TUDelft

Delft University of Technology

Rolling out the red (and green) carpet
Supporting driver decision making in automation-to-manual transitions

Eriksson, Alexander ; Petermeijer, Bastiaan; Zimmermann, Markus; de Winter, Joost; Bengler, Klaus J.;
Stanton, Neville A.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Rolling Out the Red (and Green) Carpet: Supporting Driver Decision Making in Automation-to-Manual Transitions

Alexander Eriksson ⓘ, Sebastiaan M. Petermeijer, Markus Zimmermann, Joost C. F. de Winter ⓘ, Klaus J. Bengler, and Neville A. Stanton ⓘ

Abstract—This paper assessed four types of human–machine interfaces (HMIs), classified according to the stages of automation proposed by Parasuraman *et al*. ["A model for types and levels of human interaction with automation," *IEEE Trans. Syst. Man, Cybern. A, Syst. Humans*, vol. 30, no. 3, pp. 286–297, May 2000]. We hypothesized that drivers would implement decisions (lane changing or braking) faster and more correctly when receiving support at a higher automation stage during transitions from conditionally automated driving to manual driving. In total, 25 participants with a mean age of 25.7 years (range 19–36 years) drove four trials in a driving simulator, experiencing four HMIs having the following different stages of automation: baseline (information acquisition—low), sphere (information acquisition—high), carpet (information analysis), and arrow (decision selection), presented as visual overlays on the surroundings. The HMIs provided information during two scenarios, namely a lane change and a braking scenario. Results showed that the HMIs did not significantly affect the drivers' initial reaction to the take-over request. Improvements were found, however, in the decision-making process: When drivers experienced the carpet or arrow interface, an improvement in correct decisions (i.e., to brake or change lane) occurred. It is concluded that visual HMIs can assist drivers in making a correct braking or lane change maneuver in a take-over scenario. Future research could be directed toward misuse, disuse, errors of omission, and errors of commission.

Index Terms—Augmented reality, automated driving, driver support systems, human factors, human performance, transitions of control.

I. INTRODUCTION

HIGHLY automated driving will probably be introduced onto public roads within a number of years. Vehicle manufacturer Tesla was the first to market what can be described as a basic Autopilot [1], a partially automated highway driving system (SAE Level 2) and is approaching a higher level of automation (conditionally automated, SAE Level 3) with their Autopilot 2.0 hardware update [2]. Volvo will be launching their first trial with the IntelliSafe Autopilot system as part of their Drive ME project [3], and Daimler is piloting Highway Pilot technology among truck drivers [4].

Conditionally automated vehicles enable extended periods of hands- and feet-free driving during which the driver is free to engage in nondriving tasks, but with the legal constraint that the driver has to be able to switch OFF or override the automation when required [5]. Such automated driving systems will prompt the driver, using a so-called take-over request (TOR), to resume control when the system's limits (e.g., functional or geographical) are reached.

A. Importance of HMIs in Take-Over Scenarios

In a review by De Winter *et al.* [6], it was found that drivers who have been out of the control loop for an extended period of time tend to suffer from degraded situation awareness. It has been argued that drivers need to be aware of the functional limits of the automation before these limits are reached [7]–[10]. Eriksson and Stanton [9] and Stanton [11] proposed a chatty codriver where the vehicle continually informs the driver about its state and limitations.

Furthermore, conditionally automated vehicles need to allow for a "sufficiently comfortable transition time" [12] of "several seconds" after presenting a TOR [5]. In an attempt to gain an understanding of how long drivers need to resume control from an automated vehicle, Eriksson and Stanton [13] reviewed the literature on control transitions and found that drivers take a median of 2.5 s, and in some cases up to 15 s to resume control in urgent scenarios (e.g., [14]). Their review also showed that when drivers are requested to resume control without time pressure, they take between 2.1 and 3.5 s (median) longer

than when under time pressure, depending on task engagement [13]. Moreover, they argued that only considering the "average driver" is insufficient, as this excludes a large part of the driving population due to the long tail of the reaction time distribution (see also [15] and [16]).

In summary, a challenge of conditionally automated driving is to get a driver back to the driving task in a safe manner. Human–machine interfaces (HMIs) should be designed to support a safe response of the driver during a take-over scenario [8], [9].

B. Existing HMIs that Support Take-Over Scenarios

According to Petermeijer *et al.* [19], Zeeb *et al.* [20], and Kerschbaum *et al.* [21], resuming control from an automated vehicle involves several mental and physical stages. The driver resuming control must do the following.

1) Shift visual attention from the nondriving task back to the road.
2) Scan the driving scene to cognitively process and evaluate the traffic situation and make an appropriate decision.
3) Move the hands and feet to the steering wheel and the pedals so that control inputs can be made.
4) Implement the appropriate action via the steering wheel and/or pedals.

A driver's performance during a take-over scenario can also be described at a control level and a tactical level, as per Michon [22]. For example, retaking the steering wheel and stabilizing the vehicle occur at the control level, whereas identifying obstacles and making an evasive manoeuvre are behaviors at the tactical/decision-making level.

Many previous studies on take-over scenarios have provided simple auditory and visual warning signals to convey a TOR to the driver (e.g., [20], [23], and [24]. For a review, see [13] and [25]). Auditory and vibrotactile TORs have been shown to elicit faster reaction times than visual ones [26]. These effects may be due to the fact that auditory and vibrotactile feedback compete less for perceptual resources than visual feedback [27] as driving is primarily a visual task [28]. Moreover, it has been found that presenting bimodal auditory/vibrotactile warnings yielded a slight improvement in reaction time compared to their unimodal constituents [29].

In addition to receiving a take-over warning, a driver could also be supported in making decisions. Research has indicated that drivers, after receiving a vibrotactile warning, first visually assess the outside environment [30], [31]. The vibrotactile modality is not particularly effective in conveying complex information [32], [33]. Visual and vocal messages, on the contrary, can convey complex information that is linked to the surrounding scene [33]–[35]. Thus, auditory [36] and vibrotactile signals are recommended as warnings (i.e., they are expected to attract attention and support a fast initial response), whereas visual and vocal displays are recommended for conveying semantics to the driver (i.e., they are expected to support cognitive processing and tactical decision making).

C. Automation framework to Support Decision Making

A framework proposed by Parasuraman *et al.* [37] stated that automation can be divided into the following four stages: infor-
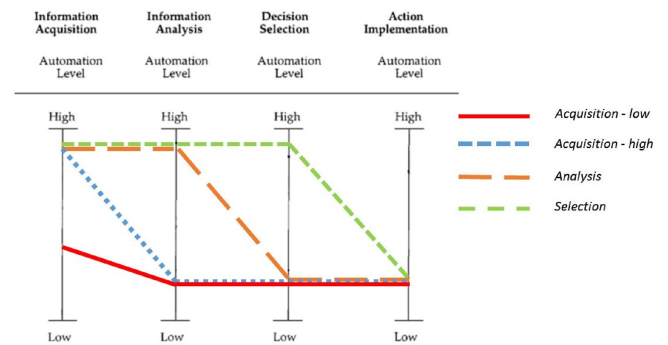
Fig. 1. Representation on the four stages of information support, namely *acquisition – low* (red solid), *acquisition – high* (blue dotted), *analysis* (orange long dashes), and *selection* (green short dashes). Adapted from Parasuraman *et al.* [37].

mation acquisition; information analysis; decision selection; and action implementation (in short, *acquisition, analysis, selection,* and *implementation*). According to Parasuraman *et al.* [37], an automated system may involve different levels of automation at each stage. Note that Parasuraman *et al.* [37] based their model on existing models of human information processing, which explains the similarities between the stages of their framework and the information processing stages in the take-over process described above.

When a conditionally automated vehicle (SAE Level 3) reaches its functional limits and presents a TOR to the driver, this inherently means that the automated system cannot safely *implement* actions anymore and requires driver intervention. Despite no longer being able to implement actions, the system could potentially still assist the driver in making decisions by means of an HMI displaying information available from the remaining three automation stages (i.e., *information acquisition*, *information analysis*, and *decision selection*).

A TOR consisting of a notification in the instrument cluster combined with an auditory signal, as in the work of Gold *et al.* [23], would be considered a low level of *acquisition* support (see Fig. 1, *acquisition—low*), because the HMI only informs the driver that she/he needs to take over (starting with scanning the environment). A higher level of *acquisition* (*acquisition—high*) would draw the attention toward important elements in the surroundings. An interface that also provides information about the surrounding traffic situation (e.g., adjacent lane is free/occupied) [38] and suggests actions (e.g., change lane/brake) [39] would score highly on information *analysis* and decision *selection,* respectively.

D. Possible Advantages and Disadvantages of Feedback and Support Systems

The benefits of feedback and support systems have been widely reported in the literature. For example, forward collision warning systems are known to decrease brake reaction times [40]–[42], and a vibrotactile gas pedal was found to improve eco-driving performance [43]. A simulator study by Israel [44] showed that visual head-up displays decreased the number of navigational mistakes at intersections. Moreover, it has

previously been shown that visual augmented feedback can be used to improve drivers' situation awareness [27], [45].

Detrimental effects of support systems have also been reported (primarily in aviation), such as complacency [46] and skill degradation [47], [48]. Another issue that arises with increasing support is automation bias, in the form of errors of omission or commission. An error of omission occurs when an operator fails to implement an appropriate action because the operator was not informed by the support system [38], [49].

An error of commission occurs when an operator implements an incorrect action suggested by the support system, without considering other indicators [50]. In a review of the literature, Mosier and Skitka [51] noted that automation bias occurs not only for untrained operators but also for experienced ones, suggesting that automation bias is a persistent problem. These forms of automation bias could lead to dangerous situations (e.g., [52]), for example when the system falsely instructs the driver to change lane whilst the target lane is occupied by other vehicles.

E. Aim of This Experiment

The aim of this experiment was to investigate driver behavior in take-over scenarios with different stages of support.

Eriksson and Stanton [53] previously used the so-called COntextual COntrol model (COCOM) [54] to explain driver performance in a take-over scenario. This model states that successful tactical decision can be invoked by giving operators more time or by enhancing the predictability of the situation. The authors used this to compare driver-paced transitions [53] (which allow for extra planning time) with transitions under time pressure (cf., [23] and [55]–[58]). In accordance with the predictions of the COCOM, we expected that improvements would occur in driver decision making by increasing the predictability of the situation through HMIs that involve different stages of automation. We expected that the HMIs assessed in this paper would not help reduce the initial reaction times (e.g., grabbing the steering wheel) after the TOR. The immediate control activity is "automatic" and requires little conscious effort to be executed (cf., [59] and [60]).

This study assessed how driver's performance was affected as a function of visual support within the following automation stages: information acquisition; information analysis; and decision selection. It was expected that drivers would implement decisions more correctly and faster when receiving a higher stage of support, as shown in Fig. 1. Moreover, by measuring head movements, we aimed to obtain insight into whether drivers are prone to automation bias and follow the HMI's suggestion without verifying the safety of the suggested action [46].

II. METHOD

A. Participants

A total of 25 participants (14 male and 11 female) with a mean age of 25.7 years (SD = 3.9, min = 19, max = 36, and $N = 24$ because one participant did not report his age) and an average driving experience of 8.3 years (SD = 4.1) took part in the study.
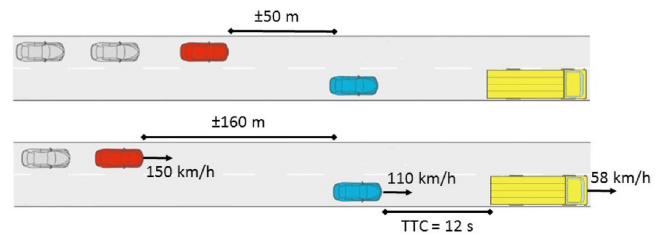

Fig. 2. Two take-over scenarios. Top: the group of cars is too close to change lane safely and the driver is expected to brake (i.e., braking scenario). Bottom: the group of cars is far enough away for the driver to overtake the vehicle safely (i.e., lane change scenario).

Two participants indicated to drive daily, 3 participants reported 4–6 days a week, 10 reported 1–3 days a week, 5 reported once a month, 4 reported less than once a month, and 1 reported they never drove in the past 12 months. The study received ethical approval from the Southampton University Ethics Committee (RGO number: 19930), and all participants provided written informed consent.

B. Apparatus

A static simulator, fixed-base, BMW 6-series mockup, operated the SILAB (version 4) software. The simulator offered a 180° front view and rear projections for every mirror (left, inner, and right), generated by six projectors. Road and engine noise was played back, and low-frequency vibrations were provided via a bass shaker in the driver seat. The automation could be toggled by pressing a button (with a diamond-shaped icon) on the steering wheel. The automation adhered to the lane centre by applying light torques on the steering wheel. The driver could still steer when the automation was active, and accordingly influence the lateral position of the vehicle. The automation disengaged when the lateral speed of the car exceeded about 1 m/s, or when the brake pedal depression exceeded 25%. An icon located between the speedometer and tachometer indicated the automation status (i.e., unavailable, active, or inactive).

The participants played "Angry Birds" as a nondriving task during the intervals of automated driving. Angry Birds was deemed suitable because it is an interruptible [61] task that does not penalize the player for switching to another task. The driver played the game on a Lenovo A7-50 7-inch tablet that was mounted in the centre console, in front of the radio.

The participants' head and gaze motion were tracked using a three-camera remote system (Smart Eye Pro 6.1). Simulation and eye tracking data were synchronized and logged at 60 Hz. The vehicle environment was modeled in the Smart Eye software to relate eye gaze and real-world objects. The windshield was defined as an area of interest.

C. Take-Over Scenarios

The automated vehicle drove in the right lane on a two-lane highway at 110 km/h (68.4 mi/h) and approached a slow-moving vehicle (e.g., truck, tractor, or moped) driving at 58 km/h (36.0 mi/h) (see Fig. 2). When the time to collision (TTC) with the slow-moving vehicle decreased below 12 s, the

Fig. 3. Illustration of the vibrotactile seat and the location of the 48 vibration motors (i.e., white circles). The motors are arranged in two matrices of 6×4 motors (one matrix in the seat back, and the other matrix in the seat bottom).

automation issued a TOR. Simultaneously, a group of other vehicles, driving at 150 km/h (93.2 mi/h), approached in the left lane. The group of vehicles was, at the moment of the TOR, either approximately 165 m behind (i.e., the first vehicle would pass in approximately 1 s) so that the driver could safely change lane (i.e., lane change scenario), or approximately 50 m behind (i.e., the first vehicle would pass in approximately 4.5 s) so that the driver was required to reduce the speed of his or her vehicle (i.e., braking scenario). In summary, drivers could safely change lane in the lane change scenario, whereas they could not safely change lanes in the braking scenario until the platoon had passed.

D. HMIs for TORs

To increase the likelihood that drivers respond successfully to a TOR, a bimodal feedback paradigm was used. The HMIs in this experiment consisted of vibrotactile stimuli in the seat, provided by vibration motors (see Fig. 3). Simultaneously, an augmented reality display (based on [39], [62], and [63]) showed warnings, information, or decision suggestions for courses of action (see Fig. 4). Depending on whether the drivers faced a braking or a lane change scenario, the information analysis and decision selection visuals were redundantly encoded by means of color (red and green; i.e., having a well-established meaning, see also [45]), shape (wide or narrow carpet), and direction (left or backward arrows).

More specifically, this study tested the following four types of information support conditions during six take-over scenarios per condition (three lane change conditions and three braking conditions) with various stages of support (see Fig. 1).

1) Information Acquisition—Low: A vibrotactile warning indicating that the driver had to resume control. The vibration seat (see Fig. 3) presented a series of three 320 ms pulses (70 ms engaged and 250 ms disengaged) in all 48 motors in the seat to inform the driver that he/she needed to resume control. No extra visuals were presented in this condition. Hence, the driver did not receive any additional information from the interface other than the vibrotactile
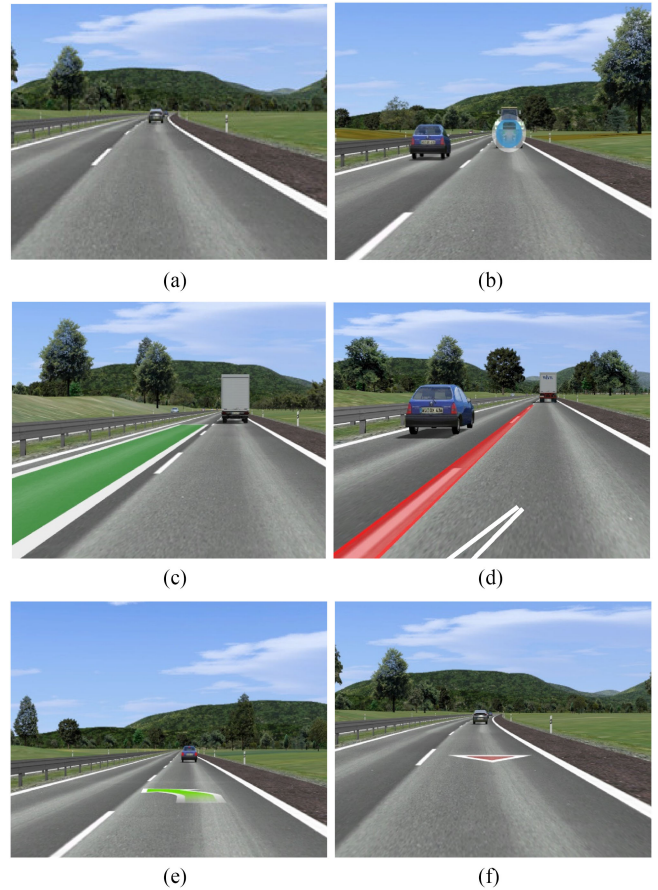
Fig. 4. Visual interface for the four levels of support. (a) Baseline condition: No visual support in both scenarios. (b) Sphere condition: A blue sphere highlighting the slow-moving vehicle ahead in both scenarios. (c) Carpet condition: A green carpet in the left lane for the lane change scenario. (d) Carpet condition: A red barrier covering the lane markings for the braking scenario. (e) Arrow condition: A green arrow pointing left for the lane change scenario. (f) Arrow condition: A red arrow pointing backward for the braking scenario.

TOR [see Fig. 4(a)]. This vibration was the *baseline* condition.

2) Information Acquisition—High: At the same moment as the TOR (i.e., the vibrotactile warning), an augmented sphere highlighted the slowly moving vehicle ahead (see Fig. 4(b); similar to [39] and [64]). Thus, in addition to the TOR, the driver received a visual cue to direct his/her attention toward the cause of the TOR (the leading vehicle). The vibrations together with the sphere overlay are referred to as the *sphere* condition.

3) Information Analysis: In addition to the vibrotactile warning, an augmented-reality overlay informed whether there was a gap in the left lane. In the lane change scenario, a wide green carpet in the left lane informed drivers about available space in the other lane [like in [39]; see Fig. 4(c)], whereas in the braking scenario, a narrow red barrier between the lanes emphasized a no-passing zone [inspired by the H-mode visuals [63]; see Fig. 4(d)]. The vibrotactile warning and visual information formed the *carpet* condition.

4) Decision Selection: At the same moment as the vibrotactile warning, augmented reality arrows at a fixed distance from the driver indicated that the driver could change lane or brake (see Fig. 4(e) and (f); see also [39]). The vibrotactile warning with the arrow is referred to as the *arrow* condition.

In all scenarios, the HMI was hidden when the host vehicle crossed into the adjacent lane. Additionally, the green carpet and green arrow disappeared when the approaching vehicles on the left lane were too close (TTC = 2 s). In the braking scenario, the HMI disappeared when the platoon had passed. Additionally, the red arrow disappeared when TTC to the lead vehicle became larger than 12 s, as this was an indication that the participant had already braked sufficiently.

E. Experimental Design and Instructions to Participants

A within-subject design was used for the different HMI conditions. The participants drove a 1.5-min practice run during which they could familiarize themselves with the automation and the take over, after which the following four trials were driven in counterbalanced order, each trial with a different HMI: information acquisition—low; information acquisition—high; information analysis; and (4) decision selection.

Participants were provided with an instruction form, which stated that they would be driving an automated car that controls speed and stays in the lane. The form also explained the automation-status icons on the dashboard, and instructed participants to have their hands off the steering wheel and their feet off the pedals while the automation is active. Participants were instructed to play Angry Birds on the tablet in the car and were informed that they did not have to look at the road. Participants were also informed that they will be approaching a slow-moving vehicle ahead, at which moment the automation will ask them to take back control of the car, via vibrations in the seat and one of four assistance systems. The form included a picture and text explaining each HMI. Finally, participants were informed that the automation will function perfectly and does not need any monitoring, except when it provides a TOR. Participants were not informed about the behavior of the approaching platoon in the left lane.

During each trial, the participant experienced six take-over procedures of which three took place in the braking scenario and three in the lane change scenario. Specifically, the lane change (LC) and brake (B) scenarios were presented in the following order: B, LC, B, LC, B, LC for the baseline and arrow trials, LC, B, LC, B, LC, B for the carpet trial, and LC, B, B, LC, LC, B for the sphere trial. Each trial lasted approximately 12 min, with a request to resume control in a braking or lane change scenario occurring about every 110 s. After each trial, participants stepped out of the vehicle to have a break and to complete two questionnaires.

F. Dependent Measures

The experiment employed several *objective* measures to capture performance and reaction times, which are as follows.

1) Success Rate: In the lane change scenario, a manoeuvre was considered successful if the driver changed lanes *before* the cars in the adjacent lane passed (thus, avoiding unnecessary, harsh braking). In the braking scenario, a manoeuvre was regarded as successful when the participant made a lane change *after* all cars in the adjacent lane had passed (thus, avoiding aggressive merging into the other lane and maintaining the speed limit on the road). The definition of a lane change was that the host vehicle's centre of gravity had crossed the lane boundary.

2) Braking Rate: The percentage of scenarios in which the participants used the brake pedal. Application of the brakes in the lane change scenario is an indication of unnecessary deceleration.

3) Eyes-on-Windshield Reaction Time: The time between the onset of the TOR to the moment the eye gaze of the driver was first detected in the windshield area.

4) Hand-on-Wheel Reaction Time: The time between the onset of the TOR to the moment the drivers put a hand back on the steering wheel, measured with induction coils in the steering wheel.

5) Steer Move Time: The time between the onset of the TOR and the first detectable steering input (i.e., above sensor noise threshold). The steer move time is equivalent to the hand-on-wheel reaction time but was measured from the steering wheel angle instead of induction coils in the steering wheel.

6) Brake Reaction Time: The time between the onset of TOR to the onset of a depression of the brake pedal [23].

7) Lane Change Time: The lane change time is the time from the onset of the TOR to the moment that the host vehicle's centre of gravity had crossed the lane boundary.

8) Head Angle: The mean and standard deviation of the angle of the head as a function of travelled distance was used to represent the direction of the participant's visual attention. The head angle was defined as the nose angle (a vector originating at the middle of the head and pointing out of the nose) in world coordinates (perpendicular to the front screen with its origin approximately at the vanishing point of the road). Thus, the head angle is zero when the driver's head is pointing straight to the road. Note that we used head movements instead of eye movements, because eye movement data was deemed less robust according to our data quality assessment. Considering that head orientation is a proxy for glance direction [65], head orientation was deemed suitable for our purpose of assessing whether the participants looked at the road or to at secondary task display.

The following two questionnaires were utilized as *subjective* measures for workload and acceptance.

1) The NASA raw TLX was used to evaluate the perceived workload per condition [66], [67]. The questionnaire consists of six items: mental demand, physical demand, temporal demand, performance, effort, and frustration. The items had a 21-tick Likert scale, ranging from "very low" to "very high," except the performance item, which ranged from "perfect" to "failure."
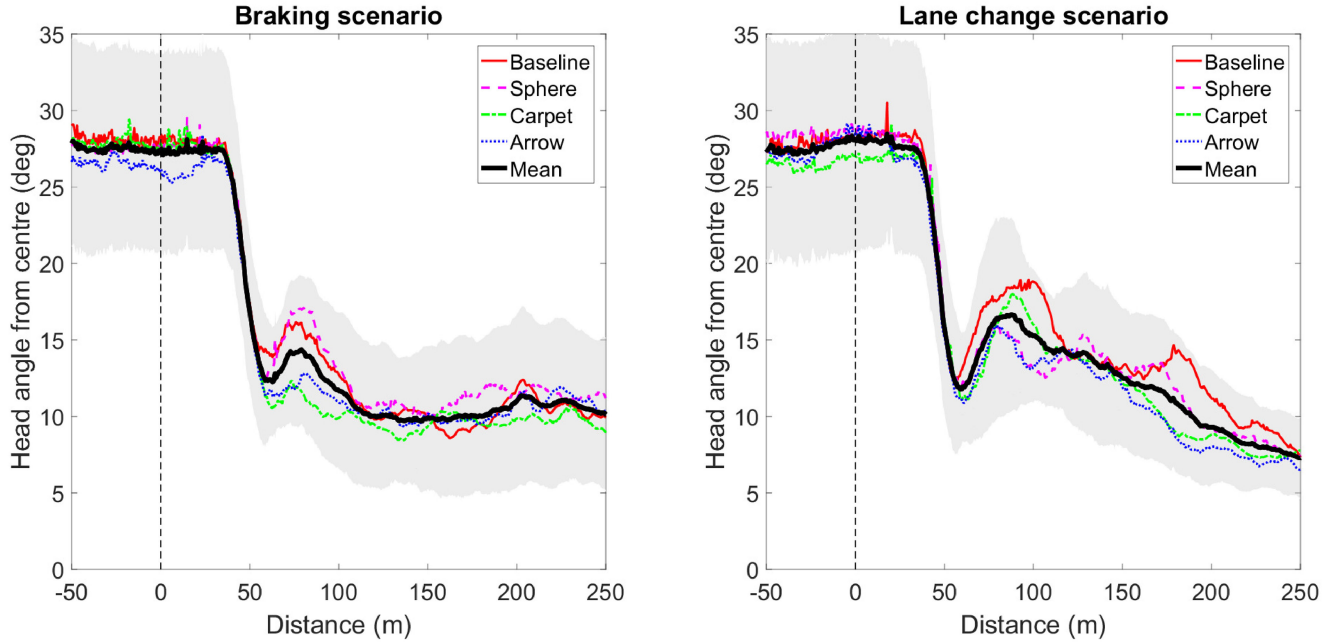
Fig. 5.   Mean head angle across participants ($N = 25$) as a function of travelled distance for braking (left) and lane change (right) scenarios. The shaded area around the (black) mean indicates the standard deviation across the means of the participants in all conditions. The vertical dashed line indicates the moment of the TOR.

2) A nine-item technology acceptance questionnaire [68] was used to measure the usefulness and satisfaction of the different support types. The usefulness score was calculated from the following five items on a semantic-differential five-point scale from -2 to $+2$: 1. useful–useless, 3. bad–good, 5. effective–superfluous, 7. assisting–worthless, and 9. raising alertness–sleep-inducing. The satisfaction score was calculated from the following four items: 2. pleasant–unpleasant, 4. nice–annoying, 6. irritating–likeable, 8. undesirable–desirable. Sign reversals were conducted for items 1, 2, 4, 5, 7, and 9, so that a higher score indicates higher useful-ness/satisfaction.

G. Statistical Analyses

Due to the expected nonnormal distribution of the response time data (these types of data are known to have a high-kurtosis distribution) [13], [18], nonparametric Friedman tests with Wilcoxon signed-rank tests (with the alpha level corrected for multiple comparisons) were used. Effect sizes of the Friedman's test were represented by Kendall's W, defined as $W = \chi^2/N(k-1)$, where χ^2 is the test statistic, N is the number of participants (25), and k is the number of conditions per participant (4).

For the Wilcoxon signed-rank tests, effect sizes were calculated as $r = |Z/N^{0.5}|$, where Z is the Z-statistic, and N is the number of participants.

For the comparison of the head movements between HMIs, Wilcoxon signed-rank tests of the head eccentricity were performed for every time sample (see Fig. 6). The level of significance was visualized as the negative base-10 logarithm of the

p-value, where large values represent small p-values in a similar fashion to the 'Manhattan' plot [53], [69]–[71]. Our use of multiple Wilcoxon signed-rank tests allows for a high temporal resolution (as opposed to using larger bin sizes and fewer tests). It must be noted that despite relatively conservative corrections of the significance level, the results should only be seen as indicative. The interpretability of the analysis has been increased through the addition of the effect size measure r. Additionally, two animated clips of the head movements represented as a heatmap over time, for the same section of road as shown in Fig. 5, have been made (see supplementary materials).

All statistical tests were performed at the level of the participant. Average values of the dependent measures per participant were calculated across the three braking or three lane-change scenarios within a trial. For an alpha of 0.05, a sample size of 25, and a medium-to-strong effect size ($d_z = 0.60$), the achieved statistical power for a two-tailed test is 80%.

III. RESULTS

From the 600 scenarios (25 participants $\times$4 trials $\times$ 6 scenarios per trial), 13 scenarios were excluded due to improper data recording or a participant already touching the steering wheel at the moment of the TOR. In 81.3% of the braking scenarios, the participants made a lane change after the cars in the adjacent lane had passed, whereas in 95.7% of the lane change scenarios the participants performed a lane change ahead of the cars (see Table I). In events that were counted as unsuccessful (e.g., braking from 120 to 58 km/h, while a lane change is possible, cannot be considered a "successful" action while driving on a highway. Furthermore, it is debatable whether such braking is safe. We argue that it is not safe, because such a major deceleration may
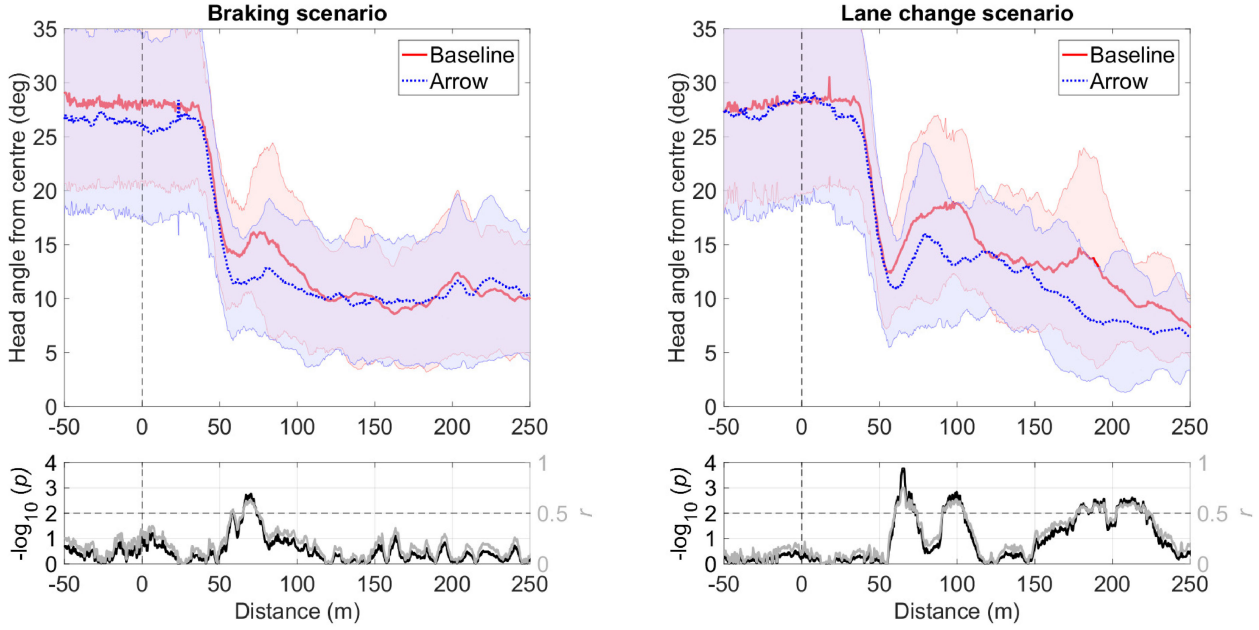
Fig. 6. Top: Mean and standard deviation (shaded area) of head angle across participants ($N = 25$) as a function of travelled distance for the baseline and arrow conditions. The vertical dashed line indicates the moment of the TOR. Bottom: *p*-values from Wilcoxon signed-rank tests for the head angle between the baseline and arrow conditions. The horizontal dashed line indicates a threshold of $p = 0.01$.

TABLE I
SUCCESS RATE AND BRAKING RATE FOR STEERING AND BRAKING SCENARIOS AS A FUNCTION OF THE SUPPORT CONDITION

	Braking scenario		Lane change scenario	
	Success rate (%)	Braking rate (%)	Success rate (%)	Braking rate (%)
Baseline	77.3	88.0	96.0	12.0
Sphere	78.7	86.7	86.7	33.3
Carpet	85.3	96.0	100	2.7
Arrow	84.0	96.0	100	6.7
Overall	81.3	91.7	95.7	13.7

TABLE II
MEDIANS AND INTERQUARTILE RANGES (IQR) FOR THE DEPENDENT MEASURES ACROSS PARTICIPANTS

Reaction time (s)		Baseline Median (IQR)	Sphere Median (IQR)	Carpet Median (IQR)	Arrow Median (IQR)
Lane change scenario	Eyes on windshield	1.42 (0.53)	1.59 (0.53)	1.47 (0.42)	1.43 (0.58)
	Hand on wheel	2.12 (1.30)	1.89 (1.22)	1.86 (0.63)	1.98 (0.87)
	Steer move	2.01 (0.88)	2.34 (1.52)	2.06 (0.62)	2.13 (0.82)
	Braking	–	–	–	–
	Lane change	6.97 (1.94)	6.65 (2.76)	5.54 (2.07)	5.54 (1.19)
Braking scenario	Eyes on windshield	1.47 (0.43)	1.51 (0.59)	1.49 (0.61)	1.41 (0.53)
	Hand on wheel	2.16 (0.95)	2.20 (1.14)	1.85 (0.61)	1.99 (1.18)
	Steer move	2.22 (0.99)	2.33 (1.11)	2.15 (1.33)	2.21 (1.21)
	Braking	3.56 (2.17)	3.14 (1.41)	3.33 (2.26)	3.04 (1.34)
	Lane change	20.46 (5.01)	20.94 (3.50)	21.36 (2.04)	20.76 (2.92)

Note. Brake reaction times are not reported in the lane change scenario because participants often did not brake (see Table I).

cause a third vehicle to collide in the back), the participants either hazardously changed lanes (causing the platoon to brake) or waited until all vehicles had passed before changing lanes (braking and lane change scenario, respectively).

Participants in the braking scenario changed lanes erroneously more often (i.e., a lower success rate) in the baseline and sphere conditions than in the carpet and arrow conditions (see Table I). In the lane change scenario, the sphere yielded fewer successful lane changes and more unneeded braking actions compared to baseline, carpet, and arrow. Furthermore, a learning effect was observed, where participants braked less often as they experienced additional lane change events (supplementary material); this observation is consistent with findings from Petermeijer *et al.* [29]. Descriptive statistics for the objective measures are given in Table II. There were no significant effects for the HMIs on the steer move times in both the braking scenario, $\chi^2(3, N = 25) = 1.42, p = 0.702$, and $W = 0.02$, and the lane change scenario, $\chi^2(3, N = 25) = 5.30, p = 0.151$, and $W = 0.07$. In the braking scenario, there were significant

differences in brake reaction time ($\chi^2(3, N = 21) = 14.2, p = 0.003$, and $W = 0.23$). The arrow, in particular, evoked a rapid and strong braking input in the braking scenario (see supplementary materials).

A significant effect of the HMIs was also found for the lane change time in the braking scenario, $\chi^2(3, N = 25) = 8.47, p = 0.037$, and $W = 0.11$. Post-hoc analysis showed a significantly

TABLE III
PAIRED COMPARISONS BETWEEN THE FOUR HMIS REGARDING THE LANE CHANGE TIME AFTER THE TOR

Lane change scenario									
	Baseline			Sphere			Carpet		
	Z	p	r	Z	p	r	Z	p	r
Sphere	0.69	0.493	0.14						
Carpet	-2.64	0.008*	0.53	-2.49	0.013	0.50			
Arrow	-3.16	0.002*	0.63	-3.86	<0.001*	0.77	-0.70	0.484	0.14
Braking scenario									
Sphere	-0.87	0.382	0.17						
Carpet	2.38	0.017	0.48	1.99	0.046	0.40			
Arrow	0.20	0.840	0.04	1.09	0.276	0.22	-2.92	0.004*	0.58

*indicates a significant difference at the Bonferroni-corrected alpha level (0.0083).

slower lane change time for the carpet as compared to the arrow (see Table III), which may be due to the red barrier which remained present until the traffic stream had passed.

A significant main effect of the HMIs was also found for the time it took to change lane in the lane change scenario, $\chi^2(3, N = 25) = 15.84$, $p = 0.001$, and $W = 0.21$. Wilcoxon signed-rank post-hoc tests showed a significantly faster execution of the lane change for the arrow compared to the sphere and baseline, and for the carpet compared to baseline (see Table III).

A. Head Movements

Fig. 5 shows the mean head angle (with respect to looking to the road straight ahead) across participants in the four driving conditions, for the lane change and braking scenarios. The shaded area represents the standard deviation across the means of participants. Before the TOR, participants exhibited a large head angle, because they were performing the secondary task located at their bottom right. After the TOR was issued, participants shifted their attention back to the road.

Fig. 6 is a detailed version of Fig. 5, which compares the baseline with the arrow condition. The bottom graphs show the results of Wilcoxon signed-rank tests between these two conditions. It is worth noting that the values in the bottom graphs are negative base-10 logarithms of p, meaning that high values represent low p-values. The horizontal dashed lines show the threshold for significant differences ($p < 0.01$). Significantly larger head angles were found for the baseline condition compared to the arrow condition, in the braking scenario around 70 m post TOR. This effect may be due to the arrow reducing the need to check the status of the left lane. Another explanation is that participants were focusing on the arrow, which appeared in the centre of the lane (see Fig. 4).

B. Satisfaction and Usefulness Scale

The results of a Friedman test showed significant differences in perceived usefulness of the different HMIs, $\chi^2(3, N = 25) = 22.72$, $p < 0.001$, and $W = 0.30$.

As shown in Fig. 7, the sphere condition yielded the lowest scores of the four HMIs. Post-hoc Wilcoxon signed-rank tests
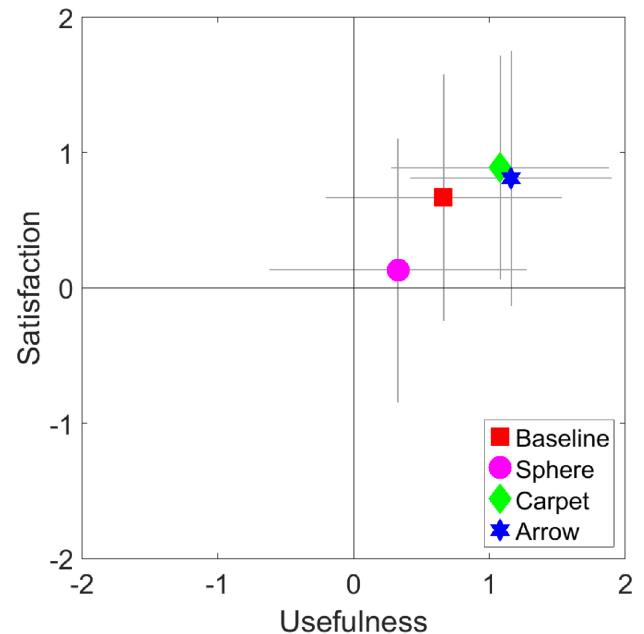
Fig. 7. Mean usefulness and satisfaction score per condition. The grey error bars indicate the mean $\pm$ 1 standard deviation across participants. Mean (SD) usefulness: 0.66 (0.87), 0.33 (0.94), 1.08 (0.80), and 1.16 (0.74), for baseline, sphere, carpet and arrow, respectively. Mean (SD) satisfaction: 0.67 (0.91), 0.13 (0.97), 0.89 (0.83), 0.81 (0.94), for the baseline, sphere, carpet, and arrow conditions, respectively.

showed that the arrow and carpet yielded a significantly higher usefulness than the baseline and sphere (see Table IV).

A Friedman test of perceived satisfaction also showed significant differences between the HMI, $\chi^2(3, N = 25) = 12.30$, $p = 0.006$, and $W = 0.16$. Post-hoc Wilcoxon signed-rank tests showed significantly higher satisfaction for the carpet compared to the sphere condition (see Table IV).

C. Overall Workload

The results of a Friedman's ANOVA showed significant differences in self-reported overall workload as measured by the NASA-TLX levels $\chi^2(3, N = 25) = 10.74$, $p = 0.013$, and

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ERIKSSON *et al.*: ROLLING OUT THE RED (AND GREEN) CARPET: SUPPORTING DRIVER DECISION MAKING 9

TABLE IV
PAIRED COMPARISONS BETWEEN THE PERCEIVED USEFULNESS AND SATISFACTION OF THE FOUR HMIS

Usefulness	Baseline			Sphere			Carpet		
	Z	p	r	Z	p	r	Z	p	r
Sphere	-1.63	0.102	0.33						
Carpet	1.81	0.071	0.36	3.39	< 0.001*	0.68			
Arrow	2.99	0.003*	0.60	3.54	< 0.001*	0.71	0.61	0.540	0.12
Satisfaction									
Sphere	-1.89	0.059	0.38						
Carpet	0.73	0.464	0.15	3.24	0.001*	0.65			
Arrow	1.09	0.276	0.22	2.40	0.016	0.48	-0.19	0.847	0.04

*indicates a significant difference at the Bonferroni-corrected alpha level (0.0083).

TABLE V
MEAN AND STANDARD DEVIATION OF THE SELF-REPORTED WORKLOAD PER
HMI CONDITION. RESULTS ARE SHOWN ON A SCALE FROM 0 (VERY
LOW/PERFECT) TO 100 (VERY HIGH/FAILURE)

	Baseline	**Sphere**	**Carpet**	**Arrow**
	M (SD)	M (SD)	M (SD)	M (SD)
Mental demand	36.8 (22.6)	37.2 (22.6)	31.0 (21.7)	30.2 (22.3)
Physical demand	20.2 (12.8)	21.8 (16.1)	17.8 (14.2)	19.6 (14.9)
Temporal demand	39.4 (20.0)	41.0 (22.3)	31.0 (21.3)	31.4 (22.6)
Performance	30.4 (17.3)	29.2 (17.1)	27.6 (19.8)	27.4 (16.5)
Effort	34.2 (19.8)	28.2 (13.6)	27.6 (18.7)	25.6 (16.5)
Frustration	32.0 (18.1)	34.2 (21.3)	29.0 (21.1)	28.8 (20.4)
Total	**32.2 (14.0)**	**31.9 (13.6)**	**27.3 (15.0)**	**27.2 (15.1)**

$W = 0.14$ (see Table V). No significant difference was observed when carrying out a Bonferroni-corrected Wilcoxon signed-rank post-hoc analysis between pairs of conditions.

IV. DISCUSSION

A. Success Rate and Braking Rate

The results showed that drivers had an overall success rate of 81.3% in the braking scenario and 95.7% in the lane change scenario. The success rates for the carpet and arrow conditions were higher than for the baseline and sphere conditions, indicating that drivers were better assisted when they received higher stages of support. Similar effects with respect to a baseline condition without visual support have been shown by Lorenz *et al.* [45], who used a green and red augmented reality carpet that indicated a safe versus a restricted way of travel, and by Zimmermann *et al.* [39], who reported higher success rates when employing a combination of a carpet and an arrow. The high success rates of the carpet and arrow may be due to the use of salient green

and red colors, which are well-established indicators of safety and danger/urgency, respectively, [72], [73]. One difference between the carpet and the arrow in the braking scenario was that the arrow yielded stronger and more immediate braking than the carpet (i.e., red barrier) (see also supplementary materials). This can be explained by the fact that the arrow represents a directive to brake (i.e., decision-selection automation).

In the lane change scenario, there was a lower braking rate for the arrow and carpet conditions as compared to the baseline and sphere conditions, indicating that drivers receiving a higher stage of support (cf., Fig. 1) made more successful lane changes. The relatively high braking percentages in the baseline and sphere conditions in the lane change scenario may be because participants were uncertain about which action to undertake, or because they intended to increase their time budget for making a decision.

Contrary to our expectations, the sphere yielded a lower success rate than the baseline condition in the lane change scenario (see Table I). An explanation could be that the meaning of the sphere was unclear or that the sphere was interpreted as a danger, thereby triggering an unneeded braking reaction. Another explanation is that the sphere masked some of the intrinsic visual cues, such as optical looming [74], which drivers normally use in braking. While this is a possibility, some of the looming information was still provided by the sphere overlay, as the sphere scaled with the distance to the lead vehicle.

B. Reaction Times

The reaction times of the eyes on the windshield and the steering wheel were not significantly affected by the visual information presented to the driver. This is expected, as these measures reflect a shift of attention, which is unrelated to the type of visual support. Lorenz *et al.* [45] and Langlois and Soualmi [64] also found no significant differences in initial reaction times between visual interfaces (i.e., augmented red or green carpets) and a control condition.

C. Head Movements

The head movements for the four HMIs were relatively similar (see Fig. 5). However, there were subtle but significant differences after the TOR: the carpet and arrow attracted visual

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

10   IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS

attention, whereas participants in the baseline condition were less likely to look straight ahead (see Figs. 5 and 6).

The increase in forward attention could be a manifestation of attentional tunnelling [75], as participants focus on the augmented feedback, while paying less attention to the rest of the scene (e.g., checking the mirrors to see whether the left lane is free).

According to theories in linguistics [7], [76], [77], effective communication is achieved when messages are relevant and not more informative than required. Extrapolating this idea to HMI design, we may recommend that HMIs should not present too much information (e.g., visual clutter and multiple symbols), as this may cause participants to focus on the HMI itself rather than the surrounding environment.

D. Subjective Data

In terms of perceived usefulness and satisfaction of the HMIs, there was no statistically significant difference between the carpet and arrow conditions, but drivers found both the carpet and the arrow more useful than the sphere. Previous research by Werneke and Vollrath [78] found that augmented feedback in the form of a bird's eye view received more positive ratings from drivers than a late warning in the form of a sphere that highlighted a dangerous vehicle. Schwarz and Fastenmeier [79] found that augmented reality warnings (i.e., scenario-specific icons accompanied by arrows coming from the direction of danger) were rated more highly than unspecific visual or auditory warnings.

The relatively poor driving performance in the sphere condition may indicate that the information it conveyed was insufficient for making a proper decision. It is possible that drivers needed time to interpret the sphere, leading to lower satisfaction and usefulness scores.

V. CONCLUSION

This study assessed four types of HMIs, classified along the stages of automation suggested by Parasuraman *et al.* [37]. We hypothesized that drivers would benefit from visual feedback on the *information acquisition*, *information analysis*, and *decision selection* stages during transitions to manual control following conditionally automated driving.

The HMIs in this paper did not benefit initial reaction times. Improvements appeared, however, in decision making, where participants had to assess whether to brake or to change lane after the request to resume control. The carpet (*information analysis*) and arrow (*decision selection*) conditions outperformed the sphere and baseline conditions in terms of manoeuver success rates. Merely highlighting an obstacle via a sphere (*information acquisition—high*) did not improve decision making, but rather increased unnecessary braking.

In our study, the HMIs always provided reliable information, yet some participants still made unsafe lane changes in the braking scenario (see Table I). Future research could investigate more authoritative HMIs (e.g., by adding speech feedback) to increase driver compliance. Future research could also investigate driver behavior when the HMI does not appear when it should (potentially resulting in an error of omission) and when

the HMI provides incorrect advice (potentially resulting in an error of commission). In our study, braking and lane changes scenarios were presented in an alternating fashion. Furthermore, future research should be carried out in a larger variety of randomized traffic situations. Finally, the participants experienced the HMIs for a rather short time (approximately 12 min), but showed learning in the form of a reduction of unneeded braking as a function of scenario number. A longitudinal study on actual roads should be performed to study whether participants develop complacency/misuse (e.g., whether the HMI causes drivers to fail to check the blind spot and change lanes when braking would be safer) and disuse (e.g., whether drivers disengage the HMI).

REFERENCES

[1] Tesla Motors, "Model S software version 7.0," 2016. [Online]. Available: https://www.teslamotors.com/presskit/autopilot

[2] Tesla Motors, "Autopilot," 2017. [Online]. Available: https://www.tesla.com/en_GB/autopilot

[3] VolvoCars, "Intellisafe autopilot." [Online]. Available: http://www.volvocars.com/au/about/innovations/intellisafe/autopilot

[4] Daimler, "The auto pilot for trucks: Highway pilot," 2016. [Online]. Available: https://www.daimler.com/innovation/autonomous-driving/special/technology-trucks.html

[5] *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*," *J3016_201609*, SAE International, Warrendale, PA, USA, 2016.

[6] J. C. De Winter, R. Happee, M. H. Martens, and N. A. Stanton, "Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 27, pp. 196–217, 2014.

[7] L. F. Cranor, "A framework for reasoning about the human in the loop," in *Proc. 1st Conf. Usability, Psychol., Secur.*, San Francisco, CA, USA, 2008, pp. 1–15.

[8] A. Eriksson and N. A. Stanton, "The chatty co-driver: A linguistics approach applying lessons learnt from aviation incidents," *Saf. Sci.*, vol. 99, pp. 94–101, 2017.

[9] A. Eriksson and N. A. Stanton, "When communication breaks down or what was that? – The importance of communication for successful coordination in complex systems," *Proc. Manuf.*, vol. 3, pp. 2418–2425, 2015.

[10] B. D. Seppelt and J. D. Lee, "Making adaptive cruise control (ACC) limits visible," *Int. J. Human Comput. Stud.*, vol. 65, no. 3, pp. 192–205, 2007.

[11] C. Poulin, N. A. Stanton, D. Cebon, and W. Epple, "Responses to autonomous vehicles," *Ingenia*, no. 62, pp. 8–11, 2015.

[12] National Highway Traffic Safety Administration, "Preliminary statement of policy concerning automated vehicles," 2013.

[13] A. Eriksson and N. A. Stanton, "Take-over time in highly automated vehicles: Non-critical transitions to and from manual control," *Human Factors*, vol. 59, no. 4, pp. 689–705, 2017.

[14] M. Körber, C. Gold, D. Lechner, and K. Bengler, "The influence of age on the take-over of vehicle control in highly automated driving," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 9, pp. 19–32, 2016.

[15] N. Merat, A. H. Jamson, F. F. C. H. Lai, M. Daly, and O. M. J. Carsten, "Transition to manual: Driver behaviour when resuming control from a highly automated vehicle," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 26, pp. 1–9, 2014.

[16] W. Payre, J. Cestac, and P. Delhomme, "Fully automated driving: Impact of trust and practice on manual control recovery," *Human Factors, J. Human Factors Ergonom. Soc.*, vol. 58, no. 2, pp. 229–241, 2016.

[17] C. D. Wickens, "Attention to safety and the psychology of surprise," in *Proc. Symp. Aviation Psychol.*, Columbus, OH, USA, 2001.

[18] A. Eriksson, V. A. Banks, and N. A. Stanton, "Transition to manual: Comparing simulator with on-road control transitions," *Accident Anal. Prevention*, vol. 102C, pp. 227–234, 2017.

[19] S. M. Petermeijer, D. A. Abbink, M. Mulder, and J. C. De Winter, "The effect of haptic support systems on driver performance: A literature survey," *IEEE Trans. Haptics*, vol. 8, no. 4, pp. 467–479, Oct.–Dec. 2015.

[20] K. Zeeb, A. Buchner, and M. Schrauf, "What determines the take-over time? An integrated model approach of driver take-over after automated driving," *Accident Anal. Prevention*, vol. 78, pp. 212–221, 2015.

[21] P. Kerschbaum, L. Lorenz, and K. Bengler, "A transforming steering wheel for highly automated cars," in *Proc. IEEE Intell. Veh. Symp.*, 2015, pp. 1287–1292.

[22] J. A. Michon, "A critical view of driver behavior models: What do we know, what should we do?" in *Human Behavior and Traffic Safety*, L. Evans and R. C. Schwing, Eds. New York, NY, USA: Plenum, 1985.

[23] C. Gold, D. Damböck, L. Lorenz, and K. Bengler, "'Take over!' How long does it take to get the driver back into the loop?" *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, vol. 57, no. 1, pp. 1938–1942, 2013.

[24] J. Radlmayr, C. Gold, L. Lorenz, M. Farid, and K. Bengler, "How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving," *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, vol. 58, no. 1, pp. 2063–2067, 2014.

[25] C. Cabrall *et al.*, "A survey of human machine interfaces in automated driving," Manuscript in preparation, 2017.

[26] S. M. Petermeijer, P. Hornberger, I. Ganotis, J. C. de Winter, and K. J. Bengler, "The design of a vibrotactile seat for conveying take-over requests in automated driving," in *Proc. Int. Conf. Appl. Human Factors Ergonom.*, 2017, pp. 618–630.

[27] W. Narzt *et al.*, "Augmented reality navigation systems," *Universal Access Inf. Soc.*, vol. 4, no. 3, pp. 177–187, 2006.

[28] M. Sivak, "The information that drivers use: Is it indeed 90% visual?" *Perception*, vol. 25, no. 9, pp. 1081–1089, 1996.

[29] S. Petermeijer, P. Bazilinskyy, K. Bengler, and J. de Winter, "Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop," *Appl. Ergonom.*, vol. 62, pp. 204–215, 2017.

[30] K. Suzuki and H. Jansson, "An analysis of driver's steering behaviour during auditory or haptic warnings for the designing of lane departure warning system," *JSAE Rev.*, vol. 24, no. 1, pp. 65–70, 2003.

[31] J. Navarro, F. Mars, J.-M. Hoc, R. Boisliveau, and F. Vienne, "Evaluation of human-machine cooperation applied to lateral control in car driving," in *Proc. 16th World Congr. Int. Ergonom. Soc.*, 2006, pp. 4957–4962.

[32] A. Gallace and C. Spence, *In Touch With the Future: The Sense of Touch From Cognitive Neuroscience to Virtual Reality*. London, U.K.: Oxford Univ. Press, 2014.

[33] F. Meng and C. Spence, "Tactile warning signals for in-vehicle systems," *Accident Anal. Prevention*, vol. 75, pp. 333–46, Feb. 2015.

[34] M. A. Nees and B. N. Walker, "Auditory displays for in-vehicle technologies," *Rev. Human Factors Ergonom.*, vol. 7, no. 1, pp. 58–99, 2011.

[35] D. C. Foyle, R. S. McCann, and S. G. Shelden, "Attentional issues with superimposed symbology: Formats for scene-linked displays," in *Proc. 8th Int. Symp. Aviation Psychol.*, vol. 98, pp. 98–103, 1995.

[36] N. A. Stanton and J. Edworthy, *Human Factors in Auditory Warnings*. Farnham, U.K.: Ashgate Publ., 1999.

[37] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Trans. Syst. Man, Cybern. A, Syst. Humans*, vol. 30, no. 3, pp. 286–297, May 2000.

[38] N. A. Stanton, A. Dunoyer, and A. Leatherland, "Detection of new in-path targets by drivers using stop & go adaptive cruise control," *Appl. Ergonom*, vol. 42, no. 4, pp. 592–601, May 2011.

[39] M. Zimmermann, S. Bauer, N. Lütteken, I. M. Rothkirch, and K. J. Bengler, "Acting together by mutual control: Evaluation of a multimodal interaction concept for cooperative driving," in *Proc. IEEE. Int. Conf. Collaboration Technol. Syst.*, 2014, pp. 227–235

[40] E. Coelingh, A. Eidehall, and M. Bengtsson, "Collision warning with full auto brake and pedestrian detection - A practical example of automatic emergency braking," in *Proc. IEEE 13th Int. Conf. Intell. Transp. Syst.*, 2010, pp. 155–160.

[41] E. Coelingh, H. Lind, W. Birk, M. Distner, and D. Wetterberg, "Collision warning with auto brake," in *Proc. FISITA World Automot. Congr.*, Oct. 22–27, 2006, Paper F2006V130.

[42] V. A. Banks, N. A. Stanton, and C. Harvey, "What the drivers do and do not tell you: Using verbal protocol analysis to investigate driver behaviour in emergency situations," *Ergonomics*, vol. 57, no. 3, pp. 332–342, 2014.

[43] R. C. McIlroy, N. A. Stanton, L. Godwin, and A. P. Wood, "Encouraging eco-driving with visual, auditory, and vibrotactile stimuli," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 5, pp. 66–672, Oct. 2016.

[44] B. Israel, *Potenziale Eines kontaktanalogen Head-up-Displays Für Den Serieneinsatz*. Göttingen, Germany: Cuvillier Verlag, 2012.

[45] L. Lorenz, P. Kerschbaum, and J. Schumann, "Designing take over scenarios for automated driving: How does augmented reality support the driver to get back into the loop?" in *Proc. Human Factors Ergonom. Soc.*, 2014, pp. 1681–1685.

[46] R. Parasuraman, R. Molloy, and L. I. Singh, "Performance consequences of automation-induced 'complacency'," *Int. J. Aviation Psychol.*, vol. 3, pp. 1–23, 1993.

[47] E. L. Wiener and R. E. Curry, "Flight-deck automation: Promises and problems," *Ergonomics*, vol. 23, no. 10, pp. 995–1011, 1980.

[48] A. Haslbeck and H.-J. Hoermann, "Flying the needles flight deck automation erodes fine-motor flying skills among airline pilots," *Human Factors, J. Human Factors Ergonom. Soc.*, vol. 58, no. 4, pp. 533–545, 2016.

[49] N. A. Stanton, M. S. Young, and B. McCaulder, "Drive-by-wire: The case of mental workload and the ability of the driver to reclaim control," *Saf. Sci.*, vol. 27, no. 2/3, pp. 149–159, 1997.

[50] L. J. Skitka, K. L. Mosier, and M. Burdick, "Does automation bias decision-making?" *Int. J. Human-Comput. Stud.*, vol. 51, no. 5, pp. 991–1006, 1999.

[51] K. L. Mosier and L. J. Skitka, "Human decision makers and automated decision aids: Made for each other," in *Proc. Automat. Human Perform., Theory Appl.*, 1996, pp. 201–220.

[52] S. N. Weingart, R. McL Wilson, R. W. Gibberd, and B. Harrison, "Epidemiology of medical error," *Western J. Med.*, vol. 172, no. 6, pp. 390–393, 2000.

[53] A. Eriksson and N. A. Stanton, "Driving performance after self-regulated control transitions in highly automated vehicles," *Human Factors*, vol. 59 no. 8, pp. 1233–1248, 2017.

[54] E. Hollnagel and D. D. Woods, *Joint Cognitive Systems Foundations of Cognitive Systems Engineering*. Boca Raton, FL, USA: CRC Press, 2005.

[55] P. A. Desmond, P. A. Hancock, and J. L. Monette, "Fatigue and automation-induced impairments in simulated driving performance," *Human Perform., User Inf. Highway Des.*, vol. 1628, pp. 8–14, 1998.

[56] F. Naujoks, C. Mai, and A. Nekum, "The effect of urgency of take-over requests during highly automated driving under distracted conditions," *Adv. Human Aspects Transp.*, vol. 7, no. pt. I, 2014, Art. no. 431.

[57] D. Damböck, K. Bengler, M. Farid, and L. Tönert, "Übernahmezeiten beim hochautomatisierten Fahren," *Tagung Fahrerassistenz. München*, vol. 15, 2012.

[58] B. Mok *et al.*, "Emergency, automation off: Unstructured transition timing for distracted drivers of automated vehicles," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, 2015, pp. 2458–2464.

[59] D. A. Norman, *Memory and Attention*. Hoboken, NJ, USA: Wiley, 1976.

[60] J. Engström, G. Markkula, T. Victor, and N. Merat, "Effects of cognitive load on driving performance: The cognitive control hypothesis," *Human Factors*, vol. 59, pp. 737–764, 2017.

[61] F. Naujoks, K. Wiedemann, and N. Schomig, "The importance of interruption management for usefulness and acceptance of automated driving," in *Proc. 9th Int. Conf. Automot. User Interfaces Interactive Veh. Appl.*, Oldenburg, Germany, 2017, pp. 254–263.

[62] N. Lütteken, M. Zimmermann, and K. Bengler, "Using gamification to motivate human cooperation in a lane-change scenario," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst.*, 2016, pp. 899–906.

[63] D. Damböck, T. Weißgerber, M. Kienle, and K. Bengler, "Evaluation of a contact analog head-up display for highly automated driving," in *Proc. 4th Int. Conf. Appl. Human Factors Ergonom.*, San Francisco, CA, USA, pp. 6011–6020, 2012.

[64] S. Langlois and B. Soualmi, "Augmented reality versus classical HUD to take over from automated driving: An aid to smooth reactions and to anticipate maneuvers," in *Proc. IEEE 19th Int. Intell. Transp. Syst.*, 2016, pp. 1571–1578.

[65] J. Lee, M. Muñoz, L. Fridman, T. Victor, B. Reimer, and B. Mehler, "Investigating the correspondence between driver head position and glance location," *PeerJ Comput. Sci.*, vol. 4, 2018, Art. no. e146.

[66] J. C. Byers, A. Bittner, and S. Hill, "Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary?," in *Proc. Adv. Ind. Ergonom. Saf. I*, 1989, pp. 481–485.

[67] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, 1988.

[68] J. D. Van Der Laan, A. Heino, and D. De Waard, "A simple procedure for the assessment of acceptance of advanced transport telematics," *Transp. Res. Part C, Emerg. Technol.*, vol. 5, no. 1, pp. 1–10, 1997.

[69] S. Petermeijer, S. Cieler, and J. de Winter, "Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat," *Accident Anal. Prevention*, vol. 99, pp. 218–227, 2017.

[70] C. Tanikawa *et al.*, "A genome-wide association study identifies two susceptibility loci for duodenal ulcer in the Japanese population," *Nature Genetics*, vol. 44, no. 4, pp. 430–434, 2012.

[71] G. Gibson, "Hints of hidden heritability in GWAS," *Nature Genetics*, vol. 42, no. 7, pp. 558–560, 2010.

[72] M. S. Wogalter, "Technology will revolutionize warnings," in *Proc. Solutions Saf. Technol. Symp.*, Scottsdale, AZ, USA, 2006.

[73] C. Baber, "Psychological aspects of conventional in-car warning devices," in *Human Factors in Alarm Design*, N. Stanton, Ed. Chicago, IL,USA: Taylor & Francis, 1994, pp. 193–205.

[74] G. Markkula, J. Engström, J. Lodin, J. Bärgman, and T. Victor, "A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies," *Accident Anal. Prevention*, vol. 95, pp. 209–226, 2016.

[75] M. Yeh, C. D. Wickens, and F. J. Seagull, "Target cuing in visual search: The effects of conformality and display location on the allocation of visual attention," *Human Factors*, vol. 41, no. 4, pp. 524–542, 1999.

[76] D. Sperber and D. Wilson, *Relevance: Communication and Cognition*. Oxford, U.K.: Blackwell, 1986.

[77] H. P. Grice, "Logic and Conversation," in *Speech Acts*, P. Cole and J. L. Morgan, Eds. New York, NY, USA: Academic Press, 1975, pp. 41–58.

[78] J. Werneke and M. Vollrath, "How to present collision warnings at intersections?—A comparison of different approaches," *Accident Anal. Prevention*, vol. 52, pp. 91–99, 2013.

[79] F. Schwarz and W. Fastenmeier, "Augmented reality warnings in vehicles: Effects of modality and specificity on effectiveness," *Accident Anal. Prevention*, vol. 101, pp. 55–66, 2017.

Markus Zimmermann received the Bachelor's degree from the Ludwig-Maximilians-Universität München, Munich, Germany, with a focus on human–computer interaction. He received the Diploma from the Center for Digital Technology and Management, Munich, Germany in 2011, both in media informatics.

Since then, he has been with the European research project "Designing Dynamic Distributed Cooperative Human-Machine Systems (D3CoS)" with the Institute of Ergonomics, Technische Universität München, Munich, Germany.

Joost C. F. de Winter received the M.Sc. degree in aerospace engineering and the Ph.D. degree (*cum laude*) in mechanical engineering from the Delft University of Technology, Delft, The Netherlands, in 2004 and 2009, respectively.

He is currently an Associate Professor with the Mechanical Engineering Department, Delft University of Technology. His research interests include human factors and statistical modeling, including the study of individual differences, driver behavior modeling, multivariate statistics, and research methodology.

Alexander Eriksson received the M.Sc. degree in cognitive science from Linköping University, Linköping, Sweden, in 2014, and the Ph.D. degree in human factors from the University of Southampton, Southampton, U.K., as part of the HF-Auto project, in 2017.

He is currently a Chief Research Engineer with the Institute of Transport Economics, Oslo, Norway, and is the Competence Area Leader in driving simulator applications with SAFER Vehicle and Traffic Safety Centre at Chalmers, Gothenburg, Sweden.

Klaus J. Bengler received the Bachelor's degree from the University of Regensburg, Regensburg, Germany, in 1991, and the Ph.D. degree from the University of Regensburg, in cooperation with BMW from the Institute of Psychology (Prof. M. Zimmermann), in 1995, both in psychology.

After receiving the Ph.D. degree, he was active on topics of software ergonomics and evaluation of human–machine interfaces. Since May 2009, he has been the Leader with the Institute of Ergonomics, Technische Universität München, Munich, Germany, which is active in research areas like digital human modeling, human–robot cooperation, driver assistance HMI, and human reliability.

Sebastiaan M. Petermeijer received the M.Sc. degree (*cum laude*) in mechanical engineering from the Delft University of Technology, Delft, The Netherlands, in April 2014, and the Ph.D. degree in human factors from the Technische Universität München, Munich, Germany, in April 2017.

He is currently a Research Fellow with the Delft Haptics lab, Delft University of Technology, focusing on haptic human–machine interaction in a highly automated vehicle.

Neville A. Stanton received the Bachelor's degree in occupational psychology from the University of Hull, Hull, U.K., in 1986, the Master's degree in applied psychology from Aston University, Birmingham, U.K., in 1989, the Doctorate degree in human factors from Aston University, Birmingham, U.K., in 1993, and the Higher Doctorate degree in human factors engineering from the University of Southampton, Southampton, U.K., in 2014.

He is currently a Professor of human factors engineering with the Faculty of Engineering and the Environment, University of Southampton.