

Response Times for Visual, Auditory and Vibrotactile Directional Cues in Driver Assistance Systems

André Lundkvist and Arne Nykänen
Luleå University of Technology

ABSTRACT

The number of advanced driver assistance systems is constantly increasing. Many of the systems require visual attention, and a way to reduce risks associated with inattention could be to use multisensory signals. A driver's main attention is in front of the car, but inattention to surrounding areas beside and behind the car can be a risk. Therefore, there is a need for driver assistance systems capable of directing attention to the sides. In a simulator study, combined visual, auditory and vibrotactile signals for directional attention capture were designed for use in driver assistance systems, such as blind spot information, parking assistance, collision warnings, navigation, lane departure warning etc. An experiment was conducted in order to measure the effects of the use of different sensory modalities on directional attention (left/right) in driver assistance systems. Attention was assessed in a driving simulator using Lane Change Task together with a secondary task, designed to measure choice response times and error rates to directional (left/right) information for multisensory signals. Different combinations of visual, auditory and vibrotactile signals were tested and compared. Visual signals alone (when captured by the driver) or in combination with other modalities provided shortest response times (570 ms on average). Auditory and vibrotactile signals captured attention equally well in terms of response time (650 ms and 740 ms on average). No significant differences in localization error rates were observed.

CITATION: Lundkvist, A. and Nykänen, A., "Response Times for Visual, Auditory and Vibrotactile Directional Cues in Driver Assistance Systems," *SAE Int. J. Trans. Safety* 4(1):2016, doi:10.4271/2015-01-9153.

INTRODUCTION

When driving, it is not only important to keep your eyes on the road. Inattention to surrounding areas beside and behind the car can be a risk, and systems to aid attention in these areas are being developed. The driving environment is also getting more and more populated with systems such as entertainment, communication, navigation, and advanced driver assistance systems. Examples include Blind-Spot Information Systems (BLIS), Parking Assistance, Cross Traffic Alert (CTA) and Lane Departure Warning (LDW). One possible way to increase attention, decrease off-road glance times and decrease response times to warnings including spatial information is to combine visual signals with other signal modalities, for instance auditory or vibrotactile. Auditory or vibrotactile signals alone could allow the driver to keep focus on the road, not forcing the driver to look anywhere else to get the information. It has also been shown that reaction time to an auditory signal is shorter than to a visual signal [1], while reaction time for touch is normally somewhere in-between. However, visual signals are easier to miss, especially in more complex situations. It has been shown that reaction time to a peripheral visual signal is extended for an increasing cognitive load, for instance in a more complex traffic environment. In such environments, peripheral signals are also more often missed [2]. A signal combining the senses can lead to reduction of response times from the driver when re-orienting their attention [3]. This could allow shorter off-road glance times and therefore increasing driver safety.

In addition, a multisensory system could be more robust by providing several information channels instead of only one. This should lead to a lower risk that a warning would be missed.

Previous research has shown that cognitive processing of a secondary task has a negative impact on driving performance [4, 5, 6, 7, 8, 9, 10, 11]. Furthermore, visual inattention was the main factor behind crashes found in a study of driving data collected in traffic, as stated by Klauer et al. [4]:

"Drivers engaging in visually and/or manually complex tasks have a three-times higher near-crash / crash risk than drivers who are attentive."

and that

"in the cases of secondary task engagement, if the task is simple and requires a single short glance the risk is elevated only slightly, if at all."

Further, it has been shown that auditory events help in locating a visual event [12]. Attention in one modality spreads to other senses, and may improve attention even if the event is irrelevant to the main task [13]. Ernst and Bühlhoff [14] stated that several congruent

modalities provide a more robust perception. Cao et al. [9] found that visual signals in combination with speech or beep sound signals in an alertness warning system, used during simulated driving, provided more safe behaviour than only speech, and that it was preferred over only visual warnings. Unsafe behaviour was defined as either an incorrect reaction to the relevant information, such as changing lane instead of breaking, too slow reaction or no reaction at all.

Based on previous research, it is expected that replacing or complementing visual signals with auditory signals should decrease choice reaction times in response tasks [1]. Combined visual and auditory signals have also been shown to result in shorter response times than pure visual signals on a gaze re-orientation task [3]. However, studies have shown that vision is faster and more accurate in spatial tasks, such as locating an object, than hearing [13]. How multisensory spatial information in e.g. a left/right localization task is best provided to a driver with respect to response time and error rate is therefore difficult to predict based on previous experiments on choice reaction times [1] and localization performance [13]. Data on how different combinations of modalities used for presentation of directional information affects response times and error rates during driving is missing. This data is fundamental for designing good human-machine-interfaces in cars, utilizing the strengths of all our senses.

In the present experiment, signals from left or right were presented randomly during simulated driving. The subject was requested to respond to the signals immediately by telling the direction using paddles on the steering wheel. Response times and error rates were measured. There is a lot to consider when designing multisensory directional information signals. A key aspect is that their direction should be easily determined. In this study, the signals were designed with that aspect in mind, and the signals should also catch immediate attention. Background noise in cars (road vibrations, wind noise etc.) were not considered, but should be considered in real implementations. The signals were visual, auditory, vibrotactile or different combinations of these modalities. The objective was to assess whether there are any differences in people's response times and abilities to distinguish signal direction between visual, auditory and vibrotactile signals, and any combination of these while driving. In addition, driving performance was measured using Lane Change Task [15] and an estimate of the cognitive workload experienced by the driver was measured with the Overall Workload scale [16].

Objective: To 1) measure how human response times and error rates for directional (left/right) cues are affected by different combinations of signal modalities and 2) assess whether the use of different signals affects simulator driving performance.

METHOD

Driving Scenario

Lane Change Task (LCT) [15] is a simple driving scenario for measuring driving performance, see Figure 1. It is commonly used to measure the impact of different types of secondary tasks [17]. While there exist several other types of simulators, LCT was selected since it is a well proven and standardized tool for secondary task evaluation. LCT is a laboratory test which provides the subject with a

task similar to driving, allowing easy and controlled measurement of the subject's performance. The simplified setting limits the generalizability of the conclusions to be drawn, but as the objective of this study was to investigate how sounds, vibrations and visual cues can be used to capture attention while driving the setup was considered suitable. The simple driving scenario does not fully represent real car driving and interaction with active safety systems and HMI designs. However, effects from using different sensory modalities for giving directional cues were expected to be fairly small, and therefore an experimental setup giving sensitive and controlled measures was required. Even though this experiment does not prove that measureable effects will be found for real driving conditions, it is likely that if significant effects can be shown using this laboratory setup there will also be differences in driver performance for real car driving. A more sophisticated setup that corresponds well to real car driving would be important for validation of the results.

In LCT the task is to change lane when instructed by signs on the side of the road. Mean deviation between a normative model and the subject driving course, a standardized measure for LCT, was used as a measure of driving performance. A simple driving simulator was used consisting of LCT, a projector, a steering wheel and a modified driver seat, see Figure 2. To allow for good control over the sound, the experiment was carried out in a sound recording studio at Luleå University of Technology.



Figure 1. Screenshot from LCT. In this example, the subject should change to the left lane, as indicated by the arrow on the signs.



Figure 2. The driving simulator in a sound recording studio with the projected driving scenario (3.7 m in front of the subject), visual and auditory indicators and the modified driver seat with vibration generators.

To get a measure of driving performance, the recorded driving path is compared to a reference driving path. According to the ISO standard [15], the mean deviation from the reference path can be used as a measure of driving performance. The measure takes lane-keeping performance into consideration, as well as reaction time to sign information. Furthermore, failing to identify a sign and changing lanes give a large impact on driving performance. Since the driving task in LCT is relatively easy, the subjects drove at 100 km/h instead of the standardized 60 km/h, in order to produce a more demanding test.

Choice Response Time Task

This experiment measured response time and error rate for attention to a signal indicating direction (left/right). The task was designed to resemble the layout of a visual Blind Spot Information System (BLIS) currently used in series-produced vehicles. The subject was requested to respond on which side the information was coming, left or right, by pressing a corresponding paddle on the steering wheel. There was one paddle located on each side. The signals were triggered from a list of 10 delay times used in random order. The delay times were uniformly distributed in the range 5580 - 16020 ms. Each delay time was used once and the order of the delay times was random for each trial. Delay time range was selected to fill the length of the track when travelling at 100 km/h.

In addition to visual signals, auditory and vibrotactile signals were also presented in different combinations. Sound was presented to the driver by loudspeakers placed at the same positions as the visual indicators, see Figure 3. Vibrotactile signals were presented by vibration shakers mounted in the back of the driver seat. The mounting can be observed in Figure 4. The diameter of the circular area of excitation was 37.5 mm.



Figure 3. Placement of the loudspeakers close to the visual light signals.



Figure 4. Mounting of shakers in the back of the driver seat for presenting vibrotactile signals to the subject. The seat padding and upholstery were in place during the experiment.

Multisensory Signals

The signals used in the experiment were designed over a set of workshops with two interaction designers from Volvo Car Corporation and the two authors. The two resulting signals were designed to primarily use vibration (Signal A) or sound (Signal B) as the main attention capture feature. The signals were designed as combined sound and vibration signals, but their sub-signals (sound and vibration parts) could be and were also used independently. Transient sound has been proven to be an effective means of capturing attention [18]. However, due to the short time a transient sound is audible, they lack in providing the ability to accurately determine the spatial position of the sound. Sanders and McCormick [19] recommended that sound signals should have duration of at least 500 ms for easy localization. In Signal A, a transient vibration signal was used to capture attention and a longer sound signal was added to provide better localization ability. The opposite was used for Signal B, where sound was used to capture attention, and vibration was used for spatial resolution.

The vibration signal was always presented also as low frequency components in the sound, played via subwoofers. This was done in order to make the two different modalities more integrated, and to form what in design theory is called a gestalt. Figures 5, 6, 7, 8 show the time signals for the different signals. The pause in the beginning of Signal B (vibration sub-signal) was removed when only using vibration, see Figure 8. Figures 9 and 10 show the spectrum for each signal's sound components. Signal A has defined tonal components, using a fundamental tone at around 587 Hz combined with a fifth at 880 Hz, followed by harmonic overtones. Signal B have less defined tonal components, because it is more transient than Signal A, and the vibration component is less protruding in higher frequencies than for Signal A. The vibration component of Signal A is defined around 60 Hz, where the vibration component of Signal B is more broadband up to around 350 Hz.