Sonification Use Cases in Highly Automated Vehicles:

Designing and Evaluating Use Cases In Level 4 Automation

Chihab Nadri[1], Sangjin Ko[1], Colin Diggs[1], Michael Winters[2], Sreehari V. K. [1], and Myounghoon Jeon[1]

[1]Grado Department of Industrial and Systems Engineering, Virginia Tech

[2]School of Music, Georgia Tech

Author Note

We have no conflict of interest to disclose.

Abstract

The introduction of highly automated driving systems is expected to significantly change in-vehicle interactions, creating opportunities for the design of novel use cases and interactions for occupants. In this study, we sought to identify and extract these novel use cases and determine preliminary auditory display recommendations for these novel situations. We developed and generated use cases for level 4 automated vehicles through an expert workshop (*N*=17) and online focus group interviews (*N*=12). Most of the use cases we generated were then tested, apart from meditation, and user opinions were collected in a driving simulator study (*N*=20). Results indicated participants were interested in functions that support their experience with both driving and non-driving related interactions in highly automated vehicles. Three categories of use cases for level 4 automated vehicles were developed: driving automation use cases, immersion use cases, and in-vehicle notification use cases. For the driving simulator study, we tested three display modalities for interaction with drivers: visual alert only, non-speech with visual, and speech with visual. In terms of situation awareness (SA), the non-speech with visual display was associated with significantly better SA for the use case consisting of a planned increase in automation level than the speech-with visual display. This study will provide guidance on sonification design to advance user experiences in highly automated vehicles.

Keywords: highly automated driving, participatory design, driving simulator, sonification

Introduction

Automation technology is changing the way drivers interact with their vehicles. As technology advances, users of automated vehicles will increasingly disengage with the driving task and perform non-driving tasks. Highly automated vehicles, defined by SAE International as level 4 vehicles (SAE-International, 2014), are expected to reduce driver stress and increase productivity (Litman, 2020). Challenges exist in the adoption and acceptance of higher levels of automation in vehicles, revolving around safety concerns with the vehicles, and can be addressed through novel interactive displays for safe transitions in automation (Jeon, 2019), such as robot agents (Lee et al., 2019) or augmented reality displays (von Sawitzky et al., 2019). Furthermore, researchers have suggested the use of adaptive auditory alerts (Šabić et al., 2019), spatial sounds (Petermeijer et al., 2017), or other auditory displays as ways to quickly provide alerts or feedback for drivers to takeover or respond to different vehicle states.

Sonification, which is transcribing data into non-speech sound (Nees & Walker, 2011), is a display method that has been suggested for automated vehicles (e.g., driving data sonification) (Landry et al., 2016). This display method has been used to transcribe vehicle states, intentions, or user emotions to increase situation awareness (SA) (Gang et al., 2018; Landry et al., 2016). SA can be defined as the perception of the 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 (Endsley, 1988a).

Attitudes and user information requirements for automated driving (Hock et al., 2017; Lee et al., 2020) indicate an interest in receiving information about the vehicle's state, intentions, or environmental information. The use of sonification could be extended to both driving and non-driving related tasks and activities as an additional display modality to improve occupant experience, trust, and acceptance of automated vehicle technology. Trust measures the degree of confidence individuals have in strangers or in technology, with trust in automation being a major focus of research (Hoff & Bashir, 2015).

However, these methods have not yet been evaluated for the operation of level 4 automated vehicles, despite the interest in developing situations for highly automated driving in the past (Bosch et al., 2018; Frison et al., 2017; Meschtscherjakov et al., 2015; Riener et al., 2016). Furthermore, it is also important to assess design and use case approaches for highly automated vehicles, as well as identify occupant preferences in such conditions, as part of participatory design (Bergold & Thomas, 2012; Jeon et al., 2011). As such, there is a gap in research regarding both the value and design of sonification in level 4 automated vehicles in both contemporary use cases that will adapt to the new automation level, as well as new prospective use cases that will be enabled from this increase in automation.

To address this research gap, we have conducted several studies to identify key use case categories, design sonification applications for these use cases, and evaluate their perceived value and effect on drivers. To do so, we conducted a workshop with experts in the automotive and audio fields, online focus group sessions with young drivers, and a preliminary driving simulation study using sonifications developed for relevant use cases in level 4 automated driving. The results of our study can help inform automotive display designers regarding evolving auditory display needs for contemporary driving use cases and novel use cases targeting immersive applications (e.g., meditation and live journey sonification).

Research process

We conducted a workshop with experts in the automotive and auditory display design fields to generate hypothetical use cases unique to highly automated driving and benefit from the perspective of experts (Pfadenhauer, 2009). Use cases were defined as situations encompassing an interaction occurring in the vehicle, whether it is driving-related or not. Use cases allow for the description of events that lead to a system doing something useful (Bittner & Spence, 2003), making them an effective means to extract important tasks and interactions in highly automated driving. We also conducted online focus group interviews (FGIs) and a driving simulator study with young drivers to evaluate the selected use cases and create general design directions. Younger drivers were recruited to collect opinions from potential users familiar with concepts of driving automation and complement expert feedback, as part of the participatory design process (Bergold & Thomas, 2012; Jeon et al., 2011). For both the online FGIs and driving simulator study, we collected quantitative and qualitative data following a mixed-methods research approach (Creswell & Clark, 2017). Detailed information about how we conducted the study and analysed the results of each study component is covered in the following sections. All research procedure conducted was approved by the Virginia Tech Institutional Review Board (VT IRB).

Expert workshop

Methods

Experts volunteered to participate in the workshop during the 25th International Conference on Auditory Display (ICAD). Workshop objectives were to define use cases and sound concepts for highly automated driving, as well as refine understanding of user needs in highly automated vehicles.

Participants

In total, 17 practitioners and researchers participated in the workshop, with eight experts from academia and nine from industry/government. Practitioners had a varied experience profile in terms of years of practice (*M*=11.24yrs, *SD*=8.36yrs).

Procedure

The workshop took a full day and lasted six hours in total, with several breaks in the middle. The full-day workshop began with an ice-breaking session before a brainstorming activity was conducted. Over the course of the brainstorming session, experts were instructed to develop scenarios, brief use case descriptions, and key elements pertaining to seven themes of sound that can contribute to the automated vehicle environment, according to previous research in the field (Kun et al., 2016). These themes included the following: 1) Safety, 2) Privacy/security, 3) Usability, 4) Situation awareness, 5) Trust, 6) Play, and 7) Work (as in, work while the vehicle takes care of driving; the mobile office concept (Kun et al., 2016)). Sticky notes for each theme were placed on a board, and experts were asked to write down the key elements and brief use cases in sticky notes that would be placed next to the relevant overarching theme sticky note. Experts were informed at the start of the session that the use cases they would discuss were highly automated vehicles and they were made aware of their functionalities.

Next, experts were divided into three teams at random and asked to develop one detailed auditory display use case in a specific theme. When developing the concept, experts were tasked with providing a storyboard, design rationale, and a short sound prototype for likely sounds used. Each team then presented their detailed use case and discussed over the next hour.

Results

Brainstorming wall

Through the affinity diagram activity, a total of 107 use cases were created by experts. We re-organized these use cases by theme, sound type, and sound source. Experts wrote use cases on designated spots according to the theme of the situation. One researcher was tasked with identifying the sound type and source based on what experts wrote.

Theme: for the seven themes used to organize the workshop scenarios (Kun et al., 2016), experts generated at least 11 use case per theme. The two most represented use case themes (Table 1) were "play" and "situation awareness".

Table 1

Use cases developed by experts during the brainstorming wall activity per theme

Use case theme	Count #
Play	22
Situation Awareness (SA)	16
Privacy/Security	15
Trust	15
Usability	15
Work	13
Safety	11

For "play" use cases, experts expected the highly automated vehicle to provide a variety of games using the vehicle surrounding traffic, such as using the steering wheel of the vehicle to play a virtual game or socializing and playing with occupants in other cars through shared-network activities. Other use cases included meditation and using surrounding environmental information and coordinates to bring outside world sounds for tourist experiences and learn about nearby cities and landmarks.

For "situation awareness" use cases, experts believed that the automation system should provide at least, low-level situation awareness data and vehicle intentions to occupants. Use cases meant to recover situation awareness were frequently mentioned in the case of waking occupants or to respond to an urgent situation. As was found for the "play" theme, experts believed that sound could be used to provide information about the surroundings, with a focus on road condition, weather, and potential hazards.

Sound type: The two most frequently mentioned sound and alert types were sonification and speech. Sonification was used to transcribe vehicle surroundings, intentions, and road conditions. Speech was used for social activities in the car with other vehicle occupants, as well as receiving information from outside the car.

Content: experts developed use cases for a variety of situations and contexts in highly automated vehicles. While the most frequent cases still focused on the driving task, experts also suggested many use cases related to road conditions and driving status. As discussed earlier, experts believed that vehicle occupants needed to know about the surrounding road condition and upcoming obstacles in the driving journey. Another frequently mentioned group of use cases focused on in-vehicle notifications, with content related to security notifications, car status, and vehicle occupant health. With the increase in automation and reduction in control of the driving task, experts commented on the importance of in-vehicle notifications to maintain appropriate trust from users. Finally, the last group of use cases were related to non-driving related tasks, with use cases enhancing travel experience by providing information and thematic sounds of landmarks being frequently mentioned. Additional use case content in this theme included video games, board games between vehicle occupants, and virtual reality environments for occupants to explore while the vehicle is driving.

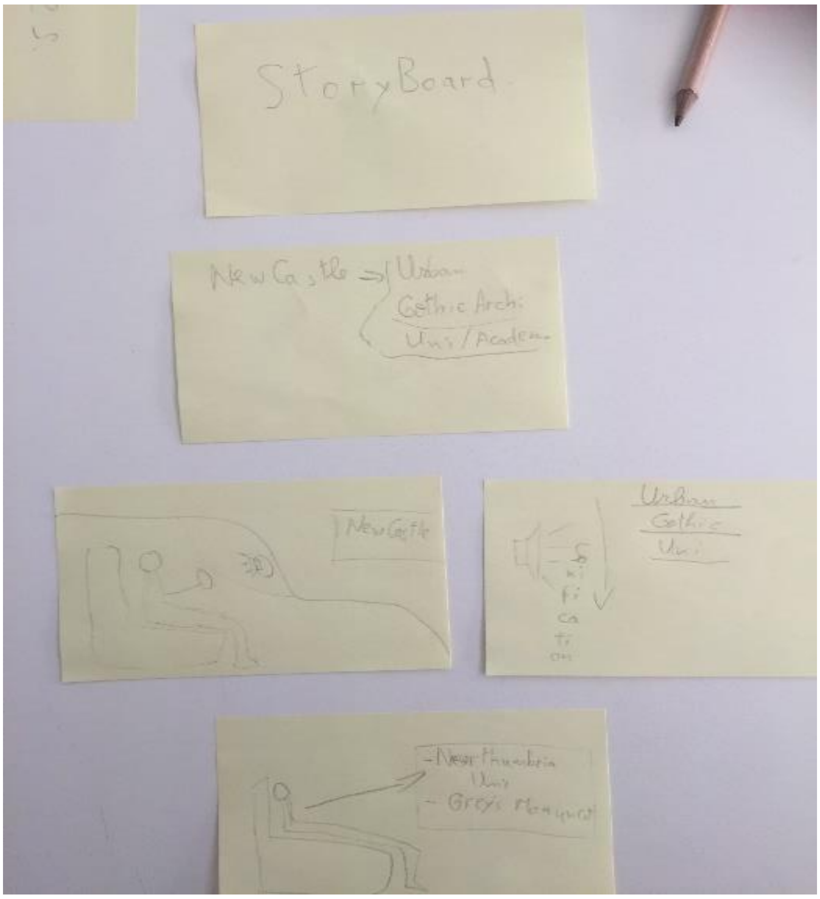
Group concept development

The three teams created during the workshop expanded the following use cases and categories:

- Regaining trust: this use case from the “trust” theme focused on providing vehicle occupants with information following an error in the automation system. Experts considered the use of both music and speech to both regulate occupant emotions (Fakhrhosseini et al., 2014) and provide sound reasoning to allow occupants to know how the error would be corrected in future instances.
- Sonic logo for places: this use case from the “play” theme sought to provide a brief sonic logo summarizing nearby landmarks and locations reached. When approaching a town such as Newcastle (Figure 1), the gothic architecture of the city could be sonified by playing music from a similar time-period.

Figure 1

Storyboard made for the sonic logo concept



- Meeting on the mobile office: this use case from the “work” theme related to the use of new auditory display tools in a highly automated vehicle. As experts assumed that the layout of the automated vehicle would be arranged differently, they suggested using window screen displays and spatial audio in the vehicle to enhance virtual conference experiences while driving.

Online focus group interviews

The aim of the FGIs was to evaluate salient use cases developed from the prior workshop and literature review and elicit opinions regarding potential user needs and preferences in highly automated driving.

Methods

We conducted focus group interviews (FGIs) online due to the COVID-19 pandemic, and three group sessions were held over Zoom.

Participants

We recruited 12 participants for the FGIs, distributed between three group sessions that were held. Participants were young university students (*M*=23.33yrs, *SD*=3.52), with five male and seven female participants.

Procedure

Participants received a video conference link via Zoom to join the online FGI session. All participants provided consent for the study by signing an informed consent form approved by the VT IRB.

The researchers started the online session by explaining the purpose of the session, explaining the functionalities afforded by highly automated vehicles, and presenting use cases generated from the previous expert workshop, as well as a subsequent review of relevant literature on the subject of automated vehicle functionalities (Gang et al., 2018; Hock et al., 2017; Landry et al., 2016). The seven use cases reflected common and important situations expected in highly automated vehicles and were categorized into three groups (Table 2). Unlike use case themes (Kun et al., 2016), which were used as a starting point for discussion with experts, use case groups were based on the analysis of feedback provided by experts.

Table 2

Use cases and use case groups identified during the online focus group

Use case groups	Use case
Driving automation	Automation level change
	Handover/takeover request
Immersion	Meditation
	Live trip sonification
In-vehicle notification	Battery status alert
	Virtual conference call notification
	Internet and network security notification

The groups were as follows:

- Driving automation use cases: following both the workshop and literature review, we identified two driving automation change situations. The first situation, labelled as an automation level change, was defined for planned transitions in automation for complex road conditions, as the automation system in level 4 automated vehicles is not expected to control the vehicle in all road types (SAE-International, 2014). We identified the second use case as a handover/takeover request to designate urgent requests from the automated system when it encounters an unexpected obstacle and requires drivers to regain control.

- Immersion use cases: following recommendations by experts to introduce novel presence use cases for vehicle occupants in the form of travel surroundings, we developed the live trip sonification use case. Immersion in this situation refers to the degree of involvement and integration of the user's sense of surroundings with an environment, as understood from the concept of game immersion (Brown & Cairns, 2004). We envision this group of use cases to seek in immersing drivers in either states not related to the vehicle or with the environment the vehicle is passing by. In the case of immersing drivers into the driving environment around them, which we called a live trip sonification, we utilize auditory icons (sounds that are related to their referent object, event, or process) (Gaver, 1986) and natural soundscapes to relay information on the environments surrounding the driver in real-time (e.g., birds chirping for a forest, seagulls and waves crashing for a beach) (Mauney & Walker, 2004). In the second use case related to immersion into a meditative state, we envisioned the case would take shape in the automated vehicle through using meditation techniques and calming soundscapes to relax drivers.
- In-vehicle notification use cases: the last group of important use cases for drivers and occupants in highly automated vehicles pertained to in-vehicle

notifications. Though not comprehensive of every notification mentioned in the workshop, three main use cases were presented, first with a battery status alert, followed by virtual conference call notifications, and lastly, Internet and network security notifications.

When presenting each of the seven use cases, we generated both speech and non-speech sounds for participants to rate. In the case of speech alerts created, a male, native English speaker voice was used. For the immersion use cases, a voice clip was played describing the area crossed by an individual in their journey (e.g., “Passing by a forest”), whereas meditation instructions were provided for the meditation use case (e.g., “Breathe slowly and relax”). Non-speech sounds were created for all seven use cases as well, and included both earcons (short abstract sounds) (Blattner et al., 1989) and in the case of the live journey sonification event, auditory icons in the form of natural sounds to be encountered when crossing the same environment (sounds of birds chirping in a forest, sounds sea waves and seagulls for a beach environment) (Gaver, 1986). The design of non-speech sounds for both environmental auditory icons and earcons followed guidelines and recommendations previously set by researchers (Brewster, 1994; Gaver, 1986).

This was followed by a group discussion regarding the situations shown. Participants were then instructed to listen to alerts made for each use case. Participants could listen to each sample as many times as possible before rating each alert, and study investigators waited for participants to rate and listen to all sounds made for each use case. After completing an evaluation of sounds for each use case, a group discussion was briefly held to gather subjective feedback and user preferences. Throughout the FGI, one of the researchers took notes of the discussion.

Results

Focus group participants indicated that automation level change (n=7), handover/takeover requests (n=4), and battery status alerts (n=1) were the most useful use cases in a highly automated vehicle. In terms of personal preference, participants selected live trip sonification, battery status alerts, and takeover/handover requests as the most favored use cases. When asked by the researchers for their recommendations on how to better implement the use cases selected, participants said the following:

- Automation level change: “I think a multimodal setup alerting the driver as to what the mode is would be helpful. The sounds are a little confusing. I'd rather see the sounds on a gradient, changing levels according to the mode” (P1 in FG2) – “Voice overs for hard-to determine vehicle functions and spearcons or auditory icons for obvious functions” (P2 in FG3).
- Takeover/Handover request: “I think it would be useful to hear when we should control or leave the control. More specifically, it would be helpful if you get asked about the automation. I like it how a machine could communicate for decisions.” (P1 in FG3) – “It's important to notify the driver of changing in levels, and being clear when giving options for those changes, etc.” (P4 in FG3).
- Live trip sonification: “I believe it can be implemented depends on the surrounding and it can be calming in high traffic areas.” (P3 in FG2) – “In traffic jams/less pleasant locations to make a car ride more enjoyable: in locations that match the noises to add experience when opening windows isn't an option/the user wants privacy” (P5 in FG2).
- Battery status alert: “Based on the scheduled trip if the alert if the driver needs to recharge the battery to complete a full trip, or if near a charging station and battery is in a certain range, alert the driver to charge.” (P5 in FG2).

Focus group comments were incorporated into the design of use case implementations within the driving simulator study. A multimodal setup using both visual and auditory elements to alert drivers was used as per focus group recommendations. Additionally, automation level change and takeover request alerts were provided with detailed, clear, and adequately urgent instructions both in terms of speech, “please take control of the vehicle” or non-speech earcons (e.g., a repetition of tones with two dominant frequencies of 880 and 1760Hz for high perceived urgency) (Jeon, 2019).

Driving simulator study

The aim of the driving simulator study was to empirically assess the effects of auditory displays and sonifications designed from the iterative needs analysis studies within a more controlled experimental setup.

Methods

Participants

For the driving simulator study, 20 participants were recruited (female = 6, male = 13, nonbinary = 1). Participants were university students and residents from the area with a valid driving license, with an average age of 24.3 years old (*SD* = 6.32). Only one participant had prior experience with automated vehicles. Participants had a valid driving license for a number of driving experiences (*M* = 5.65 years, *SD* = 4.25) when participating in the study. Out of the 20 participants recruited, 8 drove for more than 10,000 miles in the past year and 12 drove less than 10,000 miles in the past year. All participants provided consent.

Experimental design

A within-subject design was implemented with three conditions (Visual alert only (V), Non-speech with visual (N), Speech with visual (S)). In each of the three driving scenarios participants completed, seven similar events were encountered. The events reflected the use cases discussed in the focus group sessions apart from the meditation use case. We did not expose users to meditation sounds due to the nature of the short drive conducted: as participants completed the driving loop in 12 minutes or less, there would not be enough time to expose participants to a meditative session. Additionally, the focus group results indicated that the use case did not receive as much interest as other cases from participants.

The order of events reflected a single driving session with an electric vehicle, which we assume to be the type of vehicle that will possess advanced automation technology for SAE level 4 automation. Participants started their drive with an initially charged vehicle, indicated by a full battery notification. The vehicle eventually lost battery and subsequently downgraded in automation level. We provided this progression of events for the battery status and automation level change use cases as the most plausible succession of events in such a case. Meanwhile, the order of events for the video conference, takeover situation, live trip sonification, and phone call use cases was randomized to reduce familiarity effects.

Apparatus

A motion-based driving simulator (Nertech) was used in the study. Visuals were displayed on three 48” displays, and the simulator was equipped with a surrounding sound dome. During the simulated drive, participants received information from visual or audiovisual cues. Visual cues consisted of a visual icon displayed on the vehicle dashboard (Figure 2).

Figure 2

Participant’s view and dashboard in the driving simulator during the live journey event at a beach (left) and a forest (right)

Stimuli

Auditory cues were speech or non-speech messages depending on the scenario condition (Table 3). The set of non-speech sounds used in the driving simulator study originated from non-speech cues previously created and evaluated during the FGI. Selection was based on perceived recognizability and user feedback during the FGI. For the live journey sonification event, environmental sounds (birds chirping for the forest environment, sea waves crashing and seagulls for the beach environment) were used in the non-speech with visual condition. A speech message indicating the vehicle's current location was used in the speech with visual condition. All sounds lasted at most seven seconds per event.

Table 3

Summary of sound cues provided to participants according to the use cases investigated in the driving simulation study

Use case	Speech cue and length	Non-Speech cue and length
Automation level change	- "Increase in automation level" (3s) - "Decrease in automation level" (3s)	- Set of three piano notes increasing in pitch (3s) - Set of three piano notes decreasing in pitch (3s)
Handover/takeover request	"Please take control of the vehicle" (2.5s, repeat until takeover)	Two dominant frequencies (880, 1760 Hz) repeated four times (1.5s, repeat until takeover)
Live trip sonification	- "Passing by a beach" (2.5s) - "Passing by a forest" (2.5s)	- Hyundai brand soundscape with the sound of waves crashing and seagulls (15s) - Hyundai brand soundscape with the sound of birds chirping in a forest (15s)
Battery status alert	- "Battery is full" (2.5s) - "Battery is low" (2.5s)	- Six-note increasing melody ending with a high pitch (2.5s) - Six-note decreasing melody ending with a low pitch (2.5s)
Virtual conference call notification	"Incoming call" (2s)	Phone ring sound with a six-note melody (4s)

Procedure

Before the study started, participants provided consent for the study by signing an informed consent form approved by the VT IRB. Participants initially had a short driving session to test for simulation sickness (Gable & Walker, 2013). Participants were familiarized with the driving simulator and were informed that they would complete three 12-minute trials. The vehicle was assumed to be a level 4 automated vehicle. Study investigators explained level 4 automation to participants, informing them that the in-vehicle system would be able to take care of the driving task in most cases, though in some rare instances it might require takeover. Participants were told that they could perform non-driving tasks and only needed to takeover control if the vehicle requested it. After completing each run, participants answered statements on a seven-point Likert-scale regarding the influence of auditory displays on trust and their assessment of the experience.

For each drive, after each of the seven events happened (a second after the end of the auditory signal), researchers interrupted the driving simulation, which was paused and blocked out from participants' view, before querying participants about their perception, understanding, and future projection of the present situation, following the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988b). Participants then answered queries regarding the event for a 0 or 1 score for each SA component. Participants received a 1 score for all three event SA components if they (1) perceived the in-vehicle notification; (2) understood why the notification is happening; and (3) could predict subsequent events or actions that happen next. Participants could receive partial scores based on answering one or two SA components correctly. Participants filled out the electronic version of the NASA-TLX workload index at the end of each simulated run. After completing each run, participants answered an adapted version of the Trust in Automated Systems scale (Jian et al., 2000) for automated vehicles ($\alpha = .90$) with a seven point Likert-scale. SA, trust, and user experience were selected as outcome variables based on the definition of highly automated vehicles (SAE-International, 2014) and related research on requirements for fully automated driving (Lee et al., 2022) indicating that trust and acceptance will be salient factors heavily influencing the popularity of the technology, while SA is crucial to maintaining safety in level 4 driving automation.

Results

The non-parametric Friedman test was used for valence, arousal, and the "Trust in Automated Systems" scale, as the data did not follow a normal distribution. If a significant difference was found, a Wilcoxon signed-ranks test was used for pairwise comparisons, using a Bonferroni correction for comparisons based on the sound condition (with an adjusted $\alpha = 0.05/3 = 0.0167$).

Situation Awareness (SA)

Table 4 presents results of the analysis of SA scores for the different display conditions. Significant differences were found between the different display conditions for all events, except for the decrease in automation event. Pairwise comparisons showed that the audiovisual conditions induced higher SA for battery alerts, takeover requests, and live trip sonification. The non-speech with visual condition resulted in significantly higher SA than speech with visual for the increase in automation event.

Table 4

Situation Awareness scores for different display conditions (V = visual-only, N = non-speech with visual, S = speech with visual)

Event	Display condition			Sig.	Pairwise comparisons
	V	N	S		
Full Battery	0.550	0.883	0.966	$p < 0.0001$*	V < N = S
Low Battery	0.633	0.883	0.883	$p < 0.0001$*	V < N = S
Takeover Request	0.733	0.950	1.000	$p < 0.0001$*	V < N = S
Incoming Call	0.866	0.983	0.950	$p = 0.013$	V < N
Increase in Automation	0.550	0.916	0.700	$p < 0.0001$*	V = S < N
Decrease in Automation	0.816	0.833	0.933	$p = 0.081$	V = N = S
Live Trip Sonification	0.650	0.883	0.933	$p < 0.0001$*	V < N = S

* $p < 0.05$

NASA-TLX

For the NASA-TLX overall score, a main effect for display type was found $F(2, 38) = 3.939$, $p = 0.028$. The visual-only display condition ($M = 52.25$, SD = 20.40) resulted in significantly higher workload scores than the non-speech with visual display ($M = 41.26$, $SD = 17.31$), $p = 0.013$ and *marginally* higher than the speech with visual display ($M = 42.98$, $SD = 21.27$), $p = 0.033$ (note that this is not statistically significant with the adjusted alpha level, 0.0167). There were no statistically significant differences between the speech with visual display and the non-speech with visual display.

For the NASA-TLX subscale scores, a main effect for display type was found for physical demand $F(2, 38) = 3.623$, $p = 0.036$. The non-speech with visual condition (M = 20.75, SD = 13.50) resulted in significantly lower scores than the visual-only display (M = 30, SD = 22.71), $p = 0.012$. There were no statistically significant differences between the speech with visual display (M = 26.75, SD = 17.57) and the non-speech with visual display. A main effect for display type was found for effort $F(2, 38) = 5.525$, $p < 0.0001$. The non-speech with visual display (M = 31.75, SD = 17.27) resulted in significantly lower scores than the visual-only display (M = 51.25, SD = 25.80), $p = 0.002$. There were no statistically significant differences between the speech with visual display (M = 39, SD = 23.20) and the non-speech with visual display. A main effect for display type was found for frustration $F(2, 38) = 4.005$, $p = 0.026$. The speech with visual display (M = 30.25, SD = 19.57) resulted in significantly lower scores than the visual-only display (M = 46.5, SD = 29.11), $p = 0.009$. There were no statistically significant differences between the non-speech with visual display (M = 35.25, SD = 21.85) and the speech with visual display. There were no statistically significant differences among the three conditions for mental demand, temporal demand, and performance.

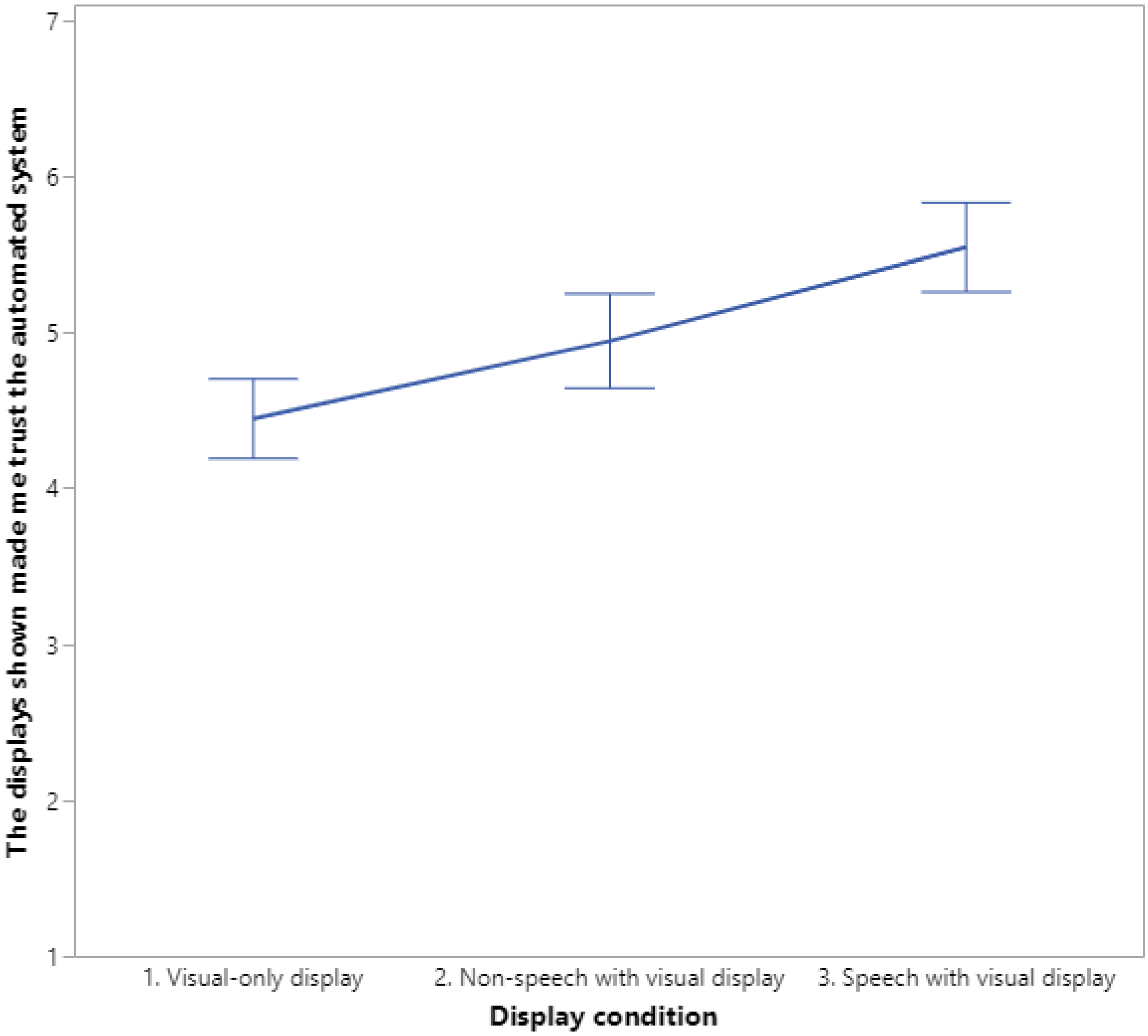
Trust

Participants responded to whether the displays generated while driving made them trust the automated system after the end of each drive.

There were significant differences in trust for display type (Figure 3). A Friedman test resulted in $\chi^2(2) = 7.433$, $p = .0243$. A Wilcoxon each pair comparison showed higher trust levels for the speech with visual display condition when compared to the visual-only display condition $Z = 2.662$, $p = .0078$.

Figure 3

Trust in automated system following introduction of displays depending on display type, error bars represent standard errors.

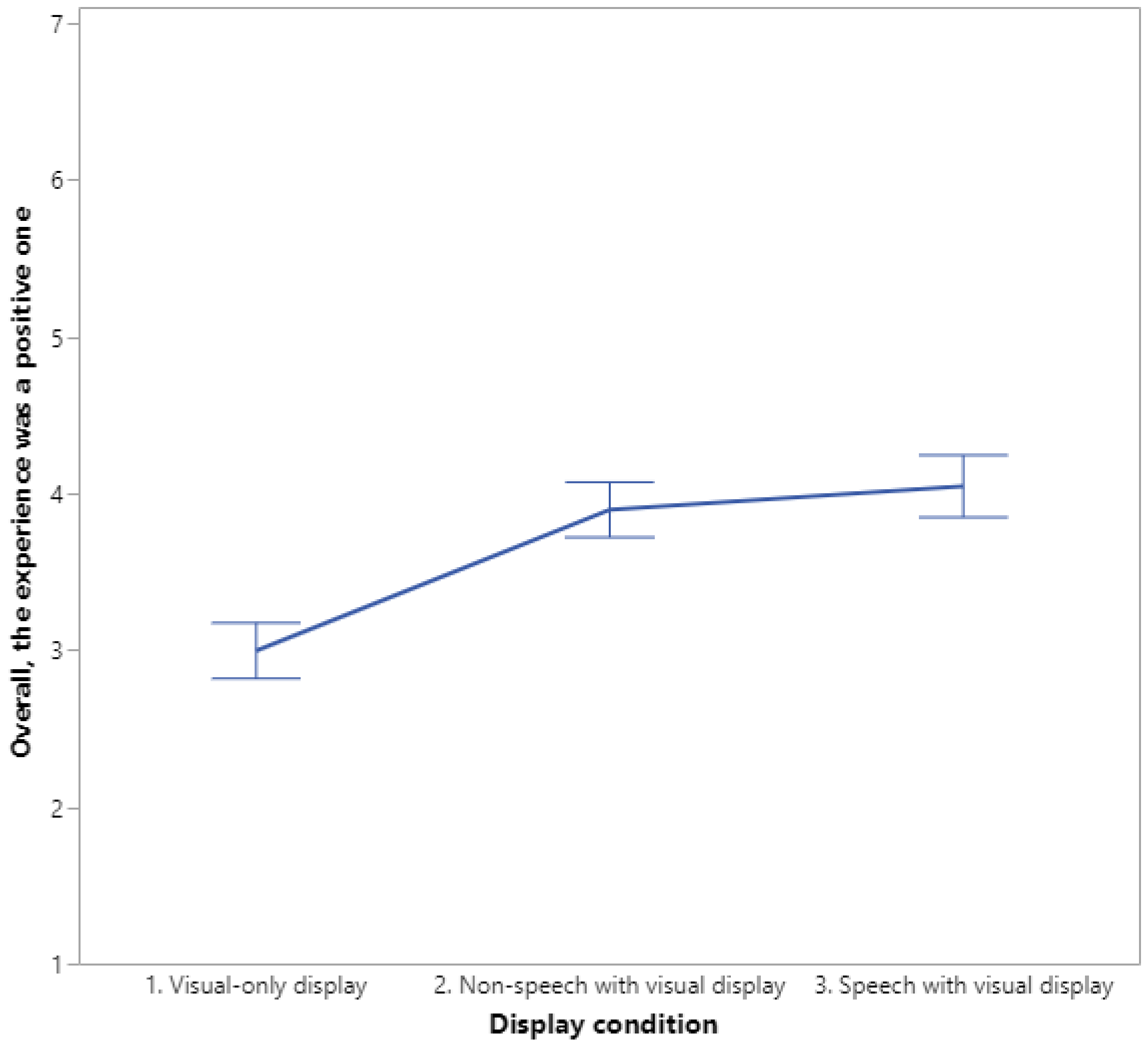


Opinion on driving experience

Participants responded to whether the driving experience was a positive one after the end of each drive. There were significant differences in opinion for display type (Figure 4). A Friedman test resulted in $\chi^2 (2) = 15.186$, $p = .0005$. A Wilcoxon each pair comparison showed participants viewed the experience more positively for the speech with visual condition when compared to the visual-only condition $Z = 3.408$, $p = .0007$. Participants viewed the experience more positively for the non-speech with visual display condition when compared to the visual-only display condition $Z = 3.241$, $p = .0012$.

Figure 4

Opinion on the automated driving experience based on display type, error bars represent standard errors.



Discussion

In this study, we developed use cases of auditory displays for highly automated vehicles and determined general design directions for potential situations in highly automated vehicles. We collected expert opinions and use cases during the first workshop activity. Through this step, we were able to identify seven use cases to evaluate in subsequent focus group interviews and driving simulator studies.

Experts generated use cases along seven tasks (Kun et al., 2016). The focus on "play" use cases highlight the expected shift in user needs in highly automated vehicles (Lee et al., 2020). As occupants have more free time in the vehicle, the scope of leisure activities was expected to increase. However, because level 4 vehicles will require occupants to take control in select situations, getting the driver back in the loop was deemed important, as it garnered the next highest amount of use cases through the "situation awareness" task. This falls in line with previous research on the subject (Du et al., 2020; Gold et al., 2013; Sanghavi, 2020), as appropriate takeover and handover influence trust in the automated vehicle. The expanded use cases developed by expert teams also helped in highlighting three main design considerations and changes relevant to a change to level 4 automated vehicles: the need for vehicle-user trust, the need to facilitate or provide non-driving tasks for leisure, and the ability to modify the vehicle environment and devices within to enhance user experience. While these considerations are present in lower levels of automation, user expectations and requirements can be addressed in different ways as drivers have less involvement in the driving task and share more similarities to passengers, with more focus on smooth short transitions into manual driving before leaving the system to take over the driving task.

During the focus group sessions, we explored drivers' subjective evaluations on the use cases developed from the expert workshop and literature review we conducted. We were able to obtain user preferences and priorities in level 4 automated driving. The importance participants indicated for the change in automation level and takeover request use cases falls in line with the previous findings on the need to provide adequate situation awareness (Hester et al., 2017; Köhn et al., 2019). To address this user need, clear and precise information will need to be displayed ahead of time for users. This is also in line with previous research identifying this as a key information need (Hancock et al., 2020; Lee et al., 2020) which could affect user trust, and in turn, acceptance of automated vehicle technology (Choi & Ji, 2015; Haspiel et al., 2018). In terms of user preference, participants' views also converged with expert recommendations and use cases during the workshop activity, as the live trip sonification use case received increased interest. During the focus group interviews, participants also recommended providing additional feedback and user choices.

When asked if the displays presented made participants more trustful of the automated system, a statistically significant effect was found for display types, with the speech-visual display having higher trust scores than the unimodal condition, which is in line with the previous studies about multimodal displays (Jeon, 2019; Liu, 2001; Petermeijer et al., 2017). While user trust in the automated system was numerically higher for the non-speech with visual condition, no statistical difference was found with the other displays. This lack of statistical significance in trust score could be investigated further through a wider participant pool.

Finally, the use of both speech and sonification were shown to improve user experience in the automated driving condition presented to participants. These findings echo previous benefits associated with the use of auditory displays and sonification on driving satisfaction and performance (Fakhrhosseini et al., 2014; Lee et al., 2019; Nadri et al., 2021). Statistically significant differences were found between the speech with visual and non-speech with visual conditions. Speech with visual was associated with lower frustration scale scores, while non-speech with visual was associated with higher situation awareness for the automation level increase. These results align with research on speech alerts (Nees & Walker, 2011) and focus group results, as speech alerts were identified as more pleasant than non-speech alternatives. In addition, non-speech earcons are associated with higher urgency levels (National Highway Traffic Safety, 2016), making them suitable for safety-critical and time-sensitive alerts. It is based on these results that we recommend the use of a non-speech with visual display for abrupt takeover alerts, and the use of a speech with visual display for planned automation level changes, because one type of display is more suited to the different level of urgency than the other.

Limitations and future work

The current study identified and designed several use cases in level 4 automated vehicles in addition to testing sonifications that could address user needs and improve user trust in the automated system. While the driving study was able to identify the effects of the display type on situation awareness, trust results in terms of statistical significance could be investigated further by conducting a study with a wider pool of participants.

Additionally, we have been able to categorize the use cases we designed into three main groups: driving automation, immersion, and in-vehicle notifications. A future set of studies could focus on each use case group separately. Subsequent work on meditation and live trip sonification within the immersion use case group can be conducted to evaluate long-term user experience benefits and requirements, and better explore more novel use cases that will be afforded as a result of an increase to level 4 automation.

Conclusion

In this study, we developed and tested use cases for level 4 automated vehicles through three iterative participatory design studies. Use cases that were developed from the expert workshop were tested in focus group sessions and a driving simulator study, and both user opinions and subjective ratings were collected. Results revealed the need to differentiate voluntary and planned automation level changes with involuntary and abrupt takeover situations for level 4 automated vehicles. Participant data and comments indicate that speech alerts are more suitable to planned changes, whereas non-speech should be used for abrupt takeovers. Additionally, three categories of use cases for level 4 automated vehicles were developed: driving automation use cases, immersion use cases, and in-vehicle notification use cases. For the driving simulator study, different display conditions were tested, with audio-visual displays improving user experience and affecting trust in automated system, with non-speech alerts being more suitable for safety-critical situations. This study will provide guidance on sonification design for highly automated vehicles.

References

Bergold, J., & Thomas, S. (2012). Participatory research methods: A methodological approach in motion. *Historical Social Research/Historische Sozialforschung*, 191-222.

Bittner, K., & Spence, I. (2003). *Use case modeling*. Addison-Wesley Professional.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, *4*(1), 11-44.

Bosch, E., Oehl, M., Jeon, M., Alvarez, I., Healey, J., Ju, W., & Jallais, C. (2018). Emotional GaRage: A workshop on in-car emotion recognition and regulation. *Adjunct Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications*, *AutomotiveUI 2018*, 44-49. https://doi.org/10.1145/3239092.3239098

Brewster, S. A. (1994). Providing a structured method for integrating non-speech audio into human-computer interfaces.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. CHI'04 extended abstracts on Human factors in computing systems,

Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*, *31*(10), 692-702.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.

Du, N., Tilbury, D. M., Robert, L. P., Zhou, F., Pradhan, A. K., Pulver, E., & Yang, X. J. (2020). Predicting Takeover Performance in Conditionally Automated Driving. *0*, 1-8. https://doi.org/10.1145/3334480.3382963

Endsley, M. R. (1988a). Design and evaluation for situation awareness enhancement. Proceedings of the Human Factors Society annual meeting,

Endsley, M. R. (1988b). Situation awareness global assessment technique (SAGAT). Proceedings of the IEEE 1988 national aerospace and electronics conference,

Fakhrhosseini, S. M., Landry, S., Tan, Y. Y., Bhattarai, S., & Jeon, M. (2014). If you're angry, turn the music on: Music can mitigate anger effects on driving performance. *AutomotiveUI 2014 - 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, in Cooperation with ACM SIGCHI - Proceedings*(September). https://doi.org/10.1145/2667317.2667410

Frison, A. K., Jeon, M., Pfleging, B., Alvarez, I., Riener, A., & Ju, W. (2017). Workshop on user-centered design for automated driving systems. *AutomotiveUI 2017 - 9th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, Adjunct Proceedings*(October), 22-27. https://doi.org/10.1145/3131726.3131734

Gable, T. M., & Walker, B. N. (2013). *Georgia Tech Simulator Sickness Screening Protocol*.

Gang, N., Sibi, S., Michon, R., Mok, B., Chafe, C., & Ju, W. (2018). *Don't Be Alarmed: Sonifying Autonomous Vehicle Perception to Increase Situation Awareness* Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,

Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, *2*(2), 167-177.

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). Take over! How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society*, 1938-1942. https://doi.org/10.1177/1541931213571433

Hancock, P. A., Kajaks, T., Caird, J. K., Chignell, M. H., Mizobuchi, S., Burns, P. C., Feng, J., Fernie, G. R., Lavallière, M., & Noy, I. Y. (2020). Challenges to human drivers in increasingly automated vehicles. *Human Factors*, *62*(2), 310-328.

Haspiel, J., Du, N., Meyerson, J., Robert, L. P., Tilbury, D., Yang, X. J., & Pradhan, A. K. (2018). Explanations and Expectations: Trust Building in Automated Vehicles. *ACM/IEEE International Conference on Human-Robot Interaction*(Dc), 119-120. https://doi.org/10.1145/3173386.3177057

Hester, M., Lee, K., & Dyre, B. P. (2017). "Driver take over": A preliminary exploration of driver trust and performance in autonomous vehicles. *Proceedings of the Human Factors and Ergonomics Society, 2017-Octob*, 1969-1973. https://doi.org/10.1177/1541931213601971

Hock, P., Kraus, J., Walch, M., Lang, N., & Baumann, M. (2017). Elaborating Feedback Strategies for Maintaining Automation in Highly Automated Driving. 105-112. https://doi.org/10.1145/3003715.3005414

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors*, *57*(3), 407-434.

Jeon, M. (2019). Multimodal Displays for Take-over in Level 3 Automated Vehicles while Playing a Game. 1-6. https://doi.org/10.1145/3290607.XXXXXXX

Jeon, M., Roberts, J., Raman, P., Yim, J. B., & Walker, B. N. (2011). Participatory design process for an in-vehicle affect detection and regulation system for various drivers. *ASSETS'11: Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*, 271-272. https://doi.org/10.1145/2049536.2049602

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International journal of cognitive ergonomics*, *4*(1), 53-71.

Köhn, T., Gottlieb, M., Schermann, M., & Krcmar, H. (2019). Improving take-over quality in automated driving by interrupting non-driving tasks. 510-517. https://doi.org/10.1145/3301275.3302323

Kun, A. L., Boll, S., & Schmidt, A. (2016). Shifting gears: User interfaces in the age of autonomous driving. *IEEE Pervasive Computing*, *15*(1), 32-38.

Landry, S., Jeon, M., FakhrHosseini, M., & Tascarella, D. (2016). Listen to Your Drive : An In-Vehicle Sonification Prototyping Tool for Driver State and Performance Data. *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '16)*, 21-26. https://doi.org/10.1109/CIMCA.2005.1631387

Lee, S. C., Ko, S., Sanghavi, H., & Jeon, M. (2019). Autonomous driving with an agent: Speech style and embodiment. *Adjunct Proceedings - 11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019*(September), 209-214. https://doi.org/10.1145/3349263.3351515

Lee, S. C., Nadri, C., Sanghavi, H., & Jeon, M. (2020). Exploring user needs and design requirements in fully automated vehicles. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems,

Lee, S. C., Nadri, C., Sanghavi, H., & Jeon, M. (2022). Eliciting user needs and design requirements for user experience in fully automated vehicles. *International Journal of Human–Computer Interaction*, *38*(3), 227-239.

Litman, T. (2020). Autonomous vehicle implementation predictions: Implications for transport planning.

Liu, Y. C. (2001). Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveler information systems. *Ergonomics*, *44*(4), 425-442. https://doi.org/10.1080/00140130010011369

Mauney, B. S., & Walker, B. N. (2004). Creating functional and livable soundscapes for peripheral monitoring of dynamic data.

Meschtscherjakov, A., Krome, S., & McCall, R. (2015). 3rd Workshop on User Experience of Autonomous Driving 3 rd Workshop on User Experience of Autonomous Driving. (September). https://doi.org/10.13140/RG.2.1.2726.1928

Nadri, C., Lee, S. C., Kekal, S., Li, Y., Li, X., Lautala, P., & Nelson, D. (2021). Effects of auditory display types and acoustic variables on subjective driver assessment in a rail-crossing context. *Transportation Research Record*.

National Highway Traffic Safety, A. (2016). Human Factors Design Guidance for Driver-Vehicle Interfaces.

Nees, M. A., & Walker, B. N. (2011). Auditory Displays for In-Vehicle Technologies. *Reviews of Human Factors and Ergonomics*, *7*(1), 58-99. https://doi.org/10.1177/1557234X11410396

Petermeijer, S., Bazilinskyy, P., Bengler, K., & de Winter, J. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied Ergonomics*, *62*, 204-215. https://doi.org/10.1016/j.apergo.2017.02.023

Pfadenhauer, M. (2009). At eye level: The expert interview—A talk between expert and quasi-expert. In *Interviewing experts* (pp. 81-97). Springer.

Riener, A., Jeon, M. P., Alvarez, I., Pfleging, B., Mirnig, A., Tscheligi, M., & Chuang, L. (2016). 1st workshop on ethically inspired user interfaces for automated driving. Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,

Šabić, E., Henning, D., & MacDonald, J. (2019). Adaptive Auditory Alerts for Smart In-Vehicle Interfaces. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 1545-1549. https://doi.org/10.1177/1071181319631404

SAE-International. (2014). Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems (J3016 Ground Vehicle Standard). In.

Sanghavi, H. (2020). *Exploring the Influence of anger on takeover performance in semi-automated vehicles*

von Sawitzky, T., Wintersberger, P., Riener, A., & Gabbard, J. L. (2019). Increasing trust in fully automated driving: route indication on an augmented reality head-up display. Proceedings of the 8th ACM International Symposium on Pervasive Displays,

Cover Letter:

A part of this study was published as a proceeding in the HFES Annual Meeting Conference 2021.

Nadri, C., Ko, S., Diggs, C., Winters, M., Sreehari, V. K., & Jeon, M. (2021, September). Novel Auditory Displays in Highly Automated Vehicles: Sonification Improves Driver Situation Awareness, Perceived Workload, and Overall Experience. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 65, No. 1, pp. 586-590). Sage CA: Los Angeles, CA: SAGE Publications.

Author Biographies:

Chihab Nadri is a Ph.D. candidate and lab manager of the Mind Music Machine Lab. His research interests include Automotive Display Design, Music Computing, and Human-Computer Interaction.

Sangjin Ko is a graduate from Virginia Tech now working as a Technical Project Manager at Bear Robotics. He was a graduate researcher at the Mind Music Machine Lab.

Colin Diggs is a graduate from Virginia Tech with a major in Industrial and Systems Engineering and currently a data science specialist at MITRE. He was an undergraduate researcher at the Mind Music Machine Lab.

Michael Winters is a graduate from Georgia Tech now working as a Postdoctoral Researcher at Microsoft. He collaborated with the Mind Music Machine Lab as an independent researcher.

Sreehari VK is a graduate from Pandit Deendayal Energy University with a Major in Mechanical Engineering. His research interests include Human-computer interaction and Sound design. He collaborated with the Mind Music Machine lab as an independent researcher.

Dr. Myounghoon Jeon is an Associate Professor of Industrial and Systems Engineering and Computer Science at Virginia Tech and director of the Mind Music Machine Lab. His research focuses on emotions and sound in the areas of Automotive User Experiences, Assistive Robotics, and Arts in XR Environments.