound to alert them to the situation. The instructions stated that once the warning sounded, a driver would have to decide upon the appropriate driving action to take. If braking was appropriate, they were told they would have to use hard braking as in the previous study. Participants were told that this depiction demonstrated that the warning sound must get the driver's attention while allowing the driver to respond appropriately. The possibility of false alarms, or instances when the warning may sound in response to non-threatening events (such as a guardrail on a sharp curve) was then described to explain that the warning sound needed to be attention-getting without being overly annoying.

The participants were then instructed that they would be listening to the candidate warning sounds and rating the extent to which they agreed with various statements made about the sounds. Participants were asked to rate the extent to which they agreed with each statement for each sound using a 7-point scale which ranged from strongly agree (3) to strongly disagree (-3), shown in the top portion of Table 9. The attributes related to each statement (which were not shown to the subjects) are shown in parentheses. The statements are listed in Table 9in the order they were presented during each evaluation session. Participants were instructed to circle a number on the scale to reflect their agreement with the statement for that sound. For example, using the practice statement "this sound is very musical," participants were told that they should circle the response on the scale that reflected the extent to which they agreed that the sound was very musical. After the practice trial, participants were encouraged to ask questions about the procedure and rating scale.

At the beginning of each trial, subjects would hear the experimenter read the statement aloud to the group. The participants then listened to each 18 candidate alert sounds examined (presented in a random order) and rated each sound on the statement. Between each sound presentation, subjects were provided ample time to make sound ratings. All the sounds were rated on a statement before the next statement was introduced. The 13 statements were presented in the order shown in Table 9. Thus, subjects rated each of the 18 sounds 13 times for a total of 234 sound rating trials.

Fifteen females and 20 males participated in the evaluation of the alert sounds. All of these individuals had previously participated in CAMP Study 1, in which they were asked to make last-second hard braking judgements while approaching the slowing or stopped CAMP surrogate (lead vehicle) target. The mean age of the participants was 49 years old (standard deviation=16 years). Participants were either in their 20s, 40s or 60s, which corresponds to the three age groups tested in CAMP Study 1. The 20s group consisted of 8 males and 1 female, the 40s group consisted of 5 males and 6 females, and the 60s group consisted of 7 males and 8 females. Eight individuals from each gender by age category were originally recruited. However, thirteen individuals (including 7 young females) did not appear for testing. All participants reported normal hearing ability. Participants received $35 for completing the 75-minute testing session.

The evaluation sessions were conducted with small groups of one to six participants, depending on participant turnout. Participants were seated in a conference room with their backs to a large table. The seating arrangement prevented participants from viewing each other's facial reactions to the sounds. The sounds were presented using a DAT player, amplifier, and headphones. Participants provided written responses to the statement ratings using clipboards.

The mean agreement rating for each candidate alert sound on each of the 13 statements is shown in Table 10. On Statement 1, which asked participants whether they agreed that a sound was a good choice for a collision alert sound, all of the sounds had a mean rating between +1 and -1. Thus, overall, none of the sounds were strongly favored on the first trial by the participants in general. The sounds which received mean ratings greater than zero, in order of the highest rating, were #7, #30, #26, #6, #8, #29, #24, #4, and #21. These sounds had mean ratings ranging between +0.09 and +0.51. On Statement 12, participants were again asked whether each sound would be a good choice for a collision warning sound. The results for this question differed from those from Statement 1. Three sounds had mean ratings greater than positive one. These sounds were #26, #24, and #25, which correspond to each of the three speech alert sounds examined. Only three other sounds (#8, #30, and #6) had positive mean ratings on this statement. It should be noted that, with respect to interpreting the absolute (as opposed to the relative) ratings provided on the 7-point scale employed, a general preference for speech alerts may have penalized the ratings for all non-speech alerts. That is, if speech alerts had not been included in the set of sounds examined, it seems quite likely that the non-speech sounds would have received higher absolute ratings on the rating scale provided.

Results from Statement 12 are considered the most informative for two primary reasons. First, by the time they completed this statement, participants had been "educated" about the desirable attributes of a collision warning. Second, by this time, participants had heard each sound 12 times, which gave them additional sound experiences to make relative comparisons between alerts, and gave them a chance to determine which alerts still "stuck out" as having alerting qualities during this somewhat lengthy, monotonous rating task. Table 11 lists the sounds in rank order according to the mean ratings on Statement 1 and Statement 12. Three sounds, #26 ("warning", "warning", "warning"), #8 and #30, were in the top five rankings as good choices for a warning at both the beginning and the end of the evaluation.

There are two striking differences between these findings and those reported in Tan and Lerner (1985). First, the top-rated sound from the Tan and Lerner (1995) study, an off-the-shelf low fuel aircraft warning (#4 in this study), fell in the middle of the pack of the sounds rated, and was rated particularly poorly on the annoyance and interference statements. This difference in studies is undoubtedly due to the difference in assumptions provided to raters across studies with respect to nuisance alert frequency. (However, overall, it should be noted that the 5 non-speech sounds carried over from the Tan and Lerner (1985) study performed quite well relative to the 18 sounds examined.) Second, the speech alert sounds in this study were rated substantially higher than the male and female synthesized and digital speech alerts examined in the Tan and Lerner study (none of which were among the top five highest total ratings in this previous study) . This is unlikely due to the relatively minor procedural differences between studies, but instead, in all likelihood is due to differences across studies in the specific nature of the speech stimuli employed.

The last statement asked participants whether a sound seemed louder than the other sounds in the evaluation. The mean ratings for the candidate sounds on the loudness statement ranged from +1.69 to -1.97. Thus, even though the decibel levels of the sounds had been adjusted in an attempt to equate them for subjective loudness prior to the evaluations, participants still reported that some sounds appeared louder or softer than others. A scatter plot, shown in Figure 1, shows the relationship between perceived loudness and participants' final rating of the candidate sounds (Statement 12). The plot incorporates a regression line, which describes the final rating for the candidate sound as a function of the loudness rating. In general, sounds located above the line were rated more highly as a choice for a collision alert than would be expected if the rating was based solely on the perceived loudness of the sound. Conversely, sounds located below the line were rated more poorly as a choice for a collision alert than would be expected if the rating was based solely on the perceived loudness of the sound. The observation that appears most striking in this scatter plot is participants' preference for speech alerts (i.e., #24, #25, and #26).

Table 10 Mean Agreement Rating for Each Candidate Crash Alert Sound Across Each of the Thirteen Sound Rating Statements

	Sound Rating Statement Number												
Sound Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1	-0.94	-0.76	1.17	-0.51	-1.53	1.30	0.64	0.75	0.45	0.58	-1.27	-1.74	-0.40
4	0.10	1.74	-0.81	2.40	1.93	-0.97	-0.79	2.51	1.97	0.87	2.03	-0.18	1.69
5	-0.76	0.43	1.10	0.57	-0.65	1.03	0.90	1.06	0.63	0.77	0.27	-1.03	-0.14
6	0.37	1.26	0.26	1.74	1.03	0.17	-0.09	1.71	1.54	0.11	1.37	0.03	1.54
7	0.51	1.43	-0.46	1.43	0.57	0.69	0.63	1.31	1.11	0.00	1.14	-0.09	1.14
8	0.31	1.53	-0.49	1.86	1.20	0.29	0.43	1.54	1.51	-0.74	1.46	0.66	1.51
12	-0.03	-0.49	1.46	-0.06	-1.73	1.48	1.71	-0.20	-0.47	-0.51	-0.44	-0.73	-0.80
19	-0.09	-0.11	0.83	0.29	-1.28	1.34	1.47	-0.20	-0.06	-0.50	-0.30	-0.87	0.17
20	-0.03	0.54	0.43	0.49	-1.09	1.60	1.54	0.03	-0.14	-0.86	0.14	-0.09	0.06
21	0.09	-0.74	1.29	0.43	-1.51	1.49	1.40	0.09	-0.17	-0.80	-0.49	-0.91	-0.54
22	-0.29	-0.74	1.00	-0.26	-1.57	1.63	1.46	-0.11	-0.86	-0.11	-0.77	-0.74	-0.97
24	0.21	1.54	-2.03	2.31	0.69	0.57	1.29	0.66	0.14	-1.49	2.40	1.77	0.91
25	-0.50	1.45	-2.03	2.37	1.13	0.46	0.77	0.97	0.46	-1.23	2.72	1.26	0.91
26	0.40	1.57	-2.03	2.27	0.23	1.06	1.26	0.29	-0.03	-1.44	2.13	1.86	0.26
27	-1.17	-1.83	1.74	-1.29	-2.00	1.49	1.57	-0.20	-0.31	0.66	-1.69	-2.20	-1.94
28	-0.29	-0.65	1.63	-0.33	-1.43	1.54	1.28	-0.07	-0.30	-0.03	-0.46	-1.40	-1.97
29	0.23	0.40	-0.31	1.06	0.10	1.31	1.20	0.40	-0.10	-1.00	0.17	-0.09	1.20
30	0.51	0.86	-0.46	1.34	-0.06	0.83	1.17	0.60	0.03	-1.00	0.69	0.26	0.86

Table 11 Sounds Ranked on Mean Ratings for Statement 1 and Statement 12

Rank	Statement 1	Statement 12
1	30	26
2	7	24
3	26	25
4	6	8
5	8	30
6	29	6
7	24	7
8	4	29
9	21	20
10	12	4
11	20	12
12	19	22
13	22	19
14	28	21
15	25	5
16	5	28
17	1	1
18	27	27

Weighted Sound Ratings

In addition to viewing participant's ratings for the sounds on each statement separately, a total score, or sum of the mean ratings on various attributes, was created for each sound. Each mean rating was weighted according to expert judgments about the importance of the attribute to an auditory alert. The weights used in this study were adapted from the Tan and Lerner (1995) study. To create the attribute weights, Tan and Lerner asked 36 experts in the human factors and safety community to rate the importance of thirteen attributes on a scale of 1 to 10. The mean of the experts' importance ratings for each attribute became the weight for the attribute. Eight attributes from the Tan and Lerner study corresponded closely to eight statements rated in the present study. Table 12 shows the attribute and weighting from the Tan and Lerner study along with the corresponding statement from the present study. A ninth statement, Statement 8, which asked whether a sound would be annoying if it occurred once a day as a nuisance alert, was also included in the set of weighted attribute statements. Because a nuisance alert rate of once a day depicts a situation where annoyance may become a critical negative attribute, this statement was set equal to the highest weight from the group of eight attributes (i.e., 9.43 / Noticeability). In this weighting analysis, the "once a week" nuisance alert assumption was assumed to correspond to the general "minimal" false alarm assumption used by Tan and Lerner (1995).

To create the weighted attribute totals, the ratings were first transformed to a scale of 0, strongly disagree, to 6, strongly agree. This was accomplished by adding 3 to each mean rating. Also, the weights for the attributes discriminability and appropriateness, which were positive weights in the Tan and Lerner (1995) study, were changed to negative weights. This change was made because, as negatively worded statements, higher ratings for Statement 3 and Statement 10 reflected more of a negative attribute for the sound. Finally, each mean rating was multiplied by its attribute weight. Two weighted attribute totals were then summed. A total of the weighted mean ratings excluding the mean rating for the annoyance - once per day statement (i.e., assuming nuisance alerts occur once a week), and a total excluding the mean rating for the annoyance - once per week statement (i.e., assuming nuisance alerts occur once a day).

Table 13 shows the two weighted mean rating totals for each sound in rank order as well as the sounds in rank order according to their mean ratings on Statement 12. The three sounds that ranked highest according to these weighted mean attribute ratings were #26, #24, and #25. These sounds were all speech alerts, corresponding to "warning", "danger", and "look out", respectively. Of the three speech sounds, #26 ("Warning", "Warning", "Warning") slightly outperformed the other speech sounds, as is evident in Table 13. In contrast to these speech alerts, the rank order of the remaining non-speech alerts was somewhat influenced by the annoyance attribute. Based on the drivers' overall ratings provided for the non-speech sounds (Statement 12), sounds #8 and #30 appear most promising, coming in fourth and fifth respectively in the final overall ratings. A closer look at the individual statement ratings (shown in Table 10) suggested that Sound #8 may more appropriate for more of an imminent-type or 1-stage alert crash sound, whereas sound #30 may be more appropriate for more of a cautionary-type crash alert.

Figure 3 Scatter Plot of Final (Statement 12) Ratings by Loudness Ratings for Each of the Candidate Alert Sounds

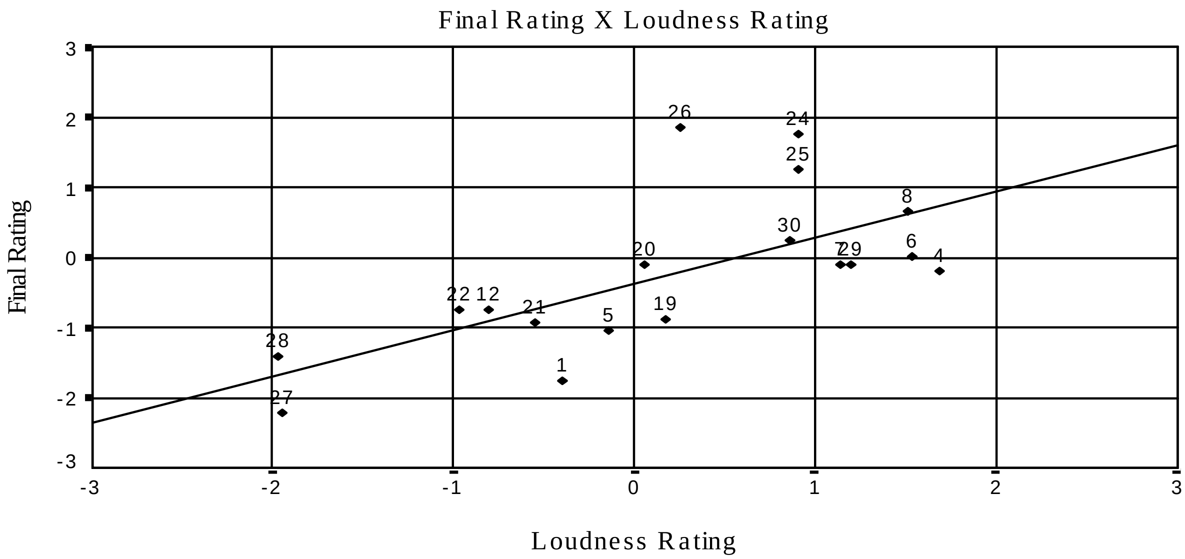

Table 12 Attribute and Weight with the Corresponding Sound Rating Statement

Weight	Attribute	Sound Rating Statement (Sound # in the current study)
9.43	Noticeability	This sound would clearly stand out and be noticeable among the other noises inside and outside the vehicle such as engine noise, the fan blowing, talking and music on the radio, horns, and sirens (#2)
-9.23	Discriminability	This sound would be confused with other sounds inside and outside the vehicle such as engine noise, talking and music on the radio, horns, sirens, car phones, or other electronic device (#3)
8.80	Urgency	This sound would get my attention immediately (#4)
-7.60	Startle	This sound would startle me, that is, cause me to blink, jump, or make a rapid reflex-like movement (#5)
8.63	Natural Response	This sound would not interfere with my ability to make a quick and accurate decision about the safest driving action take (#6)
-9.43	Annoyance	This sound would annoy me if it came on <u>once a day</u> in a situation where NO driving action was required (#8) (*Note: See text for explanation of this weighting.)*
-4.37	Annoyance	This sound would annoy me if it came on <u>once a week</u> in a situation where NO driving action was required (#9)
-5.66	Appropriateness	This sound would appear out of place as a warning in a car or truck (#10)
7.63	Emergency Relationship	This sound would clearly tell me that I'm in danger and I need to react immediately (#11)

<u>Note:</u> Statement 8 was excluded in this weighting analysis because no attribute referred to a sound's influence on the ability to perform an emergency driving action

Table 13 Sounds Ranked by Weighted Mean Rating Totals for Attribute Statements (Totals weighted mean ratings in parentheses)

	Weighting Analysis Mean Rating Totals		
	Annoyance Assumption (or Assumed Nuisance Alert Frequency)		Overall Rating (Statement 12)
Rank	**Once per week…**	**Once per day…**	
1	26 (108)	26 (90)	26
2	24 (102)	24 (82)	24
3	25 (97)	25 (75)	25
4	30 (65)	30 (45)	8
5	8 (60)	20 (41)	30
6	7 (59)	29 (37)	6
7	20 (57)	8 (37)	7
8	29 (57)	7 (36)	29
9	4 (47)	19 (25)	20
10	6 (44)	6 (20)	4
11	19 (39)	21 (17)	12
12	21 (34)	12 (16)	22
13	12 (32)	4 (16)	19
14	5 (31)	22 (11)	21
15	22 (29)	5 (8)	5
16	28 (21)	28 (5)	28
17	1 (8)	1 (-12)	1
18	27 (-9)	27 (-24)	27

Summary of Results from the Sound Selection Process

This study built upon previous work conducted by Tan and Lerner (1995), which examined 26 sounds, including various non-speech, ear con (car horn, tire skid) and speech sounds. The current study, employing nearly the identical methodology employed by Tan and Lerner, examined 15 non-speech and 3 speech sounds, including the 5 top rated sounds (which were all non-speech) from the previous Tan and Lerner study. Hence, in some sense, together, these two studies have examined 39 distinct sounds, including 22 distinct non-speech sounds, 15 distinct speech sounds (all using either the word "warning", "danger", "look out", or "hazard"), and 2 distinct ear con-type sounds (car horn, tire skid).

As a result of the current study, Sound #26 ("Warning, Warning, Warning") was used for both driver interface studies (i.e., Study 2 and Study 3) which evaluated a speech alert condition. In addition, based on the current findings, Sound #8 (which corresponds to Stimuli 10 in the earlier Tan and Lerner study) was used for all three driver interfaces studies (i.e., Study 2, Study 3, and Study 4) as the non-speech alert sound. A 1/3 octave band and time series analysis of this non-speech sound can be found in the Tan and Lerner paper (see Appendix A). This 2.1 second long non-speech sound involved repeating the exact same macro "sound pattern" (or macro sound burst) four times. Each repetition of the macro sound pattern was followed by 110 milliseconds of silence. Each macro sound pattern in turn involved repeating the exact same micro sound pattern (or micro sound burst) four times. These micro sound bursts, which are the building blocks for a macro sound burst, consisted of 2500 Hz and 2650 Hz peaks.

In conclusion, these results provided a sound empirical justification for the selection of the non-speech and speech sounds used in the follow-up, closed-course driver-interface studies.

Verbal Instructions Used in the Auditory Crash Alert Evaluation Procedure

The reason we have invited you here today is that we are in the process of trying to select sounds to use in vehicles that would serve as a collision warning sound. In a few minutes I am going to have you listen to a number of different sounds. Each sound you will hear is being considered as a collision warning sound. But before you listen to the sounds it is important that you understand the requirements of the sound.

To help give yourself some frame of reference, try to recall your experience at the General Motors Proving Grounds in Milford, MI this past fall. In one part of that study, your task was to brake at the last second possible using hard braking to avoid colliding with the lead car.

Now, imagine that instead of being on the test track you're driving on a real road. Further, suppose your distracted or not paying attention to your driving and you're rapidly approaching a slower or stopped vehicle. The collision warning sound would alert you to this dangerous situation. When you hear the warning sound you have to decide upon the appropriate driving action to take. The driving action required, for example braking or steering, would depend on your driving situation. And going back to what you did on the test track, if braking is the appropriate action, you would need to brake hard immediately.

So as you see, the warning sound needs to get the driver's attention while at the same time allowing the driver to respond appropriately.

In addition, it is also possible that the warning may sound in an inappropriate situation. In other words, when it is a "false" alarm. For our purposes, assume that false alarms could occur as often as once a day to once a week, depending on the driver. A false alarm could be caused by a non-threatening event such as, approaching a guard-rail or sign on a sharp curve. In this case, the collision warning system may mistake the guardrail or sign for a stopped vehicle. It would not be the case that the warning would sound periodically without any reason at all.

But because false alarms may occasionally occur, the warning sound needs to get the driver's attention without being overly annoying.

Okay, we are now ready to listen to the sounds. For each sound you hear, you will be asked to rate the extent to which you agree with a statement made about that sound. For example, consider the practice statement "This sound is very musical." You will hear a sound. Then you will rate that sound on a scale ranging from Strongly Agree to Strongly Disagree based on the extent to which you agree that the sound is very musical. And in just a moment we will go over the scale in more detail.

But before we begin I would like to stress upon you to remember that the warning sound needs to immediately get your attention and allow for an appropriate response but not be overly annoying when false alarms occur. Please keep this information in mind as you make your judgments about each sound.

Do you have any questions so far?

You are going to be listening to the sounds over these headphones. But wait just a few more moments until were done with the directions to put them on and adjust them.

During the session each one of you will be sitting with your back to the table. We are doing this primarily to keep the equipment and cords out of your way. But I should mention that the headphone cords are delicate so it would be very helpful if you are careful with them. The headphones are marked for right and left ear and you should wear them that way. One last thing, while you're listening to the sounds, try to avoid touching the outside of the earphones because that will distort the sounds.

When the tape begins, the first sound you will hear is the interior sound of the Taurus that you drove at the proving grounds traveling at 70 mph. This is the actual ambient noise that is present inside the vehicle while you are traveling. All of the test sounds have been recorded on top of the ambient noise so this noise will be continuous. You will hear the first warning sound a few seconds after the ambient noise begins. The warning sounds may at times appear strange but I am going to ask that you refrain from making any comments about them during the test.

Okay we are now ready to go through an actual practice run. The practice statement on your answer sheet is THIS SOUND IS VERY MUSICAL. As you hear each warning sound you should circle your response for that sound on the scale provided. That is, you should circle the response on the scale ranging from Strongly Agree to Strongly Disagree that reflects the extent to which you agree that the sound is very musical. You will hear each sound in sequence. After you have rated all the sounds on the first statement we will follow the same procedure for the second statement and so on. The sounds used for practice are the same sounds being tested. Any questions? Please put your headphones on now.