Accident Analysis and Prevention 181 (2023) 106927

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Repeated conditionally automated driving on the road: How do drivers leave the loop over time?

J. Dillmann [a,b,*], R.J.R. Den Hartigh [a], C.M. Kurpiers [b], F.K. Raisch [b], N. Kadrileev [c], R.F.A. Cox [a], D. De Waard [a]

[a] Department of Psychology, University of Groningen, Groningen, the Netherlands
[b] BMW Group Research and Development, Munich, Germany
[c] Robert Bosch GmbH, Leonberg, Germany

ARTICLE INFO

Keywords:
Automated driving
Repeated measures design
Realistic driving paradigm
Conditionally automated driving
Wizard of Oz

ABSTRACT

The goal of this on the road driving study was to investigate how drivers adapt their behavior when driving with conditional vehicle automation (SAE L3) on different occasions. Specifically, we focused on changes in how fast drivers took over control from automation and how their gaze off the road changed over time. On each of three consecutive days, 21 participants drove for 50 min, in a conditionally automated vehicle (Wizard of Oz methodology), on a typical German commuting highway. Over these rides the take-over behavior and gaze behavior were analyzed. The data show that drivers' reactions to non-critical, system initiated, take-overs took about 5.62 s and did not change within individual rides, but on average became 0.72 s faster over the three rides. After these self-paced take-over requests a final urgent take-over request was issued at the end of the third ride. In this scenario participants took over rapidly with an average of 5.28 s. This urgent take-over time was not found to be different from the self-paced take-over requests in the same ride. Regarding gaze behavior, participants' overall longest glance off the road and the percentage of time looked off the road increased within each ride, but stayed stable over the three rides. Taken together, our results suggest that drivers regularly leave the loop by gazing off the road, but multiple exposures to take-over situations in automated driving allow drivers to come back into loop faster.

1. Introduction

The incremental introduction of automated driving is likely to stay relevant for the coming decades (Fagnant & Kockelman, 2015). Whereas advanced driving assistance systems currently share control with the driver (Level 2) a next step towards traded control is conditionally automated driving (Level 3, SAE International, 2018). This step has been regulated, enabling manufacturers to certify and introduce automated systems on the road (UNECE, 2020). At this level of automation the vehicle takes over the entire driving task but the driver but must be capable of taking back control rapidly and adequately if necessary (SAE International, 2018). However, research has shown that the quality of the drivers' take-over reaction to hazard situations is lower if the driver had been driving with automation before the incident (De Waard et al., 1999; Gold et al., 2013; Louw et al., 2017). Taking back control from automation can be adversely influenced if the driver is out of the loop (Merat et al., 2018). Recently, research has investigated being out of the loop from a perception–action theory perspective (Dillmann et al., 2021a,b; Mole et al., 2019). From this perspective, drivers' being in the loop is seen as being in an active cycle of perception and action (Fajen & Devaney, 2006). In braking behavior this loop would consist of (a) drivers attuning to perceptual cues such as optical looming, and subsequently (b) calibrating their braking actions to the changing visual cues (Fajen, 2005). If drivers leave the loop during automated driving, they may need time to reattune to relevant visual information and re-calibrate their actions (Brand et al., 2017; Mole et al., 2019; Russell et al., 2016). Indeed, leaving the loop during automated driving has been associated with reduced glances at relevant sources of perceptual information (Dillmann et al., 2021a; Schnebelen et al., 2020), and deteriorating motor-perceptual calibration in driving actions when taking back control (Dillmann et al., 2021b).

To date, over 200 simulator studies have investigated how drivers come back into the loop and take back control from the conditional vehicle (de Winter et al., 2021; for a review see Zhang et al., 2019).

* Corresponding author.
E-mail address: j.dillmann@rug.nl (J. Dillmann).

https://doi.org/10.1016/j.aap.2022.106927
Received 3 June 2022; Received in revised form 7 October 2022; Accepted 8 December 2022
Available online 28 December 2022

However, de Winter and colleagues (2021) argue that such simulator studies have focused on conveniently measurable urgent take-overs, and have left self-paced interactions with automation in realistic driving situations underrepresented in the literature. Therefore, different research is necessary to gain a deeper understanding of driver behavior in self-paced take-overs in *realistic driving studies*.

Recently, studies using the Wizard of Oz methodology (Bengler et al., 2020) to study non-critical take-overs are starting to fill this gap: Rydström and colleagues (2022) found that participants' take-over reactions to non-critical take-overs on the road, became faster during the ride over approximately 45 min of automated driving, The authors argue that this change of behavior could be a learning effect similar to those that have been reported in repeated interaction with automation (Forster et al., 2019; Winkler et al., 2018). However, changes in behavior when interacting with automation have been shown to develop over several rides (Beggiato et al., 2015; Large et al., 2018). For example, on-the-road research of fifteen drivers commuting with adaptive-cruise-control (Driver Assistance System, Level 1) showed that it took 185 km or 3.5 h of driving for learning- or behavioral adaptation effects to flatten (Beggiato et al., 2015). Considering that behavioral changes occurred over several rides with driver assistance systems, the literature is lacking research with respect to whether and how behavioral changes occur over several rides in conditional vehicle automation in realistic driving conditions. This gap in the literature pertains to both self-paced take-overs which may occur frequently for standard situations such as leaving the highway (Eriksson & Stanton, 2017), as well as urgent take-overs which are likely to occur infrequently (SAE International, 2018). Therefore, new studies are needed to understand how drivers leave and come back into the perception–action loop over several rides.

A further behavioral change that could be expected to change over time is gaze behavior. In manual driving the length of gaze off the road has been shown to be a key predictor of accident involvement (Seppelt et al., 2017). In the following we will explore changes in gaze during automated driving from the perspective of perception–action theory. According to perception–action theory, perception is considered the attunement to visual information that is relevant for the current action, such as the driving task (Mathieu et al., 2017; Wilson, 2004). As tasks change, so does the visual information we search for (Sullivan et al., 2012). Experimental research on conditionally automated driving has shown that once drivers delegate the driving task to automation, they change visual strategies in a manner that includes more glances off the road (Schnebelen et al., 2020). This has been confirmed in on the road studies showing that drivers may glance at the non-driving-related task for more than 80 % of the time (Klingegard et al., 2020). On the other hand it has been demonstrated that the extent to which they keep their gaze on the road is influenced by their continued involvement in the driving task, for instance by remaining responsible for manual lane-change maneuvers (Dillmann et al., 2021a,b). Although, these studies indicate how gaze behavior can change once control is delegated to the automated system, these studies did not explore how gaze behavior changes across several rides. This represents a gap in the literature as gaze behavior has been shown to change as a function of the duration of automated driving (Feldhütter et al., 2017). Specifically, Feldhütter and colleagues (2017) found that as the duration of conditionally automated driving increases, drivers' gaze is increasingly directed off the road. However, although gaze behavior has been linked with driving safety, there is a gap in the literature as to how gaze behavior changes over time when gaining experience with driving an automated system (Mole et al., 2019). Across several rides it could be expected that both the proportion of time participants look off the road (Schnebelen et al., 2020) and the length of the glances off the road (Seppelt et al., 2017) would increase.

Taken together, previous research has suggested that take-over times and glance behavior change as control is delegated to a conditionally automated vehicle. However, it remains unknown how take-over and gaze behavior will develop as drivers leave the loop in on the road driving situations over several rides. In addition, it is not clear how drivers will react to urgent take-overs occurring after several rides. This knowledge is necessary to design systems that are safe when used on the road regularly. To fill this gap in the literature we used a Wizard of Oz Methodology to allow for L3 driving in real traffic, and measured participants on a commuting highway for three consecutive days.

2. Method

2.1. Participants

Twenty-one participants (*mean age* = 43.4 years; *SD* = 10.0 years, min = 28 years, max = 65 years, eight female) with an average yearly *driving experience* of 39,500 km (*SD* = 28,000) took part in our study. A wide spread in age was included and selection criteria consisted of using a vehicle on most or all days per week and owning a smart-phone. They were recruited and compensated by a local recruiting agency and could stop their participation in the study at any time without financial disadvantages. Upon arriving the participants filled out and informed consent and the study complied with the American Psychological Association Code of Ethics.

2.2. Apparatus

2.2.1. The conditionally automated driving Wizard of Oz vehicle

In order to maximize safety, the study was conducted in a BMW X5 xDrive50i in which conditionally automated driving was emulated with Wizard-of-Oz methodology (Dahlbäck et al., 1993, Fig. 1). The safety concept had been approved for road testing by the local authorities (Gold et al., 2017) and the study complied with the tenets of the Declaration of Helsinki. The experimental rides were completed on the German highway A92 on a return ride between the Exit Airport Munich and Exit Essenbach, with approximately 94 km and 54 min per ride. In accordance with the highway traffic flow, the sessions for each participant were recurrently scheduled during off-peak hours (8:00, 11:00, 14:00), for the initial travel direction on the highway to ensure consistent low traffic density. Both vehicle Controller-Area-Network data and video recordings of the driver were recorded on a Vigem recorder and handled in accordance with the European General Data Protection Regulation. When automation was activated, the vehicle drove a maximum of 130 km/h with a passive driving style focused on maintaining distance to other vehicles and a conservative lane change strategy. The driver in the back (wizard driver) could take over control over the vehicle and emulate automated driving. The participant was only told that the vehicle was driving automatically and was not aware of the presence of the wizard driver.

2.2.2. Human Machine interface (HMI) and activation of automation

The vehicle generally offered a standard conditionally automated driving vehicle interface (Forster et al., 2016; Manca et al., 2015) with an indication of automation availability, active automation, and driver or vehicle-initiated take-over requests. Specific to the Wizard of Oz set-up, drivers' turning on and off the automation corresponded to vehicle control being traded with the invisible wizard driver. For the transfer of vehicle control, drivers needed to perform two distinct actions (1) having both hands on the steering wheel and (2) pressing a button on the left hand side of the steering wheel. Whereas activation of automation was available when indicated on the display, deactivation was possible at all times. The HMI was developed specifically for research in the Wizard of Oz setup and also contained two 5 cm LED strips, on both sides of the steering wheel. These LED Strips shone in blue light when automation was active to increase participants' mode awareness (Kurpiers et al., 2020). Together with an auditory warning sound, a spoken audio message, and a display in the speedometer, these blinking LED strips ensured the salience of the take-over requests (see Fig. 2). Specifically, the takeover cascade started with (1) blue illumination on LED Strips near the steering wheel and mild acoustic warning signal and the text

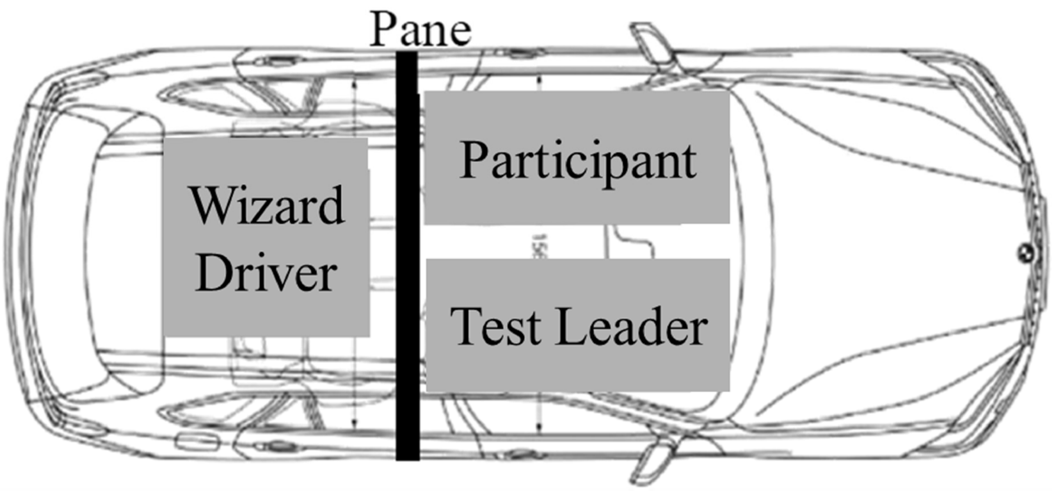


Fig. 1. Schematic presentation of the BMW Wizard of Oz X5 adapted from Gold et al. (2017). The driver in the back (wizard driver) could take over control over the vehicle and emulate automated driving. The participant was only told that the vehicle was driving automatically.

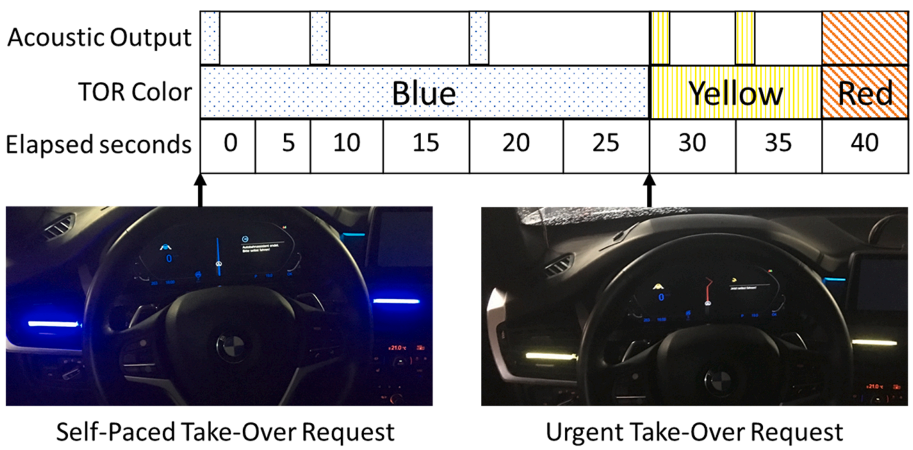


Fig. 2. Schematic presentation of the take-over human–machine interface. TOR = Take Over Request.

“Highway Assistant is ending, please drive yourself” both spoken and displayed in the speedometer in German (see Fig. 2 Blue). At this stage, both acoustic signals were repeated every 10 s. In the speedometer a countdown timer starting at 45 s was visible. After 30 s this level escalated to the next level where a yellow illumination was accompanied by the text “Drive yourself now” both spoken and displayed in the speedometer in German (see Fig. 2 Yellow). At this level of escalation, a countdown with 15 s was shown. If unanswered this moved to a final level with red illumination accompanied by a sharp audio signal and the text “immediately drive yourself”. A countdown timer with 5 s was shown at this stage. If they were to have not reclaimed control after this stage the test leader would ask them to take over. This did not occur in the current experiment.

Once participants had taken over the standard driving HMI appeared, the LED Strips turned off and participants heard the message “Highway Assistant deactivated. You have taken back vehicle control”. The take-over starting at 45 s will be referred to as self-paced take-over whereas the take-over starting at 15 s will be referred to as urgent take-over (cf. Fig. 2).

2.2.3. Non-driving-related-task

In order to allow for non-driving related tasks (NDRT) that were natural and yet standardized we allowed participants to bring their own smartphone. We required that the smartphone was placed in smartphone holder near the center console (see . 4a). This allowed participants to engage in a natural NDRT (Naujoks et al., 2018) while the device would not interfere during the take-over (Wintersberger et al., 2021), which could have been a safety issue on the road. Participants were instructed that once the automated system took control, they were free to use their smartphone as they wished. The only constraint was that the phone had to remain mounted at the middle console.

2.3. Procedure and design

Upon their arrival participants’ driver’s licenses were inspected, they filled in a questionnaire with regard to demographic data, and read information we had compiled on conditionally automated driving. In the first vehicle contact only, participants performed the activation and deactivation of automation while standing still. They were then exposed to the self-paced and urgent take-over cascades and asked to take-over control. Afterwards a practice session was conducted on a secondary road with the vehicle moving at a walking pace (see Fig. 1). Participants then drove up to the A92 where they drove three segments of automated driving for a total duration of approximately 52 Minutes (see Fig. 3). The route was in total 132 km long with 18 km of manual driving, followed by 4 km of automated driving, 4 km of driving through a road works area, 40 km of automated driving, 40 km of automated driving, 26 km of manual driving. Depending on unforeseen circumstances (e.g., construction work) participants could be asked to take back control intermittently during sequences of automated driving.

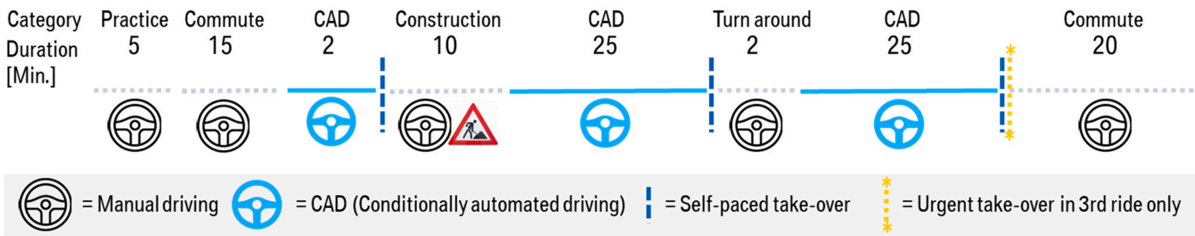


Fig. 3. Schematic of the driving actions per ride.

At the end of the final automated driving interval in ride three, the wizard safety driver was instructed to issue an urgent take-over. In order to increase comparability of these urgent take-overs across participants the take-overs were issued when the vehicle was approaching a truck or large vehicle on the ego lane. For this scenario, the professional safety driver was instructed to maintain a safe distance that did not entail more risk than everyday driving situations.

2.4. Measures

2.4.1. Take-over times

For the self-paced take-over, we analyzed the time from the start of the blue take-over cascade take-over signal to the hands on the wheel and button press combination necessary to deactivate the automated system. This occurred at the end of each automated driving interval and therefore three times in rides one and two and two times in ride three (see Fig. 3).

For the urgent take-over, we analyzed the time from the yellow take-over signal to the hands on the wheel and button press combination necessary to deactivate the automated system. This occurred at the end of ride three (see Fig. 3).

2.4.2. Percentage of gaze and maximum duration of gaze at the NDRT

In order to unobtrusively measure gaze behavior, we recorded the participants' gaze on video (see Fig. 4). Through postprocessing with a neural network (Dari et al., 2020) we were able to calculate a classification of gaze into 8 categories: left shoulder, left mirror, road ahead, NDRT, right mirror, and right shoulder, speedometer, and rear view mirror.

We excerpted the intervals in which automation was activated, and calculated the percentage of time participants gaze was oriented at the non-driving related task. Next, we calculated the longest continuous duration that participants looked off the road at the NDRT.

2.5. Statistical analysis

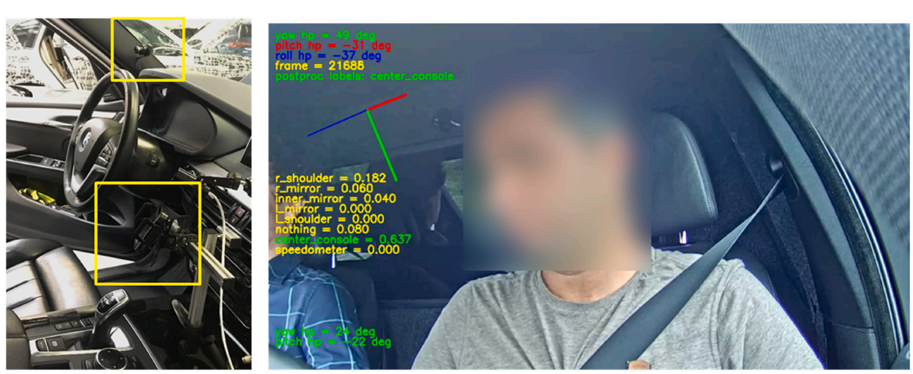
According to the experimental design 168 self-paced take-overs and 21 urgent take-overs took place. Over all 63 rides, traffic circumstances such as lawn mowing crews or special transport trucks required 30 additional take-overs to be issued. In order to gain an understanding of the 168 comparable take-overs, these unsystematically occurring take-overs were disregarded. Furthermore, two participants were excluded as they did not engage in normal driving behavior but intentionally experimented with the automated system. This led to the exclusion of 16 self-paced and two urgent take-overs. Additionally technical measurement issues occurred in three further self-paced take-overs and one urgent take-over leading to their exclusion. Therefore 149 self-paced take-overs and 19 urgent take-overs were included.

Given these missing data and repeated measures design we analyzed our data with linear mixed model with the lme4 (V. 1.1–27) package in R Statistics (V 3.5.1). This analysis is similar to a repeated measures ANOVA but is suitable for datasets with missing data (Cohen, 1988). In addition to statistical significance tests we calculated partial eta-squared effect sizes to convey the practical significance (Cohen, 1988; Sullivan & Feinn, 2012). Theses analysis were followed up with planned comparisons (Cohen et al., 2014), comparing the first with second and third interval, and the first with the second and third ride.

3. Results

3.1. Take-Over Behavior: Self-Paced Take-Overs

Table 1 displays the descriptive results of take-overs within and over rides. Regarding the time to take-over within rides, a linear mixed model controlling for participant and ride did not did not reveal clear differences $F(2,90) = 0.24$, $p = 0.78$, $\eta_p^2 < 0.01$. A linear mixed model with participant as random factor showed that take-over time decreased over the first ride one ($M = 5.97$, $SD = 1.37$), to ride two ($M = 5.53$, $SD = 1.70$), and ride three with a medium-to-large effect size ($M = 5.26$, $SD =$



a. b.

Fig. 4. Schematic presentation of (a.) vehicle interior and (b.) a frame of the eye-tracking analysis. Note. In the lower box in picture a. the mobile phone holder for the non-driving-related-task is visible. In the upper part of picture a. the camera used for gaze observation can be seen. The resulting picture and camera output and gaze analysis is displayed in figure b.

Table 1
Descriptive Statistics on the time to take-over in seconds.

	All self-Paced	Interval 1	Interval 2	Interval 3	Ride 1	Ride 2	Ride 3	Urgent Final
N	149	57	57	35	55	56	38	18
Mean	5.62	5.65	5.58	5.60	5.97	5.53	5.26	5.28
SD	1.55	1.44	1.58	1.66	1.37	1.70	1.49	2.62
Minimum	2.21	2.21	2.39	2.48	2.76	2.21	2.27	2.50
Maximum	9.88	9.88	9.88	9.88	9.88	9.88	9.88	13.11
5th percentile	2.78	3.39	2.79	2.76	4.06	2.74	2.49	2.86
25th percentile	3.84	4.78	4.45	4.52	5.18	4.53	4.46	3.55
Median	5.62	5.62	5.44	5.65	5.92	5.35	5.26	4.72
75th percentile	6.46	6.46	6.46	6.39	6.70	6.49	6.15	5.86
95th percentile	8.29	8.01	8.14	8.15	8.44	8.20	7.34	9.49

Note. Interval = the conditionally automated driving (CAD) interval across all rides, Ride = the CAD ride overall all intervals within the ride, Urgent Final = final takeover in Ride 3.

1.49, $F(2,93) = 4.07$, $p = 0.019$ $\eta_p^2 = 0.06$, see Fig. 5). To further investigate when these changes occurred, the planned contrasts showed a small effect between ride one and ride two $b = -0.44$, $t(128) = -1.92$, $p = 0.055$, $\eta_p^2 = 0.03$. However the data showed that participants took over significantly faster in ride three than in ride one with a medium effect size, $b = -0.70$, $t(-2.75) = 2.75$, $p = 0.006$, $\eta_p^2 = 0.06$.

3.2. *Take-Over Behavior: Final urgent take-over requests in ride 3*

To understand the impact of the final urgent take-over on participants take-over behavior, we calculated a linear mixed model with participant as random factor, comparing two self-paced take-overs with the final urgent takeover within ride 3. The test revealed very marginal differences in the time to take over between the first $M = 5.11$, $SD = 1.21$, second $M = 5.31$, $SD = 1.77$ and the third take-over $M = 5.28$, $SD = 2.62$, $F(2,34) = 0.07$, $p = 0.928$, $\eta_p^2 < 0.01$.

3.3. *Gaze Behavior: Maximum duration and percentage of time with glances off the road*

To understand how the maximum duration and percentage of time with glances off the road changed, we first excerpted comparable time samples at the beginning of each interval, which excluded the take-over situations. Regarding the maximum duration of glances off the road, a linear mixed model with participant and ride as random factors showed a clear medium-to-large effect of the interval $F(2,110) = 5.14$, $p = 0.007$, $\eta_p^2 = 0.08$ (see Fig. 6a). Planned contrasts showed that the first interval $M = 14.82$, $SD = 21.11$, differed significantly to the second interval $M = 22.15$, $SD = 25.70$, $b = 7.32$, $t(110) = 2.98$, $p = 0.003$, $\eta_p^2 = 0.08$, and the third interval, $M = 19.25$, $SD = 23.64$, $b = 6.22$, $t(110) = 2.50$, $p = 0.013$, $\eta_p^2 = 0.05$.

Regarding the percentage of time with glances off the road, a linear mixed model with interval as fixed and participant and ride as random factors showed a significant difference between the intervals with a medium-to-large effect size $F(2,110) = 5.64$, $p = 0.004$, $\eta_p^2 = 0.09$ (see Fig. 6b). Planned contrast showed a medium-to-large difference between the first, $M = 37.58$, $SD = 35.79$, and the second interval, $M = 53.03$, $SD = 38.10$, $b = 15.45$, $t(110) = 3.35$, $p = 0.001$, $\eta_p^2 = 0.09$, whereas the difference between the first and the third interval, $M = 44.68$, $SD = 38.74$ showed a small-to-medium effect, $b = 8.23$, $t(110) = 2.073$, $p = 0.042$, $\eta_p^2 = 0.04$. When we tested for differences in gaze behavior over rides (see Table 2 for descriptive statistics) a mixed linear model with ride as fixed and participant as random factor found no significant effect of ride on the maximum duration $F(2,36) = 1.57$, $p = 0.22$, $\eta_p^2 = 0.08$ and percentage of time with glances at the NDRT duration $F(2,36) = 1.22$, $p = 0.30$, $\eta_p^2 = 0.06$. However, both showed medium to large effect sizes suggesting practical relevance (Sullivan & Feinn, 2012).

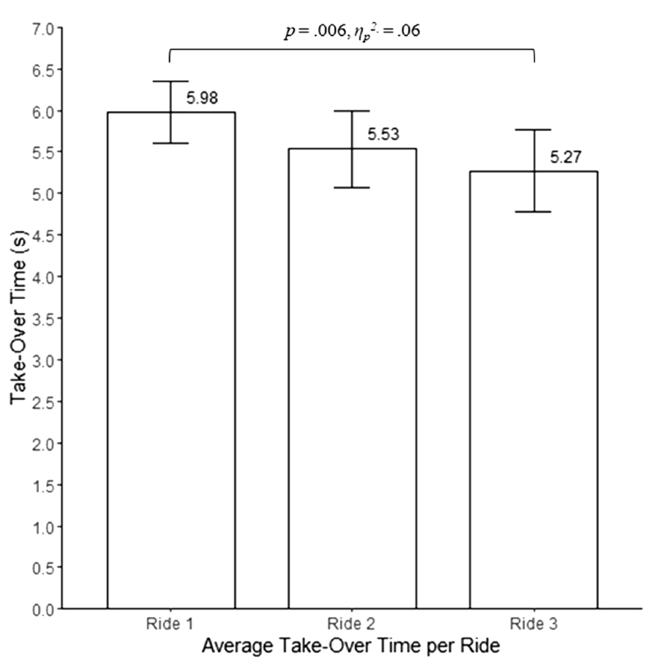


Fig. 5. Self-paced take-over times with 95% confidence intervals.

4. Discussion

The goal of the current study was to investigate how take-over and gaze behavior change over the course of three conditionally automated rides on the road. In line with research showing that changes occur when driving with ACC up to 3.5 h (Beggiato et al., 2015), the data showed that over three rides, self-paced take-overs became 0.72 s faster. It has been shown that learning processes take place as drivers interact with automation repeatedly (Forster et al., 2019). With regard to perception–action loop in driving research suggests that a variety of learning situations are necessary to improve motor-perceptual calibration in the loop (Fajen & Devaney, 2006). The real driving environment in the current study may have facilitated learning by offering both repeated and diverse learning opportunities to coming back into the motor-perceptual loop in take-over situations.

Overall, the non-critical average take-over time of 5.62 s we found, is between an on the road driving study reporting 2.70 and 5.50 s (Naujoks et al., 2019) and a similar study reporting an average of 8.70 s (Rydström et al., 2022). These results align well with the typical 6 s drivers needed to take back control from automation when engaged with a non-driving related task in a simulator study (Eriksson & Stanton, 2017). At the end of the third and final ride an urgent take-over request was issued. In this scenario participants took over within 5.28 s, on average. We did not find a significant difference in take-over time with previous self-paced take-overs in that ride. In other words, in spite of an

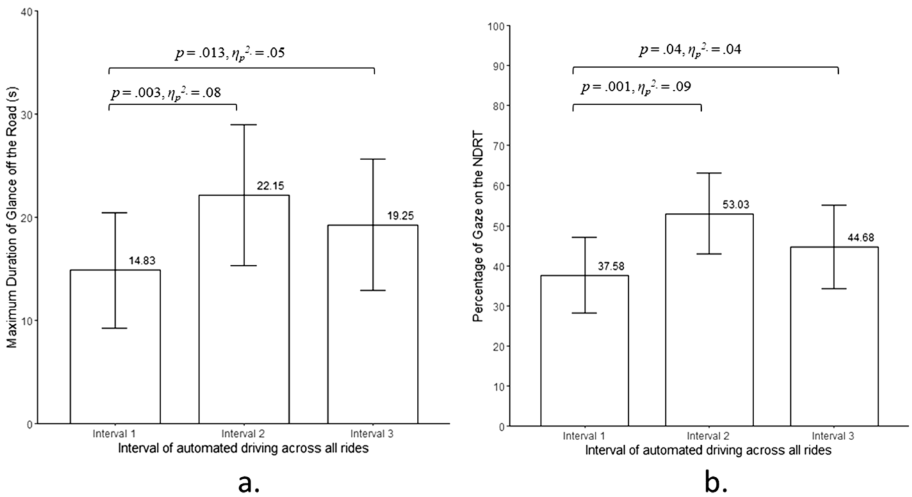


Fig. 6. Bar Plots across intervals with (a.) maximum duration of glance off the road and (b.) percentage of gaze on the NDRT with confidence intervals.

Table 2
Descriptive Statistics of gaze metrics over rides.

	Percentage NDRT (percent)				Maximum Glance Duration NDRT (s)			
	Overall	Ride 1	Ride 2	Ride 3	Overall	Ride 1	Ride 2	Ride 3
N	57	19	19	19	57	19	19	19
Mean	44.82	38.39	47.68	48.4	85.7	74.72	96.28	86.1
SD	32.37	37.56	32.25	27.28	59.48	59.47	61.98	58.23
Minimum	0.17	0.17	2.76	5.80	6.80	6.80	19.36	17.92
Maximum	96.33	96.33	92.57	85.35	247.24	247.24	219.12	243.00
5th percentile	2.58	1.43	4.82	7.1	17.82	10.33	22.96	27.14
25th percentile	10.63	4.28	19.97	27.28	42.21	27.68	51.88	48.64
Median	45.65	20.61	44.82	50.34	69.68	69.68	67.44	70.57
75th percentile	75.04	74.87	81.19	67.74	115.44	104.24	147.84	110.96
95th percentile	91.40	95.24	90.10	85.13	213.74	163.32	213.07	197.42

immediate escalation of the take-over, participants were not significantly faster in taking back control from the automated vehicle. A possible explanation from a perception–action perspective could be that visual information in the environment, such as braking vehicles looming before the automated vehicle (Markkula et al., 2016), may influence take-over time beyond what human–machine-interface design of the take-over request is showing. However, it should be noted that for safety reasons, the current study only deployed an urgent and a not a critical take-over. Thus, the results pertain only to the experimental HMI used in this study, and the take-over times found in the urgent scenario are in line with other on the road take-over times.

Regarding gaze behavior we were interested in whether participants showed signs of out-of-the-loop gaze behavior and how this develops over the three rides. In line with previous research the extent of gaze off the road increased within the duration of each ride (Feldhütter et al., 2017). The data suggest that within each ride, drivers both (a) look off the road for a higher percentage of the time and (b) look off the road for long continued time-periods. However, this effect was only found within the intervals of automated rides and not over the three rides. It should be noted that this was the case in the idealized automated system which drove 130 km/h and performed lane changes on a human driver level. As the automated system proficiently takes over full control, the visual information relevant for the driving task becomes entirely irrelevant to the drivers and they appear to focus their visual attention off the road (Sullivan et al., 2012).

The extent to which gaze is allocated off the road may be reduced if drivers remain somewhat involved in the driving task. If drivers remain responsible for lane changes, for instance, they maintain more gaze on the road during conditionally automated driving (Dillmann et al., 2021b). It is therefore unclear if the current results extend to more restricted automated systems, for example without automated lane changes and with restricted velocity (UN-ECE, 2020), as these systems may keep drivers more in the loop regarding gaze behavior.

Given the strong 0.72 s decrease of take-over times over the three rides and increased gaze off the road per ride, the data suggest that drivers leave the loop but become faster at coming back into the loop. The practical relevance of this 0.72 s faster reaction can be pointed out with a small calculation: with a vehicle speed of 130 km/h, the vehicle moves 36 m per second. A 0.72 s faster reaction time translates to the begin of a response 26 m before. In the case of an emergency brake requiring 85 m to come to a halt, the faster reaction time would allow drivers to stop at 30 % less distance. However, it should be noted that the 0.72 s show an average improvement, over three rides, across participants. Although this average is powerful in inferring a trend to the population of drivers, further research is necessary to understand which mechanisms shape the behavior of drivers that take longest to take-over (see Table 1). From a perception–action theory perspective we suggest that repeated exposure to different take-overs over the three rides improves perceptual learning and motor calibration in taking back control. More specifically, drivers become faster at coming back into the loop by more rapidly attuning to the relevant visual information for the take-over (Fajen & Devaney, 2006; Jacobs & Michaels, 2007) and more readily scaling their motor-perceptual actions to this information (Russell et al., 2016). Future research should investigate this possibility by extending dedicated research on motor-perceptual calibration in taking back control (Russell et al., 2016) into longitudinal designs including more time driven and more diverse take-over situations.

5. Limitations and Future directions

The repeated measures design in this study allows a first perspective into how real-life interaction with an automated vehicle may develop over rides, but it does not extend beyond the three studied rides. Future research is necessary to understand how gaze behavior develops beyond the three rides, in daily commuting for example. Due to the efforts and time necessary for the current study the sample size was relatively small in this study. However, we believe that the external validity is high as the sample was strong and the repeated measures design allowed us to detect relevant effects.

In this study, the goal of allowing participants to use their mounted smartphone as non-driving-related task (NDRT) during automated driving, was to create a natural NDRT setting. It should be noted that such a natural scenario has been shown lead to some intraindividual variance in NDRT use and gaze behavior (see Marberger et al., 2019). However, given the research question of how gaze behavior changes over time in a realistic setting, we argue that this gaze variance is measurement of realistic behavior and not a confounding effect.

The Wizard of Oz design had the disadvantage that it is difficult for larger samples and the driving style may not correspond perfectly to a certified conditionally automated vehicle (Bengler et al., 2020). Additionally, the current study may already reflect the full development of driver's behavioral adaptation to automated driving, given the total ride duration of three hours and 396 km of manual- and 252 km of automated driving (Beggiato et al., 2015). However, to confirm this it would be necessary to investigate this with longer on the road studies as have been performed with adaptive cruise control (Beggiato et al., 2015). A promising next step would be in depth analyses of larger groups of participants commuting with a true SAE L3 system and a fallback ready safety driver.

The variables analyzed in the present study were chosen in line with the research focus on behavior in a real highway environment. As a result, the analyses were performed on high-level variables that were expected to be robust in real-world investigations. We argue that the approach taken in this study should be complemented by more fine scaled realistic driving research in more controlled environments (cf. Wintersberger et al., 2021). This approach would permit zooming in on sensitive variables such as gaze or steering reactions to self-paced and urgent take-over requests.

6. Conclusion

In the current study we investigated how drivers' take-over and gaze behavior developed over three rides. The take-over data showed that drivers' non-critical take-overs became faster over the three rides. This shows that exposure to multiple and variable real life take-over situations allows drivers to come back into the loop faster. Unexpectedly, urgent take-overs did not evoke even faster take-overs, which may have been due to the lack of an actual critical situation. Gaze behavior indicated that drivers progressively leave the loop within each ride but show consistent out-of-the-loop gaze behavior over all three rides. The changes over the rides that we observed thereby suggest that drivers increasingly leave the loop within each ride but become faster at coming back into the loop over the rides. Future research could further explore if these findings extend over longer periods of time, more rides and larger samples. It should be investigated if drivers seek or attune to different visual information while taking-over as they gain extensive practice in this safety–critical task over repeated rides.

CRediT authorship contribution statement

J. Dillmann: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft. **R.J.R. Den Hartigh:** Conceptualization, Methodology, Formal analysis, Writing – original draft. **C.M. Kurpiers:** Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft. **F.K. Raisch:** Conceptualization, Methodology, Writing – review & editing. **N. Kadrileev:** Methodology, Investigation, Formal analysis, Writing – review & editing. **R.F.A. Cox:** Conceptualization, Methodology, Formal analysis, Writing – original draft. **D. De Waard:** Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Beggiato, M., Pereira, M., Petzoldt, T., Krems, J., 2015. Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. Transport. Res. F: Traffic Psychol. Behav. 35, 75–84. https://doi.org/10.1016/j.trf.2015.10.005.

Bengler, K., Omozik, K., Müller, A., 2020. The Renaissance of Wizard of Oz (WoOz)—Using the WoOz methodology to prototype automated vehicles. In: de Waard, D., Toffetti, A., Pietrantoni, L., Franke, T., Petiot, J.F., Dumas, C., Botzer, A., Onnasch, L., Milleville, L., Mars, F. (Eds.)Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2019 Annual Conference.

Brand, M.T., de Oliveira Ferraz, R., 2017. Recalibration in functional perceptual-motor tasks: a systematic review. Hum. Mov. Sci. 56, 54–70. https://doi.org/10.1016/j.humov.2017.10.020.

Cohen, J., 1988. Statistical power for the social sciences. Laurence Erlbaum and Associates, Hillsdale, NJ, pp. 98–101.

Cohen, P., West, S.G., Aiken, L.S., 2014. In: Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Psychology Press. https://doi.org/10.4324/9781410606266.

Dahlbäck, N., Jönsson, A., Ahrenberg, L., 1993. Wizard of Oz studies - why and how. Knowl.-Based Syst. 6 (4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N.

Dari, S., Kadrileev, N., Hullermeier, E., 2020. A Neural Network-Based Driver Gaze Classification System with Vehicle Signals. In: Proceedings of the International Joint Conference on Neural Networks, pp. 1–7. https://doi.org/10.1109/IJCNN48605.2020.9207709.

de Waard, D., van der Hulst, M., Hoedemaeker, M., Brookhuis, K.A., 1999. Driver behavior in an emergency situation in the automated highway system. Transportation Human Factors 1 (1), 67–82. https://doi.org/10.1207/sthf0101_7.

de Winter, J., Stanton, N., Eisma, Y.B., 2021. Is the take-over paradigm a mere convenience? Transp. Re. Interdiscip. Perspect. 10 (January), 100370 https://doi.org/10.1016/j.trip.2021.100370.

Dillmann, J., den Hartigh, R.J.R., Den, K.C.M., Raisch, F.K., Waard, D De, Cox, R.F.A., 2021a. Keeping the driver in the loop in conditionally automated driving : a perception-action theory approach. Transp. Res. F Psychol. Behav. 79 (April), 49–62. https://doi.org/10.1016/j.trf.2021.03.003.

Dillmann, J., den Hartigh, R.J.R., Kurpiers, C.M., Pelzer, J., Raisch, F.K., Cox, R.F.A., de Waard, D., 2021b. Keeping the driver in the loop through semi-automated or manual lane changes in conditionally automated driving. Accid. Anal. Prev. 162, 106397 https://doi.org/10.1016/j.aap.2021.106397.

Eriksson, A., Stanton, N.A., 2017. Takeover time in highly automated vehicles: noncritical transitions to and from manual control. Hum. Factors 59 (4), 689–705. https://doi.org/10.1177/0018720816685832.

Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. Transp. Res. A Policy Pract. 77, 167–181. https://doi.org/10.1016/j.tra.2015.04.003.

Fajen, B.R., 2005. Calibration, information, and control strategies for braking to avoid a collision. J. Exp. Psychol. Hum. Percept. Perform. 31 (3), 480–501. https://doi.org/10.1037/0096-1523.31.3.480.

Fajen, B.R., Devaney, M.C., 2006. Learning to control collisions: The role of perceptual attunement and action boundaries. J. Exp. Psychol. Hum. Percept. Perform. 32 (2), 300–313. https://doi.org/10.1037/0096-1523.32.2.300.

Feldhütter, A., Gold, C., Schneider, S., Bengler, K., 2017. How the Duration of Automated Driving Influences Take-Over Performance and Gaze Behavior. In: Schlick, C.M., Duckwitz, S., Flemisch, F., Frenz, M., Kuz, S., Mertens, A., Mütze-Niewöhner, S. (Eds.), Advances in Ergonomic Design of Systems, Products and Processes. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 309–318.

Forster, Y., Naujoks, F., Neukum, A.., 2016. Your turn or my turn? Design of a human-machine interface for conditional automation. AutomotiveUI 2016 - 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Proceedings, 253–260. https://doi.org/10.1145/3003715.3005463.

Forster, Y., Hergeth, S., Naujoks, F., Beggiato, M., Krems, J.F., Keinath, A., 2019. Learning to use automation: behavioral changes in interaction with automated

driving systems. Transp. Res. F Psychol. Behav. 62, 599–614. https://doi.org/10.1016/j.trf.2019.02.013.
Gold, C., Damböck, D., Lorenz, L., Bengler, K., 2013. Take over! How long does it take to get the driver back into the loop? Proc. Human Factors Ergon. Soc. 57 (1), 1938–1942.
Gold, C., Meyer, M.L., Fischer, F., 2017. Take-Over Performance in a Wizard of Oz Vehicle in Level 3 CAD. Active Safety and Automated Driving.
Jacobs, D.M., Michaels, C.R., 2007. Direct learning. Ecol. Psychol. 19 (4), 321–349. https://doi.org/10.1080/10407410701432337.
Klingegard, M., Andersson, J., Habibovic, A., Nilsson, E., Rydstrom, A., 2020. Drivers' ability to engage in a non-driving related task while in automated driving mode in real traffic. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3043428.
Kurpiers, C., Biebl, B., Hernandez, J.M., Raisch, F., 2020. Mode awareness and automated driving-What is it and how can it be measured. Information (Switzerland) 11 (5), 277–290. https://doi.org/10.3390/INFO11050277.
Large, D.R., Burnett, G., Morris, A., Muthumani, A., Matthias, R., 2018. A longitudinal simulator study to explore drivers' behaviour during highly-automated driving. Adv. Intel. Syst. Comput. 597 https://doi.org/10.1007/978-3-319-60441-1_57.
Louw, T., Markkula, G., Boer, E., Madigan, R., Carsten, O., Merat, N., 2017. Coming back into the loop: Drivers' perceptual-motor performance in critical events after automated driving. Accid. Anal. Prev. 108 (February), 9–18. https://doi.org/10.1016/j.aap.2017.08.011.
Manca, L., de Winter, J.C.F., Happee, R., 2015. Visual Displays for Automated Driving : a Survey. *Workshop on Adaptive Ambient In-Vehicle Displays and Interactions - AutomotiveUI '15*, *August*, 1–5. https://doi.org/10.13140/RG.2.1.2677.1608.
Marberger, C., Manstetten, D., Klöffel, C., 2019. Highly automated driving in the real world - a wizard-of-oz study on user experience and behavior. VDI Ber. 2019 (2360). https://doi.org/10.51202/9783181023600-109.
Markkula, G., Engström, J., Lodin, J., Bärgman, J., Victor, T., 2016. A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. Accid. Anal. Prev. 95, 209–226. https://doi.org/10.1016/j.aap.2016.07.007.
Mathieu, J., Bootsma, R.J., Berthelon, C., Montagne, G., 2017. Information-movement coupling in the control of driver approach to an intersection. Ecol. Psychol. 29 (4), 317–341. https://doi.org/10.1080/10407413.2017.1369853.
Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J.D., Johansson, E., Green, C.A., Katazaki, S., Monk, C., Itoh, M., McGehee, D., Sunda, T., Unoura, K., Victor, T., Schieben, A., Keinath, A., 2018. The "Out-of-the-Loop" concept in automated driving: proposed definition, measures and implications. Cogn. Tech. Work 21 (1), 87–98. https://doi.org/10.1007/s10111-018-0525-8.
Mole, C.D., Lappi, O., Giles, O., Markkula, G., Mars, F., Wilkie, R.M., 2019. Getting back into the loop: the perceptual-motor determinants of successful transitions out of automated driving. Hum. Factors 61 (7), 1037–1065. https://doi.org/10.1177/0018720819829594.
Naujoks, F., Befelein, D., Wiedemann, K., Neukum, A., 2018. A review of non-driving-related tasks used in studies on automated driving. Advances in Intelligent Systems and Computing 597 (October), 525–537. https://doi.org/10.1007/978-3-319-60441-1_52.
Naujoks, F., Purucker, C., Wiedemann, K., Marberger, C., 2019. Noncritical State Transitions During Conditionally Automated Driving on German Freeways: Effects of Non–Driving Related Tasks on Takeover Time and Takeover Quality. Human Factors 61 (4), 596–613. https://doi.org/10.1177/0018720818824002.
Russell, H.E.B., Harbott, L.K., Nisky, I., Pan, S., Okamura, A.M., Gerdes, J.C., 2016. Motor learning affects Car-To-Driver handover in automated vehicles. Sci. Robot. 1 (1), eaah5682. https://doi.org/10.1126/scirobotics.aah5682.
Rydström, A., Mullaart, M.S., Novakazi, F., Johansson, M., Eriksson, A., 2022. Drivers' performance in non-critical take-overs from an automated driving system—an on-road study. Hum. Factors. https://doi.org/10.1177/00187208211053460.
SAE International. (2018). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016. In *SAE International* (Vol. J3016, Issue J3016). https://doi.org/https://doi.org/10.4271/J3016_201806.
Schnebelen, D., Charron, C., Mars, F., 2020. Estimating the out-of-the-loop phenomenon from visual strategies during highly automated driving. Accid. Anal. Prev. 148, 105776 https://doi.org/10.1016/j.aap.2020.105776.
Seppelt, B.D., Seaman, S., Lee, J., Angell, L.S., Mehler, B., Reimer, B., 2017. Glass half-full: On-road glance metrics differentiate crashes from near-crashes in the 100-Car data. Accid. Anal. Prev. 107 (August), 48–62. https://doi.org/10.1016/j.aap.2017.07.021.
Sullivan, G.M., Feinn, R., 2012. Using Effect Size—or Why the P Value Is Not Enough. J. Grad. Med. Educ. 4 (3), 279–282. https://doi.org/10.4300/jgme-d-12-00156.1.
Sullivan, B.T., Johnson, L., Rothkopf, C.A., Ballard, D., Hayhoe, M., 2012. The role of uncertainty and reward on eye movements in a virtual driving task. J. Vis. 12 (13), 19.
UN-ECE. (2020). Proposal for a new UN Regulation on uniform provisions concerning the approval of vehicles with regards to Automated Lane Keeping System. *181st Session of the World Forum for Harmonization of Vehicle Regulations*, 1–63.
Wilson, M., 2004. Six views of embodied cognition. Cognition 9 (4), 1–19. http://view.ncbi.nlm.nih.gov/pubmed/12613670.
Winkler, S., Kazazi, J., Vollrath, M., 2018. Practice makes better – Learning effects of driving with a multi-stage collision warning. Accid. Anal. Prev. 117, 398–409.
Wintersberger, P., Schartmüller, C., Shadeghian-Borojeni, S., Frison, A.-K., Riener, A., 2021. Evaluation of imminent take-over requests with real automation on a test track. Human Factors: J. Human Fact. Ergon. Soc. 001872082110514 https://doi.org/10.1177/00187208211051435.
Zhang, B.o., de Winter, J., Varotto, S., Happee, R., Martens, M., 2019. Determinants of take-over time from automated driving: a meta-analysis of 129 studies. Transport. Res. F: Traffic Psychol. Behav. 64, 285–307. https://doi.org/10.1016/j.trf.2019.04.020.