Alerts for In-Vehicle Information Systems: Annoyance, Urgency, and Appropriateness

Dawn C. Marshall, John D. Lee, and **P. Albert Austria,** University of Iowa, Iowa City, Iowa

Objective: This study assesses the influence of the auditory characteristics of alerts on perceived urgency and annoyance and whether these perceptions depend on the context in which the alert is received. **Background:** Alert parameters systematically affect perceived urgency, and mapping the urgency of a situation to the perceived urgency of an alert is a useful design consideration. Annoyance associated with environmental noise has been thoroughly studied, but little research has addressed whether alert parameters differentially affect annoyance and urgency. **Method:** Three $2^3 \times 3$ mixed within/between factorial experiments, with a total of 72 participants, investigated nine alert parameters in three driving contexts. These parameters were formant (similar to harmonic series), pulse duration, interpulse interval, alert onset and offset, burst duty cycle, alert duty cycle, interburst period, and sound type. Imagined collision warning, navigation alert, and E-mail notification scenarios defined the driving context. **Results:** All parameters influenced both perceived urgency and annoyance ($p < .05$), with pulse duration, interpulse interval, alert duty cycle, and sound type influencing urgency substantially more than annoyance. There was strong relationship between perceived urgency and rated appropriateness for high-urgency driving scenarios and a strong relationship between annoyance and rated appropriateness for low-urgency driving scenarios. **Conclusion:** Sound parameters differentially affect annoyance and urgency. Also, urgency and annoyance differentially affect perceived appropriateness of warnings. **Application:** Annoyance may merit as much attention as urgency in the design of auditory warnings, particularly in systems that alert drivers to relatively low-urgency situations.

INTRODUCTION

Sensor, wireless, and computer technologies have made possible the development of many new in-vehicle information systems. These include collision warning systems that alert drivers to imminent collision situations, navigation aids that provide drivers with voice guidance, and infotainment systems that enable drivers to use E-mail and conduct business as they drive. The proliferation of in-vehicle information systems, combined with the need for drivers to keep their eyes on the road, suggests that auditory alerts may become increasingly common. Experience in other domains, such as hospital intensive care units (Meredith & Edworthy, 1995), nuclear power plants (E. Marshall & Baker, 1994; Woods, 1994), and aviation (Boucek, Veitengruber, & Smith, 1977), demonstrates the dangers of an unchecked proliferation of alerts. Frequent false alarms and the mismatch between the perceived urgency of an alert and the actual urgency of a situation can undermine alert effectiveness (Meredith & Edworthy, 1995). Perceived urgency may affect how quickly the driver will recognize and respond to an alert, whereas perceived annoyance may influence whether the driver will disable or ignore an alert, particularly in situations with many false alarms. Because perceived urgency and annoyance may govern the effectiveness of auditory alerts, this study investigates auditory characteristics that influence urgency and annoyance.

The urgency mapping principle states that the perceived urgency of an auditory alert should

Address correspondence to John D. Lee, University of Iowa, Department of Mechanical and Industrial Engineering, 2130 Seamans Center for the Engineering Arts and Sciences, Iowa City, IA 52242; jdlee@engineering.uiowa.edu. ***HUMAN FACTORS,*** Vol. 49, No. 1, February 2007, pp. 145–157.

correspond to the urgency of the situation (Edworthy & Adams, 1996; Edworthy, Loxley, & Dennis, 1991). Alerts with proper urgency mapping may help drivers to understand alerts and respond effectively. Several studies have found that perceived urgency can affect response time to alerts: People respond more quickly to alerts that sound more urgent (Burt, Bartolome, Burdette, & Comstock, 1995; Edworthy, Hellier, & Walters, 2000; Haas & Casali, 1995; Haas & Edworthy, 1996). People are also more likely to respond to high-urgency alerts (Bliss, Gilson, & Deaton, 1995). Such alerts induce psychophysiological effects that coincide with differences in reaction time (Burt et al., 1995). High-urgency alerts can also mitigate the diminished response to true alerts in the context of a high rate of false alarms – the cry wolf effect (Bliss, Dunn, & Fuller, 1995). However, high-urgency alerts can also induce inappropriate responses to false alarms (Graham, 1999). Overall, substantial evidence suggests perceived urgency of an alert influences behavior.

A substantial research base shows that auditory alert parameters communicate different levels of urgency (Haas & Edworthy, 1996; Hellier, Edworthy, & Dennis, 1993, 1995). For example, Haas and Casali (1995) found that perceived urgency increased with the intensity of the alert and decreased as the interpulse interval increased. Similarly, increasing speed (i.e., pulse repetition rate), fundamental frequency, number of repetitions, and inharmonicity all increased perceived urgency, with pulse repetition rate having the greatest influence (Hellier et al., 1993). Edworthy et al. (1991) found that increasing the urgency of an alert sometimes made it more irritating. These results suggest that the features that make an alert highly urgent may also make them more annoying and possibly undermine acceptance of the alert.

Annoyance is a subjective response to a sound based on its physical nature, emotional content, and novelty (Kryter, 1985) and is a critical factor governing driver acceptance and the ultimate success of collision warning systems (Kiefer et al., 1999). In the operating room, 70% of anesthesiologists surveyed reported that they turned off alarms and that it was the annoyance associated with the auditory characteristics of the alarms that provoked this behavior (Block, Nuutinen, & Ballast, 1999). It is important to consider the annoyance associated with an alert because annoying alerts can undermine the influence of warning systems.

The systematic study of annoyance has a long history (Laird & Coye, 1929), with most researchers focusing on annoyance associated with environmental noise. Perceived annoyance depends on the context in which a sound is being judged as well as on its acoustical properties (Fucci, Petrosino, Hallowell, Andra, & Wilcox, 1997; Fucci, Petrosino, McColl, Wyatt, & Wilcox, 1997). Some of the more influential acoustical properties affecting annoyance include loudness, sharpness, impulsiveness, roughness, harmony, and tonal components (Khan, Johansson, & Sundback, 1997); frequency spectrum, sound level, and duration of total sound (Kryter, 1985); sound duration and onset time (Berglund & Preis, 1997); and sound onset and offset (Nixon, Von Gierke, & Rosinger, 1969). Although annoyance is somewhat specific to situations and individuals, surprisingly, loudness is sometimes more dependent on the situation and the individual (Berglund & Preis, 1997). These results show that sound characteristics systematically influence annoyance, but little research has systematically addressed annoyance associated with alerts in the driving domain.

One of the few studies to consider both urgency and annoyance in the driving domain focused on auditory alerts for collision warning systems (Tan & Lerner, 1995). They asked drivers to rate alerts according to conspicuity, discriminability, meaning, urgency, annoyance and appropriateness. Appropriateness and urgency were more highly correlated with each other ($r = .89$) than either was with annoyance (urgency-annoyance, $r = .78$; appropriateness-annoyance, $r = .53$; Tan & Lerner, 1995). These results suggest that people judge highly urgent alerts as more appropriate and that highly urgent alerts also tend to be annoying.

These findings suggest that, like perceived urgency, alert parameters have a systematic effect on annoyance. However, the implicit assumption that the parameters that increase urgency will also increase annoyance might not always be valid (Edworthy & Stanton, 1995). Additionally, the context of alerts may influence the relative importance of annoyance and urgency, prompting a consideration of a trade-off between urgency and annoyance in alert design. Through three experiments, this research addresses three objectives: (a) to identify alert parameters that increase perceived urgency more than annoyance, (b), to investigate how context affects perceived urgency and annoyance, and (c) to understand how alert parameters influence what people perceive as an appropriate alert in different driving contexts.

EXPERIMENT 1: HARMONIC SERIES, PULSE DURATION, AND INTERPULSE INTERVAL

The objective of Experiment 1 was to determine how harmonic series, pulse duration, interpulse interval, and context of an alert, in the form of imagined driving scenarios, influence perceived urgency and annoyance. The harmonic series parameter was selected because a previous study found it to have a substantial influence on urgency (Stanford, McIntyer, Nelson, & Hogan, 1988). Pulse duration and interpulse interval were selected because they share some of the stimulus characteristics of pulse repetition rate, which has a strong effect on perceived urgency (Hellier et al., 1993). Increasing duration, like increasing pulse repetition rate, increases the total time sound is present per unit time. Decreasing the interval between pulses, like increasing the pulse repetition rate, decreases the time between a pulse offset and onset. Therefore, increasing pulse duration and decreasing interpulse interval are likely to have a strong influence on urgency.

Method

Participants. Participants were screened for age (between 18 and 35 years old) and for hearing impairment; those who reported impairment were not invited to participate. Each individual participated in only one of the three experiments. All participants in these experiments were undergraduate students. Because hearing and driving expertise change with age, the age of the participants limits generalization, but this sample is representative of a group of drivers that is disproportionately involved in crashes (Evans, 2004). Experiment 1 included 24 participants (12 men and 12 women). Their mean age was 21.1 years, with a standard deviation of 1.2.

Apparatus. The alerts consisted of wave files created using Sound Forge 4.5 software by Sonic Foundry; Visual Basic programs presented the alerts to the participants. Participants wore Sony Model MDR-V900 over-the-ear type headphones to ensure uniform intensity and to minimize background noise. Sound levels were measured with a Bruel & Kjaer sound level meter Model 2209 with a 1-inch (2.54-cm) condenser microphone with the headphones mounted to a general radio Type 9A (6cc) coupler. Levels were 64 dBA for Experiments 1 and 2 and 68 dBA for Experiment 3. This apparatus provides precise control of stimulus delivery, but it may neglect important factors such as differential masking that may occur in the actual driving environment.

Independent and dependent variables. The experimental design was a $2^3 \times 3$ mixed within/between factorial design. A factorial combination of two levels of each of the three within-subject variables generated eight alerts. The within-subject variables were harmonic series, pulse duration, and interpulse interval. The two levels of pulse duration (defined as the time between pulse onset and offset) were 0.025 and 0.100 s. The two levels of interpulse interval (defined as the time between the offset of one pulse and the onset of another) were also 0.025 and 0.100 s. The pulses were not grouped into bursts but were uniformly spaced within each alert. Each pulse was followed by a period of silence defined by the interpulse interval, then another pulse, then silence, and so on. The fundamental frequency of the pulses was 150 Hz, and the overall alert was approximately 2.3 s long.

Four formants augmented the fundamental frequency to define the harmonic series. Formants describe the underlying frequency spectrum of the sound and reflect anatomical properties that govern how people produce sound, such as the length and cross section of the vocal tract (Mermelstein, 1967). Formants are fundamental characteristics of natural speech sounds and seem more likely to influence the emotional content of a sound as compared with artificial frequency combinations, such as octaves. Previous research has shown that they influence the perception of alert characteristics related to perceived urgency and annoyance (Stanford et al., 1988). The formants used here were 300, 2550, 3450, and 4050 Hz for high /i/ and 600, 1350, 2550, and 3300 Hz for high /a/.

A short description of one of three scenarios defined the driving context as a between-subject variable. Participants read the description of the scenario and then imagined hearing the alert from one of three in-vehicle systems: collision avoidance, navigation, and E-mail. These three scenarios imply differing levels of hazard, importance, and time criticality. Collision avoidance is a highly urgent context, given the consequences of drivers failing to respond to an alert. Navigation is a moderately urgent context, with consequences that include inconvenience but little danger. E-mail provides drivers with a context of minimal urgency and little consequence for ignoring the alert. Each scenario included a description of one system, its purpose, when and how often the system would issue an alert, and how many of the issued alerts might be false alarms. Participants received only these descriptions of the driving scenarios as the context for interpreting alerts. Specific driving events or conditions were not described. Participants imagined how the system might interact with the driving task, but they did not actually perform any driving tasks.

The dependent variables included perceived urgency and annoyance. These were measured using subjective ratings and paired comparisons. Participants rated the perceived urgency and annoyance on a 0 to 100 scale, with 0 representing *low urgency* or *low annoyance* and 100 representing *very high urgency* or *very high annoyance.* The paired comparisons were created by combining each alert with every other alert in random order. Participants heard two sets of randomly presented alerts. With one set participants chose the more urgent alert of each pair, and with the other they chose the more annoying alert. The rating scale and paired comparison techniques are similar to techniques used by Edworthy et al. (1991) and Hellier et al. (1995) in their investigations of perceived urgency.

Experimental protocol. The participants received the written descriptions of the scenarios, which the experimenter then read. These provided a brief description of the driving scenario and meaning of the alert information (i.e., collision alert scenario, navigational scenario, or E-mail scenario). The participants then completed four tasks. First they rated the urgency and annoyance of each of the eight alerts, and then they completed two sets of paired comparisons based on urgency and annoyance. Finally, they rated the urgency and annoyance of each alert a second time. The experimenter reminded participants to imagine the scenario before each rating and paired-comparison task. There were two orders in which the paired comparisons were completed, either annoyance first or urgency first, and these were counterbalanced between participants.

In the rating task, a Visual Basic program presented participants with the eight alerts in a random order. After each alert, participants responded by rating perceived urgency and annoyance. The participants rated each alert by sliding a marker on the display. In the paired-comparison task, a Visual Basic program presented pairs of alerts. Participants clicked the button of the alert they perceived to be more urgent (or annoying). In total, participants listened to each alert 16 times: 1 time during the first rating of urgency and annoyance, 7 times during the paired comparisons of urgency, 7 times during the paired comparisons of annoyance, and 1 more time during the second rating of urgency and annoyance.

Data analysis. The paired-comparison data consisted of binary values in a matrix indicating the more annoying or urgent alert of each alert pair. These data were transformed into interval scale data for each participant by transforming the frequency of the binary values of each alert into a z score (Woodworth & Schlosberg, 1958). The z score and rating scale data were then analyzed with similar ANOVA models using the SAS Mixed procedure.

Results

Alert parameters. The longer pulse duration (0.100 s) was rated as both more annoying than the shorter pulse duration (0.025 s), $F(1,3\ 15) = 11.54$, $p < .0001$, and more urgent, $F(1, 315) = 40.38$, $p < .0001$. Likewise, it was chosen as more annoying, $F(1, 126) = 9.63$, $p = .0024$, and more urgent, $F(1, 126) = 132.07$, $p < .0001$, in the paired comparisons. Figure 1 shows how increasing the pulse duration increased the z score of perceived urgency by 0.80, compared with a change of only 0.28 for perceived annoyance. The longer interpulse interval (0.100 s) was rated as less annoying, $F(1, 315) = 30.41$, $p < .0001$, and less urgent, $F(1, 315) = 36.49$, $p < .0001$; it was also chosen as less annoying, $F(1, 126) = 47.91$, $p < .0001$, and less urgent, $F(1, 126) = 172.22$, $p < .0001$, in the paired comparisons. Similarly, the high /a/ alerts were both rated as more annoying, $F(1, 315) = 35.02$, $p < .0001$, and urgent, $F(1, 315) = 28.73$, $p <.0001$, than the high /i/ alerts and were also chosen as more annoying,

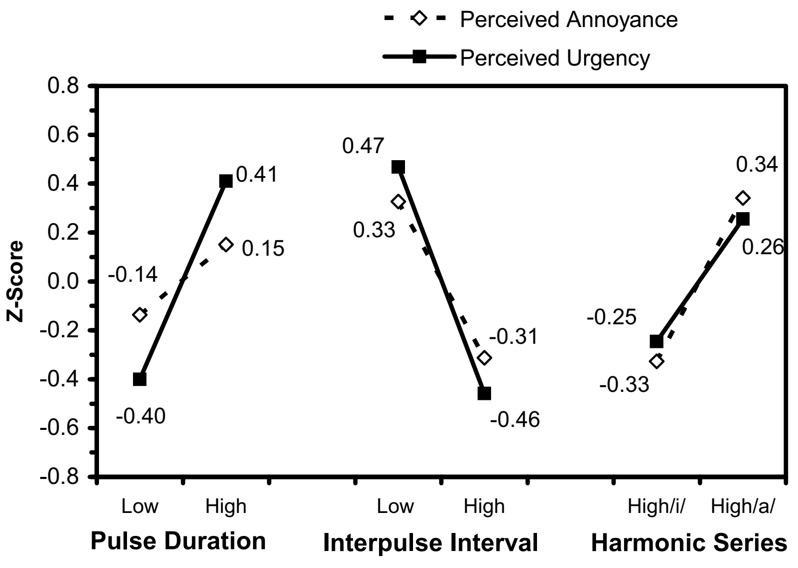

Figure 1. The effect of pulse duration, interpulse interval, and harmonic series on urgency and annoyance.

$F(1, 126) = 52.32$, $p < .0001$, and more urgent, $F(1, 126) = 50.45$, $p < .0001$, in the paired comparisons.

Driving scenarios. The driving context, as represented by the three scenarios, did not directly affect perceived urgency or annoyance. However, there was a two-way interaction between interpulse interval and scenario for perceived annoyance. The driving scenarios mediated the effect of the interpulse interval, so that it had the least effect with the E-mail scenario. This was true for both the ratings, $F(2, 315) = 6.73$, $p = .0014$, and the paired comparisons, $F(2, 126) = 4.61$, $p = .0116$.

Discussion

All three alert parameters had similar effects on urgency and annoyance; increasing urgency also tended to increase annoyance. However, as far as the defined scales are concerned, increasing pulse duration increased perceived urgency substantially more than it did annoyance. Thus pulse duration may be a useful way to increase urgency while minimizing annoyance. All three parameters affected urgency in a manner consistent with previous research (Hellier et al., 1993; Stanford et al., 1988). Of particular interest are the strong effects of pulse duration and interpulse interval. Higher urgency associated with increasing pulse duration and decreasing interpulse intervals suggests that the influence of pulse repetition rate may be attributable to the duty cycle of the alert, in which the duty cycle represents the proportion of time the sound is present. The alerts in this experiment consisted of only a series of pulses. Many alerts include groups of pulses that combine to form bursts. For this reason, the second experiment focused on burst parameters.

EXPERIMENT 2: ALERT ONSET, ALERT OFFSET, AND BURST DUTY CYCLE

The second experiment explored bursts, which consist of a group of pulses with time between groups. Pulses occur on a time scale of 100 to 300 ms, and bursts occur on a time scale of 1 to 3 s. The objective of Experiment 2 was to determine how the burst duty cycle, the onset and offset of the sound envelope that contained several bursts, and the context of an alert influence perceived urgency and annoyance. Like pulse duration and interpulse interval, duty cycle represents another factor that might contribute to the increased urgency associated with high repetition rate. Increasing the repetition rate with a fixed duration pulse increases the duty cycle. Offset and onset represent a particularly compatible way to convey urgency in a collision avoidance situation because these parameters can convey a sense of approaching and receding objects (Nixon et al., 1969). Appropriateness was added as a dependent variable to assess whether the urgency mapping principle leads to alerts that people judge appropriate in various driving contexts.

Method

Each of three experiments used a similar protocol to assess annoyance and urgency of auditory alerts. The protocols for Experiments 2 and 3 will be discussed only to the extent that they differ from the protocol used in the first experiment.

Participants. This experiment included 24 participants (12 men and 12 women; $M = 21.1$ years of age, $SD = 1.8$). None of these people participated in the other two experiments.

Independent and dependent variables. This experiment included one of the four alerts recommended for use in a rear-end collision warning system (Tan & Lerner, 1995), which has also proven effective in helping drivers respond to imminent collision situations (Lee, McGehee, Brown, & Reyes, 2002). Additional alerts were generated by modifying this basic alert. There were two levels each of alert onset, offset, and burst duty cycle. At the low level of burst duty cycle, each burst was composed of three pulses, whereas at the high level of burst duty cycle, each burst was composed of four pulses. The duration of each alert was approximately 2.3 s. Bursts were separated by 110 ms and pulses by approximately 10 ms. The prominent frequency of the sound was 2500 Hz.

For the fast offset and onset conditions, both the initial burst and the initial pulse within the initial burst were as intense as subsequent bursts and pulses. For slow onset, the intensity rose from zero to 67% of the maximum in the first quarter of the alert (575 ms) and from 67% to 100% of the maximum in the second quarter of the alert. For slow offset, the intensity dropped from 100% to 67% of the maximum in the third quarter of the alert and from 67% of the maximum to zero in the fourth quarter of the alert.

The dependent variables were the same as in Experiment 1, with the addition of perceived appropriateness as a third dependent variable. Appropriateness was measured using a rating scale similar to the one used to assess perceived urgency and annoyance. Appropriateness was not measured using paired comparisons. Appropriateness was added to investigate whether another approach to considering urgency and annoyance might complement the urgency mapping principle. The analyses of the appropriateness data for Experiments 2 and 3 are combined in a separate section.

Results

Alert parameters. The fast onset was rated as more annoying, $F(1, 315) = 17.73, p < .0001$, and more urgent, $F(1, 315) = 13.58, p = .0003$, and was judged more annoying, $F(1, 126) = 17.39$, $p < .0001$, and more urgent, $F(1, 126) = 72.58$, $p < .0001$, in the paired comparisons. The fast offset was more annoying, $F(1, 315) = 9.75, p = .0020$, and more urgent, $F(1, 315) = 27.88, p < .0001$, according to the rating scales and judged more annoying, $F(1, 126) = 5.32, p = .0227$, and more urgent, $F(1, 126) = 65.9, p < .0001$, in paired comparisons. However, Figure 2 shows that the fast offset increased perceived urgency substantially more than it did perceived annoyance, as measured by the z score derived from the paired comparisons. Higher burst duty cycle was rated as more urgent, $F(1, 315) = 13.96, p = .0002$, but not more annoying, $F(1, 315) = 2.49$, $p = .1159$, in the rating scales. In the paired comparisons, the higher burst duty cycle was chosen as both more urgent, $F(1, 126) = 32.14$, $p < .0001$, and more annoying, $F(1, 126) = 6.12$, $p = .0147$.

Alert onset had a larger effect on perceived urgency for alerts with a slow offset for the ratings of urgency, $F(1, 315) = 22.02, p < .0001$, and for the paired comparisons of urgency, $F(1, 126) = 34.64$, $p < .0001$. Participants perceived the combination of slow onset and slow offset as a particularly low urgency alert. Increasing offset increased perceived urgency by a z score of 0.66, compared with 0.22 for perceived annoyance.

Driving scenarios. As in Experiment 1, the driving context did not directly affect perceived urgency or annoyance. However, scenario did interact with alert onset to affect rated annoyance, $F(2, 315) = 7.76, p = .0005$, and paired comparisons of annoyance, $F(2, 126) = 3.71, p = .0271$. Alert onset had the largest effect for the collision avoidance scenario and little effect during the navigation and E-mail scenarios, with slow onset alerts being least annoying when paired with the collision avoidance scenario.

Discussion

As far as the defined scales are concerned, increasing offset and burst duty cycle increased

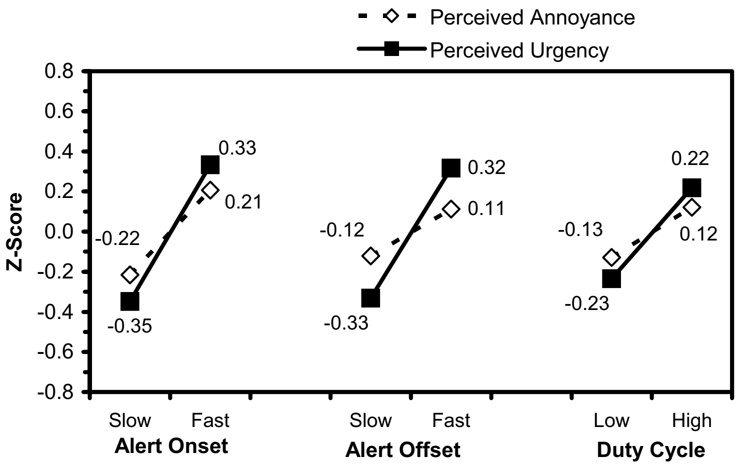

Figure 2. The effect of alert onset, alert offset, and duty cycle on urgency and annoyance.

perceived urgency substantially more than they increased perceived annoyance. This suggests that manipulating the offset and burst duty cycle may be an effective way of increasing perceived urgency while minimizing annoyance. The driving context again influenced the effect of the alert parameters, such that a slow alert onset was much less annoying for the collision warning scenario. This is similar to the larger effect of pulse duration on annoyance found in the collision warning scenario of Experiment 1. The base alert in this experiment was a relatively complex waveform compared with that in Experiment 1. The following experiment directly compares alerts composed of simple and complex waveforms.

EXPERIMENT 3: ALERT DUTY CYCLE, INTERBURST PERIOD, AND SOUND TYPE

The objective of Experiment 3 was to understand how alert duty cycle, interburst period, and sound type (frequency series and high /i/) influence perceived urgency, annoyance, and appropriateness. Duty cycle and interburst period were chosen because in Experiment 1, pulse duration and interpulse interval had strong effects on both perceived annoyance and perceived urgency. Two sound types were also chosen: a frequency series from the second experiment and the high /i/ sound from the first experiment. As in Experiment 1, the high /i/ harmonic series (300, 2550, 3450, and 4050 Hz) augmented the base frequency of 150 Hz (Stanford et al., 1988). The point of comparing the sound types was to assess the degree to which a complex waveform used in the second experiment influences annoyance and urgency, as compared with a relatively simple series of pulses based on a harmonic series.

Method

Participants. This experiment included 24 participants (12 men and 12 women) between 18 and 35 years of age ($M = 21.9$, $SD = 3.8$). As with Experiment 2, none of these people participated in the other experiments.

Independent and dependent variables. The independent within-subject variables were alert duty cycle, interburst period, and sound type. Alert duty cycle is the percentage of the alert during which sound is present, and here it was either 20% or 80% of the duration of the alert. The interburst period is the time from the beginning of one burst to the beginning of the next burst, and here it was either 378 or 227 ms. Sound type describes the base burst used to create the alerts. The dependent variables were the same as in Experiment 2.

Results

Alert parameters. Alerts with a high duty cycle were rated as more annoying than those with a low duty cycle, $F(1, 315) = 139.20$, $p < .0001$, and more urgent, $F(1, 315) = 312.15$, $p < .0001$. In the paired comparisons, high duty cycle alerts were also chosen as more annoying, $F(1, 126) = 22.52$, $p < .0001$, and urgent, $F(1, 126) = 458.09$, $p < .0001$. A shorter interburst period was rated more annoying, $F(1, 315) = 3.96$, $p = .0475$, and more urgent, $F(1, 315) = 21.74$, $p < .0001$. The shorter interburst period was also chosen as more annoying, $F(1, 126) = 11.46$, $p = .0001$, and urgent, $F(1, 126) = 64.27$, $p < .0001$, in the paired comparisons. Figure 3 shows that decreasing the interburst period decreases urgency but increases annoyance, as measured by the *z* score. The high /i/ sounds were rated more annoying than the frequency series alerts, $F(1, 315) = 35.47$, $p < .0001$, but less urgent than the frequency series alerts. High /i/ alerts were also chosen as more annoying, $F(1, 126) = 19.15$, $p < .0001$, and less urgent, $F(1, 126) = 18.05$, $p < .0001$, in the paired comparisons. High /i/ sounds had higher perceived annoyance but lower perceived urgency than the frequency series sounds. Increasing the duty cycle increased rated annoyance more for the high /i/ alerts than for the frequency series alerts, $F(1, 315) = 5.57$, $p = .00188$, and in the paired comparisons, $F(1, 126) = 6.27$, $p = .0136$.

Discussion

Increasing the duty cycle of the alert increased both the perceived urgency and annoyance and had a strong effect as compared with sound type or interburst period. Although the interburst period affected perceived urgency less than the duty cycle, it had a relatively strong effect on urgency as compared with its effect on annoyance. The frequency series alerts were rated as more urgent but less annoying. These results demonstrate that it is possible to create alerts with high levels of perceived urgency and relatively low levels of perceived annoyance by using a high duty cycle, low interburst period frequency series. In particular, carefully crafted complex sounds, such as the frequency series alert from Tan and Lerner (1995), can convey a high degree of urgency with relatively little annoyance. Unlike the first two experiments, in this experiment driving context did not directly affect perceived urgency or annoyance as a main effect or through interactions with other independent variables.

ANNOYANCE, URGENCY, AND APPROPRIATENESS

The second and third experiments included ratings of appropriateness. These ratings, in combination with the ratings of perceived annoyance and urgency, provide an opportunity to evaluate the urgency mapping principle. According to the urgency mapping principle, appropriateness should be positively correlated with ratings of perceived urgency for high-urgency situations and negatively correlated for low-urgency situations. Figures 4 through 6 show the relationship among rated urgency, annoyance, and appropriateness for the three different driving scenarios.

For the collision avoidance scenario, Figure 4 shows a strong positive relationship between perceived appropriateness and urgency, with an R^2

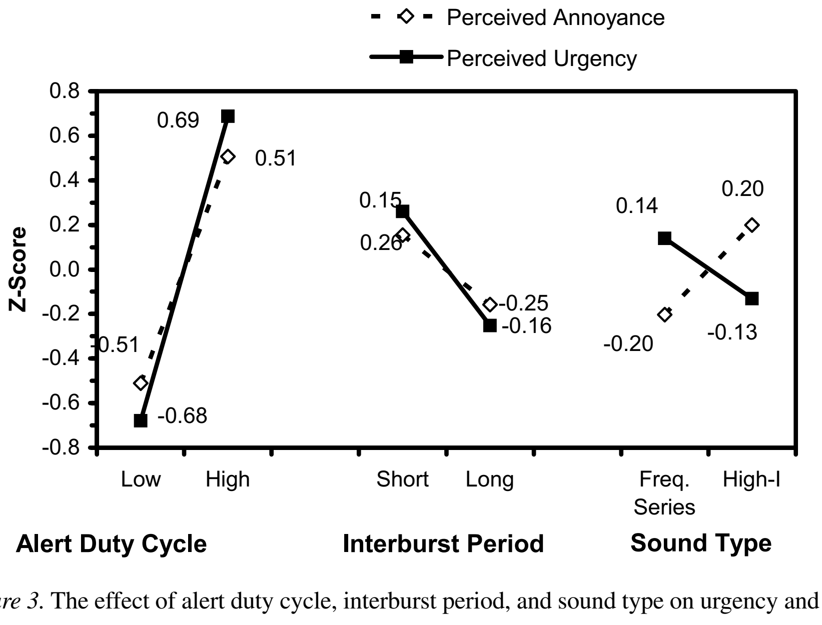

Figure 3. The effect of alert duty cycle, interburst period, and sound type on urgency and annoyance.

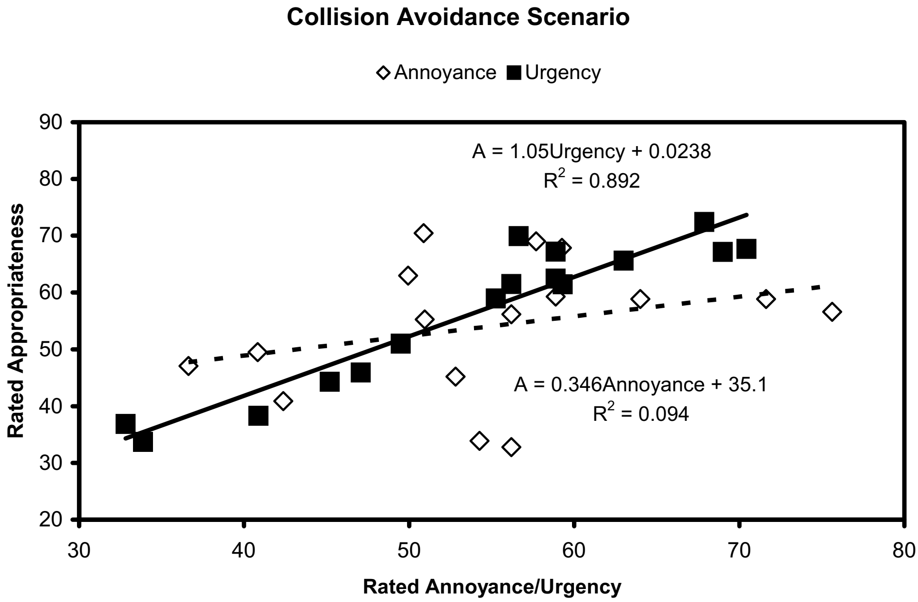

Figure 4. Rated annoyance as a function of rated urgency and rated appropriateness in the collision avoidance scenario.

value of .89. Consistent with the urgency mapping principle, alerts perceived as highly urgent are regarded as appropriate. In contrast, there is a very weak relationship between perceived annoyance and appropriateness, with an R^2 value of .094. For the navigation scenario, Figure 5 shows no clear trend for either the perceived urgency or annoyance of the alerts. The ratings also remain within a narrow range, with an R^2 value below .20 in both cases. For the E-mail scenario, Figure 6 shows less appropriate alerts had high levels of perceived urgency and annoyance and the more appropriate alerts had lower levels of perceived urgency and annoyance. The data for the E-mail scenario differ from the collision avoidance data in two important respects. First, consistent with the urgency mapping principle, the coefficient that relates perceived urgency

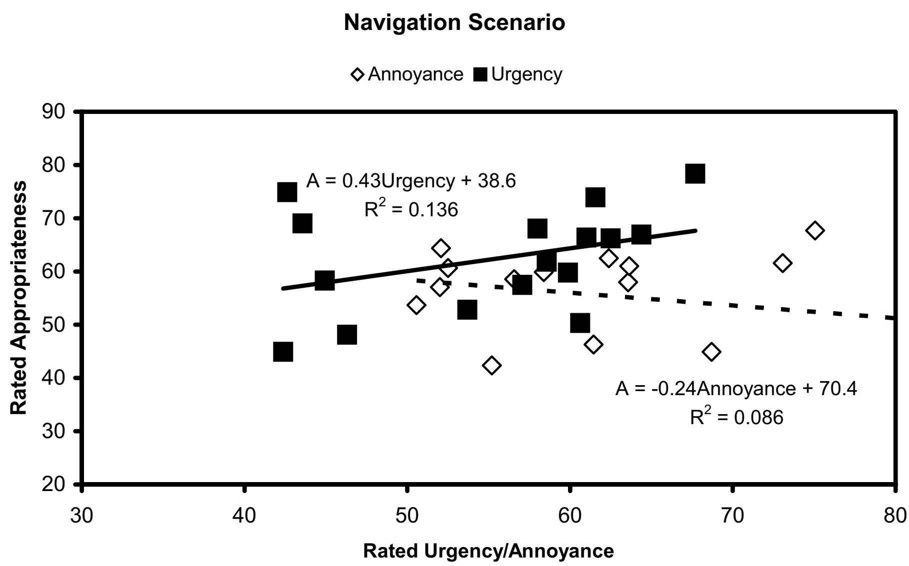

Figure 5. Rated annoyance as a function of rated urgency and rated appropriateness in the navigation scenario.

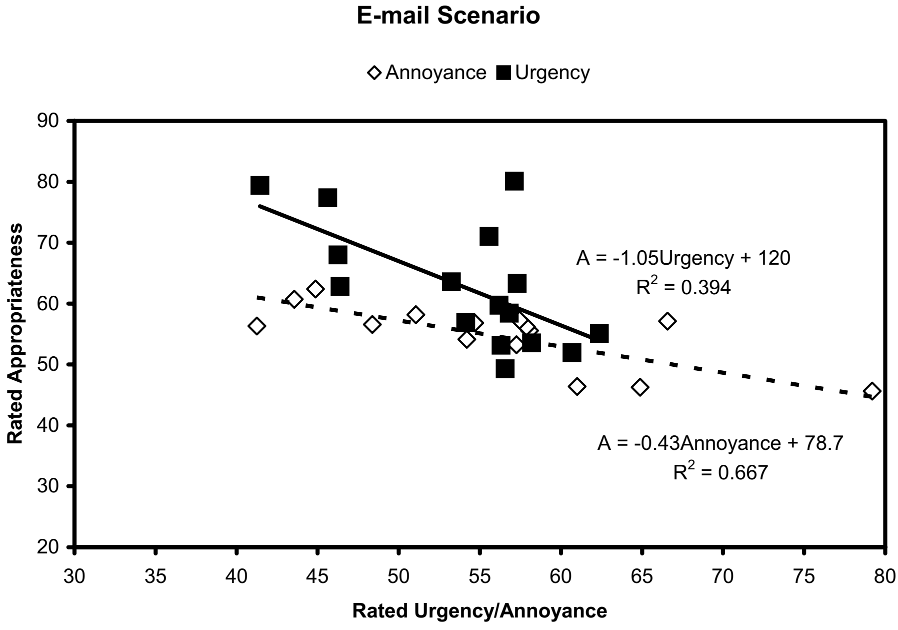

Figure 6. Rated annoyance as a function of rated urgency and rated appropriateness in the E-mail scenario.

and annoyance to appropriateness is negative. Second, appropriateness is more strongly associated with perceived annoyance (R^2 = .67) than with perceived urgency (R^2 = .39). This suggests considerations of annoyance need to complement the urgency mapping principle in developing alerts for noncritical situations, such as E-mail systems.

GENERAL DISCUSSION

This study achieved three objectives. First, it showed that certain alert parameters increase perceived urgency more than annoyance. Second, it showed that context affects how alert parameters affect perceived urgency and annoyance. Third, it showed that the criteria used to judge the appropriateness of alerts depend on the driving context. Because this study used only two levels of each alert parameter, considered a limited range of driving scenarios, and did not engage the participants in any actual driving activity, the results must be viewed as preliminary and in need of further verification before firm design guidelines can be developed.

Alert characteristics affect perceived urgency and annoyance to different degrees. In the case of sound type, investigated in Experiment 3, even the direction of the effect was different. These results suggest alert parameters can be adjusted to increase perceived urgency with relatively little effect on perceived annoyance. Many previous studies of urgency mapping have adopted the implicit assumption that adjusting alert parameters to increase perceived urgency will increase annoyance to the same degree. The results found here indicate this assumption is not always valid. Table 1 suggests that pulse duration, interpulse interval, alert offset, alert duty cycle, and sound type may be particularly promising parameters for increasing urgency with relatively little effect on annoyance.

Sound parameters that affect perceived annoyance of alerts are similar to those associated with environmental noise. For example, Khan et al. (1997) investigated annoyance attributable to diesel engine noise and found that intensity, sharpness, and harmonic ratio had a strong effect on annoyance. Sharpness and harmonic ratio influenced people in a manner that is generally similar to the effects of onset, offset, and harmonic series in these experiments. Likewise, “approaching” sounds were judged as more annoying in a way that is similar to the effects of alert onset and offset in this study (Nixon et al., 1969). A critical caveat regarding Table 1 is that it reflects only two levels of each alert parameter. Systematic exploration of the full range of parameter values

TABLE 1: Effects of Parameters on Perceived Urgency and Annoyance

Characteristic	Effect on Urgency	Effect on Annoyance
Pulse duration	Longer more urgent (0.80)	Longer more annoying (0.28)
Interpulse interval	Longer less urgent (0.92)	Longer less annoying (0.62)
Harmonic series (formants)	High /a/ more urgent (0.50)	High /a/ more annoying (0.66)
Alert onset	Fast more urgent (0.66)	Fast more annoying (0.42)
Alert offset	Fast more urgent (0.66)	Fast more annoying (0.22)
Burst duty cycle	High more urgent (0.44)	High more annoying (0.24)
Alert duty cycle	High more urgent (1.36)	High more annoying (1.0)
Interburst period	Shorter more urgent (0.50)	Shorter more annoying (0.30)
Sound type	Frequency series more urgent (0.26)	Frequency series less annoying (–0.40)

Note. The difference in *z* scores is shown in parentheses. Parameters having a differential effect on the *z* score of urgency and annoyance greater than 0.3 are in italics.

and power law quantification of their effects is needed. Such experiments have been conducted to quantify the effects of alert parameters on perceived urgency (Hellier & Edworthy, 1999), and the results of this study suggest that a similar effort to quantify the effects on annoyance may be worthwhile.

Consistent with previous findings regarding annoyance (Fucci, Petrosino, Hallowell, et al., 1997), this study shows that perceived urgency and annoyance depend on both the auditory characteristics of the alert and the context. The weak implementation of driving context in this study, through imagined scenarios with no actual driving activity, suggests that driving context may play an important role in how people respond to alerts. If imagined scenarios affect responses, then real scenarios are likely to have a much larger effect.

The frequency of false alarms may be a critical element of the driving context affecting how people respond to alerts. A mildly annoying alert may become extremely annoying if many false alarms occur. Context effects, such as the rate of false alarms, may dominate perceived annoyance (Tan & Lerner, 1995). The results of this study provide a promising, although tentative, approach to this challenge. Likelihood alarm displays, in which the alert is graded according to the likelihood that a response is needed, have long been advocated as a response to the false alarm problem (Sorkin, Kantowitz, & Kantowitz, 1988). The urgency mapping principle provides one approach to implementing graded alerts (Bliss, Dunn, et al., 1995; Hellier & Edworthy, 1999). This study shows that certain parameters may increase urgency with a relatively small increase in annoyance, which may further reduce the annoyance associated with false alarms. Additional research needs to verify whether the auditory parameters identified in this study are able to increase urgency with minimal effect on annoyance in the context of frequent false alarms in actual driving situations.

In addition to the problem of frequent false alarms, the proliferation of in-vehicle systems represents an important design challenge. This study shows that the influence of perceived urgency and annoyance on judgments of alert appropriateness depends on the driving context and the relevant in-vehicle information system. Similar to the results of Tan and Lerner (1995), ratings of urgency in this study were highly correlated with ratings of appropriateness for a collision warning system, but annoyance was not. In contrast, urgency was weakly related to appropriateness for an E-mail system, but annoyance was highly correlated with appropriateness. The analysis of urgency and annoyance relative to appropriateness for the three driving scenarios suggests that the urgency mapping principle should be augmented with a consideration of annoyance. Specifically, minimizing annoyance should take precedence over urgency mapping for less critical alerts. These results complement those of a recent simulator study that showed a strong positive association between subjective workload and annoyance as compared with a weaker association between urgency and subjective workload (Wiese & Lee, 2004). The degree to which annoyance dominates perceived appropriateness of alerts for low-urgency scenarios needs to be assessed in a greater range of actual driving scenarios.

This study examined only auditory parameters, but other modalities, such as haptic interfaces, may be able to convey greater urgency with less annoyance (Lee, Hoffman, & Hayes, 2004; Sklar & Sarter, 1999). Multimodal alerts may offer a useful means of increasing urgency while minimizing annoyance. The degree to which parameters affecting urgency and annoyance of auditory alerts also affect multimodal alerts remains an open question. Preliminary evidence suggests that increasing the magnitude of a brake pulse used as a collision warning increased drivers' braking response, much like increasing the intensity of an auditory alert increases perceived urgency and decreases reaction time (Tijerina et al., 2000). Likewise, recent research on cross-modal links among visual, auditory, and tactile sensory systems suggests multimodal investigation of annoyance and urgency may be fruitful (Spence, 2002).

ACKNOWLEDGMENTS

We would like to thank Dr. John Campbell for his comments during all phases of this research. This paper was based on experiments described in a thesis (Austria, 2000), a technical report (Lee, Marshall, Austria, Wiese, & Williams, 2002), and a conference paper (D. Marshall, Lee, & Austria, 2001). The paper presents research conducted in collaboration with Battelle's Human Factors Transportation Center under Contract No. DTFH61-97-C-00061 to the Federal Highway Administration. Nazemeh Sobhi was the government technical representative for the contract until January 1999 and was followed by Tom Granda. We would also like to thank Emily Wiese for her comments and Professor Christopher Turner at the University of Iowa Speech Pathology and Audiology Department for his assistance in measuring sound levels.

REFERENCES

Austria, P. A. (2000). *Perceived urgency and perceived annoyance of in-vehicle auditory warning signals.* Iowa City: University of Iowa.

Berglund, B., & Preis, A. (1997). Is perceived annoyance more subject-dependent than perceived loudness? *Acoustica, 83*(2), 313–319.

Bliss, J., Dunn, M., & Fuller, B. S. (1995). Reversal of the cry-wolf effect: An investigation of two methods to increase alarm response rates. *Perceptual and Motor Skills, 80,* 1231–1242.

Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behavior in response to alarms of varying reliability. *Ergonomics, 38,* 2300–2312.

Block, F. E., Nuutinen, L., & Ballast, B. (1999). Optimization of alarms: A study on alarm limits, alarm sounds, and false alarms, intended to reduce annoyance. *Journal of Clinical Monitoring and Computing, 15,* 75–83.

Boucek, G. P., Jr., Veitengruber, J. E., & Smith, W. D. (1977). *Human factors guidelines for aircraft alerting systems: Aircraft alerting systems criteria study: Vol. II* (No. FAA-RD-76-222, II). Washington, DC: U.S. Department of Transportation.

Burt, J. L., Bartolome, D. S., Burdette, D. W., & Comstock, J. R. (1995). A psychophysiological evaluation of the perceived urgency of auditory warning signals. *Ergonomics, 38,* 2327–2340.

Edworthy, J., & Adams, A. (1996). *Warning design: A research perspective.* Bristol, PA: Taylor & Francis.

Edworthy, J., Hellier, E., & Walters, K. (2000). The relationship between task performance, reaction time, and perceived urgency in nonverbal auditory warnings. In *Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 3.674–3.677). Santa Monica, CA: Human Factors and Ergonomics Society.

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design – Relationship between warning sound parameters and perceived urgency. *Human Factors, 33,* 205–231.

Edworthy, J., & Stanton, N. (1995). A user-centered approach to the design and evaluation of auditory warning signals: 1. Methodology. *Ergonomics, 38,* 2262–2280.

Evans, L. (2004). *Traffic safety.* Bloomfield Hills, MI: Science Serving Society.

Fucci, D., Petrosino, L., Hallowell, B., Andra, L., & Wilcox, C. (1997). Magnitude estimation scaling of annoyance in response to rock music: Effects of sex and listeners' preference. *Perceptual and Motor Skills, 84,* 663–670.

Fucci, D., Petrosino, L., McColl, D., Wyatt, D., & Wilcox, C. (1997). Magnitude estimation scaling of the loudness of a wide range of auditory stimuli. *Perceptual and Motor Skills, 85,* 1059–1066.

Graham, R. (1999). Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics, 42,* 1233–1248.

Haas, E. C., & Casali, J. G. (1995). Perceived urgency of and response time to multi-tone and frequency-modulated warning signals in broadband noise. *Ergonomics, 38,* 2313–2326.

Haas, E. C., & Edworthy, J. (1996). Designing urgency into auditory warnings using pitch, speed and loudness. *Computing and Control Engineering Journal, 7*(4), 193–198.

Hellier, E. J., & Edworthy, J. (1999). On using psychophysical techniques to achieve urgency mapping in auditory warnings. *Applied Ergonomics, 30,* 167–171.

Hellier, E. J., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design – Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors, 35,* 693–706.

Hellier, E. J., Edworthy, J., & Dennis, I. (1995). A comparison of different techniques for scaling perceived urgency. *Ergonomics, 38,* 659–670.

Khan, M. S., Johansson, O., & Sundback, U. (1997). Development of an annoyance index for heavy-duty diesel engine noise using multivariate analysis. *Noise Control Engineering Journal, 45,* 157–167.

Kiefer, R., LeBlanc, D., Palmer, M., Salinger, J., Deering, R., & Shulman, M. (1999). *Development and validation of functional definitions and evaluation procedures for collision warning/avoidance systems* (No. DOT HS 808 964). Washington, DC: Crash Avoidance Metrics Partnership.

Kryter, K. D. (1985). *The effects of noise on man.* London: Academic Press.

Laird, D. A., & Coye, K. (1929). Psychological measurements of annoyance related to pitch and loudness. *Journal of the Acoustical Society of America, 1,* 156–163.

Lee, J. D., Hoffman, J. D., & Hayes, E. (2004). Collision warning design to mitigate driver distraction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 65–72). New York: ACM Press.

Lee, J. D., Marshall, D., Austria, P. A., Wiese, E., & Williams, E. (2002). *In-vehicle display icons and other information elements: Task E, Experiment 3. Urgency and annoyance of auditory alerts* (No. DTFH6197-C-00061). Iowa City: University of Iowa.

Lee, J. D., McGehee, D. V., Brown, T. L., & Reyes, M. L. (2002). Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. *Human Factors, 44,* 314–334.

Marshall, D., Lee, J. D., & Austria, A. (2001). Annoyance and urgency of auditory alerts for in-vehicle information systems. In *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 1627–1631). Santa Monica, CA: Human Factors and Ergonomics Society.

Marshall, E., & Baker, S. (1994). Alarms in nuclear power plant control rooms: Current approaches and future design. In N. Stanton (Ed.), *Human factors in alarm design* (pp. 183–191). London: Taylor & Francis.

Meredith, C., & Edworthy, J. (1995). Are there too many alarms in the intensive-care unit: An overview of the problems. *Journal of Advanced Nursing, 21,* 15–20.

Mermelstein, P. (1967). Determinism of the vocal-tract shape from measured formant frequencies. *Journal of the Acoustical Society of America, 41,* 1283–1294.

Nixon, C. W., Von Gierke, H. E., & Rosinger, G. (1969). Comparative annoyances of "approaching" versus "receding" sound sources. *Journal of the Acoustical Society of America, 45,* 843–853.

Sklar, A. E., & Sarter, N. B. (1999). Good vibrations: Tactile feedback in support of attention allocation and human automation coordination in event-driven domains. *Human Factors, 41,* 543–552.

Sorkin, R. D., Kantowitz, B. H., & Kantowitz, S. C. (1988). Likelihood alarm displays. *Human Factors, 30,* 445–459.

Spence, C. (2002). Multisensory attention and tactile information-processing. *Behavioural Brain Research, 135,* 57–64.

Stanford, L. M., McIntyer, J. W. R., Nelson, T. M., & Hogan, J. T. (1988). Affective responses to commercial and experimental auditory alarm signals for anaesthesia delivery and physiological monitoring equipment. *International Journal of Clinical Monitoring and Computing, 5,* 111–118.

Tan, A., & Lerner, N. (1995). *Multiple attribute evaluation of auditory warning signals for in-vehicle crash warning systems* (No. DOT HS 808 535). Washington, DC: National Highway Transportation Safety Administration.

Tijerina, L., Johnston, S., Parmer, E., Pham, H. A., Winterbottom, M. D., & Barickman, F. S. (2000). *Preliminary studies in haptic displays for rear-end collision avoidance system and adaptive cruise control applications.* Washington, DC: National Highway Transportation Safety Administration.

Wiese, E. E., & Lee, J. D. (2004). Effects of multiple auditory alerts for in-vehicle information systems on driver attitudes and performance. *Ergonomics, 47,* 965–986.

Woods, D. (1994). Cognitive demands and activities in dynamic fault management: Abductive reasoning and disturbance management. In N. Stanton (Ed.), *Human factors in alarm design* (pp. 63–92). London: Taylor & Francis.

Woodworth, R. S., & Schlosberg, H. (1958). *Experimental psychology.* New York: Henry Holt.

Dawn C. Marshall is a researcher scientist at the National Advanced Driving Simulator, Coralville, Iowa. She received her M.S. in industrial engineering from the University of Iowa in 2004.

John D. Lee is a professor in the Department of Mechanical and Industrial Engineering at the University of Iowa. He received his Ph.D. in mechanical engineering from the University of Illinois in 1992.

P. Albert Austria is a research engineer at Toyota Technical Center, Ann Arbor, Michigan. He received his M.S. in industrial engineering from the University of Iowa in 2000.

Date received: October 30, 2003
Date accepted: November 21, 2005