Investigating the Design of Information Presentation in Take-Over Requests in Automated Vehicles

MARK COLLEY, Institute of Media Informatics, Ulm University, Germany
LUKAS GRULER, Institute of Media Informatics, Ulm University, Germany
MARCEL WOIDE, Institute of Psychology and Education, Dept. Human Factors, Ulm University, Germany
ENRICO RUKZIO, Institute of Media Informatics, Ulm University, Germany

In (partially) automated vehicles, users will sometimes have to take over control due to system failure or reach of an operational driving domain end. In such scenarios, the user becomes the driver and has to quickly gain situational awareness. With advanced sensory, an automated vehicle could aid the user in building situational awareness by providing information. A literature analysis found no commonly conveyed information. Therefore, to evaluate the effects of four different abstraction levels (high, medium-high, medium, and low) and used modalities (visual vs. visual+auditory) on situational awareness and accompanying usability scores, we evaluated eight systems and a baseline without information display. In the between-subjects online monitor-based study (*N*=225), we found that while subjective measures are higher and a warning is required, providing abstract information does not improve objective situational awareness, and only being provided with visual information was preferred.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: Takeover; interaction design; information visualization; feedback.

ACM Reference Format:
Mark Colley, Lukas Gruler, Marcel Woide, and Enrico Rukzio. 2021. Investigating the Design of Information Presentation in Take-Over Requests in Automated Vehicles. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction (MobileHCI '21), September 27-October 1, 2021, Toulouse & Virtual, France.* ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3447526.3472025

1 INTRODUCTION

The role of the user inside a vehicle will change from an active driver to a passive passenger with increasing automation. However, automated vehicles (AVs) are not yet able to handle all traffic situations on their own. There are still traffic situations and environmental conditions in which these vehicles reach their intended operational driving domain. circumstances [66, 73]. In such situation, the role of the user changes from a passive passenger to an active driver. Until the technology is fully mature and the autonomous vehicle can handle any traffic situation completely on its own, TORs will continue to be necessary both in scheduled and unscheduled scenarios [53]. Therefore, these transfers of control and their requests (Take-over request — TOR) were extensively studied to understand potential benefits and pitfalls. Numerous aspects have been studied such as priming [5], effects of arousal and emotional valence [17], non-driving related tasks (NDRTs) [84], or age differences [9, 36]. AVs are equipped with numerous sensors such as radars, cameras,

---

and lidars [61]. The gathered data could be used to aid the passenger in the take-over process in forming sufficient situation awareness [19].

The most important factors in the take-over process "are the timely prediction of the handover action" [15, p. 1]), "an informative human-machine interface" (HMI) [15, p. 1]), and "proper training of the driver" [15, p. 1]). An informative HMI could aid users of AVs to quickly (re-)gain an appropriate situation awareness, especially after being engaged in NDRTs. Thus, to safely transfer control of the vehicle in the take-over process, it is necessary to know which information should be conveyed to the AV's user during the TOR, so that the driver can quickly build up extensive situational awareness and take-over safely.

A literature review of existing HMIs that supports the formation of driver situational awareness in general and specifically in TORs was conducted. Of particular interest was which and especially at which level of detail (from specific to abstract) the information and the reason for the TOR were presented. The result was that there are different approaches and ideas about what information should be presented in a certain way. However, almost no concepts could be found that relate the presentation of information to the formation of the driver's situational awareness. Therefore, we designed and conducted an online study (N=225) simulating a take-over at a construction site. We systematically varied the level of detail of the information *information abstraction* (abstract, less abstract, less specific, specific) and *modality* (visual and visual+auditory) for the take-over information presentation. Conveying higher abstraction of information is based on Endsley's argument that it is crucial to convey information, not data [18] and based on the Multiple Resource Theory by Wickens [88] which postulates that using multiple input channels can lead to increased performance. Roche and Brandenburg formulated this expectation with respect to the Multi Resource Theory in the context of TORs as: "Additional information or warnings are perceived faster if they are presented auditory or tactile " [64, p. 1035].

Contribution statement: This work enhances the knowledge on the design of the take-over process from AVs to manual driving. In a monitor-based simulation study with N=225 participants, we found that providing highly abstract information worsens the situation awareness formation process. Also, providing information visually and auditorily does not improve situation awareness.

2 TAKE-OVER REQUESTS

This work builds on the extensive research on take-overs and enhances it towards the actual implementation in AVs.

Take-overs can be broadly distinguished into *scheduled* and *unscheduled* as well as *system-initiated* or *user-initiated* [53]. Another categorization from Lu and de Winter [48] depends on which party initiated and which party is in control after the TOR. There are four such control authority transition categories: automation-initiated automation control (AIAC), automation-initiated driver control (AIDC), driver-initiated automation control (DIAC), and driver-initiated driver control (DIDC). The procedure can be distinguished between *immediate*, *stepwise*, *driver monitored*, and *system monitored* [78]. The take-over mechanism is mostly executed with acoustic and/or visual cues to fade in the take-over prompt [49]. In general, the higher the complexity of the scene and the more demanding NDRTs underwent by the passenger, the more impaired is the take-over performance [81].

2.1 Take-Over Performance Measurements

No standardized method exists for recording the time required to take control of the automated vehicle, resulting in vehicle manufacturers reporting varying mean values from 0.84 s to 3.06 s, or merely stating < 1.00 s and maximum values from < 1.00 s to 21.00 s [49].

The performance of take-overs was mostly assessed with driving behaviors after TORs [16]. These can be broadly distinguished into timeliness and quality [16]. Timeliness includes measurements of response timing to the TOR. Take-over quality metrics include speed, time to collision statistics, crash rate, lane derivation, amongst others. Subjective measurements about the take-over include usability and situation awareness ratings [16]

Köhn et al. [35] showed that interrupting NDRTs repeatedly (every 30s) and showing the scene in front on the screen improves take-over quality. Improved quality in their study meant *increased subjective situational awareness* measured with the SART [71] and *improved reaction times* (time from TOR signal to move steering wheel or press foot pedal) [35]. However, cognitive load was higher.

According to Clark et al. [10], directability through vocal interaction can lead to higher SA. For this, they used some of 10 potential questions, e.g., "What colour is the vehicle in front?" They employed this technique to make sure that the SA is high enough for a planned handover to the human user. The proposed system could also be employed in such a way: the user has to select the correct option from a randomized set of answers to a presented object before the user is allowed to take over control.

Work on estimating driver situational awareness was published by Louie et al. [45]. They used a vision-based system that identifies visual cues to determine their situational awareness: head pose, eye pupil position, average head movement rate and visual focus of attention. An experiment to test the technical capabilities was conducted within a stationary vehicle.

2.2 Post-Automation Effects

While take-overs can be utilized to benefit from the advantages of automation, concerns were also mentioned. Situations that are difficult for an AV to handle are likely to be difficult for a human as well, especially when after a longer period of not driving. Take-overs have shown to have post-automation effects such as unstable lateral control [55] or reduced distance between vehicles after platooning [6]. Therefore, some researchers propose to, instead of shifting control, use a cooperative approach where the passenger aids the AV in the failing task [76, 77, 79, 81]. Humans can help AVs to recognize unforeseen situations and decide on the appropriate action [82], predict pedestrians behavior [80], and approve the execution of maneuvers [83]. Nonetheless, if, in a cooperative approach, a situation also arises where the vehicle needs support from the driver in order to be able to react adequately to a situation, it is still highly important that the driver has adequate situational awareness in order to be able to successfully support the system.

2.3 Auditory Cues

Kutchek and Jeon used three different sounds to indicate a TOR: known sounds from other real vehicles, start and stop sounds, and sounds to simulate a real soundscape in vehicles [38]. The authors concluded that familiar warning tones (e.g., Hyundai IONIQ EV) with a high frequency and a high repetition rate, lead to a shorter reaction time after the take-over message [38]. Thus, an auditory warning should generate as much attention as possible and sound urgent [38] to achieve a low response time. Additionally, the sounds should also be short as isolated participants waited for an auditory warning to end and only then took over control, which, however, led to a virtual accident in the scenario [38].

Roche and Brandenburg analyzed whether an intrusive tone (auditory TOR) better represents an urgent acute situation by matching the urgency of the TOR and the situation, thus achieving better scores in performance and subjective perception of the driver [65]. The study concluded that matching urgency of auditory TOR and take-over of control in the corresponding situation does **not** lead to improved driver performance in terms of take-over time [65]. In general, however, a TOR with high urgency led to a slower reaction time after the TOR [65].

A driving simulation under realistic conditions revealed that spatial sound cues, which reflect, for example, the position and distance of an overtaking vehicle, can assist the user in building up situational awareness so that the user achieves higher situational awareness [87]. In complex situations, the 3D sounds, however, were declared to be confusing, therefore, support should be provided through visual cues [87].

2.4 Visual Cues

As already mentioned, research has developed numerous visual concepts for the take-over interface. Politis et al. [63] for example evaluated language-based multimodal displays. They found that with higher urgency, quicker transitions were achieved. On the other hand, Walch et al. [78] investigated whether including an alert ("Fog") improved the take-over quality [78]. They found that the take-over took longer but participants had higher comfort with the alert. Yang et al. [89] proposed a LED-based concept to improve driver situation awareness. The LED strip was used to indicate maneuvers, hazards, and a TOR. Compared to a *Baseline* with only indicating the TOR, presenting information during the automation lead to higher take-over quality and improved gaze behavior. Roche and Brandenburg found no evidence that the urgency of the TOR should match the urgency of the situation [64]. However, to get a more holistic picture of these visual interfaces, we conducted a literature analysis.

2.5 Literature Analysis Process

A literature review of existing HMIs that support the formation of driver situational awareness in general and specifically in TORs was conducted. Of particular interest was which and especially at which level of detail (from specific to abstract) the information and the reason for the TOR were presented. The result was that there are different approaches and ideas about what information should be presented in a certain way. However, almost no concepts were found that relate the presentation of information to the formation of the driver's situational awareness. We included the databases Google Scholar, IEEE, PubPsych, PsychInfo, and ScienceDirect for the time-frame of 2010 to 2020. The search criteria is included in Appendix A. We categorized the resulting **32** papers with respect to (1) information types, (2) information classification, (3) information display technology, and (4) information abstraction (see Table 1).

For the *information abstraction*, we used four levels (abstract, less abstract, less specific, specific), a common distinction for information comprehension[1]. Abstract information refers to simply stating that a take-over is necessary and the time constraint (see Figure 1(1)). Specific information refers to visualizing raw data, presumably in the scene via Augmented Reality (AR) technology or the display of specific information (such as the distance to the vehicle in front) in the HUD or other screens (see Figure 1(4)). Less abstract information adds the reason for the take-over (see Figure 1(2)) and the less specific information includes data such as distance to take-over, presence of other vehicles, and current speed limits (see Figure 1(3)).

We only categorized used visual display technologies and broadly distinguished Head-Down Displays (HDDs), Head-Up Displays (HUDs), and AR. HDD refers to all instrument cluster visualizations. The categorization was done independently by two authors. Differences were resolved via discussions.

2.6 Results

We found that there was no common approach to designing the interface with regards to *information abstraction*. The distribution of the approaches was approximately one third for abstract and less abstract while less specific (5 times)

---

[1] https://www.psychologydiscussion.net/social-psychology-2/language/comprehension-meaning-and-types-psychology/1394; Accessed: 27-JANUARY-2021

Paper	Abstract	Less Abstract	Less Specific	Specific	Information Classification	Display Technology	Information types
[2]	X				Information + Warning	HDD	which modes are activated or can be activated (system limits), request for action for TO, suggestions (lane change).
[3]	X				System Uncertainty	HDD	Uncertainty/ limits of automation or system limit (automation uncertainty), e.g., fog and unclear in which lane leading vehicle is.
[31]	X				System Uncertainty	HDD	Inability of the system to maintain autonomous driving
[39]	X				Navigation + Information	HDD	System activation, action request for transfer of control or take-over (TO) when leaving the platoon.
[40]	X				Navigation + Information	HDD	System activation and availability, display of the included, subordinated modes
[43]	X				Warning	HDD	Road or vehicle related warnings, e.g., obstacle, flat tire
[44]	X				Navigation	AR	Red = avoid corridor, Green = corridor can be navigated safely
[46]	X				Information + Warning	HDD	System activation - availability and uncertainty, request for action on TO, collision warning
[54]	X				Information + Warning	HDD	Activated mode, activation, road type, traffic situation, TO-necessity.
[57]	X				Warning	HDD	Call to action on TO
[78]	X				Information + Warning	HUD	Call to action on TO after information about weather conditions (fog warning)
[24]	X				Information + Warning	AR	Possible threats; system classifies threat risk
[90]	X				Information + Warning	HDD	System for autonomous driving (highway): not available, available, activated, TOR (deactivated)
[85]	X				Navigation + Information	HDD	Enabled Mode, Predicted Distance to TO/Mode Change
[23]		X			Navigation + Information	HDD	Current status/maneuver
[91]		X			Navigation + Information	AR	3 traffic-related system phases: Request, Proposal, Action, and marking other road users
[34]		X			Information + Warning	HDD	Virtual shadow: Anticipated walking path of pedestrians; Border: Pedestrians detected by the system.
[1]		X			Navigation + Information	HDD	System status, action request to TO, Navi hints , traffic hints, display Ego-car and possible maneuvers.
[58]		X			Information + Warning + System Uncertainty	HDD	Mode- availability, activation, road type, traffic situation, TO-necessity.
[63]		X			Information + Warning	AR	Depending on urgency, request for action on TO or request whether TO desired
[68]		X			Information + Warning	HDD	Display of traffic events, in case of attention: warning and hint, in case of interpretation: warning, hint, and specific warning in situation of possible events: Warning, hint, and specific warning in situation before possible events
[32]		X			Navigation + Information	HDD	System availability, activation, speed specifications, navigation direction (curve, etc.)
[70]		X			Navigation + Information	HDD	Current system state to longitudinal automation, system actions, traffic rules, TO- necessity.
[42]		X			Navigation + Information	AR	System status, reasons for TO
[4]			X		Navigation + Information	HDD	Lane change indications (Ego-car, indication free lane and potential dangers [occupied lane])
[21]			X		Navigation + Information	HUD	Input and display of driving maneuvers, suggestions of the vehicle, visual feedback
[29]			X		Warning	AR	Warning symbols that draw attention to dangers on the road
[86]			X		Information + Warning	HDD	System Status, Warnings, System Uncertainty, Maneuver Options, TO-Option, Trip Related Information (e.g. Applicable Speed)
[56]			X		Navigation	AR	Study 1: Target parking space; Position of the car; Study 2: Pedestrian detection
[8]				X	Navigation	AR	Distances to vehicles, collision warning, own lane, course of curves, traffic jam and slow-moving traffic
[26]				X	Navigation + Information	AR	Automation intention (allow vehicle to merge), alternative (accelerate); traffic situation
[37]				X	Navigation + Information	AR	Strategic display (trip, distance, map), traffic situation, markers and obstacles, convoy "colleagues", entertainment)

Table 1. Classification of publications regarding visualizations.

and specific (3 times) abstraction information regarding the take-over was scarcely given. Regarding the Information classification, 13 concepts presented *navigation & information* while 11 had *information & warning*. Purely displaying a warning (3 times) or system uncertainty (once) was done significantly less. As no common understanding seems to be present on which *information abstraction* is appropriate, we varied the levels systematically in the subsequent study.

Mostly, HDDs were used (22 times) followed by (simulated) AR technology (10 times). A HUD was only used two times. While HUDs were rarely used in the investigated papers, in our experiment, we used these as they are becoming more common in current vehicles and have been shown to increase performance by reducing the need of diverting attention from the street compared to HDDs [67].

2.7 HMI Design

The aim of the study was to review which levels of abstraction of information representation and which modalities are useful in the formation of adequate situational awareness in TORs. For this purpose, the findings of the literature review as well as the recommendations and results of the authors of the visual and auditory cues section were taken into account. Based on this, several prototypes were developed and internally evaluated. After several iterative runs, 4 levels of abstraction of the information presentation were implemented for the two modalities (visual + audiovisual) and based on the research of information processing. To test the developed HMIs, an online study was conducted.

3 EXPERIMENT

To evaluate the effects of *modality* and *information abstraction* in a take-over process, we designed and conducted a between-subject online study with *N*=225 participants. The study was guided by the research question:

> How do the *information abstraction* and the *modality* affect (1) situation awareness, (2) cognitive load, (3) acceptance, (4) usability, (5) and objective information intake?

3.1 Participants

Before the experiment, we computed the required sample size via an a priori power analysis using G*Power [20]. To achieve a power of .95, with an alpha level of .05, 158 participants should result in a medium effect size (0.25 [22]) in a One-Way Independent Samples ANOVA.

Participants were recruited via prolific.co. Due to failed attention checks, 50 datasets had to be rejected. The average age of all participants was *M*=31.46 years (*SD*=10.11), the mean of individual conditions ranges from *M*=29.84 to *M*=33.36. A Kruskal-Wallis test showed no significant differences for the age of the subjects between the nine conditions (χ^2 (8)=3.39, *p*=.908).The participants were compensated with € 1.

Pearson's Chi-square tests showed that there were no significant differences for gender (χ^2(8, N = 225) = 10.88, *p*=.209), education level (χ^2(32, N = 225) = 27.35, *p*=. 701), occupational situation (χ^2(40, N = 225) = 41.99, *p*=.385), driving frequency (χ^2(40, N = 225) = 39.39, *p*=.498), and driving distance (χ^2(32, N = 225) = 32.32, *p*=.451) of participants between conditions. Another Kruskal-Wallis test showed that there were no significant differences between conditions in terms of how long participants had held a driver's license (χ^2(8) = 7.90, *p*=.443). Among the participants were 86 women (38%) and 139 men (62%).

Participants reported medium interest in AVs (*M*=4.04, *SD*=1.15), medium belief that AVs will ease their lives (*M*=3.94, *SD*=1.19), and were unsure about whether AVs would become reality by 2030 (*M*=3.99, *SD*=1.11). Kruskal-Wallis tests were conducted to determine differences in participants' interest and attitude toward autonomous vehicles between conditions. No significant differences were found for interest ($\chi^2(8)$ = 8.65, *p*=.372), expectation that automated vehicles will make one's life easier ($\chi^2(8)$ = 8.43, *p*=.393), and will be a reality in the next 30 years ($\chi^2(8)$ = 9.27, *p*=.320).

The MiniDBQ [52] was used to analyze the driving behavior of the study participants in road traffic. Questions could be answered on a scale from 0 (never) to 5 (almost always). The calculated average value for *violations* is *M*=1.18 (*SD*=1.19), the value for *errors* is *M*=1.20 (*SD*=1.06), and the value for *omissions* is *M*=1.44 (*SD*=1.12). A Kruskal-Wallis test was performed to examine the resulting *Aberrant Driver Behavior Score* for significant differences between conditions. No statistical significance was shown ($\chi^2(8)$ = 7.99, *p*=.434).

3.2 Materials

This section describes how the different conditions were implemented in the HMI and in which scenario they were tested. Additionally, it describes how the measurement of objective and subjective data was operationalized.

Condition	V_A	V_{LA}	V_{LS}	V_S	VA_A	VA_{LA}	VA_{LS}	VA_S
Steering wheel lights	X	X	X	X	X	X	X	X
Countdown	X	X	X	X	X	X	X	X
Warning symbol	X	X			X	X		
Textual Warning	X	X			X	X		
Reason for TOR		X				X		
Permanent HUD			X	X			X	X
Top-Down Display			X	X			X	X
Distance Display			X	X			X	X
Blind spot Display			X	X			X	X
Extending steering wheel			X	X			X	X
AR highlights				X				X

Table 2. Visual functions of TOR per condition; baseline excluded as it does not include TOR. V_A stands for visual abstract, V_{LA} for visual less abstract, V_{LS} for visual less specific, and V_S for visual specific information. VA_A to VA_S add the auditory information.

3.2.1 Visual HMI. The *information abstraction* has four levels: abstract (Figure 1(1)), less abstract (Figure 1(2)), less specific (Figure 1(3)), and specific (Figure 1(4)). The higher the abstraction, the less specific information is given (Table 2). In the lowest abstraction, relevant information was highlighted with a simulated AR-HUD. We highlighted data that are reasonable to assume that object detection algorithms will be able to detect, i.e., traffic signs.

3.2.2 Auditory HMI. For the auditory component, we included both warning sounds and, depending on the condition, information regarding the situation (see Table 3). Depending on the information content, the frequency of auditory component was adjusted: i.e. with abstract information, the frequency is 1 Hz, with less abstract and less specific information it is 0.66 Hz and specific was 0.31 Hz. We included the warning sound in the visual only conditions to not artificially worsen the concept.

3.2.3 Scenario. A typical scenario for unscheduled [53] take-overs could be reaching a construction site. Various information can be displayed such as current speed limit, presence of vehicles on the other lane, and distance to the latest take-over time. We implemented a scenario resembling an unscheduled system-initiated take-over [53]. A construction

Fig. 1. Simulation. (1) shows the abstract, (2) the less abstract, (3) the less specific, and (4) the specific information presentation according to Table 2.

Condition	V_A	V_{LA}	V_{LS}	V_S	VA_A	VA_{LA}	VA_{LS}	VA_S
Warning sound	X	X	X	X	X	X	X	X
Voice based prompt					X	X	X	X
Reason for TOR						X	X	X
Distance to TO							X	X
Speed Limit							X	X
Voice Message					"Please take over control."	+ "Roadwork ahead."	+ "… in one mile. Speedlimit is 30 miles per hour."	see VA_{LS}

Table 3. Auditory functions of TOR per condition; baseline excluded because it does not include TOR.

Fig. 2. Top down views on the construction site.

site was chosen due to prevalence on streets today, the need to take action and change the lane, and the likely presence of other (vulnerable) road users. The AV drove on a two-lane highway with several other vehicles present (see Figure 1).

The AV drove 60 mp/h (approximately 100 km/h) and then reduced the speed to 30 mp/h before a construction site (see Figure 2). The vehicle drove 1 min 01 s. We recorded 9 videos of the simulation in Unity [72]. The videos show a ride on a highway with several other vehicles. The ego vehicle is on the left lane. The TOR is triggered as soon as the vehicle approaches the construction site; in the videos, this occurs at the 50th second. The participant is prompted to take control of the vehicle by this take-over notification with various audible and visual displays and information visualizations (see Figure 1 and Table 2). The driver thus has 11 seconds to become aware of the cues and become aware of the environment (build situational awareness). The scenario was implemented with Unity version 2019.4.1f1.

3.3 Measurements

3.3.1 Subjective Measurements. Cognitive workload was measured using the raw NASA-TLX [28], usability of the system was measured using the system usability scale (SUS) [7], acceptance using van der Laan's acceptance scale [74] and intention to use was measured using the questionnaire of Venkatesh et al. [75]. Perceived situation awareness was measured using the situation awareness rating technique (SART) [71]. The SART was used to assess the perceived quality of situation awareness [19] which may be a predictor of "how a person will choose to act on that SA" [19, p. 86]. Additionally, self-developed items were used. These measured participants' agreement to statements regarding the aspect of timeliness ("The automated vehicle asked me to take over control in time."), control take-over capability ("I could have taken over control after the request and steered the vehicle through the construction site."), warning sufficiency ("The automated vehicle has warned me sufficiently."), information presentation ("The automated vehicle showed me all relevant information."), capabilities of the vehicle ("The automated vehicle recognized signs in time.", "The automated vehicle recognized signs correctly."), and usability ("I can imagine that this automated vehicle is used in reality.") all self-developed items used a 5 point Likert scales (1 = Not at all to 5 = Definitely). Also, the driving style was rated on a 7 point Likert scale (1 = completely safe to 7 = completely dangerous).

3.3.2 Objective Measurements. Participants were asked to report several objective data related to their perception of the scene. The questions were:

- How many people on the construction site could you perceive?
- What was the speed limit in the area of the construction site?
- Has your vehicle overtaken a truck? (yes/no/Don't remember)
- What color was the vehicle in front of you?
  (silver/blue/green/yellow/black/Don't remember)

Finally, participants were asked for open feedback regarding the take-over concept.

3.4 Study Design

The experiment was designed as a 2 × 4 between-subjects online study. The independent factors were *modality* (visual vs. visual+auditory) and *information abstraction* (abstract, less abstract, less specific, specific), resulting in 8 conditions. Additionally, a *Baseline* was administered. In the baseline, no information is given. This represents a silent automation failure (see [47]).The participants were randomly assigned to the conditions.

3.5 Procedure

Each session started with a brief introduction, agreeing to the consent form, and a demographic questionnaire. Participants were instructed to use a laptop or a computer. The introduction was given as follows (including boldness):

You will see a ***video*** *of a* ***highly automated ride*** *through a simulated environment. You are in an automated vehicle on a highway with a speed limit of 60 mph. Helpful visualizations are displayed in the HUD (Head-up-Display) or on the screen in the center console if necessary. The vehicle takes over the steering and accelerating/braking (lateral and longitudinal guidance). There is the possibility that the vehicle has to* ***hand over the control to you****, because of a zone that the vehicle can't pass automated. You are supposed to follow the entire journey attentively and then assess it.*

After the introduction, the participants were randomly assigned to one of the eight conditions or the baseline group, without their knowledge. Then they watched the video of their condition. After the video, the objective questions were asked first followed by the SART. Afterwards, the SUS, the intention to use, the NASA TLX, and finally the acceptance questionnaires were administered. After they answered all the questions there was a written explanation about the objectives of the study.

A script was running in the background that ensured window maximization and that participants could not skip or rerun the video (ensuring equal exposure time). An HD monitor and loudspeakers were required.On average, a session lasted 10 min. Participants were compensated with € 1.

4 RESULTS

Participants took, on average, *M*=19:58 (min:sec, *SD*=14:19; fastest: *M*=14:57, *SD*=07:15; slowest: *M*=27:49, *SD*=30:32). First, we tested the data's distribution (normally/parametric vs. non-normally/non-parametric). As we obtained non-parametric data [69], we used Kruskal-Wallis tests to compare the nine conditions. We investigated main and interaction effects of *modality* × *information abstraction.* For our non-parametric data, we used the factorial non-parametric analysis of variance (NPAV) provided by Lüpsen [50]. For post-hoc tests, Bonferroni correction was used. We used Version 4.0.5 of R with all packages up-to-date as of May 2021. RStudio Version 1.4.1103 was used. For the figures, we used the package *ggstatsplot* [59]. These include the mean or median (blue dot), the density plots, the boxplots, and the data points.

4.1 Situation Awareness, Usability, and Cognitive Load

A Kruskal-Wallis test showed a significant difference for situation awareness between the concepts ($\chi^2(8) = 20.84$, $p = .008$, $\epsilon^2 = .093$). A post-hoc Dunn test showed significant differences between the *Baseline* (*M*=16.04, *SD*=7.69) and *Visual Specific* (*M*=22.48, *SD*=6.07). No significant effects were shown by non-parametric analysis of variance (NPAV) for the four levels of *information abstraction* ($\chi^2(3, N = 200) = 7.72$, $p = .052$).

A Kruskal-Wallis test showed a significant difference between the usability ratings ($\chi^2(8) = 16.19$, $p = .040$, $\epsilon^2 = .072$). The Dunn post-hoc test showed that this significant difference was between the *Baseline* *M*=59.40 (*SD*=17.97), and *Visual Abstract* (*M*=77.80 (*SD*=15.63, $p = .017$). The NPAV found no significant effects of *information abstraction* ($\chi^2(3, N = 200) = 1.43$, $p = .698$) or *modality* ($\chi^2(1, N = 200) = 3.09$, $p = .079$) and no significant interaction effect of *modality* × *information abstraction* ($\chi^2(3, N = 200) = 7.47$, $p = .058$) on the usability rating.

Neither a one-way ANOVA nor a two-way ANOVA found significant effects on the combined cognitive load score. Kruskal-Wallis tests found no significant differences for any of the subscales of the NASA-TLX. A NPAV found no significant interaction effect of *modality* × *information abstraction* on the mental workload subscale ($\chi^2(3, N = 200) = 6.96$, $p = .07$; see Figure 3). Mental workload went up for the less specific and less abstract information but increased for the abstract and specific information when visual and auditory communication was used.

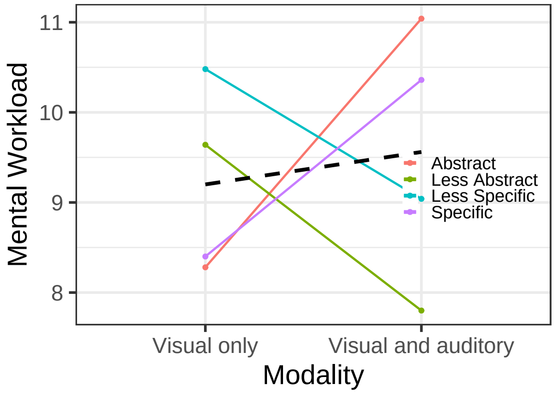
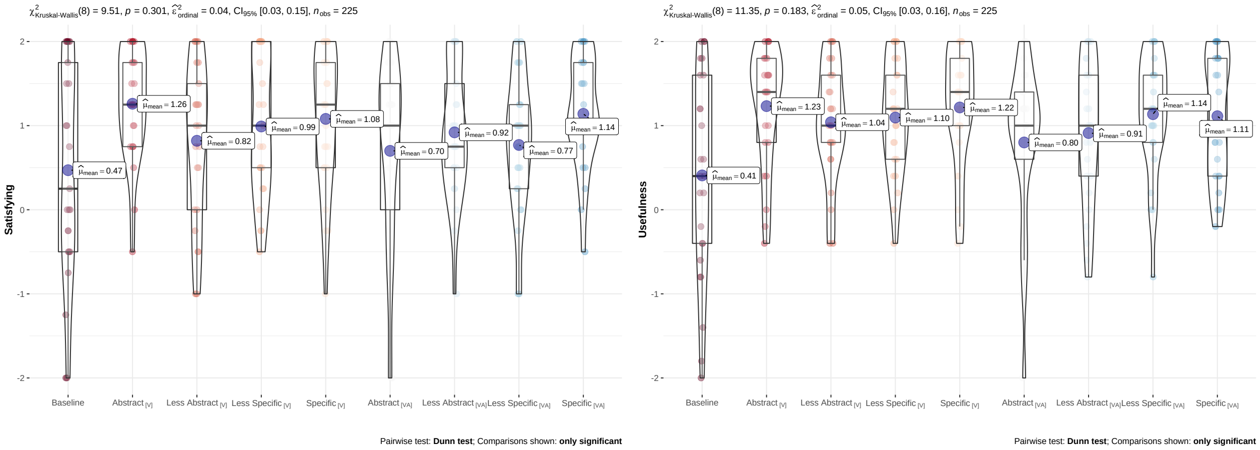


Fig. 3. Non-significant interaction effect of *modality* × *information abstraction* on mental workload. The dotted black line shows the non-significant main effect of *modality*.

4.2 Acceptance and Intention to Use



(a) Satisfying scores.

(b) Usefulness scores.

Fig. 4. Results of the Acceptance scales.

Kruskal-Wallis tests (see Figure 4) found no significant differences neither for Usefulness nor the Satisfying score [74]. Also, NPAVs found no significant effects on either of these scales.

Kruskal-Wallis tests found no significant difference for the intention to use [75]. A NPAV found no significant interaction effect (χ^2(3, N = 200) = 6.69, p = .08; see Figure 5). Intention to use increased when using visual and auditory communication only in the less specific conditions, but stayed constant (less abstract information) or fell compared to only using visual cues.

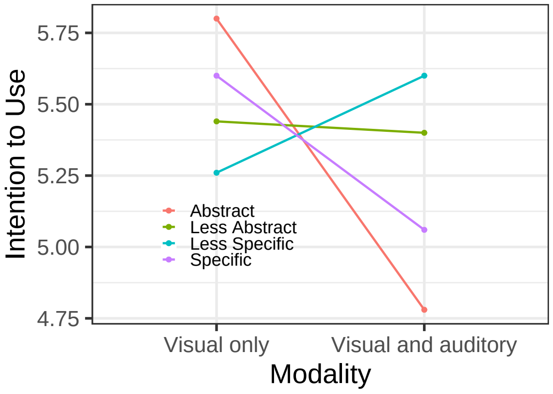
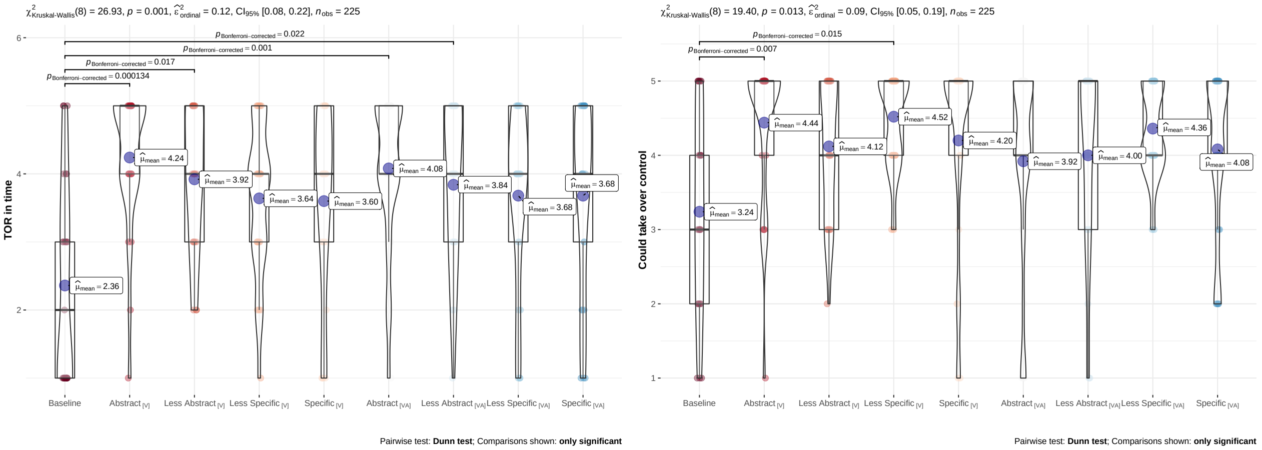


Fig. 5. Interaction effect of *modality* × *information abstraction* on intention to use.

4.3 Timeliness, Control Take-over, Warning Sufficiency, Information Presentation, Driving Style, and Capabilities



(a) Timeliness.

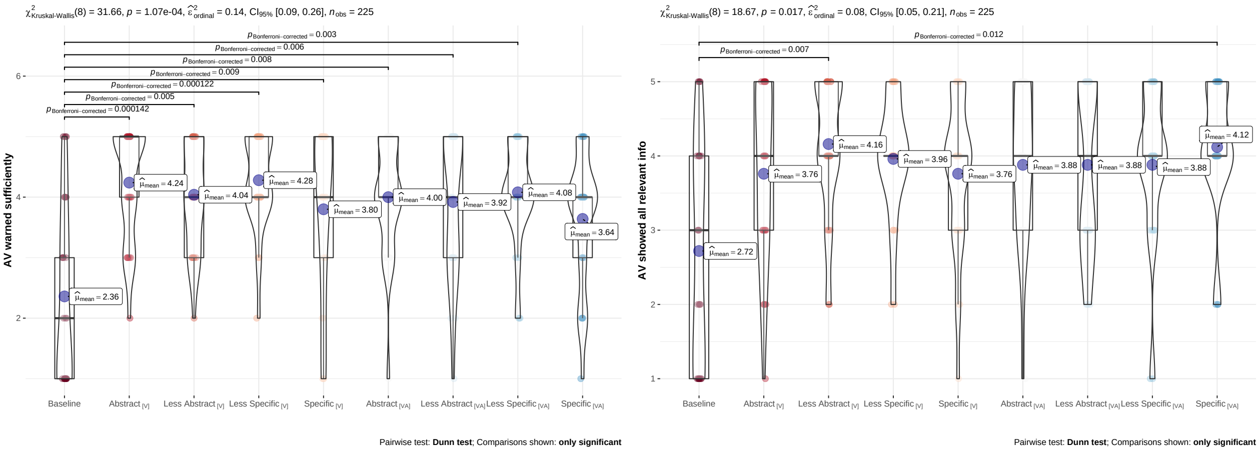
(b) Capability for control take-over.

Fig. 6. Differences in timeliness and control take-over.

A Kruskal-Wallis test showed significant differences for the timeliness ($\chi^2(8) = 26.93$, $p < .001$, $\varepsilon^2 = .12$; see Figure 6a). Post-hoc tests showed the assumed difference to the *Baseline* only for the abstract and less specific information systems. However, the NPAV found no significant effect (*information abstraction*: $\chi^2(3, N = 200) = 6.22$, $p = .10$).

A Kruskal-Wallis test showed significant differences for the control take-over ($\chi^2(8) = 19.40$, $p < .01$, $\varepsilon^2 = .09$; see Figure 6b). The *Baseline* received significant lower ratings compared to *Visual Abstract* and *Visual Less Specific*. The NPAV found no significant effects.

A Kruskal-Wallis test showed significant differences for warning sufficiency ($\chi^2(8) = 31.66$, $p < .001$, $\varepsilon^2 = .14$; see Figure 7a). The NPAV found no significant effects on warning sufficiency.



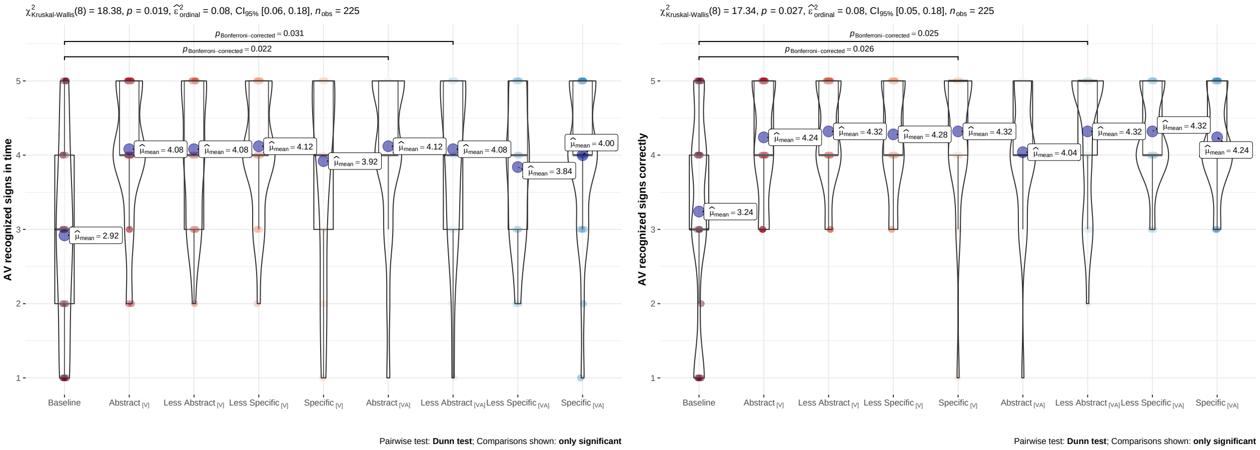
(a) Warning sufficiency.

(b) Information presentation.

Fig. 7. Differences in warning sufficiency and information presentation.

A Kruskal-Wallis test showed significant differences for the information presentation ($\chi^2(8)$ = 18.67, p = .017, ϵ^2 = .08; see Figure 7b). The NPAV found no significant effects on information presentation.

Neither a Kruskal-Wallis nor an NPAV found significant differences for the driving style (*M*=2.95, *SD*=1.54).



(a) Sign detection timely.

(b) Sign detection correct.

Fig. 8. Results of the sign detection capability scales.

A Kruskal-Wallis test showed a significant difference in the assessment of the timeliness of sign detection ($\chi^2(8)$ = 18.38, p = .019, ϵ^2 = .08; see Figure 8a) and of the correctness ($\chi^2(8)$ = 17.34, p = .027, ϵ^2 = .08; see Figure 8b). In both cases, the *Baseline* was rated worst. NPAVs found no significant effects.

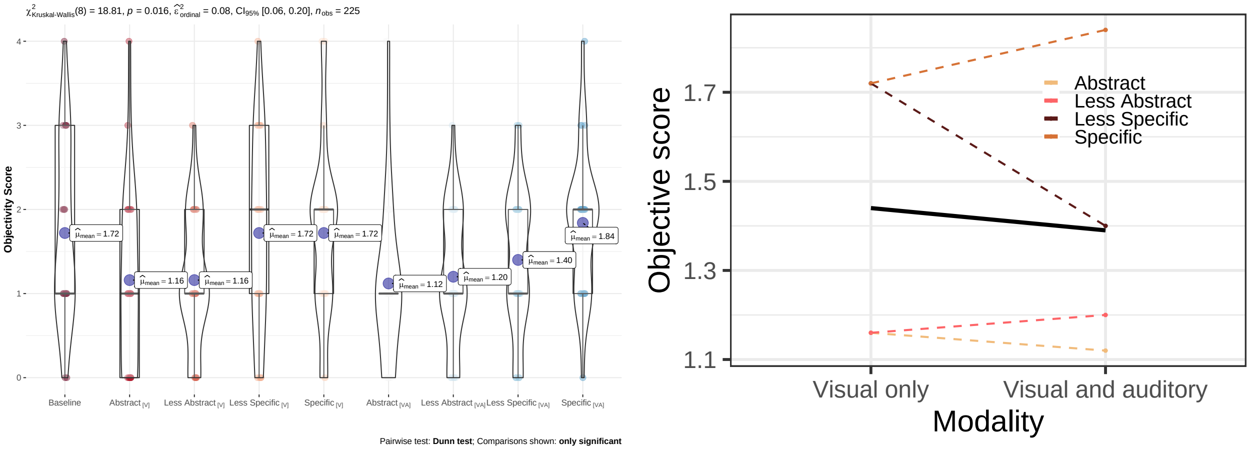
4.4 Objective Measurements

For the objective measurements, we compared all conditions with each other with Chi-square or Kruskal-Wallis tests. Of all participants (*N*=225), only 16% (37 participants) correctly reported the number of people present on the construction site. A Chi-square test showed no significant differences of the number of correct statements about the number of persons (χ^2(8, N = 225) = 5.08, p = .749).

Regarding the assessment of the speed limit, a Chi-Square test showed a significant difference (χ^2(8, N = 225) = 25.16, p = .001). A post-hoc test showed significant differences between condition *Visual Less Abstract* (8% or 2 correct) and *Visual Less Specific* (64% or 16 correct) with p = .035 and between *Visual Less Abstract* and *Visual-Auditory Specific* (64% or 16 correct) with p = .035. Overall, 41% of participants (93) reported a correct value.

Two Chi-square tests revealed no significant differences on the question about the color of the vehicle ahead (χ^2(8, N = 225) = 4.91, p = .767; correctly given color between 6 and 10 per condition) and whether one passed a truck (χ^2(8, N = 225) = 3.06, p = .930; correctly given values between 12 and 19 per condition).

The sum over the number of correct answers to the four objective questions was calculated with *M*=1.45 (*SD*=.99; see Figure 9a). A Kruskal-Wallis test showed significant differences (χ^2(8) = 18.81, p = .016, ε^2 = .084), however, post-hoc tests showed no significance. A NPAV showed a significant effect of *information abstraction* on the score (χ^2(3, N = 200) = 16.12, p = .001). Dunn's post-hoc test showed that data from the "low abstraction" level of abstraction were significantly different from the "abstract" (p = .003) and "less specific" (p = .014) levels of abstraction (see Figure 9b).



(a) Objective total score.

(b) Main effect of *modality* on objective score.

Fig. 9. Objective scores.

4.5 Effect of Self-Reported Driving Behavior on Measurements

To investigate whether the driving behavior of the study participants in the real world influences the study results on the video simulation, the subjective questionnaire MiniDBQ [52] additionally formed an overall driving behavior score ("Aberrant Driver Behavior") as the mean of all MiniDBQ questions. The average of these scores from all participants is *M*=1.27 (*SD*=.97). This value was used to divide participants into two categories: those whose aberrant driving behavior (ADB) score is below the average (hereafter "ADB-low") and those whose score is above (hereafter "ADB-high").

We used Mann-Whitney U tests for comparison which also works with unequal group sizes [51] and adjusted the p-value to correct the alpha level (division through 9 as there are 9 conditions). We report the adjusted p-values.

The Mann-Whitney U test found no significant influence of the ADB score on the reported mental workload in condition *Visual Less Abstract* (W = 125, Z = -1.74, p = .08, r = -.349) as well as *Visual-Auditory Less Abstract* (W = 119.5, Z = -1.822, p = .068, r= -0.364). In *Visual Less Abstract*, the mean of the ADB-low category (*M*=7.14, *SD*=4.24) differs from ADB-high (*M*=12.82, *SD*=3.68). The data of condition *Visual-Auditory Less Abstract* show a similar trend: the mean of ADB-low *M*=5.25 (*SD*=4.07) is lower than that of ADB-high with *M*=12.33 (*SD*=5.89). The same trend could be observed in the results of the complete TLX score: Mann-Whitney U tests showed significant differences between ADB groups in conditions *Visual Less Abstract* (W = 130, Z = -2.09, p = .036, r = .419) and *Visual-Auditory Less Abstract* (W = 129.5, Z = -2.54, p = .01, r = -.508). The mean in condition *Visual Less Abstract* (ADB-low: *M*=39.86, *SD*=19.13; ADB-high: *M*=62.91, *SD*=12.14) and *Visual-Auditory Less Abstract* (ADB-low: *M*=31.06, *SD*=15.77; ADB-high: *M*=69.78, *SD*=26.14) indicate a clear tendency: participants in the ADB-high category reported significantly higher stress than those in the ADB-low category in these conditions. The mean of all these conditions for the ADB-low group is *M*=36.52 (*SD*=20.65), and that of the ADB-high is *M*=64.86 (*SD*=23.28).

A Mann-Whitney U test showed a significant difference in condition VA_S between the ADB-low and ADB-high groups for the objectivity score formed from the four objective questions (W = 28, Z = -2.18, p = .029, r = -.436). In *Visual-Auditory Specific* (ADB-low: *M*=2.31, *SD*=0.75; ADB-high: *M*=1.33, *SD*=0.65). The ADB-low group had a significantly higher average score and thus better reception of information.

5 DISCUSSION

This work shows the necessity of conveying a take-over warning in advance. We found few significant differences both in subjective as well as in objective measurements.

5.1 Necessity of Information Presentation

We found a few positive effects of presenting information to the participants in terms of subjective ratings. The *Baseline* was rated worse in terms of timeliness (see Figure 6a), control take-over (see Figure 6b), and warning sufficiency (see Figure 7a). However, there were **no** significant differences in these subjective measurements between the **eight** systems (i.e., excluding the *Baseline*). Therefore, we conclude that a warning seems sufficient prior to a take-over. This warning should contain information regarding the necessity to take-over control (see [78]), however, additional data does not necessarily enhance situation awareness or improve perceived usability. This is especially relevant as the acquisition of such information is not trivial. While Waymo claims to "identify [...] [unmapped] stop signs greater than 500 meters away" [33], even this distance is driven in 15s with, for example, 120 km/h. Challenges such as weather conditions [73] are not yet taken into account into this distance. Also, misperceptions could lead to the display of wrong information (e.g., the speed limit [27]). In fact, if the system had all information, it would possibly be able to maneuver through the situation, rendering the take-over unnecessary. Therefore, we argue that an attention guiding system for relevant information or areas (e.g., the mirrors), especially **if** perceived objects in the scene are known, is less prone to errors and is better suited to enable appropriate situation awareness.

The *Visual Specific* and *Visual-Auditory Specific* represented an attention guidance system, which was rated generally well. Yang et al. [89] implemented an attention guiding system using LED strip to highlight potentially hazardous objects. They report that this system improves the time of eyesight on the street as one part of their situation awareness assessment. One problematic detail of their system, however, could be the distracting nature of the moving LED light. Future work should investigate how attention can be guided without such distraction.

5.2 Modality Usage

We expected that using two modalities (audio and visual) would increase situation awareness as two channels are used to gather data. However, as seen in Figure 9b, the opposite was the case. Visual and auditory information combined decreased the objective information significantly. This can not generally be explained with increased mental workload (see Figure 3). Only for the *Visual-Auditory Specific* condition were the objective scores marginally higher (see Figure 9b; difference for less abstract is negligible). Nonetheless, mental workload (see Figure 3) and intention to use (see Figure 5) were lower for visual and auditory information presentation. Additionally, we found no evidence of any significantly increased measurement with adding auditory information. Therefore, we conclude in line with [5] that auditory signals should be used solely as a trigger signal. Bi- and trimodal interfaces were shown to improve reaction time in general [14] and, more specific, take-over time [62, 63]. Therefore, multimodal interfaces should be utilized to signal a necessary take-over but not expected to increase situation awareness or higher usability.

5.3 Effect of Self-Reported Driving Behavior

A few effects were found that showed that participants with low ADB had lower mental workload and perceived more information. However, this was only observed in some of the conditions and no overall pattern can be observed from the data. In general, it is difficult to assess the validity of these results due to biases such as the social desirability bias [25] which could explain the very low ratings regarding ADB (*M*=1.27; *SD*=.97).

5.4 Practical Implications

In general, it seems unfeasible to equip an AV with the technology to detect all relevant information and to display this to the user (i.e., specific information). Take-overs are mainly relevant at the end of operational driving domains or unscheduled, that is, unforeseen, situations.

This work shows that providing a warning is necessary (see Figure 7a). However, there were few statistically significant differences between the systems. The *Baseline* was rated lower but still high in most of the subjective measurements (e.g., Figure 4). Regarding the objective score (see Figure 9a), the *Baseline* even performed (non-significantly) better than the abstract and less abstract information systems and was on par with the others. Therefore, we conclude that providing the information is only feasible when specific data is presented. Highly abstracted information (as suggested by Endsley [18]) had no positive effect on the objective situation awareness.

Damböck et al. [12] showed that with higher take-over times (8 vs 4 or 6s), drivers make fewer errors in take-over scenarios. Drivers were even capable of driving as well as in a baseline drive [12]. In our scenario, participants had a time budget of 11s. However, the objective situation awareness was only medium. Therefore, our data indicate that driving performance is not necessarily associated with objective situation awareness as we measured it.

5.5 On the Study Design

Our study design has several shortcomings. As a monitor-based study was used, external validity is difficult to assess and probably low. As we had no way to check attention during the videos, we included attention checks to sufficiently rate the given answers as reliable. However, it remains unclear whether participants would have been able to still maneuver through the construction site. While we included some objectively assessable ratings such as construction site speed limit, most assessments were subjective (mental workload, usability, situation awareness). Therefore, these results have to be treated with caution. Additionally, our sample was rather young, therefore, findings will most likely be even worse for less tech-savvy participants.

However, the chose study design also has some advantages. First, the number of participants was high. Additionally, less bias exists in recruiting participants (recruiting from an ad-hoc young university population). Additionally, while the aspect of sitting in front of a screen reduces the external validity of having to take over a vehicle, this also more resembles a scenario in which users of AVs will likely become adapted to AVs and, therefore, trust these. Already people are trusting, for example, Teslas enough to sleep [30]. This was also the case when Waymo first let colleagues test their technology [11] forcing them to rethink their strategy. de Winter et al. [13] criticized research on take-overs as "unrealistic". While we agree with the basic notion of this criticism, we believe that this aspect of our study is rather well-designed.

Therefore, we conclude that while our study design has some drawbacks, it enabled us to study a more realistic approach to how users will behave in AVs.

6 CONCLUSION

In conclusion, in a monitor-based online study (N=225), this work investigated the effect of *information abstraction* and *modality* on the take-over process in a highly automated vehicle. The levels of abstraction were chosen based on a systematic literature review. In the monitor-based online study, we found that differences in providing relevant information had almost no effect on subjective ratings such as mental workload, usability, or subjective situation awareness as well as an objective situation awareness ratings. Presenting information only visually was preferred over conveying it bi-modally (visually and auditorily). Therefore, we conclude that in take-over processes, our data suggest that merely conveying the need to take-over is sufficient. This is important as take-overs will most likely occur in demanding situations in which an AV would not have access to all relevant information.

ACKNOWLEDGMENTS

The authors thank all study participants. This work was conducted within the project 'SituWare' funded by the Federal Ministry for Economic Affairs and Energy (BMWi).

A SEARCH CRITERIA

Databases:

- Google Scholar
- IEEE
- PubPsych
- PsychInfo
- ScienceDirect

Keywords in different combinations (partly Boolean search):

- Anticipation
- Situation Awareness / SA
- Automation
- Display

- Design
- Interface design
- Traceability
- Navigation
- Platoon
- Human machine cooperation
- Automated driving
- HMI

NOT included:

- Flight
- Aircraft
- Aviation

Strict inclusion criteria:

- Display which can be used in a vehicle
- Displays are described
- scientific articles

Conditional inclusion criteria:

- Theoretical foundation of the display design
- Driving simulation or real traffic
- Graphical representation of the displays
- Empirical testing of the (developed or existing) displays
- Journal article
- Current articles (from 2012 until 2020)

Strict exclusion criteria:

- Studies without display use
- No description of the displays
- Technically obsolete displays (no longer installed in current vehicles or technically more advanced alternatives available, e.g. ACC)
- not related to takeovers (e.g., [41, 60])

REFERENCES

[1] Martin Albert, Alexander Lange, Annika Schmidt, Martin Wimmer, and Klaus Bengler. 2015. Automated driving–Assessment of interaction concepts under real driving conditions. *Procedia Manufacturing* 3 (2015), 2832–2839.

[2] Victoria A Banks and Neville A Stanton. 2016. Keep the driver in control: Automating automobiles of the future. *Applied ergonomics* 53 (2016), 389–395.

[3] Johannes Beller, Matthias Heesen, and Mark Vollrath. 2013. Improving the driver–automation interaction: An approach using automation uncertainty. *Human factors* 55, 6 (2013), 1130–1141.

[4] Mohamed Amir Benloucif, Jean-Christophe Popieul, and Chouki Sentouh. 2016. Architecture for multi-level cooperation and dynamic authority management in an automated driving system-a case study on lane change cooperation. *IFAC-PapersOnLine* 49, 19 (2016), 615–620.

[5] Shadan Sadeghian Borojeni, Lars Weber, Wilko Heuten, and Susanne Boll. 2018. From Reading to Driving: Priming Mobile Users for Take-over Situations in Highly Automated Driving. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile*

Devices and Services (Barcelona, Spain) *(MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, Article 14, 12 pages. https://doi.org/10.1145/3229434.3229464

[6] S. Brandenburg and E. M. Skottke. 2014. Switching from manual to automated driving and reverse: Are drivers behaving more risky after highly automated driving?. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, New York, NY, USA, 2978–2983. https://doi.org/10.1109/ITSC.2014.6958168

[7] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[8] Vassilis Charissis and Stylianos Papanastasiou. 2010. Human–machine collaboration through vehicle head up display interface. *Cognition, Technology & Work* 12, 1 (2010), 41–50.

[9] Hallie Clark and Jing Feng. 2017. Age differences in the takeover of vehicle control and engagement in non-driving-related activities in simulated driving with conditional automation. *Accident Analysis & Prevention* 106 (2017), 468–479.

[10] Jediah R Clark, Neville A Stanton, and Kirsten MA Revell. 2019. Directability, eye-gaze, and the usage of visual displays during an automated vehicle handover task. *Transportation research part F: traffic psychology and behaviour* 67 (2019), 29–42.

[11] CleanTechnica. 2017. Google/Waymo Stopped Testing Level 3 Self-Driving Tech After Testers Literally Fell Asleep While Using It, Switched To Full Autonomy. https://cleantechnica.com/2017/11/01/googlewaymo-stopped-testing-level-3-self-driving-tech-testers-literally-fell-asleep-using-switched-full-autonomy/. [Online; accessed: 25-JANUARY-2021].

[12] Daniel Damböck and Klaus Bengler. 2012. Übernahmezeiten beim hochautomatisierten Fahren. , 12 pages.

[13] Joost de Winter, Neville Stanton, and Yke Bauke Eisma. 2021. Is the take-over paradigm a mere convenience?

[14] Adele Diederich and Hans Colonius. 2004. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception & psychophysics* 66, 8 (2004), 1388–1404.

[15] Dániel András Drexler, Árpád Takács, Péter Galambos, Imre J Rudas, and Tamás Haidegger. 2018. Handover process models of autonomous cars up to level 3 autonomy. In *2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI)*. IEEE, IEEE, New York, NY, USA, 307–312. https://doi.org/10.1109/CINTI.2018.8928199

[16] Na Du, X Jessie Yang, and Feng Zhou. 2020. Psychophysiological responses to takeover requests in conditionally automated driving. *Accident Analysis & Prevention* 148 (2020), 105804.

[17] Na Du, Feng Zhou, Elizabeth M Pulver, Dawn M Tilbury, Lionel P Robert, Anuj K Pradhan, and X Jessie Yang. 2020. Examining the effects of emotional valence and arousal on takeover performance in conditionally automated driving. *Transportation research part C: emerging technologies* 112 (2020), 78–87.

[18] Mica R Endsley. 2006. Situation awareness. *Handbook of human factors and ergonomics* 3 (2006), 528–542.

[19] Mica R Endsley, Stephen J Selcon, Thomas D Hardiman, and Darryl G Croft. 1998. A comparative analysis of SAGAT and SART for evaluations of situation awareness. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 42. SAGE Publications Sage CA: Los Angeles, CA, SAGE Publications, Los Angeles, CA, USA, 82–86.

[20] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160.

[21] Frank Ole Flemisch, Klaus Bengler, Heiner Bubb, Hermann Winner, and Ralph Bruder. 2014. Towards cooperative guidance and control of highly automated vehicles: H-Mode and Conduct-by-Wire. *Ergonomics* 57, 3 (2014), 343–360.

[22] David C. Funder and Daniel J. Ozer. 2019. Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science* 2, 2 (2019), 156–168. https://doi.org/10.1177/2515245919847202 arXiv:https://doi.org/10.1177/2515245919847202

[23] Nick Gang, Srinath Sibi, Romain Michon, Brian Mok, Chris Chafe, and Wendy Ju. 2018. Don't Be Alarmed: Sonifying Autonomous Vehicle Perception to Increase Situation Awareness. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Toronto, ON, Canada) *(AutomotiveUI '18)*. Association for Computing Machinery, New York, NY, USA, 237–246. https://doi.org/10.1145/3239060.3265636

[24] Andreas Gregoriades and Alistair Sutcliffe. 2018. Simulation-based evaluation of an in-vehicle smart situation awareness enhancement system. *Ergonomics* 61, 7 (2018), 947–965.

[25] Pamela Grimm. 2010. *Social Desirability Bias.* American Cancer Society, Atlanta, Georgia, USA, 1. https://doi.org/10.1002/9781444316568.wiem02057 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781444316568.wiem02057

[26] Chunshi Guo, Chouki Sentouh, Jean-Christophe Popieul, Jean-Baptiste Haué, Sabine Langlois, Jean-Jacques Loeillet, Boussaad Soualmi, and Thomas Nguyen That. 2019. Cooperation between driver and automated driving system: Implementation and evaluation. *Transportation research part F: traffic psychology and behaviour* 61 (2019), 314–325.

[27] Isobel Hamilton. 2020. Hackers stuck a 2-inch strip of tape on a 35-mph speed sign and successfully tricked 2 Teslas into accelerating to 85 mph. https://www.businessinsider.com/hackers-trick-tesla-accelerating-85mph-using-tape-2020-2?r=DE&IR=T [Online; accessed 24-JANUARY-2020].

[28] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, Amsterdam, The Netherlands, 139–183.

[29] Renate Häuslschmid, Sven Osterwald, Marcus Lang, and Andreas Butz. 2015. Augmenting the Driver's View with Peripheral Information on a Windshield Display. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (Atlanta, Georgia, USA) *(IUI '15)*. Association for Computing Machinery, New York, NY, USA, 311–321. https://doi.org/10.1145/2678025.2701393

[30] Andrew Hawkins. 2020. Tesla owner in Canada charged with 'sleeping' while driving over 90 mph. https://www.theverge.com/2020/9/18/21445168/tesla-driver-sleeping-police-charged-canada-autopilot. [Online; accessed: 25-JANUARY-2021].

[31] Tove Helldin, Göran Falkman, Maria Riveiro, and Staffan Davidsson. 2013. Presenting System Uncertainty in Automotive UIs for Supporting Trust Calibration in Autonomous Driving. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Eindhoven, Netherlands) *(AutomotiveUI '13)*. Association for Computing Machinery, New York, NY, USA, 210–217. https://doi.org/10.1145/2516540.2516554

[32] Michael Heymann and Asaf Degani. 2013. Automated Driving Aids: Modeling, Analysis, and Interface Design Considerations. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Eindhoven, Netherlands) *(AutomotiveUI '13)*. Association for Computing Machinery, New York, NY, USA, 142–149. https://doi.org/10.1145/2516540.2516549

[33] Satish Jeyachandran. 2020. Introducing the 5th-generation Waymo Driver: Informed by experience, designed for scale, engineered to tackle more environments. https://blog.waymo.com/2020/03/introducing-5th-generation-waymo-driver.html [Online; accessed 24-MARCH-2020].

[34] Hyungil Kim and Joseph L Gabbard. 2019. Assessing distraction potential of augmented reality head-up displays for vehicle drivers. *Human factors* 0, 0 (2019), 14 pages.

[35] Thomas Köhn, Matthias Gottlieb, Michael Schermann, and Helmut Krcmar. 2019. Improving Take-over Quality in Automated Driving by Interrupting Non-Driving Tasks. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 510–517. https://doi.org/10.1145/3301275.3302323

[36] Moritz Körber, Christian Gold, David Lechner, and Klaus Bengler. 2016. The influence of age on the take-over of vehicle control in highly automated driving. *Transportation research part F: traffic psychology and behaviour* 39 (2016), 19–32.

[37] S. Krupenia. 2015. Methods for designing future autonomous systems (MODAS). http://www2.vinnova.se/PageFiles/751290059/2012-03678_EN.pdf.

[38] Kyle Kutchek and Myounghoon Jeon. 2019. Takeover and Handover Requests Using Non-Speech Auditory Displays in Semi-Automated Vehicles. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3313078

[39] Maider Larburu, Javier Sanchez, and Domingo José Rodriguez. 2010. Safe road trains for environment: Human factors' aspects in dual mode transport systems. , 12 pages.

[40] Pontus Larsson, Emma Johansson, Mikael Söderman, and Deborah Thompson. 2015. Interaction design for communicating system state and capabilities during automated highway driving. *Procedia Manufacturing* 3 (2015), 2784–2791.

[41] Talia Lavie, Tal Oron-Gilad, and Joachim Meyer. 2011. Aesthetics and usability of in-vehicle navigation displays. *International Journal of Human-Computer Studies* 69, 1-2 (2011), 80–99.

[42] Shuo Li, Phil Blythe, Weihong Guo, Anil Namdeo, Simon Edwards, Paul Goodman, and Graeme Hill. 2019. Evaluation of the effects of age-friendly human-machine interfaces on the driver's takeover performance in highly automated vehicles. *Transportation research part F: traffic psychology and behaviour* 67 (2019), 78–100.

[43] Yung-Ching Liu and Jing-Wun Jhuang. 2012. Effects of in-vehicle warning information displays with or without spatial compatibility on driving behaviors and response performance. *Applied ergonomics* 43, 4 (2012), 679–686.

[44] Lutz Lorenz, Philipp Kerschbaum, and Josef Schumann. 2014. Designing take over scenarios for automated driving: How does augmented reality support the driver to get back into the loop?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 58. SAGE Publications Sage CA: Los Angeles, CA, SAGE Publications Sage CA, Los Angeles, CA, USA, 1681–1685.

[45] Wing-Yue Geoffrey Louie, Iyad Mansour, et al. 2019. Towards a Driver Monitoring System for Estimating Driver Situational Awareness. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, IEEE, New York, NY, USA, 1–6.

[46] Tyron Louw, Georgios Kountouriotis, Oliver Carsten, and Natasha Merat. 2015. Driver inattention during vehicle automation: How does driver engagement affect resumption of control?. In *4th International Conference on Driver Distraction and Inattention (DDI2015), Sydney: proceedings*. ARRB Group, ARRB Group, Sydney, New South Wales, Australia, 16 pages.

[47] Tyron Louw, Jonny Kuo, Richard Romano, Vishnu Radhakrishnan, Michael G Lenné, and Natasha Merat. 2019. Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures. *Transportation research part F: traffic psychology and behaviour* 62 (2019), 870–882.

[48] Zhenji Lu and Joost CF de Winter. 2015. A review and framework of control authority transitions in automated driving. *Procedia Manufacturing* 3 (2015), 2510–2517.

[49] Chen Lv, Dongpu Cao, Yifan Zhao, Daniel J Auger, Mark Sullman, Huaji Wang, Laura Millen Dutka, Lee Skrypchuk, and Alexandros Mouzakitis. 2017. Analysis of autopilot disengagements occurring during autonomous vehicle testing. *IEEE/CAA Journal of Automatica Sinica* 5, 1 (2017), 58–68. https://doi.org/10.1109/JAS.2017.7510745

[50] Haiko Lüpsen. 2020. R-Funktionen zur Varianzanalyse. http://www.uni-koeln.de/~luepsen/R/. [Online; accessed 25-SEPTEMBER-2020].

[51] Henry B Mann and Donald R Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* 18, 1 (1947), 50–60. http://www.jstor.org/stable/2236101

[52] Laila M. Martinussen, Timo Lajunen, Mette Møller, and Türker Özkan. 2013. Short and user-friendly: The development and validation of the Mini-DBQ. *Accident Analysis and Prevention* 50 (2013), 1259 – 1265. https://doi.org/10.1016/j.aap.2012.09.030

[53] Rod McCall, Fintan McGee, Alexander Mirnig, Alexander Meschtscherjakov, Nicolas Louveton, Thomas Engel, and Manfred Tscheligi. 2018. A taxonomy of autonomous vehicle handover situations. *Transportation Research Part A: Policy and Practice* 124 (2018), 507–522. https://doi.org/10.1016/j.tra.2018.05.005

[54] Vivien Melcher, Stefan Rauh, Frederik Diederichs, Harald Widlroither, and Wilhelm Bauer. 2015. Take-over requests for automated driving. *Procedia Manufacturing* 3 (2015), 2867–2873.

[55] Natasha Merat, A. Hamish Jamson, Frank C.H. Lai, Michael Daly, and Oliver M.J. Carsten. 2014. Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour* 27 (2014), 274 – 282. https://doi.org/10.1016/j.trf.2014.09.005

[56] Coleman Merenda, Hyungil Kim, Kyle Tanous, Joseph L Gabbard, Blake Feichtl, Teruhisa Misu, and Chihiro Suga. 2018. Augmented reality interface design approaches for goal-directed and stimulus-driven driving tasks. *IEEE transactions on visualization and computer graphics* 24, 11 (2018), 2875–2885.

[57] Frederik Naujoks, Christoph Mai, and Alexandra Neukum. 2014. The effect of urgency of take-over requests during highly automated driving under distraction conditions. *Advances in human aspects of transportation* 7, Part I (2014), 431.

[58] Frederik Naujoks, Christian Purucker, Alexandra Neukum, Stefan Wolter, and Reid Steiger. 2015. Controllability of partially automated driving functions–does it matter whether drivers are allowed to take their hands off the steering wheel? *Transportation research part F: traffic psychology and behaviour* 35 (2015), 185–198.

[59] Indrajeet Patil. 2021. Visualizations with statistical details: The 'ggstatsplot' approach. *Journal of Open Source Software* 6, 61 (2021), 3167. https://doi.org/10.21105/joss.03167

[60] William Payre, Julien Cestac, and Patricia Delhomme. 2016. Fully automated driving: Impact of trust and practice on manual control recovery. *Human factors* 58, 2 (2016), 229–241.

[61] Juraj Peršić, Ivan Marković, and Ivan Petrović. 2019. Extrinsic 6DoF calibration of a radar–LiDAR–camera system enhanced by radar cross section estimates evaluation. *Robotics and Autonomous Systems* 114 (2019), 217–230.

[62] Ioannis Politis, Stephen Brewster, and Frank Pollick. 2014. Speech Tactons Improve Speech Warnings for Drivers. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Seattle, WA, USA) *(AutomotiveUI '14)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/2667317.2667318

[63] Ioannis Politis, Stephen Brewster, and Frank Pollick. 2015. Language-Based Multimodal Displays for the Handover of Control in Autonomous Cars. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Nottingham, United Kingdom) *(AutomotiveUI '15)*. Association for Computing Machinery, New York, NY, USA, 3–10. https://doi.org/10.1145/2799250.2799262

[64] Fabienne Roche and Stefan Brandenburg. 2018. Should the urgency of auditory-tactile takeover requests match the criticality of takeover situations?. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, New York, NY, USA, 1035–1040.

[65] Fabienne Roche and Stefan Brandenburg. 2018. Should the urgency of auditory-tactile takeover requests match the criticality of takeover situations?. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, IEEE, New York, NY, USA, 1035–1040. https://doi.org/10.1109/ITSC.2018.8569650

[66] Steven E Shladaver. 2016. The truth about "self-driving" cars. *Scientific American* 314, 6 (2016), 52–57.

[67] Missie Smith, Joseph L. Gabbard, and Christian Conley. 2016. Head-Up vs. Head-Down Displays: Examining Traditional Methods of Display Assessment While Driving. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ann Arbor, MI, USA) *(Automotive'UI 16)*. Association for Computing Machinery, New York, NY, USA, 185–192. https://doi.org/10.1145/3003715.3005419

[68] Patrick Stahl, Birsen Donmez, and Greg A Jamieson. 2016. Supporting anticipation in driving through attentional and interpretational in-vehicle displays. *Accident Analysis & Prevention* 91 (2016), 103–113.

[69] Laerd statistics. 2016. Mixed ANOVA using SPSS Statistics. https://statistics.laerd.com/spss-tutorials/mixed-anova-using-spss-statistics.php. [Online; Accessed: 12-SEPTEMBER-2019].

[70] Sonja Stockert, Natalie Tara Richardson, and Markus Lienkamp. 2015. Driving in an increasingly automated world–approaches to improve the driver-automation interaction. *Procedia Manufacturing* 3 (2015), 2889–2896.

[71] Richard M Taylor. 2017. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational awareness*. Routledge, Abingdon, UK, 111–128.

[72] Unity Technologies. 2020. *Unity*. Unity Technologies.

[73] Jessica Van Brummelen, Marie O'Brien, Dominique Gruyer, and Homayoun Najjaran. 2018. Autonomous vehicle perception: The technology of today and tomorrow. *Transportation research part C: emerging technologies* 89 (2018), 384–406.

[74] Jinke D. Van Der Laan, 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 (1997), 1 – 10. https://doi.org/10.1016/S0968-090X(96)00025-3

[75] Viswanath Venkatesh and Fred Davis. 2000. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science* 46 (02 2000), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

[76] Marcel Walch, Mark Colley, and Michael Weber. 2019. CooperationCaptcha: On-The-Fly Object Labeling for Highly Automated Vehicles. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3313022

[77] Marcel Walch, Mark Colley, and Michael Weber. 2019. Driving-Task-Related Human-Machine Interaction in Automated Driving: Towards a Bigger Picture. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 427–433. https://doi.org/10.1145/3349263.3351527

[78] Marcel Walch, Kristin Lange, Martin Baumann, and Michael Weber. 2015. Autonomous Driving: Investigating the Feasibility of Car-Driver Handover Assistance. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Nottingham, United Kingdom) *(AutomotiveUI '15)*. Association for Computing Machinery, New York, NY, USA, 11–18. https://doi.org/10.1145/2799250.2799268

[79] Marcel Walch, David Lehr, Mark Colley, and Michael Weber. 2019. Don't You See Them? Towards Gaze-Based Interaction Adaptation for Driver-Vehicle Cooperation. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings* (Utrecht, Netherlands) *(AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 232–237. https://doi.org/10.1145/3349263.3351338

[80] Marcel Walch, David Lehr, Mark Colley, and Michael Weber. 2019. Don't You See Them? Towards Gaze-Based Interaction Adaptation for Driver-Vehicle Cooperation. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings (AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 232–237. https://doi.org/10.1145/3349263.3351338

[81] Marcel Walch, Kristin Mühl, Johannes Kraus, Tanja Stoll, Martin Baumann, and Michael Weber. 2017. From Car-Driver-Handovers to Cooperative Interfaces: Visions for Driver–Vehicle Interaction in Automated Driving. In *Automotive User Interfaces: Creating Interactive Experiences in the Car*, Gerrit Meixner and Christian Müller (Eds.). Springer International Publishing, Cham, 273–294. https://doi.org/10.1007/978-3-319-49448-7_10

[82] Marcel Walch, Tobias Sieber, Philipp Hock, Martin Baumann, and Michael Weber. 2016. Towards Cooperative Driving: Involving the Driver in an Autonomous Vehicle's Decision Making. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ann Arbor, MI, USA) *(Automotive'UI 16)*. Association for Computing Machinery, New York, NY, USA, 261–268. https://doi.org/10.1145/3003715.3005458

[83] Marcel Walch, Marcel Woide, Kristin Mühl, Martin Baumann, and Michael Weber. 2019. Cooperative Overtaking: Overcoming Automated Vehicles' Obstructed Sensor Range via Driver Help. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '19)*. Association for Computing Machinery, New York, NY, USA, 144–155. https://doi.org/10.1145/3342197.3344531

[84] Bernhard Wandtner, Nadja Schömig, and Gerald Schmidt. 2018. Effects of non-driving related task modalities on takeover performance in highly automated driving. *Human factors* 60, 6 (2018), 870–881.

[85] Bernhard Wandtner, Nadja Schömig, and Gerald Schmidt. 2018. Secondary task engagement and disengagement in the context of highly automated driving. *Transportation research part F: traffic psychology and behaviour* 58 (2018), 253–263.

[86] Jiao Wang and Dirk Söffker. 2018. Bridging gaps among human, assisted, and automated driving with DVIs: a conceptional experimental study. *IEEE Transactions on Intelligent Transportation Systems* 20, 6 (2018), 2096–2108.

[87] MinJuan Wang, Sus Lundgren Lyckvi, Chenhui Chen, Palle Dahlstedt, and Fang Chen. 2017. Using Advisory 3D Sound Cues to Improve Drivers' Performance and Situation Awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2814–2825. https://doi.org/10.1145/3025453.3025634

[88] Christopher D Wickens. 2002. Multiple resources and performance prediction. *Theoretical issues in ergonomics science* 3, 2 (2002), 159–177.

[89] Yucheng Yang, Burak Karakaya, Giancarlo Caccia Dominioni, Kyosuke Kawabe, and Klaus Bengler. 2018. An HMI Concept to Improve Driver's Visual Behavior and Situation Awareness in Automated Vehicle. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, IEEE, New York, NY, USA, 650–655.

[90] Kathrin Zeeb, Axel Buchner, and Michael Schrauf. 2016. Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident analysis & prevention* 92 (2016), 230–239.

[91] Markus Zimmermann, Stefan Bauer, Niklas Lütteken, Iris M Rothkirch, and Klaus J Bengler. 2014. Acting together by mutual control: Evaluation of a multimodal interaction concept for cooperative driving. In *2014 international conference on collaboration technologies and systems (CTS)*. IEEE, IEEE, New York, NY, USA, 227–235.