ARTICLE POST-PRINT

JOURNAL OF INTELLIGENT TRANSPORTATION SYSTEMS

Takeover request (TOR) effects during different automated vehicle failures

Chelsea A. DeGuzman, Dina Kanaan, Samantha A. Hopkins, and Birsen Donmez*

Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada

Takeover request (TOR) effects during different automated vehicle failures

Research on driving automation has investigated the use of takeover requests (TORs) to warn drivers about automation failures that require their intervention. Such failures can occur with information in the environment that drivers can use to anticipate them (e.g., system-limit failures) or without such information (e.g., system-malfunction failures). There is a lack of research comparing the effectiveness of TORs prior to these different failure types. We conducted a simulator study with 19 participants to investigate whether the effect of a TOR on drivers' monitoring and takeover performance differed by failure type. Drivers were trained on automation limits so that they could identify upcoming system-limit failures. We evaluated gaze behaviors starting from 6 seconds before the failure (corresponding to TOR onset in TOR drives and the equivalent point in no-TOR drives) until drivers took over. The effect of TORs on monitoring the roadway was significant only for system-malfunction failures, with participants looking more at the roadway in TOR drives compared to no-TOR drives. For system-limit failures, the TOR did not provide any benefit in terms of visual attention to the roadway, likely because participants were already looking at the roadway because they could anticipate the failure. However, having a TOR for system-limit failures was associated with faster takeover time than not having a TOR. Although the TOR may not have had monitoring benefits when environmental information was available, our findings suggest it was still useful as a confirmation of impending failures and a prompt to take over control of the vehicle.

Keywords: automated driving; automation failures; driver behavior; eye-tracking; vehicle automation.

Introduction

Recent research on automated driving systems has extensively investigated the use of takeover requests (TORs) to warn drivers in advance about automation failures that require their intervention (Gold, Damböck, Lorenz, & Bengler, 2013; Melcher, Rauh, Diederichs, Widlroither, & Bauer, 2015; Yoon, Kim, & Ji, 2019). By providing warnings, TORs may reduce the need for drivers to monitor the driving environment and the automated driving system in order to anticipate or respond to automation failures. An aggregation of the results of six studies (reported as part of a meta-analysis on the effects of various factors on driving performance in automated vehicles) suggests that drivers' reaction time to failures (i.e., takeover time) is reduced by around 0.58 seconds in the presence of TORs (Zhang, De Winter, Varotto, Happee, & Martens, 2019). Although the reported studies highly varied in terms of failure types and various other methodological details (e.g., how and when the TOR was presented and whether drivers were performing a non-driving task in addition to operating the automated driving system), they provide evidence that TORs can support takeover performance. Research has also examined different factors that may contribute to or moderate drivers' behaviors in response to TORs. These factors include aspects of the TOR design, such as its lead time (how long before a failure that the TOR is issued, e.g., Gold et al., 2013; Wan & Wu, 2018) or its modality (Roche, Somieski, & Brandenburg, 2018; Yoon et al., 2019). Other studies have investigated factors related to the driving context, such as the complexity of the driving environment (e.g., Brandenburg & Chuang, 2019), or related to the driver, such as prior training on automation failures and TORs (Hergeth, Lorenz, & Krems, 2017) and the type of non-driving task the driver performs while operating the vehicle (Roche et al., 2018; Wan & Wu, 2018).

Another important consideration that has received little attention in the research on TORs is whether the effects of TORs are different depending on whether there is information in the environment that would allow a driver to anticipate a failure. Information can be provided about the automation itself (e.g., displays that convey information about the automation's reliability; Helldin, Falkman, Riveiro, & Davidsson, 2013; Kraft, Naujoks, Wörle, & Neukum, 2018; Seppelt & Lee, 2019), but there are also situations in which the driver can predict the automation's behaviour based on information in the environment. For example, a known limitation of lane keeping systems is that they do not work well to keep the vehicle within its lane when lane markings are faded, as was the case at the site of the fatal collision between a Tesla vehicle and a road divider in California in March 2018 (National Transportation Safety Board, 2020). Drivers who are aware of the automation's limitations may be able to anticipate such a failure based on their knowledge and cues in the environment (e.g., faded lane markings) and take over before the failure occurs. We refer to these failures as *system-limit failures* as they are associated with known functional limitations of the automation (DeGuzman, Hopkins, & Donmez, 2020). In contrast, the automation might fail with no observable reason or indicators in the environment that can help drivers anticipate them (e.g., due to sensor or algorithmic errors), and thus drivers can only react after the failure occurs unless they are provided with a TOR. For example, in 2019, Mazda recalled over 35,000 vehicles due to a programming error in its automatic emergency braking system, which was falsely detecting obstacles and applying the brakes for no apparent reason (Mazda, 2019). We refer to these failures as *system-malfunction* failures (DeGuzman et al., 2020). In this study, we used a system-limit failure as an example situation in which there is information in the environment that the driver can use to anticipate the failure, and a system-malfunction failure as an example situation in which there is no such information.

We previously reported the findings of a driving simulator study with participants who used adaptive cruise control (ACC) and lane keeping assist (LKA) systems, controlling longitudinal and lateral movement, respectively (DeGuzman et al., 2020). This simulator study is the same data source for the analyses we conduct in this paper, but in our prior analysis we used a subset of the data analyzed in the current paper and combined it with additional data. In our prior analysis, we quantified the differences in drivers' performance and monitoring when they experience failures with (system-limit) and without (system-malfunction) environmental information indicating a potential failure, without the presence of TORs. We found that when drivers experienced system-limit failures, they exhibited improved monitoring of the roadway and faster takeover time in response to the failure compared to when they experienced system-malfunction failures. In a simulator study, Dogan et al. (2017) also used system-limit and system-malfunction failures to investigate drivers' monitoring and takeover performance when experiencing what they termed "anticipated" and "unanticipated" failures, respectively. Participants were informed that the automated driving system would fail at a speed greater than 50 km/h (thus the speedometer provided failure-relevant information); however, in some trials, the automation unexpectedly failed at 30 km/h. TORs were used to warn participants about both types of failures. Drivers were required to take over immediately after a TOR was issued, but the TOR was issued 3 seconds before the automation was deactivated. The authors found that participants looked more at the speedometer before anticipated events than unanticipated events; however, no differences in takeover performance were observed between the two failure types.

Dogan et al. (2017) implemented TORs in their investigation of system-limit and system-malfunction failures, however they did not have baseline conditions (i.e., drives with no TORs) to assess the effect of the TOR. To the best of our knowledge, no research has systematically compared TOR effectiveness based on whether there is information available in the environment that would allow the driver to anticipate the failure. We expect that TORs would be more beneficial for failures without information available in the environment (system-malfunction failures in our study), as drivers would have no other way of anticipating such failures. Although it may not always be feasible to warn drivers in advance about unexpected system malfunctions, this type of scenario has been previously implemented in simulator studies (Dogan et al., 2017; Payre, Cestac, Dang, Vienne, & Delhomme, 2017; Richardson, Flohr, & Michel, 2018). An advance warning may be feasible to implement in situations when the system can warn the driver in advance of an impending disengagement of the automated driving system due to a decline in, for example, sensor reliability.

In this paper, we report an analysis of driving simulator study data that compared the effects of TORs on situations with and without environmental information available to anticipate failures, for an automated driving system that controlled the vehicle both laterally and longitudinally. As mentioned previously, in our study, a system-limit failure was used as an example of a situation in which information was available in the environment, and a system-malfunction failure was used as an example of a situation in which no such information was available. The objective of the paper is to investigate the effect of TORs on takeover performance and monitoring when experiencing these different failure types. Drivers were well-trained on the limitations of ACC and LKA shortly before the experimental drives, so they were prepared for potential failures. In reality, drivers may receive less training (if any) on automation limitations and will operate the system long after their initial training. However, our goal was to ensure that drivers knew the limitations so that they could successfully anticipate the system-limit failure, in order to assess the effectiveness of TORs depending on the failure type. The study focused on LKA failures given previous survey findings suggesting that 74% of unexpected events encountered by Tesla Autopilot users were due to issues with lane detection (Dikmen & Burns, 2016). Further, it has been demonstrated that automating the lateral and longitudinal control of the vehicle decreases drivers' awareness of the environment and makes them more likely to allocate their attention to non-driving tasks (Carsten, Lai, Barnard, Jamson, & Merat, 2012; de Winter, Happee, Martens, & Stanton, 2014). Thus, participants were asked to perform a self-paced visual-manual non-driving task so that we could assess how they divide their visual attention between monitoring the environment and the non-driving task.

Methods

Participants

Participants were recruited using posters, online postings (e.g., on Facebook), and through emails sent to listservs at the University. Participants were required to have a Canadian driver's license or equivalent for at least 2 years, drive at least several times a month, and have normal or corrected-to-normal vision. A screening questionnaire was used to recruit participants who had no experience with ACC and LKA systems and who were not prone to simulator sickness. Participants with no experience with ACC and LKA were chosen to limit individual differences in experience with automation, and so that their understanding of the systems used in the study would be based primarily on our training. Thus, our study reflects the initial experience with ACC and LKA after training, and can provide insight into how new users may interact with these systems in their early stages of using them.

Nineteen participants (11 males, 8 females) ranging in age from 25 to 30 (M = 27.53, SD = 1.68) completed the study. It should be noted that the no-TOR data for eight of the participants from this sample (plus ten additional participants not included in this sample) was used in our earlier analysis of differences between system-limit and system-malfunction failures (DeGuzman et al., 2020; Hopkins, DeGuzman, & Donmez, 2019). The main objective of the current analysis is to investigate the interaction effect between TOR presence and failure type.

Participants were compensated at a rate of C$14/hour (Canadian dollars), with a C$8 “bonus” which all participants received regardless of their performance level. The study took approximately 3 hours, thus participants received about C$50 total for their participation. All procedures were approved by the University of Toronto’s Research Ethics Board.

Experiment Design

The experiment was a 2x2 within-subjects design. The two independent variables were: failure type (system-limit versus system-malfunction) and TOR presence (no-TOR versus TOR). Each participant completed four drives with blocking on TOR presence. That is, participants could either have their first two drives with the TOR (one drive with system-limit and one drive with system-malfunction failure) and the last two with no TOR (one drive with system-limit and one drive with system-malfunction failure), or vice versa. The order of presentation of both the TOR presence blocks and the failure types within each block were counterbalanced as presented in Table 1. In order to assess whether block order (whether participants saw the TOR display in block 1 or 2) impacted the results, block was included as a third independent variable in the analysis.

Table 1. Counterbalanced order of the four drives. Each participant completed one possible sequence. L= system-limit, M = system-malfunction. Grey shading indicates drives with the TOR; no shading indicates drives with no TOR.

Possible Sequence	Drive 1	Drive 2	Drive 3	Drive 4
1	L	M	L	M
2	M	L	L	M
3	L	M	M	L
4	M	L	M	L
5	L	M	L	M
6	M	L	L	M
7	L	M	M	L
8	M	L	M	L

Apparatus

The experiment was conducted on a National Advanced Driving Simulator (NADS) MiniSim fixed-base simulator, which consists of three 42" widescreen displays that create a 130° horizontal and 24° vertical field of view at a viewing distance of 48" (Figure 1). A self-paced visual-manual secondary task (i.e., non-driving task) was presented on a Microsoft Surface Pro 2 tablet to the right of the dashboard, at a downward viewing angle of 40°, and participants wore a head-mounted Dikablis Glasses 3 eye tracker which collected gaze data. While we were primarily interested in drivers' attention to the roadway (Front Screen) and the non-driving task, we created eight areas of interest (AOIs) to determine whether there were any other areas that participants were looking at. The eight AOIs were: Left Screen, Front Screen, Right Screen, Rear View Mirror, Non-driving Task, Dashboard, Steering Wheel, and TOR Display (only for the TOR drives).

Figure 1. Simulator set up. Arrow indicates the location of the buttons to engage/disengage the automation. (1) Front Screen AOI (2) Non-driving task AOI; task presented on a Surface Pro 2 tablet (3) Dikablis eyetracker

Driving Scenarios and Automation

All experimental drives took place on a rural two-lane highway that was 12 feet (3.66 m) wide, with a yellow line separating opposing lanes of traffic, and a posted speed limit of 50 mph (80.47 km/h). The traffic in the simulator was designed to be light (9-11 cars per minute). Participants were told to follow a lead vehicle, which maintained a 3-second gap between the ego-vehicle's front bumper and the lead vehicle's rear bumper. The lead vehicle continued straight at a steady pace (i.e., it did not slow down unless the participant slowed down, in which case it slowed down to maintain the 3-second gap). There were no stop signs or traffic lights that required the participant to slow down or stop the vehicle. Each drive took approximately 6 minutes.

The driving automation consisted of ACC and LKA systems similar to those that have been used in previous studies (Shen & Neyens, 2017; Zeeb, Buchner, & Schrauf, 2016). The ACC and LKA was set to be on at all times, and participants could decide when to engage and disengage the automated systems. The ACC and LKA were engaged using buttons on the steering wheel. The ACC could be disengaged using the cancel button on the steering wheel or by pressing the brake. The LKA could be disengaged using the LKA on/off button on the steering wheel, or by turning the steering wheel at least 5° in either direction (it could not be disengaged by braking). Participants were told to drive safely and disengage the automation whenever they thought it was necessary, but to leave it engaged as much as possible.

Non-driving Task

A self-paced visual-manual non-driving task was used, which simulated searching through options on an infotainment system. Participants had to scroll through phrases to find one that matched a target phrase and click "Submit" once they found the matching phrase (see Figure 2). The non-driving task was adapted from Donmez, Boyle, and Lee (2007) and was shown to degrade driving performance in multiple simulator studies (Chen, Hoekstra-Atwood, & Donmez, 2018; Donmez et al., 2007; Merrikhpour & Donmez, 2017). The task was available throughout all the experimental drives, and participants could choose when they wanted to engage in the task. Participants were told that a C$8 bonus was possible depending on their non-driving task performance, but also that there would be deductions from the bonus for poor driving performance, and they were instructed to drive safely. All participants received the bonus regardless of performance. We implemented a self-paced task as drivers choose when to engage in non-driving tasks when in their own vehicles, however, making the non-driving task self-paced introduces the possibility that some participants might engage more in the task than others. In our study, the percentage of participants in each condition that were looking at the non-driving task 6 s prior to the failure (equivalent to TOR onset in TOR drives) was comparable (ranging from 53-63%), thus it is unlikely that our results were significantly impacted by differences in non-driving task engagement.

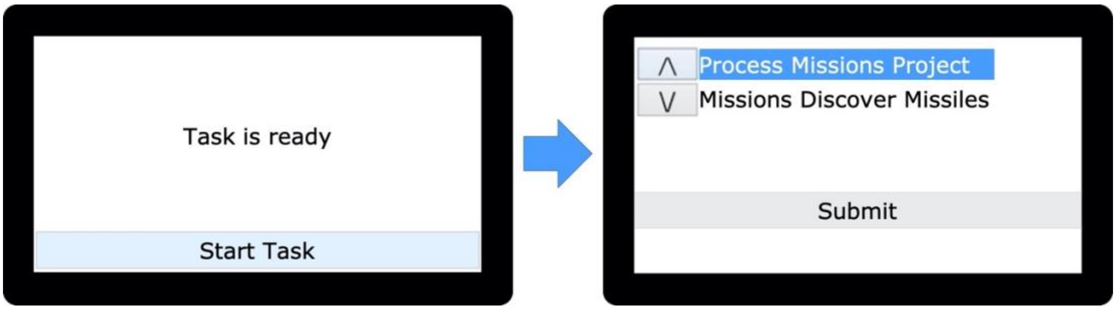

Figure 2. Non-driving task "Discover Project Missions", adapted from Donmez et al. (2007)

Automation Failures

LKA failures occurred in the form of lane departure events. The system-limit failure occurred at a T-intersection, whereby instead of following the lead vehicle through the intersection, the automation would follow the white fog line marking the road edge and veer into the right-turn lane (see Figure 3). Participants were taught during training that the LKA system did not function as well at intersections due to the quick changing of the lane markings. The system-malfunction failure occurred on a straight section of road, with the vehicle veering to the right in a similar manner to the system-limit failure (see Figure 3). However, in this case, there were no changes in the lane markings and nothing in the roadway to indicate that the automation may fail. Two failures occurred in each drive, but both were of the same type (i.e., either both were due to a system limit or both were due to a system malfunction). Due to issues with the implementation of the second failure event in the simulator, only the first failure event from each drive was analyzed. It should be noted that the rate of failures in our study is relatively high compared to a real-world setting in which failures may be rare, and thus our participants may have been more prepared to respond to failures in our study than in a real-world setting. However, the study was designed this way in order to collect enough data for statistical analysis without requiring participants to commit to a substantially longer study. Similar failure rates have been adopted in various simulator studies on driving automation (e.g., Gaspar, Carney, Shull, & Horrey, 2020; Louw et al., 2019).

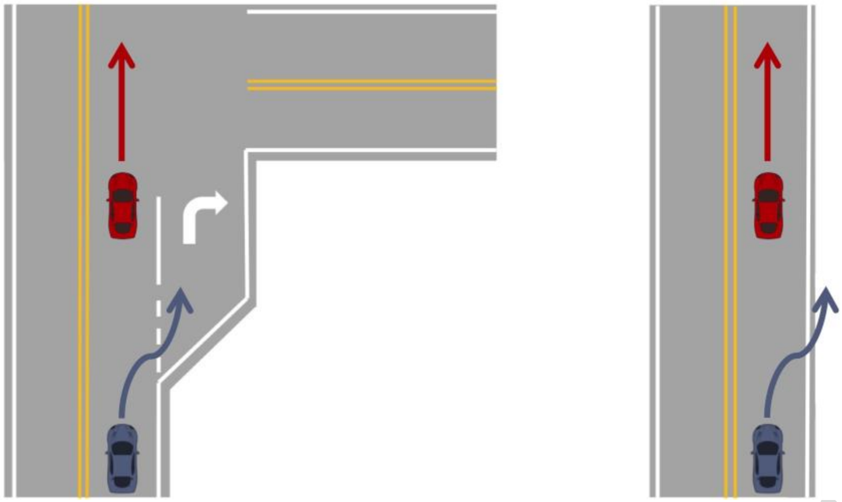

Figure 3. Left: Path of the vehicle at the system-limit failure event. Right: Path of the vehicle at the system-malfunction failure event. Blue car = ego-vehicle, red car = lead vehicle.

TOR Display

In the no-TOR drives, participants were responsible for monitoring the roadway and determining when it was safe to use the automation and when they should take over control of the vehicle. In the TOR drives, the TOR alerted participants of an upcoming failure six seconds ahead of the failure event, which would inform them that they should take over control of the vehicle. A review by Eriksson and Stanton (2017) found the mean TOR lead time (time before failure) used in the literature to be approximately 6 seconds, and thus we used that value in this experiment. Based on previous research (Dogan et al., 2017; Payre et al., 2017), the TOR we implemented for the system malfunction represented a warning about an unforeseen error before deactivating the automation 6 seconds later. Thus, the “failure” we refer to for the system-malfunction failure is the point at which the LKA deactivates, causing the vehicle to swerve out of the lane if the driver did not take over control of the steering.

The TOR was designed to have a visual and auditory component as drivers have been found to react faster to warnings that use both modalities compared to warnings that are only visual (Naujoks, Mai, & Neukum, 2014). The auditory component was a loud 2-second beep. The visual component was an icon (see Figure 4) that was inspired by previous studies that used depictions of hands and a steering wheel (Eriksson & Stanton, 2017; Melcher et al., 2015; Naujoks et al., 2014). The TOR icon appeared in the bottom right corner of the front screen (center monitor) of the simulator (see Figure 4) and remained on the screen for 20 seconds after the failure had occurred to prevent participants from trying to re-engage the automation immediately after the failure. TOR onset is defined as the time point when the auditory and visual alert first occurred.

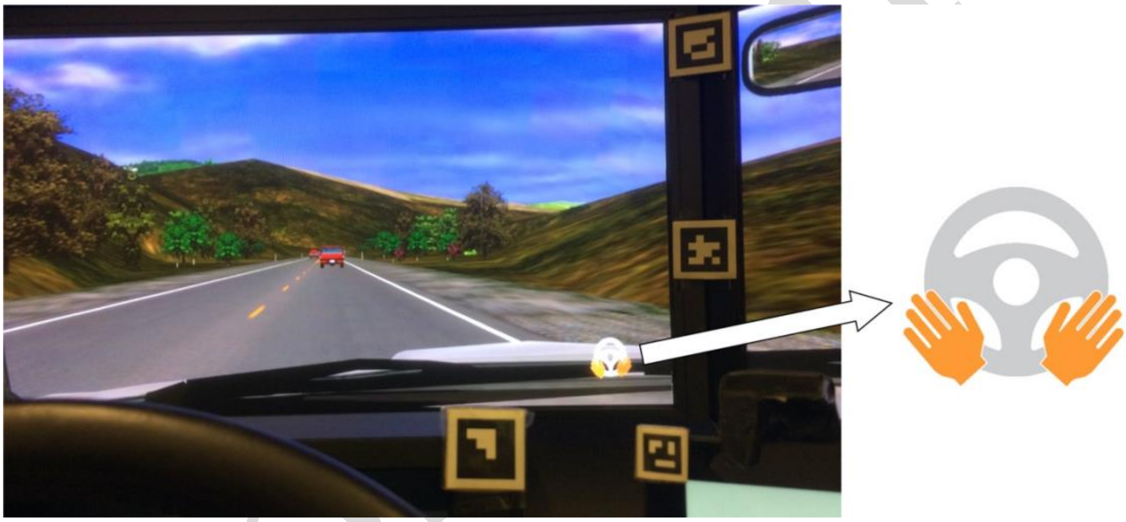

Figure 4. The view of the TOR display on the front screen in the simulator

Procedure

After receiving information about the study and providing informed consent, participants were introduced to the simulator (e.g., location of the pedals, mirrors, etc.). They were told that the simulated vehicle was slower to accelerate and decelerate and that the steering wheel may be more sensitive than in most cars they have driven, and then completed an initial training drive to practice driving in the simulator. They were then trained on the limits of the automation (both longitudinal and lateral control) so that they would have the knowledge required to identify system-limit failures. Participants were told that the ACC only had 30% braking power, which meant that it would not be able to stop in time if the lead vehicle stopped abruptly, and that it did not reliably detect pedestrians, motorcycles, or other small objects on the road. They were told that the LKA system did not function well if the lane markings were faded or missing or if the lane markings changed quickly, like at intersections or in construction zones, and that both the LKA and ACC system did not work well in poor weather conditions. Participants were also told that sensor errors or algorithmic failures could cause “seemingly random issues” with the ACC and LKA so that they were aware of potential system-malfunction failures. After this training, they were asked a series of questions about the automation’s limits, to which they responded verbally. They were provided with feedback regarding whether each answer was correct, and if it was incorrect, the correct answer was explained to them in order to try to ensure that all participants understood the limits before completing the experimental drives. Once participants were trained on the automation, they were trained on how to perform the non-driving task and then completed a second training drive in which they drove with the ACC and LKA engaged, while performing the non-driving task. Finally, participants completed a third training drive, similar to the second training drive except that they experienced a system-limit and a system-malfunction failure (the same as those in the experimental drives). This third training drive was the same length as the experimental drives (approximately 6 minutes) and participants were not told that this was a training drive – it was presented to be the first experimental drive. The purpose of this training drive was to expose participants to the failures so that their first true experimental drive would not be impacted by shock or confusion with encountering an automation failure for the first time. After training, participants completed the four experimental drives.

Analysis

To assess takeover performance, we analyzed takeover time (seconds), i.e., when, relative to the failure, the driver disengaged the LKA. Positive values indicated that participants took over after the failure occurred, whereas negative values indicated that participants took over before the failure. To assess monitoring behavior, we analyzed the period of time from six seconds before the failure event occurring to the time the participants disengaged the automation (see Figure 5). The start (six seconds before failure) was a fixed point corresponding to the TOR onset for the TOR drives, so that we could compare how drivers' monitoring changed approaching a failure when a TOR was present versus absent. While the failure point was fixed in all drives (and thus so was the TOR onset point), the end point of our analysis period was when participants disengaged the automation, which was different for each drive. As shown in Figure 7, in no TOR drives, participants took over around 1 second after the failure, thus this time period was approximately 7 seconds, on average. In TOR drives, participants took over around 1 second before the failure, thus this time period was approximately 5 seconds, on average. For this time period, we extracted the percent of time looking at an AOI before takeover. Because this period before takeover was short, it was not feasible to measure rate or number of glances, thus we measured the percent of time that participants looked at different AOIs.

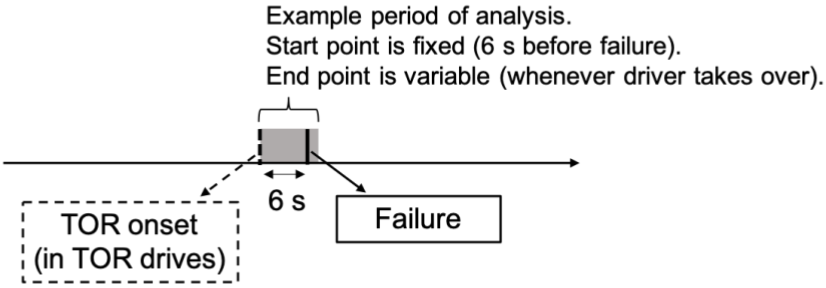

Figure 5. Timeline of events within a drive. Within each drive, the failure occurred at the same point for every participant. For drives with a system malfunction, the point of failure is the point at which the automation deactivated. In TOR drives, the TOR appeared 6 seconds before the failure. The grey shading indicates an example period of analysis. The start point was always 6 seconds before the failure, but the end point was whenever the participant disengaged the automation, thus it could be before or after the failure.

Initial inspection of the percent of time participants spent looking at each AOI before they took over due to the failure (see Figure 6) indicated that participants mainly looked at the front screen/roadway (*M*: 37.4%, *SD*: 17.8%) and non-driving task (*M*: 52.9%, *SD*: 21.0%). Thus, statistical analyses were focused on these two areas. Given that the front screen displayed the roadway ahead, we will refer to this AOI as the roadway from this point forward.

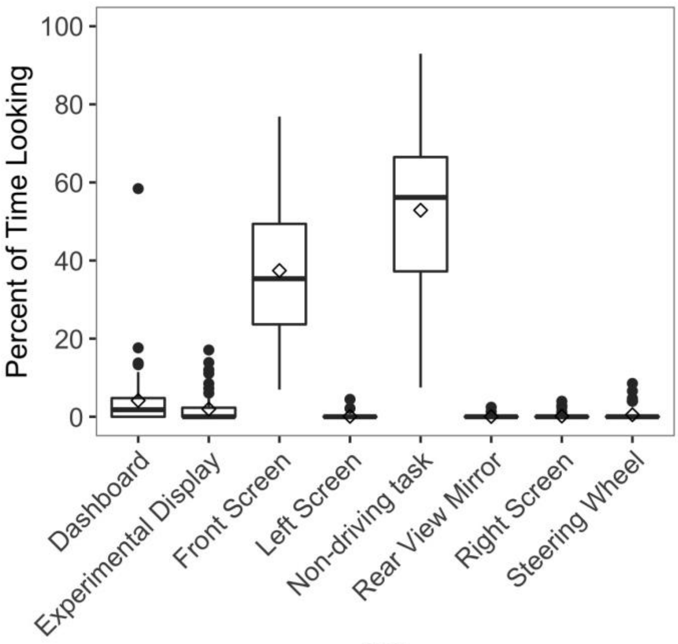

Figure 6. Percent of time participants spent looking at each AOI

Each of the 19 participants completed four drives, resulting in 76 observations. Eight observations were removed, either because the participant did not have the automation engaged for the 20 s leading up to the takeover situation or they did not follow experimenter instructions, resulting in 68 observations for the takeover time analysis. In addition, due to errors with the eye-tracking software, eye-tracking data was missing for an additional 6 drives, leaving 62 observations for the monitoring analysis. The breakdown of observations by condition can be found in Table 2.

Table 2. Number of observations by condition

	System-limit		System-malfunction		
Analysis	**No TOR**	**TOR**	**No TOR**	**TOR**	**Total**
Takeover	17	18	16	17	68
Monitoring	16	16	15	15	62

Due to the missing data, we used mixed linear models, which does not exclude participants due to missing observations (as is the case for repeated measures ANOVA). Mixed linear models also allowed us to use an unstructured variance-covariance matrix to correct for heteroscedasticity in the takeover time data. For mixed linear models (like with repeated measures ANOVAs), F-tests are conducted to assess significance of the main effects and interactions and will be reported. Initially, all independent variables (failure type, TOR presence, block order) and interactions (second-order and third-order) were included in the models. The third-order interaction was not significant in any of the models, so it was removed from all models. In addition, block order and its interactions were removed as they were not significant in any of the models. Thus, only models with main effects of failure type and TOR presence, and the interaction of these two variables (failure type × TOR presence), will be reported below. For significant interactions, we conducted four follow-up contrasts to test whether the main effect of failure type differed by TOR presence and whether the main effect of TOR presence differed by failure type (see Table 4 for these contrasts and associated results). Significance was assessed using the Benjamini-Hochberg approach (Benjamini & Hochberg, 1995) to control the false discovery rate for multiple comparisons.

Results

Takeover Time

There was a significant effect of TOR presence and failure type on takeover time (Table 3). Participants took over 0.81 s sooner in TOR drives compared to no-TOR drives (95% CI: -1.25, -0.37), and 0.59 s sooner for the system-limit failure compared to the system-malfunction failure (95% CI: -1.00, -0.17). Inspection of the raw data in Figure 7 shows that when there was no TOR, participants took over sooner for the system-limit failure, but almost all participants took over after the failure occurred. In contrast, the average takeover time for the TOR drives was before the failure. In addition, Figure 7 shows a higher variability in takeover time for the TOR group.

Table 3. Main effects and interaction results for all models. (DF = Degrees of Freedom).

	DF	*F-value*	*p-value*	ω^{2*}	*pseudo*-ω^{2**}
Takeover Time					
TOR presence	1, 18	14.82	**.001**	-	0.16
Failure type	1, 18	8.71	**.009**	-	0.09
TOR presence x Failure type	1, 18	1.47	.24	-	0.01
Percent of Time Looking at Roadway					
TOR presence	1, 16	7.87	**.01**	0.07	0.08
Failure type	1, 16	14.54	**.002**	0.10	0.16
TOR presence x Failure type	1, 10	7.24	**.02**	0.05	0.07
Percent of Time Looking at Non-driving Task					
TOR presence	1, 16	53.17	**< .001**	0.36	0.43
Failure type	1, 16	10.42	**.005**	0.04	0.08
TOR presence x Failure type	1, 10	13.12	**.005**	0.05	0.10

Note: bold text indicates significance based on the Benjamini-Hochberg adjusted alpha (Benjamini & Hochberg, 1995).
* Widely used standardized effect sizes, such as ω^2, have not been defined for mixed linear models, thus we recreated our mixed linear models as general linear models to obtain ω^2 where possible. Given that the takeover time model had an unstructured variance-covariance matrix for repeated measures, its mixed linear model cannot be specified within the general linear model framework, thus making it not possible to obtain ω^2.
** Pseudo-effect sizes were calculated following the method in Tippey and Longnecker (2016) outlined for mixed linear models. As noted by the authors, these pseudo-effect sizes can only be directly compared to each other and not to effect sizes calculated using other methods.

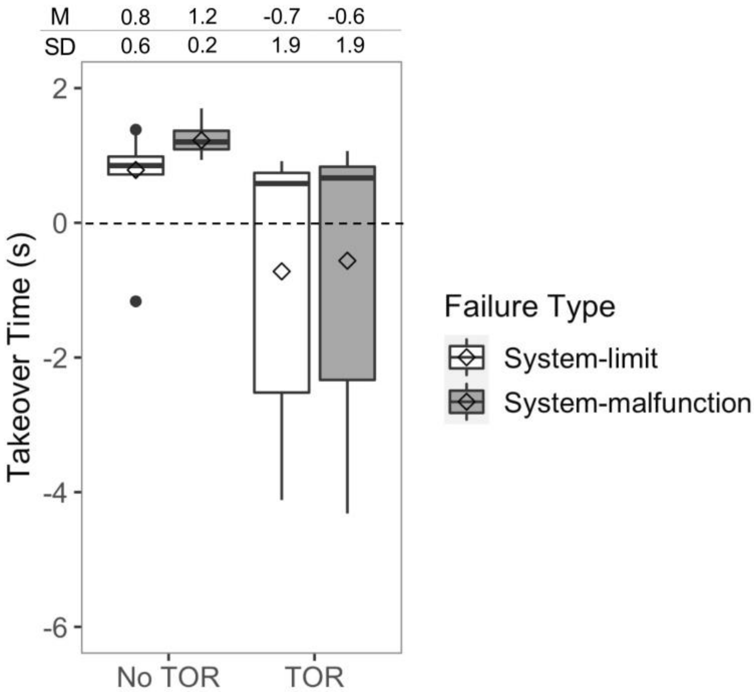

Figure 7. Raw data for takeover time by TOR presence and failure type. Dashed line indicates when the failure occurred. In this figure and all subsequent figures, diamonds indicate the mean. Mean (M) and standard deviation (SD) are presented along the top of the figure.

Monitoring

Based on visual inspection of the raw data (Figure 8), in TOR drives, participants spent more time looking at the roadway than the non-driving task, for both system-limit and system-malfunction failures. In the no-TOR drives, participants spent more time looking at the roadway than the non-driving task before the system-limit failure, but the opposite was true for the system-malfunction failure – they looked more at the non-driving task than the roadway.

The percent of time spent looking at the roadway was significantly affected by failure type, TOR presence, and their interaction (see Table 3), with small to medium effect sizes (Field, Miles, & Field, 2012). Follow-up contrasts reported in Table 4 revealed that there was a significant effect of the TOR for system-malfunction failures, but not system-limit failures. Before the system-malfunction failure, participants looked at the roadway 24% more in TOR drives compared to no-TOR drives. In addition, there was a significant effect of failure type for no-TOR drives, but not for TOR drives. In no-TOR drives, participants looked at the roadway 28% more before the system-limit failure compared to the system-malfunction failure.

For percent of time looking at the non-driving task, there was also a significant main effect of failure type and TOR presence, and a significant interaction (Table 3). The interaction effect and main effect of failure type had small effect sizes, while TOR presence had a large effect size (Field et al., 2013). The results of the follow-up contrasts were slightly different than those for percent of time looking at the roadway. The TOR was associated with a lower percentage of time looking at the non-driving task for both system-limit and system-malfunction failures (16.8% and 50.4% lower, respectively; see Table 4). However, the effect of failure type was only significant for no-TOR drives (participants looked at the non-driving task 31.7% less before the system-limit failure), similar to the findings for percent of time looking at the roadway.

For the system-limit failure, participants looked less at the non-driving task when there was a TOR (compared to no TOR), but did not spend more time looking at the roadway. In addition, for the system-malfunction failure, percent of time looking at the non-driving task was 50% lower but percent of time looking at the roadway was only 24% higher. Inspection of the raw data in Figure 8 indicates that in TOR drives, participants were allocating more attention to the dashboard, and were also allocating attention to the TOR.

Table 4. Results of contrasts for monitoring models. (SE = Standard Error, DF = Degrees of Freedom).

	Estimate	**SE**	**DF**	***t-value***	***p-value***	**95% confidence interval**	
Percent of Time Looking at Roadway (%)							
For limit failure: TOR vs. No TOR	0.41	6.05	10	.07	.95	-13.06	13.89
For malfunction failure: TOR vs. No TOR	23.98	6.29	10	3.81	**.003**	9.97	37.99
For no TOR: Limit vs. Malfunction	28.35	6.17	10	4.60	**.001**	14.61	42.09
For TOR: Limit vs. Malfunction	4.78	6.17	10	0.78	.46	-8.96	18.53
Percent of Time Looking at Non-driving Task							
For limit failure: TOR vs. No TOR	-16.80	6.42	10	-2.62	**.026**	-31.10	-2.51
For malfunction failure: TOR vs. No TOR	-50.43	6.67	10	-7.56	**< .001**	-65.29	-35.57
For no TOR: Limit vs. Malfunction	-31.69	6.54	10	-4.84	**< .001**	-46.26	-17.11
For TOR: Limit vs. Malfunction	1.94	6.54	10	0.30	.77	-12.64	16.51

Note: bold text indicates significance based on the Benjamini-Hochberg adjusted alpha (Benjamini & Hochberg, 1995).

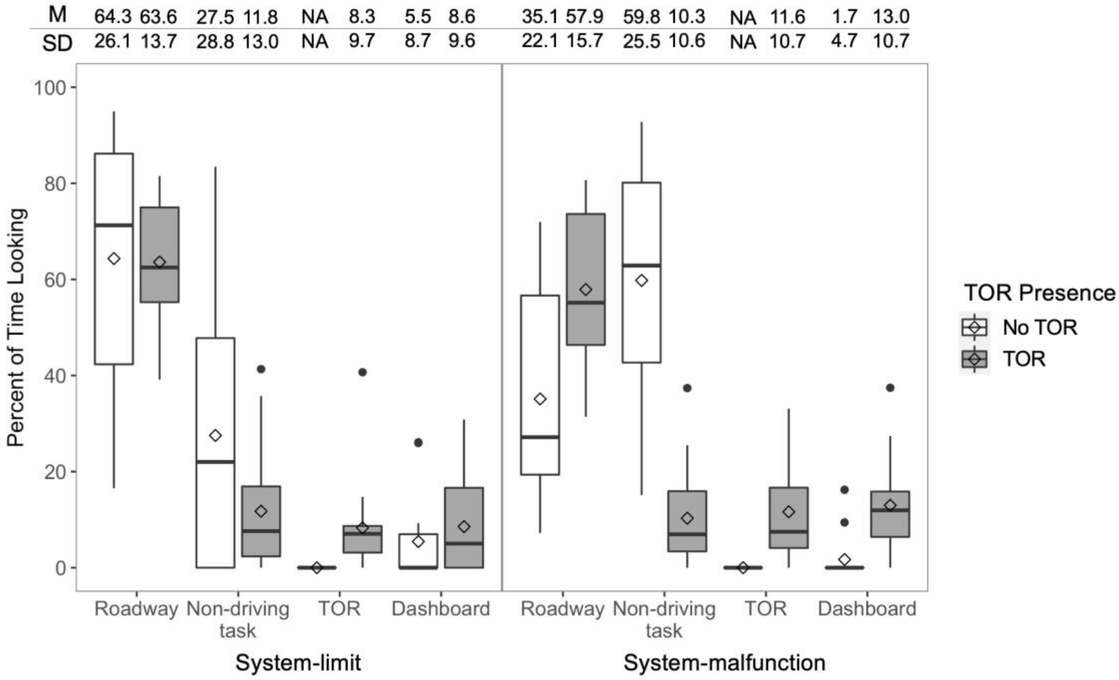

Figure 8. Raw data for percent of time looking at the roadway, non-driving task, TOR, and dashboard, by failure type and TOR presence. M and SD are presented along the top of the figure.

Discussion

We conducted a driving simulator study to analyze the effects of a TOR display on takeover time and drivers' gaze allocation (to the roadway and a non-driving task displayed on an in-vehicle display) during two types of automated vehicle failures: one in which there is information in the environment that indicates a potential failure (system-limit) and one in which there is no such information (system-malfunction). To the best of our knowledge, this is the first study to systematically compare the effects of TORs on drivers' monitoring and takeover performance based on whether or not there is environmental information that drivers can use to anticipate the failure.

Before system-malfunction failures, participants paid more attention to the roadway and less attention to the non-driving task, and took over sooner, when there was a TOR, compared to when there was no TOR. This result was expected as the purpose of the TOR was to reorient drivers' attention back to the roadway and let them know that they needed to take over, and for system-malfunction failures, there was no way for participants to anticipate the failure without being alerted through a TOR. For the system-limit failure, having a TOR did not significantly increase visual attention to the roadway, suggesting that there was no added benefit of the TOR in terms of monitoring for system-limit failures in this experiment. This result was presumably because participants were able to identify the potential occurrence of a failure based on their knowledge of the limitations and had already changed their monitoring behavior to look more at the roadway in order to prepare for it, and thus did not need a TOR to redirect their attention to the roadway. We cannot rule out the possibility that for some of the participants, increased visual attention to the roadway before the system-limit failure could be attributed to the additional visual information at the location of the system-limit failure compared to the system-malfunction failure (i.e., intersection vs. straight section of road), rather than an understanding of the automation's limitations. However, anecdotal evidence from the experimenter's observations indicates that many of the participants anticipated the failure at the intersection and were preparing to take over by placing their hands closer to the wheel even in drives without a TOR.

Although the TOR was not associated with increased monitoring of the roadway before system-limit failures, it was associated with a decrease in visual attention towards the non-driving task. An inspection of Figure 8 suggests that participants may have been paying more attention to the dashboard, as well as the TOR display, before a system-limit failure in TOR drives. Taken together, these findings suggest that participants were directing their attention towards driving-relevant information after the TOR, potentially because they were trying to locate further information on the status of the automation. In the case of the system-limit failure in our study, participants were already monitoring the roadway, so a TOR that only aimed to reorient their attention to the roadway had no added benefit in terms of monitoring. Displays presenting information about the automation status (e.g., Helldin et al., 2013; Seppelt & Lee, 2019) or the surrounding traffic (e.g., He, Kanaan, & Donmez, in press) have been found to be helpful to drivers in responding to and anticipating automation failures that require their intervention. Further research should be conducted on how best to present information from the surrounding environment to drivers in an automated driving context to support them in taking over control of the vehicle.

While the TOR only affected monitoring of the roadway for system-malfunction failures, it resulted in drivers taking over sooner for both the system-limit and system-malfunction failure. On average, participants took over 0.81 s sooner when there was a TOR, an even greater benefit than the 0.58 s observed in Zhang et al. (2019), which combined failures that had environmental information indicating a potential failure and those that did not. Of the studies reported in Zhang et al. (2019) that we could understand (one study was in German), all used more critical failures which would result in a collision with an obstacle had the participants not taken over. In our study, the failure was imminent, but there was no obstacle or vehicle in the driver's path that constituted a hazard. Future work could also investigate the effect of TORs in more critical scenarios with and without environmental information indicating a potential failure. In our study, for the system-limit failure, it is possible that without the TOR, participants suspected that a failure might occur but were unsure about whether it would actually happen, and the TOR confirmed that a failure was imminent, resulting in them taking over sooner. Thus, although the TOR may not have had additional benefits in terms of monitoring the roadway when there was environmental information that allowed drivers to anticipate the failure, our results suggest it was still useful as a confirmation that a failure may occur and a prompt to take over control of the vehicle.

For both system-limit and system-malfunction failures, we found higher variability in takeover times when a TOR was provided, which may be due to individual differences in reliance on TORs (e.g., some participants may rely on the TOR and only take over after the TOR, whereas others may take over as soon as they notice a potential upcoming failure) or strategies for using TORs (e.g., some people may take more time to take over and use the automation as long as possible because they know they have 6 s before a failure, whereas others may choose to take over immediately). Further research should investigate the factors that affect takeover time in the presence of TORs depending on whether environmental information indicating an upcoming failure is available, such as trust (Körber, Baseler, & Bengler, 2018) and prior experience with automation (Hergeth et al., 2017).

When a TOR was provided, we did not find any differences in monitoring between system-limit vs. system-malfunction failures, contrary to findings by Dogan et al. (2017). We suspect that this discrepancy is due to the fact that participants of Dogan et al.'s (2017) study were increasing their monitoring of the speedometer before system-limit failures as the speedometer was their source of information relevant to the system limit (i.e., speed above 50 km/h), whereas in our study, participants were increasing their monitoring of the roadway before system-limit failures as cues from the roadway (e.g., upcoming intersection) provided information relevant to system limits.

When no TOR was provided, participants took over faster and paid more attention to the roadway when experiencing system-limit compared to system-malfunction failures, which not surprisingly, is consistent with our previous findings (DeGuzman et al., 2020) based on a sample containing some of the same participants as in the current analysis. This result suggests that being able to anticipate a failure based on an awareness of the automation's limitations (if possible) and relevant information in the environment may be an effective way to increase driver monitoring of the roadway when TORs are not available. However, it should be noted that participants in this experiment received training on the vehicle's automation limitations so that they could anticipate system-limit failures. Had the participants not been trained on the limitations, they would not have been able to anticipate the system-limit failure. This is particularly important as it has been reported that a majority of drivers of automated vehicles do not read the vehicle manuals and are unaware of ACC limitations (Beggiato & Krems, 2013; Jenness, Lerner, Mazor, Osberg, & Tefft, 2008), rendering many of the automation's system-limit failures unpredictable to them. Thus, efforts are needed to improve the accuracy of drivers' mental models of the automation for them to experience the positive effects of training on monitoring behavior. Moreover, participants in our study had no prior experience with driving automation and were trained and tested on the limitations in the same session (lasting approximately 3 hours). Our results therefore reflect drivers' performance in their initial encounters with driving automation and TORs, in a best-case scenario in which the information learned in training is likely still retained by the participants. Even if drivers were required to complete a similar training exercise as in our study, research shows that knowledge about system limitations decays over time, particularly for limitations that are never experienced (Beggiato & Krems, 2013; Beggiato, Pereira, Petzoldt, & Krems, 2015). Thus, the effect of training in real-world situations may be reduced over time or limited only to a period shortly after the training. Future work is needed to investigate the long-term effects of knowledge of automation limits on monitoring behavior and takeover performance. In addition, failures in our study occurred more frequently (two failures during each six-minute drive) than they are likely to occur in real-world settings, and thus participants were likely more prepared for potential failures than a typical driver would be. Research is needed to determine how drivers respond to TORs depending on the information available in the environment when these failures occur at a frequency more similar to that of a real-world setting.

Finally, this study was conducted with a limited sample of 19 participants whose ages ranged from 25 to 30 years old, and thus our results are specific to a small portion of the population and should be interpreted with caution given the small sample size. Future work should investigate whether the observed findings can be generalized to different age groups and larger sections of the population. Further, as the study was conducted using a driving simulator like other studies on the emerging technologies of automated driving, the findings might not reflect the actual behaviors of drivers on the road. On-road and naturalistic driving studies can help provide further insights about drivers' monitoring of the environment and their responses to different types of automation failures.

In conclusion, we have investigated the effects of TORs on drivers' monitoring and takeover performance in automated vehicles when experiencing failures with and without environmental information that can be used to anticipate them (system-limit and system-malfunction failures, respectively, in our study). We found that TORs were associated with improved monitoring of the roadway only before the system-malfunction failure, and faster takeover time for both the system-limit and system-malfunction failure. Our findings suggest that when drivers are trained on automation limits, their visual attention is already directed to the roadway in the moments leading up to a system-limit failure, and so providing a TOR has no added benefit for monitoring when there is information in the environment that indicates a potential failure. However, we have noted that drivers may not always be well trained about automation limits or may forget about them if they are not experienced often. In that case, future research should focus on combining TORs with displays that inform drivers about system limitations. In the case of the system-malfunction failure, for which there was no environmental information indicating a potential failure, providing a warning to the driver that the automation will be disengaged can improve monitoring; however, providing such warnings may not always be feasible for all system malfunctions. For situations in which it is not feasible to provide a warning for an upcoming system malfunction, drivers may benefit from in-vehicle displays that provide information about hardware reliability and functioning, in addition to more transparent information about the operation of the automated system's software.

Acknowledgements

This work was supported by an NSERC Discovery grant awarded to Birsen Donmez. We would like to thank Yueqi Hu for helping to process the raw gaze data. The authors acknowledge that they have no conflicts of interest for this work.

References

Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*, *18*, 47–57.

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. *Transportation Research Part F: Traffic Psychology and Behaviour*, *35*, 75–84.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Brandenburg, S., & Chuang, L. (2019). Take-over requests during highly automated driving: How should they be presented and under what conditions? *Transportation Research Part F: Traffic Psychology and Behaviour*, *66*, 214–225.

Carsten, O., Lai, F. C. H. H., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control task substitution in semiautomated driving: Does it matter what aspects are automated? *Human Factors*, *54*(5), 747–761.

Chen, H. Y. W., Hoekstra-Atwood, L., & Donmez, B. (2018). Voluntary- and involuntary-distraction engagement: An exploratory study of individual differences. *Human Factors*, *60*(4), 575–588.

de Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196–217.

DeGuzman, C. A., Hopkins, S. A., & Donmez, B. (2020). Driver takeover performance and monitoring behavior with driving automation at system-limit versus system-malfunction failures. *Transportation Research Record: Journal of the*

Transportation Research Board, *2674*(4), 140–151.

Dikmen, M., & Burns, C. M. (2016). Autonomous driving in the real world: Experiences with Tesla Autopilot and Summon. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interative Vehicular Applications* (pp. 225–228). Ann Arbor, MI.

Dogan, E., Rahal, M. C., Deborne, R., Delhomme, P., Kemeny, A., & Perrin, J. (2017). Transition of control in a partially automated vehicle: Effects of anticipation and non-driving-related task involvement. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 205–215.

Donmez, B., Boyle, L. N., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. *Accident Analysis and Prevention*, *39*(3), 581–590.

Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: Noncritical transitions to and from manual control. *Human Factors*, *59*(4), 689–705.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London: SAGE Publications.

Gaspar, J., Carney, C., Shull, E., & Horrey, W. J. (2020). *The Impact of Driver's Mental Models of Advanced Vehicle Technologies on Safety and Performance*. Washington, DC. Retrieved from https://aaafoundation.org/wp-content/uploads/2020/06/AAAFTS-SaferSim_Iowa_Yr1_Report-FINAL.pdf

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). Take over! How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1938–1942.

He, D., Kanaan, D., & Donmez, B. (in press). In-vehicle displays to support driver anticipation of traffic conflicts in automated vehicles. *Accident Analysis and Prevention*, *149*, 105842.

Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013). Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 210–217. Eindhoven, The Netherlands: ACM.

Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. *Human Factors*, *59*(3), 457–470.

Hopkins, S. A., DeGuzman, C. A., & Donmez, B. (2019). The Effect of failure predictability on driver takeover performance and monitoring behavior in partially automated vehicles. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 63*(1), 2109–2109.

Jenness, J. W., Lerner, N. D., Mazor, S., Osberg, J. S., & Tefft, B. C. (2008). *Use of advanced in-vehicle technology by young and older early adopters: Survey results on adaptive cruise control systems* (Report No. DOT HS 810-917). U.S. Department of Transportation, National Highway Traffic Safety Administration. Retrieved from https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/dot-hs-810-917.pdf

Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, *66*, 18–31.

Kraft, A. K., Naujoks, F., Wörle, J., & Neukum, A. (2018). The impact of an in-vehicle display on glance distribution in partially automated driving in an on-road experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *52*, 40-50.

Louw, T., Kuo, J., Romano, R., Radhakrishnan, V., Lenné, M. G., & Merat, N. (2019). Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures. *Transportation Research Part F: Traffic Psychology and Behaviour*, *62*, 870–882.

Mazda. (2019). *Statement on Recall of Certain 2019-2020 MY Mazda3 Vehicles*. Retrieved from https://insidemazda.mazdausa.com/press-release/statement-on-recall-of-certain-2019-2020-my-mazda3-vehicles/

Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., & Bauer, W. (2015). Take-over requests for automated driving. *Procedia Manufacturing*, *3*, 2867–2873.

Merrikhpour, M., & Donmez, B. (2017). Designing feedback to mitigate teen distracted driving: A social norms approach. *Accident Analysis and Prevention*, *104*, 185–194.

National Transportation Safety Board. (2020). *Collision between a sport utility vehicle operating with partial driving automation and a crash attenuator, Mountain View, California, March 23, 2018* (Highway Accident Report NTSB/HAR-20/01). Retrieved from https://www.ntsb.gov/investigations/AccidentReports/Reports/HAR2001.pdf

Naujoks, F., Mai, C., & Neukum, A. (2014). The effect of urgency of take-over requests during highly automated driving under distraction conditions. In N. Stanton, S. Landry, G. Di Bucchianico, & A. Vallicelli (Eds.), *Advances in Human Aspects of Transportation: Part I* (pp. 431–438). AHFE Conference.

Payre, W., Cestac, J., Dang, N.-T., Vienne, F., & Delhomme, P. (2017). Impact of training and in-vehicle task performance on manual control recovery in an automated car. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 216–227.

Richardson, N. T., Flohr, L., & Michel, B. (2018). Takeover requests in highly automated truck driving: How do the amount and type of additional information influence the driver–automation interaction? *Multimodal Technologies and Interaction*, *2*(4), 68.

Roche, F., Somieski, A., & Brandenburg, S. (2018). Behavioral changes to repeated takeovers in highly automated driving: Effects of the takeover-request design and the nondriving-related task modality. *Human Factors, 61*(5), 839-849.

Seppelt, B. D., & Lee, J. D. (2019). Keeping the driver in the loop: Dynamic feedback to support appropriate use of imperfect vehicle control automation. *International Journal of Human-Computer Studies*, *125*, 66–80.

Shen, S., & Neyens, D. M. (2017). Assessing drivers’ response during automated driver support system failures with non-driving tasks. *Journal of Safety Research*, *61*, 149–155.

Tippey, K. G., & Longnecker, M. T. (2016). An ad hoc method for computing pseudo-effect size for mixed models. In *Proceedings of the South Central SAS Users Group Educational Forum*. San Antonio, TX.

Wan, J., & Wu, C. (2018). The effects of lead time of take-over request and nondriving tasks on taking-over control of automated vehicles. *IEEE Transactions on Human-Machine Systems*, *48*(6), 582–591.

Yoon, S. H., Kim, Y. W., & Ji, Y. G. (2019). The effects of takeover request modalities on highly automated car control transitions. *Accident Analysis and Prevention*, *123*, 150–158.

Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? the impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis and Prevention*, *92*, 230–239.

Zhang, B., De Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 285–307.