

Evaluating Informative Auditory and Tactile Cues for In-Vehicle Information Systems

Yujia Cao, Frans van der Sluis, Mariët Theune, Rieks op den Akker, Anton Nijholt

Human Media Interaction Group

University of Twente

P.O. Box 217, 7500 AE, Enschede, The Netherlands

{y.cao, f.vandersluis, m.theune, infrieiks, a.nijholt}@utwente.nl

ABSTRACT

As in-vehicle information systems are increasingly able to obtain and deliver information, driver distraction becomes a larger concern. In this paper we propose that informative interruption cues (IIC) can be an effective means to support drivers' attention management. As a first step, we investigated the design and presentation modality of IIC that conveyed not only the arrival but also the priority level of a message. Both sound and vibration cues were created for four different priority levels and tested in 5 task conditions that simulated possible perceptual and cognitive load in real driving situations. Results showed that the cues were quickly learned, reliably detected, and quickly and accurately identified. Vibration was found to be a promising alternative for sound to deliver IIC, as vibration cues were identified more accurately and interfered less with driving. Sound cues also had advantages in terms of shorter response time and more (reported) physical comfort.

Categories and Subject Descriptors

H.5.2 [Information interfaces and presentation]: User Interfaces, Auditory feedback, Haptic I/O

Keywords

Multimodal interfaces, interruption management, in-vehicle information systems

1. INTRODUCTION

In-vehicle information systems (IVIS) are primarily intended to assist driving by providing supplementary information to drivers in real time. IVIS have a wide variety of functions [19], such as route planning, navigation, vehicle monitoring, traffic and weather updates, hazard warning, augmented signing and motorist service. The development of Car2X communication technology¹ will allow many more functions to become available in the near future. Moreover, when in-car computers have access to wireless internet,

¹Car2X technology uses mobile ad hoc networks to allow cars to communicate with other cars and infrastructures [13].

IVIS can also assist drivers in tasks that are not driving related, such as email management [10]. However, these IVIS functions could be potentially harmful to driving safety because they impose extra attention demands on the driver and cause distraction to a certain extent. According to a recent large-scale field study conducted over a period of one year [16], 78% of traffic collisions and 65% of near collisions were associated with drivers' inattention to the road ahead, and the main source of this inattention was found to be secondary tasks distraction, such as interacting with IVIS. Therefore, the design of IVIS should aim to maximize benefits and minimize distraction. To this end, IVIS need to interrupt in a way that supports drivers' attention allocation between multiple tasks.

Supporting users' attention and interruption management has been a design concern of many human-machine systems in complex event-driven domains. One promising method is to support pre-attentive reference² by providing informative interruption cues (IIC) [5, 6, 18]. IIC differ from non-informative interruption cues, because they do not only announce the arrival of events but also (and more importantly) present partial information about the nature and significance of the events in an effort to allow users to decide whether and when to shift their attention. In a study using an air traffic control task [6], IIC were provided to convey the urgency and modality of pending tasks. Results showed that the IIC were highly valued by participants. The modality of a pending task was used to decide when to perform it in order to reduce visual scanning costs associated with the concurrent air traffic control tasks. Another study on monitoring a water control system applied IIC to present the domain, importance and duration of pending tasks [5]. Participants used the importance to decide whether and when to attend to pending tasks. These findings demonstrate that IIC can be used successfully by operators in complex event-driven environments to improve their attention and interruption management.

In the automotive domain, a large number of studies have been carried out on the design and presentation modality of IVIS messages (e.g. [3, 7, 9, 11, 12, 21]). However, using pre-attentive reference to help drivers selectively attend to these messages has rarely been investigated. As IVIS are increasingly able to obtain and deliver information, we propose that IIC could be an effective means to minimize inappropriate distractions. IIC inform drivers about the arrival of new IVIS messages and their *priority levels* associated with urgency and relevance to driving. The perception and understanding of IIC should require minimum time and attention resources. Upon reception of IIC, drivers can have control over

²Pre-attentive reference is to evaluate attention directing signals with minimum attention [26].

whether to attend to, postpone or ignore the messages, depending on the driving demand at the moment. Some messages can be delivered “on demand”, meaning that they are not delivered together with IIC but only when the driver is ready (or willing) to switch attention. In this way, the system supports drivers to manage their attention between multiple tasks. To evaluate this proposal, we first need to obtain a set of IIC that meets the criteria of pre-attentive reference [26] and is suited for the driving environment. Then, we need to investigate whether drivers can indeed make use of these IIC to better manage their attention between driving and IVIS messages. This paper presents a study that served as the first step – the design and evaluation of IIC.

Based on the criteria of pre-attentive reference [26], a set of requirements was collected for the design and evaluation of IIC. First, IIC need to be picked up in parallel with ongoing tasks and activities. Since driving is mainly a visual task, auditory or tactile cues can be better perceived in parallel with driving, because they consume separate perception resources [24]. Second, IIC should provide information on the significance and meaning of the interruption, which in our case is the priority of an IVIS message. Third, IIC should allow for evaluation by the user with minimum time and attention. This means that regardless of the message priority a cue intends to convey, the cue should always be interpreted quickly and easily. Finally, IIC need to be effective in all driving conditions. The attention demand of driving may differ with road, traffic, and weather conditions. The driver can be distracted by radio, music, or conversation with other passengers. Noise in the driving environment may also hinder the detection of cues. Therefore, the effectiveness of our cues needs to be evaluated in various conditions.

Sound has been a preferred modality for alerts and interruption cues [25]. However, in environments with rich surrounding sounds, sound cues can go unnoticed. Alternatively, the tactile modality may be more effective for cue delivery [17, 20]. In automotive studies, tactile modalities such as vibration and force were typically used to present alerts and directions (e.g. [22, 23]). Such tactile signals served as messages but not IIC. Besides, the main informative factor was the location where signals were provided. For example, a vibration signal on the left side of the steering wheel warned the driver of lane departure from the left side [22]. In this study, we intended to apply vibration signals as IIC and convey message priority by the *pattern* of vibrations rather than the location.

The objective of this study was twofold: 1) to evaluate the design of our cues, including whether they are easy to learn and whether they can be quickly and accurately identified under various types of cognitive load that drivers can encounter during driving; and 2) to compare sound and vibration, aiming to find out which modality is more suitable under which conditions. The remainder of this paper presents the design of the sound and vibration cues, describes the evaluation method, discusses the results and finally presents our conclusions from the findings.

2. DESIGN OF SOUND AND VIBRATION CUES

We first set up 4 priority levels, numbered from 1 (the highest level) to 4 (the lowest level), and then associated each priority level with certain types of IVIS messages. In fact, any application could make its own associations. In this study, levels 1 and 2 were associated with driving-related information. Level 1 could be assigned to hazard warnings and other urgent messages about traffic or vehicle condition. Level 2 covered less urgent driving-related information,

such as a low fuel level or low air pressure in the tires. Levels 3 and 4 were associated with IVIS functions that were not related to driving, such as email, phone calls and in-vehicle infotainment. Then, the aim of cue design was to construct intuitive mappings between features of sound/vibration and the associated priority levels. In other words, the signal-priority mappings should be natural so that they can be learnt with minimum effort.

Sound cues. Sounds are commonly categorized into two groups: *auditory icons* that are environmental sounds imitating real-world events and *earcons* that are abstract and synthetic sound patterns. For this study, we chose earcons for the following reasons: 1) earcons offer the flexibility to create different patterns for expressing different meanings, 2) parameters of earcons are known to be associated with perceived urgency, and 3) earcons can share common patterns with vibrations, which allows a better comparison between sound and vibration. Previous studies on the relation between sound parameters and perceived urgency [1, 4, 14, 15] commonly suggest that sound signals with higher pitch, more pulses, and faster pace (shorter inter-pulse interval) are generally perceived as more urgent.

We did not rely on only one feature to convey priority. To reinforce the effectiveness, we combined pitch, number of beeps and pace, and manipulated them in a unified manner. The four sound cues are illustrated in the right column of Figure 2. From priority 1 to 4, the pitch of the sounds was respectively 800Hz, 600Hz, 400Hz and 300Hz. We kept the pitches in this range because lower pitches were not salient and higher pitches might induce discomfort. All sounds were designed with the same length because duration was not a manipulated feature in this study, and also because this way reaction times to the cues could be better evaluated. Given the fixed duration, the number of pulses and pace were two dependent features, that is to say more pulses in the same duration leads to faster pace. From priority 1 to 4, the number of pulses was respectively 8, 6, 4 and 3, resulting in decreasing paces.

Vibration cues. Several studies have investigated the relation between vibration parameters and perceived priority [2, 5, 8]. Results showed that signals were perceived as more important/urgent when they had higher intensities, a higher number of pulses, and a higher pulse frequency (number of pulses per second). As with sound, we also combined three features of vibration: intensity, number of pulses, and pace. Vibration signals were provided by a vibration motor mounted beneath the seat of a chair (Figure 1). This location was chosen because the seat is always in full contact with driver’s body. The vibration motor was taken from an Aura Interactor³. It is a high force electromagnetic actuator (HFA), which takes a sound stream as input and generates a fast and precise vibration response according to the frequency and amplitude of the sound. Four sound input signals were created for the vibration cues (the left column of Figure 2). They had the same frequency (50Hz) and length, but different amplitudes that led to different vibration intensities. The intensity for priority 1, 2 and 3 was respectively 2.25, 1.75 and 1.25 times the intensity for priority 4. The number of pulses was also 8, 6, 4 and 3, resulting in decreasing paces from priority 1 to 4.

3. EVALUATION METHOD

At this step, the evaluation was only focused on the design of cues. It did not involve the IVIS messages or the attention management

³Aura Interactor. <http://apple2.org.za/gswv/a2zine/GS.WorldView/Resources/A2GS.NEW.PRODUCTS/Sound.Interactor.html>



Figure 1: Vibration motor mounted beneath the seat of an office chair.

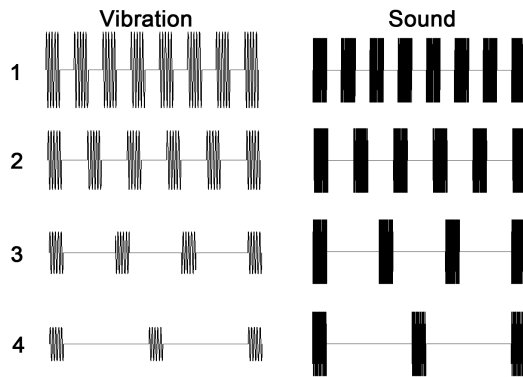


Figure 2: Illustration of the design of sound cues and vibration cues. Numbers indicate priority levels. The x and y axes of each oscillogram are time and amplitude. The duration of each cue is 1.5 seconds. The duration of each pulse is 100 milliseconds.

between driving and the messages. Therefore, we chose to mimic the driving load in a laboratory setting, in order to more precisely control the manipulated variances of task load between conditions. The task set mimicked various types of cognitive load that drivers could encounter during driving. Although this did not exactly resemble a driving situation, it did ensure that all conditions are the same for all participants, leading to more reliable results for the purpose of this evaluation.

3.1 Tasks

Visual tracking task. Since driving mostly requires visual attention, this task was employed to induce a continuous visual perception load (to keep the eyes on the “road”). Participants were asked to follow a moving square with a mouse cursor (the center blue square in Figure 3). The size of the square was 50 pixel × 50 pixel on a 20” monitor with 1400×1050 resolution. Most of the time, the square moved in a straight line along the x and y axis. At random places, it made turns of 90 degrees. To provide feedback on the tracking performance, the cursor turned into a smiley when it entered the blue square area.

There were 10 tracking trials in total and each of them lasted for 2 minutes. The participants were instructed to maintain the tracking performance throughout the whole experiment, just as drivers should continue to drive on the road. Although this task does not involve vehicle control, it does share common characteristics with driving – the performance may decrease when people pay less attention to watching the “road”, and the visual perception demand of the task can vary in different conditions.

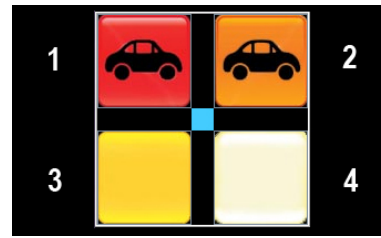


Figure 3: The tracking target square and the 4 answer buttons for cue identification response. The numbers have been added to indicate priority levels and were not present in the experiment.

Table 1: The five task conditions applied in the experiment.

Condition Index	1	2	3	4	5
Cue Identification	×	×	×	×	×
Low-Load Tracking	×		×	×	×
High-Load Tracking		×			
Radio			×		
Conversation				×	
Noise					×

Cue identification task. During the course of tracking, cues were delivered with random intervals. Upon the reception of a cue, participants were asked to click on an answer button as soon as they identified the priority level. This task aimed to show how quickly and accurately participants could identify the cues. Four answer buttons were made for this task, one for each priority level (Figure 3). Buttons were color-coded to intuitively present different priority levels. A car icon was placed on the buttons of the driving-related levels. There were always 8 cues (2 modalities × 4 priorities) delivered in a randomized order during a 2-minute tracking trial. Note that the four buttons always moved together with the center tracking square. In this way, cue identification imposed minimal influence on the tracking performance.

3.2 Task conditions

We set up 5 task conditions (summarized in Table 1), attempting to mimic 5 possible driving situations.

Condition 1 (low-load tracking) attempted to mimic an easy driving situation with a low demand on visual perception. The tracking target moved at a speed of 50 pixels per second. During a 2-minute course, the target made 8 turns of 90 degrees, otherwise moving in a straight line. The turning position and direction differed in each course. This tracking task setting was also applied to conditions 3, 4 and 5.

Condition 2 (high-load tracking) attempted to mimic a difficult driving situation where the visual attention was heavily taxed, such as driving in heavy traffic. The tracking target moved at a speed of 200 pixels per second. During a 2-minute course, the target made 32 turns of 90 degrees, otherwise moving in a straight line. The turning position and direction differed in each course.

Condition 3 (radio) attempted to mimic driving while listening to the radio. Two recorded segments of a radio program were played in this condition. Both segments contained a conversation between a male host and a female guest about one kind of sport (marathon and tree climbing). Participants were instructed to pay attention

to the conversation while tracking the target and identifying cues. They were also informed about receiving questions regarding the content of the conversation later on.

Condition 4 (conversation) attempted to mimic driving while talking with other passengers. Four topics were raised during a 2-minute trial via a text-to speech engine. They were all casual topics, such as vacation, favorite food, weather and the like. Participants had about 25 seconds to talk about each topic. They were instructed to keep talking until the next topic was raised and generally succeeded in doing so.

Condition 5 (noise) attempted to mimic driving in a noisy condition or on a rough road surface. For auditory noise, we played recorded sound of driving on the highway or on a rough surface. The signal to noise ratio was approximately 1:1. Tactile noise was generated by sending pink noise⁴ into the vibration motor. The tactile noise closely resembled the bumpy feeling when driving on a bad road surface, which was verified with a pilot study using 3 subjects. The signal to noise ratios for priorities 1 to 4 were approximately 3:1, 2:1, 1:1 and 0.6:1. Both auditory and tactile noise were always present in this condition.

3.3 Subjects and Procedure

Thirty participants, 15 male and 15 female, took part in the experiment. Their age ranged between 19 to 57 years old (mean = 31.6, SD = 9.6). None of them reported any problem with hearing or tactile perception. The experiment consisted of two sessions: a cue-learning session and a task session. After receiving a short introduction, participants started off with learning the sound and vibration cues. They could click on 8 buttons to play the sounds or trigger the vibrations, in any order they wanted and as many times as needed. Learning ended when participants felt confident in being able to identify each cue when presented separately from the others. Then, a test was carried out to assess how well they had managed to learn the cues. Feedback on performance was given to reinforce learning. At the end of this session, participants filled in a questionnaire about their learning experience. In the task session, each participant performed 10 task trials (2 modalities × 5 conditions) of 2 minutes each. The trial order was counterbalanced. Feedback on cue identification performance was no longer given. During the short break between two trials, participants filled in questionnaires assessing the task load and the two modalities in this previous trial. At the end of the experiment, participants filled in a final questionnaire reporting physical comfort of the signals and the use of features.

3.4 Measures

For the cue-learning session, two performance and one subjective measures were applied. 1) *Amount of learning*: the number of times participants played the sounds and triggered the vibrations before they reported that they were ready for the cue identification test. 2) *Cue identification accuracy*: the accuracy of cue identification in the test. 3) *Association of features with priorities*: the subjective judgements on how easy/intuitive it was to infer priorities from each feature. Participants rated the 6 features separately on a Likert scale from 1 (very easy) to 10 (very difficult). Although the pace and the number of pulses were two dependent features in our setting, we still introduced and evaluated both of them, because they

⁴Pink noise is a signal with a frequency spectrum such that the power spectral density is inversely proportional to the frequency.

were different from a perception standpoint. Drivers may find one of them more effective and reliable than the other.

For the task session, four performance measures were employed: 1) *Tracking error*: distance between the position of the cursor and the center of the target square (in pixels). Instead of taking the whole course, average values were only calculated from the onset of each cue to the associated button-click response. In this way, this measure better reflected how much cue identification interfered with tracking. 2) *Cue detection failure*: the number of times a cue was not detected. 3) *Cue identification accuracy*: the accuracy of cue identification in each task trial. 4) *Response time*: the time interval between the onset of a cue to the moment of the button-click response.

Four subjective measures were derived from the between-trial questionnaire (1 and 2) and the final questionnaire (3 and 4). 1) *Ease of cue identification*: how easy it was to identify cues in each task trial. Participants rated each task trial on a Likert scale from 1 (very easy) to 10 (very difficult). 2) *Modality preference*: which modality was preferred for each task condition. Participants could choose between sound, vibration and either one (equally prefer both). 3) *Physical comfort*: how physically comfortable the sound and vibration signals made them feel. Participants rated the two modalities separately on a Likert scale from 1 (very comfortable) to 10 (very annoying). 4) *Features used*: which features of sound and vibration participants relied on to identify the cues. Multiple features could be chosen.

Table 2: Summary of measures. P: performance measures; S: subjective measures.

Session	Measures
Learning	Amount of learning (P)
	Cue identification accuracy (P)
	Association of features with priorities (S)
Task	Tracking error (P)
	Cue detection failure (P)
	Cue identification accuracy (P)
	Response time (P)
	Ease of cue identification (S)
	Modality preference (S)
	Physical comfort (S)
Features used (S)	

4. RESULTS

4.1 Cue Learning Session

Amount of learning. The number of times participants played the cues ranged from 12 to 32. On average, cues were played 18.7 times, 9.4 times for sounds and 9.7 times for vibrations. Comparing the four priority levels, ANOVA showed that participants spent significantly more learning effort on level 2 and 3 (5.6 and 5.3 times) than on level 1 and 4 (4.0 and 3.8 times), $F(3, 27) = 20.0, p < .001$.

Cue identification accuracy. Participants showed high performances in the cue identification test. Twenty-five participants did not make any error in the 16 tasks. The other 5 participants made no more than 3 errors each, mostly at priority levels 2 and 3. On average, the identification accuracy reached 97.8% for sound cues, 99.2% for vibration cues and 98.5% overall.

Association of features with priorities. The average rating scores of the 6 features all fell below 3.5 on the 10-level scale (Figure 4), indicating that participants found it fairly easy and intuitive to infer priorities from these features. Sound features were rated as more intuitive than vibration features ($F(1, 30) = 6.0, p < .05$). For both sound and vibration, pace was rated as the most intuitive feature (sound: mean = 2.2; vibration: mean = 2.7). Number of pulses was rated as significantly less intuitive than pace (mean = 3.3 for both sound and vibration).

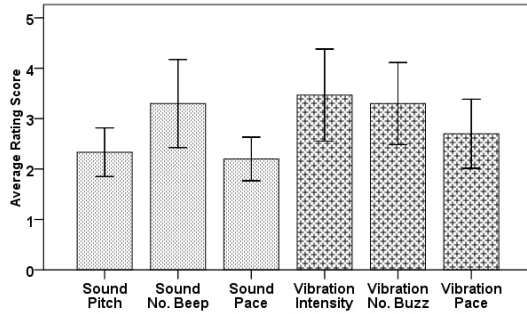


Figure 4: Rating scores on the easiness of associating variations in each feature with the corresponding priority levels. Error bars represent standard errors. (1 = easiest, 10 = most difficult)

4.2 Task Session

Tracking error. Average tracking errors in each condition are shown in Figure 5. A three-way repeated-measure ANOVA showed that modality, condition and priority level all had a significant influence on the tracking performance (modality: $F(1, 29) = 10.6, p < .01$; condition: $F(4, 26) = 81.5, p < .001$; priority: $F(3, 27) = 34.4, p < .001$). Tracking error was significantly lower when cues were delivered by vibration than by sound. Comparing the 5 conditions, tracking error was significantly the highest in the high-load tracking condition. When tracking load was low, the conversation condition induced significantly higher tracking error than the other 3, between which no difference was found. Among the 4 priority levels, tracking was significantly more disturbed by cues at the higher two priority levels than by cues at the lower two priority levels. This result makes sense because the cues at higher priority levels are more salient and intense; therefore they are more able to drag attention away from tracking during their presentation. Finally, there was also an interaction effect between modality and condition ($F(4, 26) = 5.9, p < .01$), indicating that vibration was particularly beneficial in the low-load tracking condition.

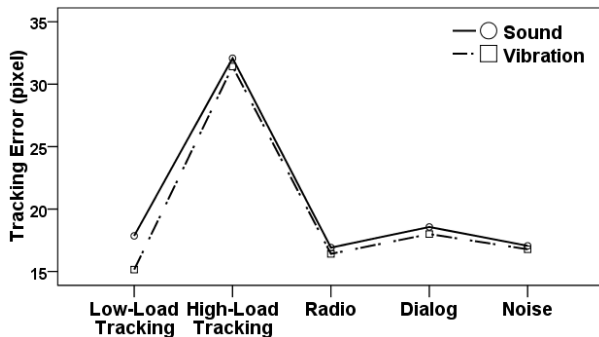


Figure 5: Average tracking error by task condition and modality.

Cue detection failure. Cue detection failure occurred 6 times, which was 0.25% of the total number of cues delivered to all participants in the task session ($30 \times 10 \times 8 = 2400$). One failure occurred in the conversation condition. The missed cue was a vibration cue at priority level 2. All the other 5 failures occurred in the noise condition. The missed cues were all vibration cues at the lowest priority level. However, considering the fact that the signal-to-noise ratio for level-4 vibration was below 1, only 5 misses (8.3%) is still a reasonably good performance.

Cue identification accuracy. The average accuracy over all task trials was 92.5% (Figure 6), which was lower than the performance in the learning test when cue identification was the only task to perform. A three-way repeated-measure ANOVA revealed that modality, task condition and priority level all had a significant influence on the cue identification accuracy (modality: $F(1, 29) = 5.5, p < .05$; condition: $F(4, 26) = 11.7, p < .001$; priority: $F(3, 27) = 9.3, p < .001$). Identifying vibration cues was significantly more accurate than sound cues. Taking the low-load tracking condition as a baseline (99.0% accurate), all types of additional load significantly decreased accuracy, among which conversation decreased accuracy the most (12.7% less). Comparing priority levels, levels 1 and 4 were identified more accurately than levels 2 and 3, presumably because their features were more distinguishable. We further analyzed the error distribution over the four priority levels. It turned out that errors only occurred between two successive levels, such as between levels 1-and-2, 2-and-3, and 3-and-4. This result reveals that in design cases where only two priority levels are needed, it would be very promising to apply the current cue design from any two disconnected levels (e.g. 1-3, 1-4, 2-4).

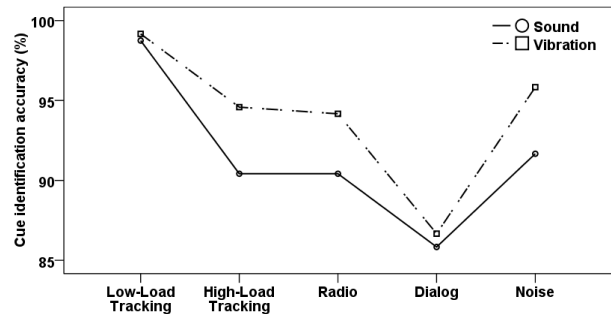


Figure 6: Cue identification accuracy by task condition and modality.

Response time. All identification responses were given within 5 seconds from the onset of cues (min. = 1.6s, max. = 4.5s, mean = 2.8s). A three-way repeated-measure ANOVA again showed that modality, task condition and priority level all had a significant influence on this measure (modality: $F(1, 29) = 5.9, p < .05$; condition: $F(4, 26) = 10.6, p < .001$; priority: $F(3, 27) = 36.9, p < .001$). Generally, identifying sound cues was significantly faster than identifying vibration cues. However, one exception was in the radio condition (see Figure 7), causing an interaction effect between modality and task conditions ($F(4, 26) = 4.1, p < .01$). Comparing conditions, response was the fastest in the low-load condition and the second fastest in the noise condition. No significant difference was found between the other 3 conditions. Regarding priority levels, there was a general trend of “higher priority, faster response”, suggesting that participants indeed perceived the associated levels of priority/urgency from the cues. Level-1 cues were identified significantly faster than the others. The level-1 sound was particularly

able to trigger fast responses, possibly because it closely resembled alarms.

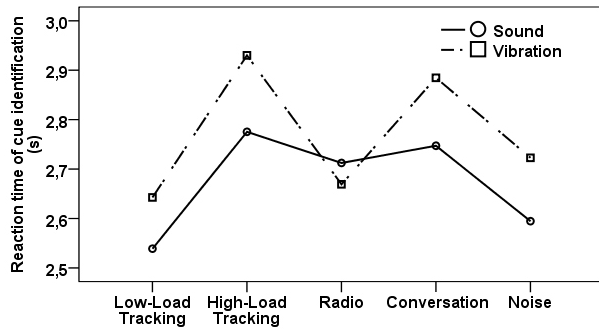


Figure 7: Response time by priority level and modality.

Ease of cue identification. Average rating scores for each condition are shown in Figure 8. ANOVA showed that rating scores were significantly influenced by task condition ($F(4, 26) = 35.3, p < .001$) but not by modality ($F(1, 29) = 3.4, p = n.s.$). Not surprisingly, cue identification was significantly the easiest in the low-load tracking condition. In contrast, identifying cues while having a conversation was significantly the most difficult, which was in line with the lowest accuracy in this condition (Figure 6). There was also a significant difference between the radio and the high-load condition.

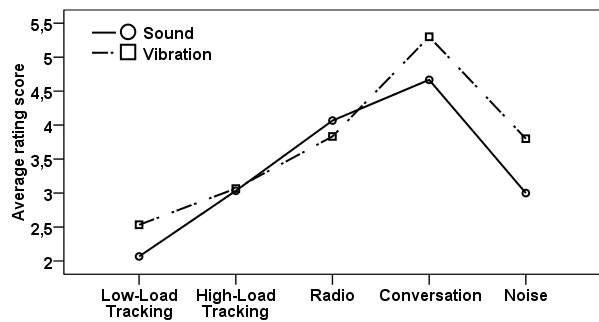


Figure 8: Subjective ratings on the ease of cue identification in each task trial. (1 = easiest, 10 = most difficult)

Modality preference. Table 3 summarizes the number of participants who preferred sound or vibration or either of these in each task condition. Except in the radio condition, sound was preferred by more participants than vibration.

Table 3: Number of participants who preferred sound or vibration or either of these in each task condition.

	Sound	Vibration	Either one
Low-Load Tracking	19	4	7
High-Load Tracking	15	7	8
Radio	13	15	2
Conversation	12	11	7
Noise	19	9	2

Physical comfort. The average scores of sound and vibration both fell below 4 on the 10-level scale (sound: mean = 2.9, sd = 1.5; vibration: mean = 3.8, sd = 1.8). A paired-sample t-test revealed a significant difference between the two modalities ($t(29) = 2.2,$

$p < .05$), indicating that participants found sound cues more comfortable than vibration cues.

Features used. Participants made multiple choices on the feature(s) they relied on to identify the cues. As Table 4 shows, a majority of participants (90%) made use of more than one sound feature and more than one vibration feature. This result suggests that combining multiple features in a systematic way was useful. This design choice also provided room for each participant to selectively make use of those features that he/she found most intuitive and reliable. Table 5 further shows how many participants used each feature.

Table 4: The number of features used to identify cues.

No. of features used	Sound			Vibration		
	1	2	3	1	2	3
No. of participants	3	20	7	3	22	5

Table 5: Number of participants who used each feature to identify cues.

Pitch	Sound		Vibration		
	No. of beeps	Pace	Intensity	No. of pulses	Pace
19	26	17	14	26	22

4.3 Discussion

With respect to our research objectives, the results are discussed from two aspects: the effectiveness of cue design and the modality choice between sound and vibration.

Effectiveness of cue design. Various measures consistently showed that the design of our cues could be considered effective. They also indicated directions in which further improvements could be achieved. First, in the learning session, all participants spent less than 5 minutes on listening to or feeling the cues, before they felt confident enough to tell them apart from each other. In the cue identification test afterwards, accuracy reached 97.8% for sound cues and 99.2% for vibration cues. These results clearly show that the cues were very easy to learn. Participants also found it fairly easy and intuitive to infer priorities from the 6 features.

Second, cues were successfully detected in 99.8% of cases. Due to a low signal-to-noise ratio (< 1), the detection of the level-4 vibration cue was affected by the tactile noise that mimicked the bumpy feeling when driving on a bad road surface. This is probably also due to the fact that both signal and noise were provided from the chair. Signal detection in a bumpy condition can be improved by providing vibrations to other parts of the body which are not in a direct contact with the car, such as the abdomen / the seat belt.

Third, cues were identified quickly and accurately. All cues were identified within 5 seconds from the onset. The average response time was 2.8 seconds (1.3 seconds after offset). The higher the priority level, the faster the response, suggesting that participants indeed perceived the associated levels of priority from the cues. The average identification accuracy over all task trials reached 92.5%. Compared to the learning test where the only task was to identify cues, the accuracy was not decreased by the low-load tracking task alone (99.0%). However, all types of additional load harmed the accuracy to some extent. Having a conversation had the largest impact, resulting in the lowest accuracy (86.3%). This result suggests

that cognitive distractions induced by talking to passengers or making phone calls are the most harmful to the effectiveness of cues.

Fourth, combining multiple features in a systematic way was found to be useful, because 90% of the participants relied on more than one feature of sound and vibration to identify cues, indicating a synergy between the combined features. This design also provided room for each participant to selectively make use of the features that he/she found more intuitive and reliable. Pace was rated as the most intuitive to convey priorities. However, number of pulses was used by the largest number of participants (86.7%). The reason might be that number of pulses is a clearly quantified feature, and thus is more reliable when the user is occupied with other concurrent activities. However, care should be taken for this interpretation, because due to the synergy between features, participants might not be able to exactly distinguish which feature(s) they had used.

Finally, the distribution of cue identification errors over four priority levels showed that more errors occurred at levels 2 and 3 than at 1 and 4. Moreover, all errors occurred between two successive levels, such as 1-2, 2-3 and 3-4. To further improve the effectiveness of cues, the features of sound and vibration need to have a larger contrast between two successive levels. In the current design, the contrast between any two disconnected levels (e.g. 1-3, 1-4, 2-4) can be a good reference for creating highly distinguishable cues.

Sound vs. Vibration. The comparison between sound cues and vibration cues is not clear-cut. On the one hand, vibration interfered less with the on-going visual tracking task than sound cues. Vibration cues were identified more accurately than sound cues in all task conditions and at all priority levels. The advantage of vibration over sound was particularly pronounced in the radio condition, where it was shown by all measures (though not always significantly). These findings show that vibration is certainly a promising modality for delivering IIC. On the other hand, several measures also showed an advantage of sound over vibration. In all task conditions except one (the radio condition), sound cues were identified faster and were reported as easier to distinguish than vibrations. The response to the level-1 sound cue was particularly fast. Sound was also preferred by a higher number of participants than vibration. Moreover, participants felt more physically comfortable with sound cues than with vibration cues.

The advantages of sound found in this study might be related to the fact that sound has been a commonly used modality to deliver alerts, notifications and cues. People are very used to all kinds of sound signals in the living environment and are trained to interpret them. For example, a fast paced and high pitched sound naturally reminds people of danger alarms. In contrast, the tactile modality is still relatively new to human-machine interaction, so people have relatively less experience in distinguishing and interpreting the patterns in tactile signals. This might explain why participants in this experiment spent more time and (reported) effort on identifying vibration cues, though they performed more accurately.

Overall, our results suggest that both sound and vibration should be included as optional modalities to deliver IIC in IVIS. When to use which modality is a situation dependent choice. Based on our results, vibration seems to be more suitable when 1) the driver is listening to radio programs while driving, 2) sound in the car (e.g. music) or surrounding noise is loud, 3) the presented message is not highly urgent. Vibration might also be better when there are

other passengers in the car, because it is private to the driver. On the other hand, sound might be more suitable when 1) the message to be presented is highly urgent, and 2) the road is bumpy. Furthermore, there might be some situations where using both modalities is necessary. For example, both sound and vibration can be used when the driver is actively involved in a conversation, because the effectiveness of a single modality could be significantly decreased by cognitive distractions. Highly urgent messages such as local danger warnings can also be cued via both modalities, so that both fast response and correct identification can be achieved. When using both modalities, signals from the two channels need to be well synchronized in order not to cause confusion. The tentative suggestions proposed here need to be further validated in a driving task setting.

5. CONCLUSIONS

We designed a set of sound and vibration cues to convey 4 different priorities (of IVIS messages) and evaluated them in 5 task conditions. Experimental results showed that the cues were effective, as they could be quickly learned (< 5 minutes), reliably detected (99.5%), quickly identified (2.8 seconds after onset and 1.3 seconds after offset) and accurately identified (92.5% over all task conditions). Vibration was found to be a promising alternative for sound to deliver informative cues, as vibration cues were identified more accurately and interfered less with the ongoing visual task. Sound cues also had advantages over vibration cues in terms of shorter response time and more (reported) physical comfort. The current design of cues seems to meet the criteria of pre-attentive reference and is effective under various types of cognitive load that drivers can encounter during driving. Therefore, it is a promising (but by no means the only or the best) solution to convey the priority of IVIS messages for supporting drivers' attention management.

Real driving environments can be more complex and dynamic than the ones investigated in this study. For example, drivers may have radio, conversation, and noise all at the same time. We predict that cue identification performance will decrease when driving load increases or more types of load are added. To make the cues more effective in high load conditions, the features of sound and vibration need to have a larger contrast between different priority levels. The contrast between two disconnected levels in the current design can be a good reference. Our next step is to apply the cues to a driving task setting in order to further evaluate their effectiveness and investigate their influence on drivers' attention management.

6. ACKNOWLEDGEMENT

This work was funded by the EC Artemis project on Human-Centric Design of Embedded Systems (SmarcoS, Nr. 100249).

7. REFERENCES

- [1] G. Arrabito, T. Mondor, and K. Kent. Judging the urgency of non-verbal auditory alarms: A case study. *Ergonomics*, 47(8):821–840, 2004.
- [2] L. Brown, S. Brewster, and H. Purchase. Multidimensional tactions for non-visual information presentation in mobile devices. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 231–238, 2006.
- [3] Y. Cao, A. Mahr, S. Castronovo, M. Theune, C. Stahl, and C. Müller. Local danger warnings for drivers: The effect of modality and level of assistance on driver reaction. In *International Conference on Intelligent User Interfaces*

(*IUI'10*), pages 134–148. ACM, 2010.

- [4] J. Edworthy, S. Loxley, and I. Dennis. Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors*, 33(2):205–231, 1991.
- [5] S. Hameed, T. Ferris, S. Jayaraman, and N. Sarter. Using informative peripheral visual and tactile cues to support task and interruption management. *Human Factors*, 51(2):126, 2009.
- [6] C. Ho, M. Nikolic, M. Waters, and N. Sarter. Not now! Supporting interruption management by indicating the modality and urgency of pending tasks. *Human Factors*, 46(3):399–409, 2004.
- [7] C. Ho, N. Reed, and C. Spence. Multisensory in-car warning signals for collision avoidance. *Human Factors*, 49(6):1107–1114, 2007.
- [8] C. Ho and N. Sarter. Supporting synchronous distributed communication and coordination through multimodal information exchange. In *Human Factors and Ergonomics Society Annual Meeting Proceedings*, volume 48, pages 426–430. Human Factors and Ergonomics Society, 2004.
- [9] W. Horrey and C. Wickens. Driving and side task performance: The effects of display clutter, separation, and modality. *Human Factors*, 46(4):611–624, 2004.
- [10] A. Jamson, S. Westerman, G. Hockey, and O. Carsten. Speech-based e-mail and driver behavior: Effects of an in-vehicle message system interface. *Human Factors*, 46(4):625, 2004.
- [11] C. Kaufmann, R. Risser, A. Geven, and R. Sefelin. Effects of simultaneous multi-modal warnings and traffic information on driver behaviour. In *Proceedings of European Conference on Human Centred Design for Intelligent Transport Systems*, pages 33–42, 2008.
- [12] J. Lee, J. Hoffman, and E. Hayes. Collision warning design to mitigate driver distraction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 65–72. ACM, 2004.
- [13] T. Leinmüller, R. Schmidt, B. Böddeker, R. Berg, and T. Suzuki. A global trend for car 2 x communication. In *Proceedings of FISITA 2008 World Automotive Congress*, 2008.
- [14] D. Marshall, J. Lee, and P. Austria. Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness. *Human Factors*, 49(1):145–157, 2007.
- [15] T. Mondor and G. Finley. The perceived urgency of auditory warning alarms used in the hospital operating room is inappropriate. *Canadian Journal of Anesthesia*, 50(3):221–228, 2003.
- [16] V. Neale, T. Dingus, S. Klauer, J. Sudweeks, and M. Goodman. An overview of the 100-car naturalistic study and findings. Technical Report 05-0400, National Highway Traffic Safety Administration of the United States, 2005.
- [17] N. Sarter. The need for multisensory interfaces in support of effective attention allocation in highly dynamic event-driven domains: the case of cockpit automation. *The International Journal of Aviation Psychology*, 10(3):231–245, 2000.
- [18] N. Sarter. Graded and multimodal interruption cueing in support of preattentive reference and attention management. In *Human Factors and Ergonomics Society Annual Meeting Proceedings*, volume 49, pages 478–481. Human Factors and Ergonomics Society, 2005.
- [19] B. Seppelt and C. Wickens. In-vehicle tasks: Effects of modality, driving relevance, and redundancy. Technical Report AHFD-03-16 & GM-03-2, Aviation Human Factors Division at University of Illinois & General Motors Corporation, 2003.
- [20] C. Smith, B. Clegg, E. Heggstad, and P. Hopp-Levine. Interruption management: A comparison of auditory and tactile cues for both alerting and orienting. *International Journal of Human-Computer Studies*, 67(9):777–786, 2009.
- [21] C. Spence and C. Ho. Multisensory warning signals for event perception and safe driving. *Theoretical Issues in Ergonomics Science*, 9(6):523–554, 2008.
- [22] K. Suzuki and H. Jansson. An analysis of driver's steering behaviour during auditory or haptic warnings for the designing of lane departure warning system. *Review of Automotive Engineering (JSAE)*, 24(1):65–70, 2003.
- [23] J. Van Erp and H. Van Veen. Vibrotactile in-vehicle navigation system. *Transportation Research Part F: Psychology and Behaviour*, 7:247–256, 2004.
- [24] C. Wickens. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2):159–177, 2002.
- [25] C. Wickens, S. Dixon, and B. Seppelt. Auditory preemption versus multiple resources: Who wins in interruption management. In *Human Factors and Ergonomics Society Annual Meeting Proceedings*, volume 49, pages 463–467. Human Factors and Ergonomics Society, 2005.
- [26] D. Woods. The alarm problem and directed attention in dynamic fault management. *Ergonomics*, 38(11):2371–2393, 1995.