

Using Advisory 3D Sound Cues to Improve Drivers' Performance and Situation Awareness

Minjuan Wang¹, Sus Lundgren Lyckvi^{1,3}, Chenhui Hong², Palle Dahlstedt^{1,3}, Fang Chen¹

¹Interaction Design, Department of Applied IT,

²Systems, Control and Mechatronics, Department of Signals and Systems
Chalmers University of Technology, Gothenburg, Sweden

³University of Gothenburg, Gothenburg, Sweden

[minw; sus.lyckvi; palle; fang.chen]@chalmers.se, chenhui@student.chalmers.se

ABSTRACT

Within vehicle Human Machine Interface design, visual displays are predominant, taking up more and more of the visual channel for each new system added to the car, e.g. navigation systems, blind spot information and forward collision warnings. Sounds however, are mainly used to alert or warn drivers together with visual information. In this study we investigated the design of auditory displays for advisory information, by designing a 3D auditory advisory traffic information system (3DAATIS) which was evaluated in a drive simulator study with 30 participants. Our findings indicate that overall, drivers' performance and situation awareness improved when using this system. But, more importantly, the results also point towards the advantages and limitations of the use of advisory 3D-sounds in cars, e.g. attention capture vs. limited auditory resolution. These findings are discussed and expressed as design implications.

Author Keywords

3D Auditory Advisory Traffic Information System; auditory display; in-vehicle design; drive behavior

ACM Classification Keywords

H.5.2 [Information interfaces and presentation]: User Interfaces.

INTRODUCTION

In everyday life, we extract a wealth of information from sounds to obtain a continuous, holistic situation awareness of our surroundings. Humans are well adapted for listening to simultaneous sounds and also have the ability to focus on the more interesting information, filtering out the rest [30]. The capability to quickly and correctly judge the position of a threat has been of vital importance for our survival. Even though the accuracy in sound localization is comparatively

poor, about 4-10 degrees [4], our ability to simultaneously interpret both the direction and content of a sound has been recognized and successfully applied in early interactive information systems [11]. In addition, the omnidirectionality and attention capture aspects of sounds are important advantages of auditory displays. Hence, many computer applications utilize the advantages of sound, using it not only for warnings, but even more so for feedback or as a means to convey system state, games being the prime example of the latter.

In cars however, sound is commonly thought of as a way to warn about critical situations, or to enhance the visual display by providing auditory feedback. There are however several concerns when using auditory displays [14,20]: Sounds can be annoying, and since they are good at attracting the users' attention, repeated or continuous sounds can cause fatigue and reduced attentiveness. Even some types of music can cause accidents, as shown by Brodsky [7]. In addition, there is the issue of privacy and disturbance, since sound is omnidirectional and cannot be directed at a single user. Lastly, other sounds may interfere with, or mask, the system sound. Despite these limitations, sound has a range of information capacities and advantages that makes it particularly useful to convey information in certain contexts. It can be used to provide an additional information channel for users whose eyes are busy attending to a different task, like for instance airplane pilots [3] and drivers [23]. And, since humans are very good at hearing patterns in sound [6], audio is a well-known approach for the monitoring of systems in visually intensive environments, since it can reduce the visual overload and can be heard over the background of other sounds, such as noise or music.

Currently, new Advanced Driver Assistance Systems (ADAS) are introduced into the car environment together with visual displays placed in the central console, the Digital Information Display (DIM) or the Head-up Display (HUD). This despite the fact that the main activities of driving implies that users must look at, and focus on, the road. As a result, drivers may miss important information because their visual system is overloaded – they simply have too much to look at. A way to avoid this overload

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CHI 2017, May 06 - 11, 2017, Denver, CO, USA

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-4655-9/17/05...\$15.00

DOI: <http://dx.doi.org/10.1145/3025453.3025634>

could be to present some information in the form of sounds instead, in this sharing the information load between senses.

In this study, we explore *how driver performance and Situation Awareness can be increased by using a 3D Auditory Advisory Information System (3DAATIS)*, in this investigating the benefits and drawbacks that come with the use of spatial sound cues. We will describe the study and the results of it, including the design implications we have found when it comes to using sound as information carrier in a normal driving situation.

RELATED WORK

Within HCI, three types of auditory signs can be used to translate symbolic visual artifacts into auditory artifacts: auditory icons, spearcons and earcons [14]. Auditory icons are based on a natural sound coupled to what they are supposed to represent. Spearcons consist of a spoken phrase, which has been speeded up until it is not recognized as speech. Earcons are symbolic sounds made up of rhythm, pitch, timbre, register (the relative high/low of the pitch) and the dynamics (the relative loudness) [5]. One example is a pinging sound to signal that a new email has arrived). McGookin and Brewster [21] have studied identification of concurrently presented earcons, testing up to 4 simultaneous earcons. Hardly surprising they found that it was easier to identify fewer icons than many, but when the earcons had different timbre, and were slightly offset in time, accuracy increased. Despite the fact that they are the hardest to learn and map, earcons are the most commonly used signs in cars as well as in everyday software, this due to the lack of natural sounds representing the activities they represent.

Auditory Displays Supporting Situation Awareness (SA)

Summala has stated that a driver's goal is to drive without discomfort and stay within his or her comfort zone [25]. Hence, drivers are required to have a good Situation Awareness (SA) in regards to the surrounding traffic environment, making it possible for them to foresee potential hazards in advance. According to Endsley [9], SA is the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future. Thus, a good strategy to enhance driver's SA is to present information regarding other road users' location and danger in relation to drivers and how the situation is evolving over time.

In one of the earlier studies regarding sound in cars, Tan and Lerner [1] did an extensive study of how subjects perceived warning sounds in a car environment. The main conclusions from this study were that albeit sound is promising for warnings in cars, even the better-performing combinations of sounds and speaker directions did not result in a great accuracy. Their experiment featured 16 speakers and six different sounds. Subjects observed a naturalistic video of driving and their task was to specify the direction of the sounds they heard, and their response

times and accuracy were tested, among other things; 90% of the subjects responded within the right quadrant (i.e. within 45 degrees in either direction from the exact location), but there were always some cases of perceptual reversal, i.e. that subjects pointed 90 degrees or more off. It is notable that the study dealt only with immediate responses to warning sounds (as opposed to advisory sounds), and that there was no real driving task involved.

Recent studies have demonstrated that auditory information improves safety in driving, shortens response time, enhances accuracy and increases drivers' SA. For example, studies show that looming warning signals (i.e. sounds whose intensity increase as the distance between the driver's vehicle and the lead vehicle decreases) and its visual analogue provide very powerful signals to the human perceptual system and that they also capture attention [18]. Gray [13] performed a comparative study and found that looming auditory warnings produced the best combination of response speed and accuracy. In the study, participants experienced four non-looming auditory warnings (constant intensity, pulsed, ramped, and car horn), three looming auditory warnings (veridical, signals shorter than actual time to collision (TTC) time, signals longer than actual TTC time) under a no-warning condition. However, in their study, only a rear end collision scenario was studied, and it focused on warning the drivers – how looming auditory cues works in other types of non-critical traffic scenarios needs to be studied further.

Exploring auditory signs, Fagerlönner and Alm [12] conducted an experiment to examine the use of different types of auditory signs to support truck drivers' situation awareness. Their results showed that abstract (earcon-like) sounds required significantly longer learning times and longer response performances in comparison with sounds that have a natural meaning in the driving context. In the study, the participants were asked to memorize different types of auditory signs and their corresponding road users. Then, their ability to map them correctly was tested, giving the sound as cue. However, this cognitive process is quite different from a real driving situation, since drivers who get a sound cue typically also visually assess the situation.

Liu and Jhuang [19] conducted a driving simulator study to evaluate the effects of five in-vehicle warning information displays (one visual; two auditory displays using static and spatialized sound respectively; and two audiovisual displays using static and spatialized sound respectively). The results showed that a spatialized auditory display significantly improved drivers' performance in reacting to the divided attention task and making accurate decisions. In a similar vein, Baillie et al [2] conducted a simulator study where five separate auditory feedback methods were compared during both autonomous and manual driving scenarios. Their results showed that the spatialized auditory presentation method was superior to the other methods, and in addition that drivers felt significantly more in control

during scenarios containing sound vs. no sound. In both of these studies, the focus was to evaluate drivers' response to different types of displays. How they utilized the information for decision making when interacting with different road users was not studied.

Summary

Together, these studies indicate that the concept of giving *spatialized auditory advisory traffic information* is a promising approach to enhance driver's situation awareness. However, most of these studies were either focusing on the perception of the sounds per se [1,12] or how drivers respond to sounds in drive simulator drives lacking complex traffic situations [1,11]. Watson and Sanderson [29] have stated that an auditory display's effectiveness at communicating information should be evaluated according to its context of use; how drivers interact with auditory information may differ in a more realistic driving situation with different traffic scenarios. Since drivers typically visually confirm the sound information, the interaction with the auditory display, as well as the information needs may be very different. In this study we thus tested our system under as real as possible driving conditions, in which it differs from the others mentioned here.

BACKGROUND: OUR PREVIOUS WORK

This particular study is a part of a larger research project, which aims to explore the design of advisory traffic information systems in cars. The particular study described in this paper was thus preceded by two other studies upon which this one is being based.

The first study [26,27] was a focus group study which studied drivers' requirements on 3DAATIS, as well as their expressed information needs. The results showed that users were positive to the concept, and gave a quite detailed view on in which situations they felt that they needed auditory information. The result of the study was a list of design recommendations in regards to sound design, information prioritization, road user categorization, etc.

This was followed up by a second study, an expert focus group study with five sound experts: one in-vehicle acoustic researcher, two acoustic designers from Swedish automotive OEMs and two in-vehicle interaction designers. The purpose was to narrow down the design requirements collected from the first study. The framing was that it should be possible to use these recommendations for a sound system in vehicles being produced five years from when the workshop was held. The result from the workshop was a number of design decisions.

1. Sound cues should be presented on the *critical level* (warning sounds), and the *safety level* (advisory sounds which direct drivers attention to potential hazard).
2. Only *one type of sound* was to be used, since the intention of this system is to direct drivers attention to potential hazards, which they can then confirm visually.

3. *The sound was to convey urgency, direction and movement.* Direction and movement quite naturally go together, since the sounds are displaying moving objects after all. The urgency parameter distinguished whether a sound was advisory or warning.

The results from the focus group lay the foundation for the design of the system tested in this study.

Within the project, studies on visual advisory traffic information systems were carried out in parallel. In this study we reused the information presentation thresholds for advisory and warning levels. These thresholds were tested and adjusted iteratively according to the characteristics of the drive simulator used in both studies [28].

RESEARCH AIM

In this particular study we set out to test our 3DAATIS focusing on the following questions:

How can a 3D Auditory Advisory Information System influence driver's situational awareness and driving performance?

From this, what can we learn about the design of 3DAATIS?

METHOD

The study utilized a within-subject design, treatment conditions (without 3DAATIS / with 3DAATIS), and five traffic incidents scenarios (see below) as independent variables. In this section we will describe participants, hardware, software, sound design, scenarios, procedure and data collection in detail, starting with a brief overview.

The Study in Short

For this study, we developed a prototype of a 3D sound system, which aims to provide relevant advisory information regarding road users in the vicinity of the own vehicle. This information includes in which direction the road users move and which risk level they are at (in terms of time/ distance to collision).

The study was a simulator study. Participants came in one by one, making two drives, one baseline, one with 3DAATIS, in random order. The drives featured the same eight scenarios, again in randomized order. Twice during each drive, the simulation was frozen and the participant's situation awareness was measured. In addition, driver performance data was collected. Finally, the participants' opinions on the system were collected in a post-interview.

Participants

In all, 30 participants took part in the study (18 men and 12 women; mean age 40 years, range 21-63 years), and their average driving experience was 18.3 years (ranging from 3 to 45 years). The participants were recruited around a science park in an urban area, this since highly educated, well-paid people can afford a car with different ADAS functions – thus they were in the right segment of buyers.

System Design

The aim of the 3DAATIS was that participants should be able to identify the *origin and moving location* of a sound cue. For example, when a car comes up from rear and takes over, participants should hear a sound “coming” from the back and “moving” closer, “passing” on the left side as the approaching car takes over. Moreover, they should be able to *map an increasing risk level with changes in the sound*. There were two levels: advisory and warning sounds.

At this point, we decided on using 3D-sound (i.e. surround sound), since we found it to be the best solution. The alternative would have been to use binaural sound. But, when using speakers with binaural sound, the complex acoustics of the car interior blurs directionality. If using head phones instead, absolute sound localization (relative to the outside world and not the listeners head) becomes confusing as soon as users move their head. Lastly there was also a pragmatic argument for using 3D-sound: we aimed to develop a solution feasible for cars in the close future, i.e. it should not rely on an expensive, hard-to-calibrate system, or on all drivers wearing head phones.

The sound design process for the 3DAATIS was based on the results from the expert focus group session described in “Background: Our Previous Work”. Following Meera’s recommendation, an earcon was designed by sound experts [5]. The sound sample was designed based on the OEM’s sound attributes which had tested well in regards to user acceptance. The sound is somewhat similar to a plucked string, with a strong attack, followed by a decaying tail consisting primarily of the fundamental, with a few overtones. Overtones were added to decrease the risk of masking by other sounds, and a new exponentially decaying tail helped decrease annoyance. This is also what makes it somewhat similar to the plucking of a string, which is generally perceived as a pleasant sound. The sound sample was tested with several sound experts as well as within the research group, and several design iterations were done in order to make sure that the sound design was appropriate for the purpose of advisory signals. Figure 1 shows a sonogram of the used sound sample.

To allow for continuous localization, the sound is looped. As long as the other road user (i.e. the subject of the sound cues) is within the advisory level, the repetitive sound is projected from the accurate relative angle, continuously tracking the object as it moves. We used a form of “fake

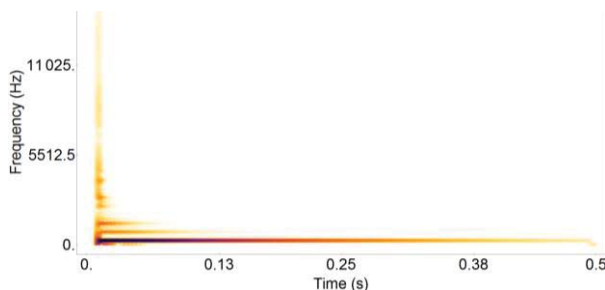


Figure 1: A sonogram of the sound used

Doppler” effect to communicate a rapidly changing distance to other road users. Real Doppler effects are mostly detectable at quite high relative speeds and if the observer is somewhat distant from the sound source, which makes it less useful in traffic situations. However, as humans are very sensitive to pitch changes, we added a slight gradual increase in pitch with closing distance (from 268Hz to 288Hz), to call for increased attention, as well as an increase in the repetition rate (from 667ms to 620ms).

When the situation gets critical, the system increases the pitch of the sound (up to 445Hz-469Hz, again varying with distance) while further shortening the loop (198ms-187ms), since this increases the perceived urgency while keeping continuity from the ongoing advisory sound. If the car is surrounded by more than one road user, several sounds are played, one per subject, from the accurate relative angle, with the correct urgency level. The different pitches and repetition rates help distinguish the sounds from each other.

The activation thresholds of *advisory* and *warning* information levels are based on two free parameters:

- *Time-to-collision* (TTC): The assumed reaction time of the driver of the following vehicle in seconds.
- *Safety margin* (SM): The distance of the closest point between the participant’s vehicle and other approaching vehicles in meters when they are in parallel lanes.

The thresholds had been tested in our previous studies [28]; this was necessary since the perception of distance is different when driving in a real car vs in the simulator, and thus we adapted the thresholds accordingly; see Table 1.

Software

The prototype was developed using a combination of the dataflow programming environment Pure Data and Matlab. The latter was used to receive simulator data and send direction, distance, and time-to-collision data to the Pure Data patch, which in turn generated the relevant sounds (i.e. warning or advisory) and projected them in a 5.1-channel surround sound system. Sound directionality was achieved through equal energy-panning between speaker pairs. The 3DAATIS information was presented at 65 dB with a traffic ambient noise at 45 dB in the background.

Hardware

A PC running STISIM Drive® software was utilized as driver simulator. A HD projector was used to project the simulated drive scenarios on the front wall. Two webcams were installed to record what the drivers saw on the road, as

| Information Level | Sound Effects | TTC | SM |
|-------------------|--|----------|-------|
| Advisory | Original sound sample | 3 to 6 s | < 3 m |
| Critical | Increased pitch and frequency of looping | < 3 s | < 2 m |

Table 1. Thresholds for warnings in terms of Time to Collision (seconds) or Safety Margin (meters).

well as their reactions to the incidents, e.g. steering and braking. This video data was synchronized with drivers' simulation data to better allocate the starting point of drivers' reactions to the incidents. A Logitech G25 Racing Wheel, which included gas and brake pedals and a gearbox was installed in front of an actual car seat.

Another PC ran the 3DAATIS system as described in the previous section. In addition, a 5.1-channel surround-sound system (Logitech Z-5500) was used for the auditory system. The arrangement of loudspeakers and seating position of participants were calibrated according to Dolby 5.1 home theatre speaker guidelines. However, the centre speaker was not used in this experiment (in a 5.1 setup it is primarily intended for dialogue sound in movies). Sound-absorbing curtains were installed on three sides of the test area to ensure a good surround sound effect.

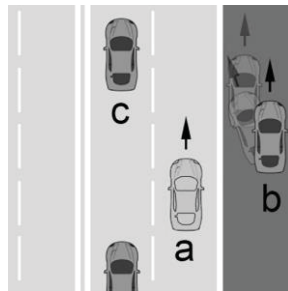
The Traffic Incidents Scenarios Studied

In order to design test incident scenarios that represents real traffic situations, the research team had, in an earlier study [28] observed over 100 naturalistic driving videos. For this particular study, we chose scenarios that cover common critical situations found in the naturalistic drive videos.

To test the effects of the design, we chose three scenarios involving vehicles, and two involving pedestrians. The order of the scenarios and the physical environments, e.g. buildings, trees and the brands of the event cars were randomized for each new test drive. Below we will describe the five scenarios we actually studied. **The car (a) represents the driver's car in all scenarios.**

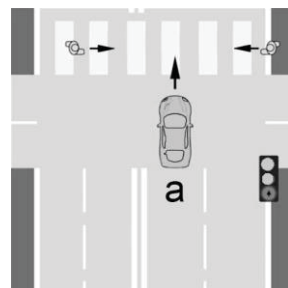
Cut In

In this event, car b (parked on the sidewalk on the right) attempts to start up and cut into the driver's lane, then it sees the driver's car in the rear view mirror, so it steers back to the sidewalk. On the lane to the left of car a (the driver's car), there are another two cars; one is in the front and behind. Hence, changing lanes is not a good strategy. The considered safe approach in this scenario is to *brake or steer slightly to the left*.



Intersection

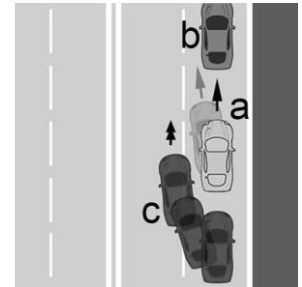
Here, the driver (in car a), has a green light. However, two pedestrians start crossing the street right in front of the driver anyway. One pedestrian is walking from right to left quite slowly, whereas the other is walking from left to right relatively fast. Of these two, the faster one starts walking



first, and is partly hidden in the A pillar blind spot, whereas the slower one starts later. In this scenario, the considered safe behavior is to *brake and wait for the pedestrians to pass*.

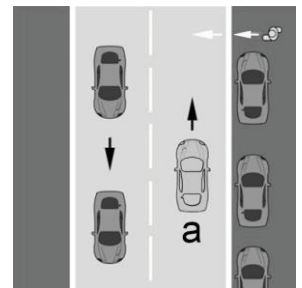
Overtake

One car (b) drives in front of the driver's car (a) at a slower speed and force the drivers to change lane. Another vehicle (c) appears from behind, having a very high speed, overtaking both a and b at a close distance. Here, the safest possible way to avoid the accident is stop change lane and *keep one's position behind car b*.



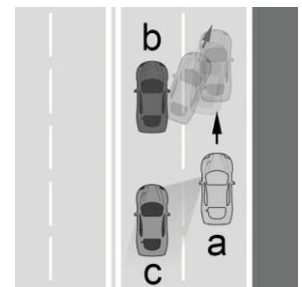
Pedestrian

Here, a single lane in a residential area, a pedestrian suddenly walks out between the cars on the right side of the road. The driver (in car a) has a visual disadvantage in this scenario; visual contact with the pedestrian is made at a very short distance. Thus, the safest possible way to avoid the accident is to *brake and steer at the same time*.



Red Cab

Here, two slow cars (b and c) drive in the left lane. Car b suddenly cuts into the driver's lane very fast directly in front of the driver's car (car a). At this moment, the other car (c) is in the blind spot of car a. In this scenario, the considered safe approach is to *slow down in order to avoid an accident*.



Overall Procedure

First, the participant was introduced to the study and filled out a demographic questionnaire on hearing capabilities, driving experience etc. Thereafter, the participant was given instructions regarding the test drives. The researchers explained the purpose and the procedure of the experiment in general, and instructed drivers to keep the speed at 60 km/h, follow traffic regulations, and stay in the right lane if possible. After getting some instructions on how to drive in the simulator and how the 3DAATIS works, each subject got to train until they fully understood the simulator and 3DAATIS features. Three questions were asked after training to make sure the participants understood the 3D localization sound. When participants were satisfied and fully understood the principles of sound system, the main test started.

Each subject drove twice, once with the 3DAATIS and once baseline. As mentioned, the 3DAATIS information was presented at 65 dB with a traffic ambient noise at 45 dB in the background, in the baseline only the latter. This as the focus of this study was to understand how drivers utilize 3DAATIS under different traffic scenarios. The drives featured eight scenarios of which five were tested; the remaining three were added in order to reduce learning effects between test drives under different conditions. Both the order of conditions and the order of the scenarios was randomized. At two points during each drive (at the same scenarios) the simulation was frozen and the screen turned off, and the participant had to fill out a SAGAT form on Situation Awareness. In addition to this data, the drivers' performance data was also logged during the drives. After the driving, a post interview was conducted with each participant, collecting their feedback regarding the system. Finally, the participants were rewarded with two movie tickets.

DATA COLLECTION

The study featured collection of both quantitative and qualitative data, since we not only measured participants' SA and performance, but also their attitude to the system and how they experienced it.

Measuring Situation Awareness

The drivers' situation awareness was measured using a query technique called Situation Awareness Global Assessment Technique (SAGAT) [10]. It is a freeze on-line probe technique, designed for real-time, human-in-the-loop simulations and provides diagnostic information regarding how well the system in question supports the operator's various SA requirements. The simulation is frozen and subjects are queried as to their perception of the situation at that instant. SAGAT queries cover data criteria corresponding to the three levels of SA: *perception*, *comprehension* and *projection*.

In our study, the drive simulation froze twice per condition (baseline vs. with 3DAATIS) directly after two randomly selected scenarios (*Cut In*, *Red Cab*, *Pedestrian* and/or *Overtake*). The *Intersection* scenario was not measured – in that scenario, participants are waiting for the traffic lights and can then already see the pedestrians at a distance.

During the freeze, the drive simulator displays were blanked while the participants were queried as to their perception of the traffic scenario at that instant. They were asked to draw the locations of other road users in relation to the own car, as well as their direction. This measured the first two levels of SA, *perception* and *comprehension*. Then, they were also asked to mark the road user they perceived to pose a potential danger to them in terms of causing an accident, measuring the *projection*-level of SA, i.e. how well they understood the situation and what was going to happen. A standard predesigned SAGAT questionnaire was used.

The subject's perceptions were then directly compared to the real situation from the frozen simulation scenarios. This was done by two researchers, who also scored the answers, giving 1 for a correct perception, and 0 for the wrong perception. Scores were given twice, once for *perception* and *comprehension* once for *projection*.

Measuring Driving Performance

To understand how the drivers responded to the events, participants' driving behavior was collected in terms of steering wheel angle, and the cars' acceleration and deceleration due to the gas and brake pedals respectively, all measured via sensors in the pedals and steering wheel (i.e. simulator log data), and double-checked with video observations. Collectively, these measures provide a clear description of how the 3DAATIS influenced drivers' response to events. The detailed measurements are as follows:

- **Video observations** of participants' simulator drive. One camera focused on the throttle and brake pedals, another on the driver's hand on the steering wheel. Camera video data were synchronized with simulator driving data.
- **Number of collisions:** Whether a participant manages a scenario with or without a collision.
- **Response time (RT):** The remaining time to a (potential) collision when drivers first react to the event (e.g. brake, release gas pedal or steer away). This measure reflects driver's safety awareness regarding the situations. (seconds)
- **Steering profile** reflects drivers' steering behaviors.
 - *Steering wheel angle (Mean):* How much the steering wheel is turned (degrees)
 - *Steering wheel angle (Range):* The difference between maximum and minimum steering angles when reacting to the event. (degrees)
- **Gas profile:** Longitudinal acceleration due to the throttle (ft/s^2); basically how much the drivers accelerate or decelerate when releasing or pressing the gas pedal
 - *Acceleration (Mean):* The average acceleration of the vehicle from initial gas pedal depression until the driver either brakes, collides or passes the other road user.
 - *Acceleration (Range):* The difference between maximum and minimum acceleration during reaction period. It reflects the amplitude of gas pedal response.
- **Brake profile:** The deceleration due to the brakes (ft/s^2)
 - *Deceleration (Mean):* The average deceleration of the vehicle from initial brake depression until the car either stops, collides, or passes the other road user.
 - *Deceleration (Range):* The difference between maximum and minimum deceleration value during each reaction period. It reflects the amplitude of brake response.

Measuring Usefulness and Experience

In order to understand how participants think about the 3DAATIS, a standard Van der Laan's System Acceptance Measurement was collected. This a 5 Likert scale, ranging from -2 to 2, including nine items. It measures two dimensions, *usefulness* and *satisfaction* [16]. In addition we also carried out semi-structured post-interviews after the participants had finished the driving task. Here, we aimed to investigate when the participants found the sounds to be helpful vs confusing, and what general feedback they had on the system and its sounds.

DATA ANALYSIS

For the SAGAT score, the number of collisions and the acceptance analysis, a Chi-square test was conducted to see whether there was a significant difference between conditions. The significance of such difference was corrected by the Fisher Exact test, when the expected frequency of an even spread of responses for each scale is less than 5 in some cases.

For driver performance data, a paired t-test was used to determine whether there was a statistically significant difference in drivers' performance between baseline and when using 3DAATIS. Only the measurements with significant effects were reported here.

As for the post-interviews section, its semi-structured interviews covered topics like, sound design, perceived useful sectors and information redundancy. Interviews were recorded and transcribed, using a thematic analysis method.

RESULTS

Drivers' Situation Awareness

No statistically significant differences were observed from the results of SAGAT. As seen in Figure 2, the results showed a tendency that participants increase both how well they could perceive the *position* of surrounding road users (SA P), as well as their ability to *understand* the situation, identifying road users with which they might have potential conflicts (SA U). This is especially evident in the two scenarios where the participants had a visual disadvantage (*Pedestrian* and *Overtake*). In the *Red Cab* scenario, however, the participants' SA accuracy rate was slightly

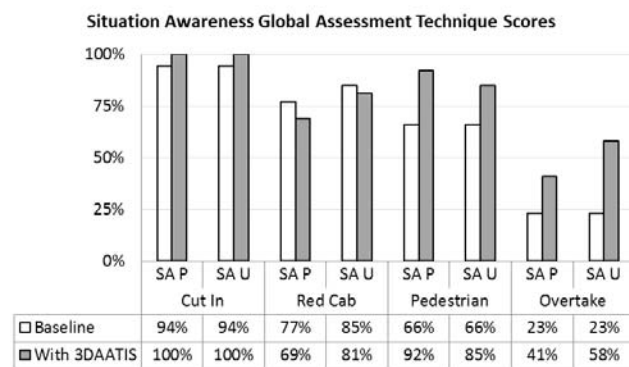


Figure 2: How respondents Situation Awareness changed between baseline and 3DAATIS.

decreased when using 3DAATIS in both regards to both perception and understanding of the situation.

Number of Collisions

The number of collisions is one of the most important indicators to assess driving performance. The results showed that when driving with the 3DAATIS, the number of collisions decreased significantly in *Cut In* and *Pedestrian* ($P < 0.05$). However, in the *Intersection*, *Red cab* and *Overtake* scenario, the number of collisions did not change significantly between conditions, see detailed numbers in Figure 3.

Driving Performance

Table 3 shows the detailed numbers, but below we will discuss the significant findings for each scenario.

Cut-In. In the *Cut In*-scenario, the response time with 3DAATIS was 1.05 seconds earlier than the baseline and the mean and range of acceleration was significantly reduced, when driving with 3DAATIS. This means that the driver's gas pedal changes in a more deliberate manner, i.e. that the subject is prepared for the situation.

Red Cab. In the *Red Cab*-scenario, the response time was 1.07 seconds earlier than the baseline, however, the drivers performed less good with 3DAATIS than baseline, which is in line with our findings in regards to Situation Awareness and Collisions. When driving with the 3DAATIS, the mean and range of acceleration was significantly higher, which means that participants' response was rash. This may indicate that drivers were unprepared and perhaps startled or confused by the sound.

Intersection. Here, no significant differences were observed on any measurements, except for the response time which increased around 1.89 seconds with the 3DAATIS. One possible explanation is that the natural response when becoming aware of pedestrians in front is to break; under the 3DAATIS condition this simply happened sooner but did not change the drivers' behaviors otherwise.

Pedestrian. When driving with the 3DAATIS, participants responded 0.66 seconds in advance and were more accurate and efficient, in terms of greater changes in acceleration (i.e. a quick release of the gas pedal) and greater steering wheel angle. Video records showed that when participants drove without 3DAATIS, most of them hit the full brake when they saw the pedestrian walking out, but when they

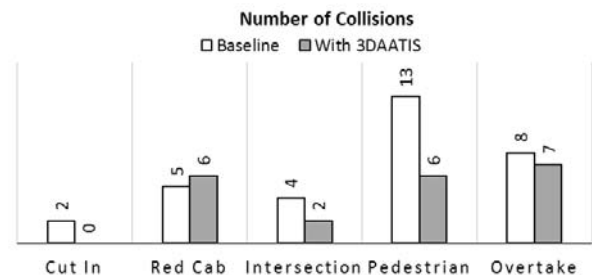


Figure 3: Number of Collisions

| Scenarios No. | Measures | Baseline Mean (SD) | With 3DAATIS Mean (SD) | df | t |
|---------------|--|--------------------|------------------------|----|---------|
| Cut In | Response time (s) | 2.27 (0.74) | 3.32 (1.75) | 26 | -3.01* |
| | Acceleration Range (ft/s ²) | 8.53 (3.74) | 5.59 (2.61) | 17 | 2.67* |
| | Acceleration Mean (ft/s ²) | 0.42 (0.15) | 0.33 (0.11) | 16 | 2.08* |
| Red Cab | Response time (s) | 2.25 (0.71) | 3.32 (1.77) | 28 | -2.92* |
| | Acceleration Range (ft/s ²) | 5.14 (2.30) | 7.02 (2.29) | 19 | -3.75** |
| | Acceleration Mean (ft/s ²) | 0.34 (0.13) | 0.48 (0.18) | 19 | -2.78* |
| Intersection | Response time (s) | 3.84 (1.51) | 5.73 (1.75) | 22 | -4.12** |
| Pedestrian | Response time (s) | 1.24 (0.35) | 1.9 (0.13) | 26 | -3.94** |
| | Acceleration Range (ft/s ²) | 8.16 (5.94) | 14.28 (12.53) | 23 | -2.08* |
| | Acceleration Mean (ft/s ²) | 0.73 (0.50) | 1.11 (0.57) | 20 | -2.36* |
| | MSteering wheel Range (degree) | 17.61 (10.66) | 31.70 (31.52) | 20 | -2.36* |
| Overtake | Response time (s) | 0.57 (0.56) | 2.80 (1.14) | 24 | -2.35* |

*P<.05. ** P<.001

Table 2. Driving performance measurements during the response period.

drove with 3DAATIS, they were more prepared and decelerated in advance, leading to a smoother driving.

Overtake. This scenario was somewhat different from the others, since the overtaking vehicle moves so fast. Thus, when the 3DAATIS started to alert the driver, the sounds started directly at the warning level, giving participants only 2-3 seconds to react. No significant differences were observed in any of the performance measures, and from video observations, we also found that when the warning started, the subjects adjusted the steering wheel or gas pedal instantly but in a non-sufficient way. Therefore, even though the response time was 2.8 seconds ahead, the number of collisions still remained the same when driving with the 3DAATIS.

Drivers' Subjective Feedback to 3DAATIS

The results from the Van der Laan's system acceptance measurement were very positive for *usefulness*. It was measured on 5 scales, yielding mean results from 0.47 to 1.13, with a total mean value of 0.77 (SD: 0.77). Since the scale runs from -2 to 2, these are positive results.

As for *satisfaction*, mean results on four scales ranged between -0.4 and 0.23, with a total mean value of -0.18 (SD 0.91). This could have indicated that the sounds were disturbing or annoying, but the post-interviews disproved this; participants quite liked the sounds. Possibly the ratings are neutral simply because the system aims to be transparent and non-disturbing, it is not meant to evoke any feelings, positive or negative. A paired T-test shows that the usefulness rate was significantly higher than satisfaction ($t_{29} = 8.42, p < 0.001$). It indicates that the participants thought the system is more useful than satisfying.

The post interviews started with the participants marking in which sectors around the car they found the sound information useful. Then, a follow-up discussion was carried out to collect their motivations. We also asked for their perceived pros and cons with the system and the sounds.

As is shown in Figure 4, most of the participants considered information from the rear blind spots sectors to be the most useful. Overall, they appreciated advisory information for any situation when they could not see for themselves. Cases they mentioned were quickly overtaking cars, other cars in lane change situations, playing children or other pedestrians in blind spots, or when the view was being blocked by something, as in the case of *Pedestrian*. One participant stated that “*I have too much to look for, sound will be good to make sure that I don't miss this information.*”

Even from the front zone, participants wanted information, if it was related to an emergency or an unexpected event. In addition, participants wanted the system to be able to notify them under conditions when their visibility was otherwise reduced too, e.g. when driving in fog, heavy rain or during a snowfall. The latter implies some more intelligence from the system in adapting to these conditions.

Positive Aspects of the System

Some participants stated that the sounds made it easy to understand which urgency state they were currently in. Some mentioned that they like the auditory information, since it made them feel protected and taken care of.

Negative Aspects of the System

In some cases, participants found the sounds to be confusing or misleading. Several participants mentioned front-rear-confusion – a known problem in sound design; it

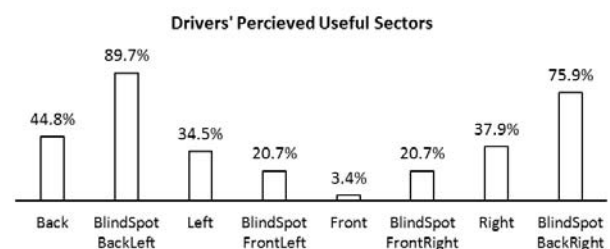


Figure 4. In which areas around the car the respondents found the information to be helpful.

is hard to distinguish the directions of sounds coming from one's exact front or back. For instance, some participants explained how they in the *Overtake*-scenario (with the fast car approaching from behind) didn't understand where the sound information came from. Therefore, they didn't know how to respond. Yet others had issues with the 3d-aspects, and suggested different sounds for right vs left side of the car, so that they'd know where to look.

Another source of confusion was when there were multiple road users in the same sector, then it was distracting and difficult to distinguish the exact location of each road user by sound. Some participants suggested that the system should only provide one sound cue from the same sector, just pointing drivers' towards the general direction of road users without communicating how many, leaving it to the driver to make a visual confirmation. Others instead wanted different sounds for different types of road users (which does not really solve the problem if there are two road users of the same kind). In contrast to this, and in line with the recommendations from the sound experts, other participants said that they felt distracted and stressed when hearing two sound cues at the same time.

DISCUSSION

In this experiment, we *studied whether spatialized sound cues can be used to present situational information in such a way that it supports situation awareness and driving performance in a real driving context*. In general, the results were promising. Participants perceived the system as being useful, response times were better, and in a majority of situations SA and driving performance increased. Interesting findings were obtained in those situations the system supported SA and performance, and those it did not, resulting in a number of design implications, as being discussed below.

As for acceptance, the 3DAATIS received notable positive scores on perceived usefulness, which means that **participants are open for spatialized sound cues**. Moreover, we used only one sound displaying three dimensions (direction, movements and urgency), and participants did not have any problem understanding this underlying concept, meaning that it was **easy to learn**. To some extent these findings go against the common car-industry notion that auditory displays should be avoided in the vehicle compartment, due to distraction and annoyance when the number of sound increases. However we did consider these negative effects in several ways. Firstly, our focus was on **advisory** sounds (being more pleasant and less urgent than warning sounds). Secondly we aimed to design the system as to only make sounds when really needed, as shown by our thresholds (table 1 and related text). The fact that the sound cues (65 dB) worked together with ambient traffic sounds (45 dB) is promising.

The most important finding is that **spatial sound cues work ideally in situations involving single road users in the front and side sectors**. In less complex situations, or

when single road users occur in front or in the front side sectors, drivers can quickly perceive and associate the localized sound cues with what they see from the traffic situation outside of the car; feedback is instant. Therefore, even though sound localization resolution is quite low, drivers' performance and Situation Awareness in these situations was still increased, as exemplified in the scenarios *Cut In*, *Intersection* and *Pedestrian*.

The second was that **if the situation is complex, spatial sound cues can be confusing**. If, for instance, there are several road users around the car, and/or sound cues come from the back, the limited auditory localization resolution and/or front-back confusion makes the information ambiguous, confusing drivers. This was clearly shown in the scenarios *Red Cab* and *Overtake* in terms of number of collision, response time and driving performance. In the *Red Cab*-scenario neither the system nor the human capability to discern the exact direction of the sound could help the drivers to tell the two vehicles' sounds apart; the angle between them was too small. Others have made similar findings. Tan et al [1] carried out an extensive test on sounds and sound localization accuracy by placing participants in a car equipped with 16 speakers (as opposed to the 4 we used), concluding that localization in general is accurate enough to orient the listener towards the direction of sound source, but with mean errors about 10 to 20 degree. In addition, several cases of perceptual reversals (i.e. more than 90 degrees off) were observed. From this, we take that **in these situations, supporting sound cues with visual information to avoid confusion ought to be beneficial**.

Like other car-related studies [12,24], our study also concluded that it **is beneficial to support drivers with both advisory/situational information sound cues and warning sounds**. The results from the SAGAT enquiry showed increased understanding and projection of the future situation in three (*Cut In*, *Pedestrian* and *Overtake*) of the four scenarios tested for this. Again, *Red Cab* was the outlier. In addition, drivers' response time improved in all scenarios as compared to baseline. These results suggest that drivers, whose attention has already been directed towards a potential hazardous situation, either avoid it all together, or react faster to the warning sound. It is worth noting especially, that in the *Overtake*-scenario, the overtaking car came up so fast, that the system went directly into warning-mode, skipping over the advisory level, which also indicates that the advisory level really adds to SA and performance. For *Overtake*, drivers did react faster, but not necessarily in the right way; the front-rear confusion came into play as mentioned above.

We believe that the last finding is particularly interesting in the light of semi-automated driving. As more and more ADAS are introduced, bringing even more displays into the car and overloading the visual channel, we believe that an auditory display could be good candidate to offload the

visual channel and highlight the most important situational information. Moreover, as drivers might switch between driving and other secondary tasks during semi-autonomous driving, jumping directly to the warning level might be too abrupt. Instead, **advisory level sound cues should be used to bring drivers back to into the loop with good situation awareness**. For instance, Lee et al [17] investigated how users, working with another task, responded to collision warnings, showing that graded alerts led to a greater safety margin and a lower rate of inappropriate responses to nuisance warnings. In addition, they found that graded alerts were more trusted than single stage alerts. (One could argue that if the car is truly self-driving it should not be necessary for drivers to take over, but until the car industry has solved the ethical dilemmas that may arise in acritical situation, there still needs to be a way for humans to quickly take back control.)

Here, we believe that the car industry has a lot to learn from the game industry, since it has long experimented with the use of ambient sounds as a means to provide players with a general sense of what is going on in the environment – this is especially relevant for any location based games where players move around in the real world [8,15,22].

Sound Design Implications

Our experiment was not specifically designed to give conclusions on the sound design, but from the post-interview and test results, we can still draw some conclusions.

The sound designed was based on an OEM car manufacturer's tested sound library. Keeping the aesthetics, we designed custom sounds devoid of the harsh and inharmonious qualities associated with warning sounds, which we assumed would work better for advisory purposes. Despite the fact that the sounds now were no longer typical, attention-grabbing, warning sounds, drivers still observed and responded to them. Thus, one conclusion is that **non-obtrusive, harmonious sounds work well for advisory driver information**.

Another implicit conclusion is that the setup of the speaker system is crucial, as is calibration of panning algorithms and acoustic preparation of the listening environment. Our rather basic setup was a simple proof of concept, and can naturally be much improved. In a real car, a larger number of speakers must probably be used for improved precision, but, resolution of directional localization will always be somewhat limited because of how human perception works. Similarly, it is worthwhile to investigate the use of different timbre as suggested by [21].

POTENTIAL LIMITATIONS AND FUTURE WORK

A potential limitation of this research is that the test only ran for an hour. The subjective feedback and driving performance could be different after a long period of use. A second limitation that is the current 3DAATIS only used time/distance to collision and trajectory as system

parameters, which resulted in a number of false alarms. One last limitation is that our system was presented through a 5.1 surround sound system, ergo the resolution for sound localization was not optimized. It could have influenced the results in scenarios like *Overtake* and *Redcab* negatively.

In addition to addressing these limitations, there are many possible directions for future studies. The first would be to conduct the study in a higher fidelity drive environment and follow the drivers' behavior and acceptance changes over a period of time. The second would be to integrate other parameters into the system algorithm to enhance the information accuracy, i.e. current acceleration, drivers' attentional status etc. One other possibility is to increase the fidelity of the sound system. Further, toying with the sound attributes, like timbre, is also worthwhile. Lastly, it would be worth to explore how the 3DAATIS can work in tandem with other modalities for conveying information in complex situations.

CONCLUSION

This study is a proof of concept, showing that advisory, spatialized sound cues can be successfully used in a 3D Auditory Advisory Traffic Information System in order to improve drivers' situation awareness and performance. Unlike previous studies this one included a simulator drive under as natural conditions as possible, featuring other road users and complex situations.

The most important implications of this research are that:

- Spatial sound cues work ideally in situations involving single road users in the front and side sectors
- In complex situations, spatial sound cues can be confusing; in these cases, supporting them with visual information ought to be beneficial.
- It is beneficial to support drivers with both advisory information sound cues and warning sounds.
- Non-obtrusive, harmonious sounds work well for advisory driver information.

A lot of work remains before this can be implemented in real cars, but our results point in a positive direction and show that this will be a worthwhile addition for drivers, increasing driver safety and situation awareness.

ACKNOWLEDGEMENTS

This project is a part of Swedish FFI ADAS presentation project. We would like to thank Fredrik Hagman (Volvo Cars) who designed the sound samples, as well as Samuel Palm (Volvo Cars) who provided continuous support and design insights throughout the entire project. Last but not least, thanks also to all our study participants.

REFERENCES

1. Tan A.K. and Lerner Neil D. 1996. Acoustic localization of in-vehicle crash avoidance warnings as a cue to hazard direction. U.S. Department of Transportation National Highway Traffic Safety

- Administration. Retrieved from <http://trid.trb.org/view.aspx?id=474030>
2. David Beattie, Lynne Baillie, Martin Halvey, and Rod McCall. 2014. What's around the corner? Enhancing Driver Awareness in Autonomous Vehicles via In-Vehicle Spatial Auditory Displays. Proceedings of the 8th Nordic Conference on Human-Computer Interaction Fun, Fast, Foundational - NordiCHI '14: 189–198. <http://doi.org/10.1145/2639189.2641206>
3. Durand R Begault. 1993. Head-up auditory displays for traffic collision avoidance system advisories: A preliminary investigation. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 35, 4: 707–717. <http://doi.org/10.1177/001872089303500409>
4. Durand R Begault. 1999. Auditory and non-auditory factors that potentially influence virtual acoustic imagery. Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction. Retrieved from <http://www.aes.org/e-lib/browse.cfm?elib=8052>
5. Meera M Blattner, Denise A Sumikawa, Robert M Greenberg, et al. 1989. Earcons and icons: Their structure and common design principles. *Human-Computer Interaction* 4, 1: 11–44.
6. Sara Bly. 1982. Presenting information in sound. Proceedings of the 1st Conference on Human Factors in Computing Systems (CHI '82), 371–375. <http://doi.org/10.1145/800049.801814>
7. Warren Brodsky. 2015. Driving with music: cognitive-behavioural implications. Ashgate Publishing, Ltd.
8. Inger Ekman, Laura Ermi, Jussi Lahti, Jani Nummela, Petri Lankoski, and Frans Mäyrä. 2005. Designing sound for a pervasive mobile game. Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology, 110–116.
9. Mica R Endsley. 1988. Design and evaluation for situation awareness enhancement. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 32, 2: 97–101. <http://doi.org/10.1177/154193128803200221>
10. Mica R Endsley. 1988. Situation awareness global assessment technique (SAGAT). Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National, 789–795. Retrieved November 6, 2014 from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:A+methodology+for+the+objective+measurement+of+pilot+situation+awareness#0>
11. Mark A. Ericson and Richard L. McKinley. 2001. The intelligibility of multiple talkers spatially separated in noise (No. AFRL-HE-WP-SR-2001-0009).
12. Johan Fagerlön and Alm Håkan. 2010. Auditory signs to support traffic awareness. *IET Intelligent Transport Systems* 4, 4: 262–269. <http://doi.org/10.1049/iet-its.2009.0144>
13. Rob Gray. 2011. Looming auditory collision warnings for driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 53, 1: 63–74. <http://doi.org/10.1177/0018720810397833>
14. Philip Kortum. 2008. HCI beyond the GUI: Design for haptic, speech, olfactory, and other nontraditional interfaces. Morgan Kaufmann.
15. Jason Kurczak, T. C. Nicholas Graham, Claire Joly, and Regan L. Mandryk. 2011. Hearing is believing: evaluating ambient audio for location-based games. Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, 32. <http://doi.org/10.1145/2071423.2071463>
16. Van Der Laan, Jinke D, Adriaan Heino, and Dick De Waard. 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies* 5, 1: 1–10.
17. John D Lee, Joshua D Hoffman, and Elizabeth Hayes. 2004. Collision warning design to mitigate driver distraction. Proceedings of the SIGCHI Conference on Human factors in Computing Systems, 65–72. Retrieved March 24, 2012 from <http://dl.acm.org/citation.cfm?id=985701>
18. Fabrizio Leo, Vincenzo Romei, Elliot Freeman, Elisabetta Ladavas, and Jon Driver. 2011. Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds. *Experimental brain research* 213, 2–3: 193–201. <http://doi.org/10.1007/s00221-011-2742-8>
19. Yung-Ching Liu and Jing-Wun Jhuang. 2011. Effects of in-vehicle warning information displays with or without spatial compatibility on driving behaviors and response performance. *Applied ergonomics*: 1–8. <http://doi.org/10.1016/j.apergo.2011.10.005>
20. Tara Matthews, Tye Rattenbury, and Scott Carter. 2007. Defining, designing, and evaluating peripheral displays: An analysis using activity theory. *Human-Computer Interaction* 22, 1–2: 221–261. <http://doi.org/10.1080/07370020701307997>
21. David K McGookin and Stephen A Brewster. 2004. Understanding concurrent earcons: applying auditory scene analysis principles to concurrent earcon recognition. *ACM Transactions on Applied Perception (TAP)* 1, 2: 130–155. <http://doi.org/10.1145/1024083.1024087>
22. Natasa Paterson, Katsiaryna Naliuka, Soren Kristian Jensen, Tara Carrigy, Mads Haahr, and Fionnuala Conway. 2010. Design, implementation and evaluation of audio for a location aware augmented reality game. Proceedings of the 3rd International Conference on Fun and Games - Fun and Games '10, 149–156. <http://doi.org/10.1145/1823818.1823835>
23. Miguel a. Recarte and Luis M. Nunes. 2003. Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied* 9, 2: 119–137. <http://doi.org/10.1037/1076-898X.9.2.119>

24. Nash Stanton, Roger Lew, Nolan Boyle, et al. 2011. An Implementation of a Graded Deceleration Display in Brake Light Warning Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 55, 1: 1573–1577.
<http://doi.org/10.1177/1071181311551328>
25. Heikki Summala. 2007. Towards understanding motivational and emotional factors in driver behaviour: Comfort through satisficing. In *Modelling driver behaviour in automotive environments*. Springer London, 189–207. Retrieved November 17, 2014 from http://link.springer.com/chapter/10.1007/978-1-84628-618-6_11
26. MinJuan Wang, Yi Ci Li, and Fang Chen. 2012. How can we design 3D auditory interfaces which enhance traffic safety for Chinese drivers? *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ACM Press, 77–83.
<http://doi.org/10.1145/2390256.2390268>
27. Minjuan Wang, Sus Lundgren Lyckvi, and Fang Chen. 2016. Same , Same but Different : How Design Requirements for an Auditory Advisory Traffic Information System Differ Between Sweden and China. *8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*.
28. Minjuan Wang, Sus Lundgren Lyckvi, and Fang Chen. 2016. Why and How Traffic Safety Cultures Matter when Designing Advisory Traffic Information Systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
<http://doi.org/10.1145/2858036.2858467>
29. Marcus O Watson and Penelope M Sanderson. 2007. Designing for attention with sound: challenges and extensions to ecological interface design. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49, 2: 331–346.
30. Deborah Withington. 2000. The Use of Directional Sound to Improve the Safety of Auditory Warnings. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 44, 22: 726–729.
<http://doi.org/10.1177/154193120004402260>