It’s about time! Earlier take-over requests in automated driving enable safer responses to conflicts

Transportation Research Part F: Psychology and Behaviour 86 (2022) 196–209

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

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

It's about time! Earlier take-over requests in automated driving enable safer responses to conflicts

Linda Pipkorn [a,*], Emma Tivesten [b], Marco Dozza [a]

[a] *Chalmers University of Technology, Göteborg, Sweden*
[b] *Volvo Cars Safety Centre, Volvo Cars, Göteborg, Sweden*

A R T I C L E I N F O

Keywords:
Automation
Automated driving
Take-over request
Response process
Driver response

A B S T R A C T

Automated driving (AD), which takes full responsibility for the driving task in certain conditions, is currently being developed. An important concern in AD is how to design a take-over request (TOR) that mitigates automation effects (e.g., delayed responses to conflict scenarios) that previous literature from simulator experiments has shown can occur. To address this concern, this study aims to investigate and compare driver responses to TORs and a lead-vehicle cut-out scenario under three conditions: (1) after a period of AD with a TOR issued early (18 s time-to-collision), (2) same as (1) except with a TOR issued late (9 s time-to-collision), and (3) baseline, with adaptive cruise control (ACC). This paper also compares the results to those of a previous study using the same conflict scenario but with near-perfect assisted driving system (SAE Level 2). The lead-vehicle cut-out scenario was encountered on a test track after 30 minutes driving with either ACC or AD. In AD the TOR was issued prior to the conflict object was revealed to the participants when the lead vehicle performed the cut-out (at conflict onset). This TOR strategy differed from previous driving-simulator studies that issued the TOR at conflict onset. The participants had to respond by steering and/or braking to avoid a crash. Our findings show that, independent of TOR timing, the drivers required similar amounts of time to 1) direct their first glance to the human–machine interface, 2) look forward, 3) end their secondary task, 4) put their hands on the steering wheel, and 5) deactivate automation. However, when the TOR was issued early rather than late, they started to brake earlier (even before conflict onset). All participants successfully managed to avoid crashing with the object, independent of the condition. AD with an early TOR resulted in the earliest response, while ACC drivers responded slightly earlier than the drivers in AD with the late TOR. Our findings do not support the findings of severe automation effects in previous driving-simulator studies. One reason for the difference is that when a TOR is issued prior to conflict onset, drivers are given the time needed for their preparatory actions (e.g., placing hands on the wheel, deactivating AD) that is not needed when driving with ACC or in manual driving (baseline), before having to respond to the conflict scenario. Thus, at conflict onset the drivers in AD are as ready to act (hands on wheel, eyes forward) as the drivers in the baseline and can perform an avoidance manoeuvre similar as to the baseline drive. Overall, the present study shows that AD does not need to end up in a highly critical situation if the TOR is issued early enough. In fact, AD with an early TOR may be safer than driving with ACC, because in the former drivers are more likely to brake earlier in preparation for the conflict. Finally, a TOR clearly communicates the need for drivers to resume manual control and handle potential events

* Corresponding author at: Chalmers University of Technology, Department of Mechanics and Maritime Sciences, Division of Vehicle Safety, Hus Saga, Hörselgången 4, floor 4, 412 96 Gothenburg, Sweden.
E-mail address: linda.pipkorn@chalmers.se (L. Pipkorn).

https://doi.org/10.1016/j.trf.2022.02.014
Received 16 April 2021; Received in revised form 2 January 2022; Accepted 13 February 2022
Available online 8 March 2022
1369-8478/

when AD has been deactivated. In our study, once the drivers had taken control, they clearly understood their responsibilities to respond to the conflict, in contrast to a previous study with a similar, near-perfect assisted driving system.

1. Introduction

Whereas vehicles with SAE Level 1 driving automation (e.g., Adaptive Cruise Control: *ACC*) have been on our roads for some time, more advanced SAE Level 2 driving automation (*assisted driving* systems) is becoming increasingly available in commercial vehicles. While ACC can support the driver with longitudinal control, an assisted driving system can support the driver with both longitudinal and lateral control simultaneously. Importantly, ACC and assisted driving systems are only assisting the driver with parts of the driving task: the driver always remains responsible for safe driving (SAE International, 2018). In contrast, the next generation of *automated driving* (AD) is intended to take full responsibility for the driving task when activated, without the need for driver supervision (Thatcham Research, 2019). However, the driver is assumed to respond appropriately and resume manual driving when a take-over request (TOR) is issued. A crucial concern for designing AD is how to ensure that the TOR will guide drivers to resume manual driving safely. Addressing this concern is important to develop safe AD and ensure safe AD through regulation (e.g., coming regulations on "Automated Lane Keeping System"; United Nations Economic Comission for Europe: UNECE, 2021).

1.1. Assessing automation effects in assisted and automated driving

As passenger vehicles are increasingly automated, there is a need to understand and address the human factors challenges in order to achieve safe AD (Seppelt & Victor, 2016). One such challenge is the deteriorated driver responses due to the automation itself, known as *automation effects* (e.g., delayed response or crashing in conflict scenarios during assisted driving or after AD). For example, Victor et al. (2018) found that about one-third of the drivers crashed with a conflict object in a lead-vehicle (LV) cut-out scenario ("Euro NCAP - Cut-out scenario," 2021) after having supervised a near-perfect assisted driving system for 30 min. The drivers who crashed reported that they had expected the automated vehicle to act, and therefore did not realize they needed to—or realized it too late (Gustavsson et al., 2018; Victor et al., 2018). Importantly, the conflict scenario was not preceded by any vehicle warning (such as a forward collision warning) about the need to act. It is unclear whether these results apply to AD; although drivers do not need to supervise the system, they would receive a TOR prior to the need to resume manual control. In fact, these effects may depend on the *driving mode* (e.g., ACC, Assisted driving, or AD), since different system capabilities change the driver's responsibilities in different ways (SAE International, 2018).

Drivers' responses to TORs have been extensively studied in virtual environments (fixed-based or moving-based driving simulators) within the last decade (McDonald et al., 2019; Zhang et al., 2019). The most used metric to assess drivers' response to TORs is the take-over time, the time needed to deactivate automation in response to the TOR by braking, steering, or pressing buttons (Zhang et al., 2019; McDonald et al., 2019). Some studies have also included other response times as metrics, such as the time after a TOR needed to: (a) redirect the gaze away from a non-driving related task (e.g., Gold et al., 2013), (b) redirect the gaze toward the road ahead (Eriksson et al., 2018; Gold et al., 2013; Zeeb et al., 2017), (c) place hands on the steering wheel (Eriksson et al., 2018; Gold et al., 2013; Wandtner et al., 2018), and (d) glance towards mirrors (Gold et al., 2013; Vogelpohl et al., 2018). However, TOR response times alone are not sufficient to detect and quantify automation effects: a driver who deactivates automation rapidly may still be slow to respond to an event (Louw et al., 2017). Whereas many studies have investigated the influence of certain factors on take-over times and driving performance after the TOR, only a few studies have included a baseline of manual driving in order to explicitly investigate the presence of automation effects (McDonald et al., 2019). These few studies reported delayed response and degraded manual driving performance compared to the manual baseline (Gold et al., 2013; Happee et al., 2017; McDonald et al., 2019). In a study of AD by Gold et al. (2013), for example, when drivers were given 5 or 7 s to respond to a conflict scenario, their degraded performance included lane excursions, delayed steering response, and/or accelerations increased two to three times. Understanding how to mitigate or avoid these automation effects is crucial to obtaining safe vehicle automation, in which drivers perform at least as well as they do when driving manually.

1.2. The influence of take-over request time budgets and take-over time on automation effects

One important factor affecting take-over times and automation effects is the take-over request time budget (the time from the TOR until the conflict object is reached). In a literature review, McDonald et al. (2019) conclude that the lower the time budget, the greater the automation effects. The authors also suggest that the take-over time increases linearly as a function of the TOR time budget: a 1-s increase in time budget corresponds to a 0.27-s increase in take-over time. Importantly, most of the studies in the review issued a TOR when the situation was already critical. In fact, most studies issued a TOR, with a 7-s time budget, at the same time the conflict object appears to the drivers (the *conflict onset*; McDonald et al., 2019; Zhang et al., 2019). Most simulator studies do not have a deactivation strategy that is separate from the TOR strategy: drivers' conflict response (braking or steering) deactivates the AD system. In other words, it is unclear whether this relationship between take-over time and take-over time budget exists for longer take-over time budgets (i.e., if a TOR is issued before the situation is critical) and deliberate deactivation strategies (i.e., when a driver needs to deactivate the system before responding to an event).

When taking over the driving task from AD, drivers need to perform certain *preparatory actions* before they are ready to respond to a conflict scenario. Examples of these actions are: looking towards the forward roadway, placing the hands on the steering wheel and/or feet on the pedals, and deactivating AD. In contrast, manual drivers can typically respond immediately at the time of conflict onset since they are likely to have their hands on the wheel and their eyes on the road already. The need to perform these preparatory actions is an important confounding factor, because of the relationship that exists between the take-over time required to perform them and the time remaining for drivers to respond to the conflict. As a result, the avoidance manoeuvre is more delayed—and harsher. In fact, some researchers have pointed out that the effects of AD on drivers' manual driving performance may be due solely to the longer take-over times in AD compared to manual driving (see Happee et al., 2017; McDonald et al., 2019). In other words, the most important contributing factor behind the automation effects observed in previous studies may be the more advanced preparatory actions needed to become ready-to-act after automation compared to manual. No previous studies have explicitly investigated the presence of automation effects when drivers are explicitly given extra time for these preparatory actions before conflict onset. Further, most previous studies collected TOR data in driving simulators that have not been validated with real world data for this application.

1.3. Research needs, aims, and questions

To better understand the influence of driving mode, TOR timings, and TOR strategy on automation effects, we performed a test-track experiment to determine whether automation effects would be present for: (a) an adaptation of the LV cut-out scenario performed by Victor et al. (2018), with AD as the driving mode instead of the near-perfect assisted driving system they tested and (b) two different TOR timings (i.e., 9-s or 18-s time budget) for a TOR strategy that gave drivers time for their preparatory actions before conflict onset. We hypothesized that such a strategy would mitigate the automation effects, resulting in drivers who respond to the conflict similar to those in the ACC baseline. Further, the present study aims to (1) address the dearth of TOR studies conducted in realistic environments by utilizing a test track rather than a simulator and (2) consider the complete driver response process, including the response time to the TOR for several driver actions and the subsequent manual driving performance, instead of a single take-over time.

Three specific research questions were asked:

1) What is the influence of early vs late TOR timing (i.e., time budget of 9 vs 18 s) on the driver response process when the TOR is issued before conflict onset?
2) Are automation effects present for AD when compared to an ACC baseline, and how does this compare to previous driving-simulator studies?
3) Given the conflict situation described by Victor et al. (2018), are the automation effects (crashing) previously observed for a near-perfect assisted driving system also present for AD?

2. Methods

2.1. Participants

The participants (N = 56) were Volvo Cars employees who had no work duties associated with AD development, did not work as test drivers, and had not been part of a similar study before. Further, all participants had driven at least 5,000 km during the last year. Eight participants were excluded from the analysis due to missing data or an inability to maintain the time headway selected for the experimental protocol (a fixed time headway between the test and lead vehicles was required to obtain the same conflict criticality at the TOR). The final sample size consisted of 48 participants (45% females), aged between 22 and 56 years (M = 38, SD = 10). Each participant was randomly assigned to one of the three conditions: 15 participants drove with ACC, 17 drove with AD and received the late TOR, and 16 drove with AD and received the early TOR. All participants signed a consent form before participation. The study was reviewed and approved by the national ethical review board in Gothenburg, Sweden (Dnr: 2019–01827).

2.2. Testing environment and equipment

The testing environment was a rural-road test track with two lanes, located outside Gothenburg, Sweden (AstaZero, 2020). Two vehicles were used in the study: a Volvo XC90 as the test vehicle (TV) and a robot-controlled Volvo XC60 as the lead vehicle (LV). The

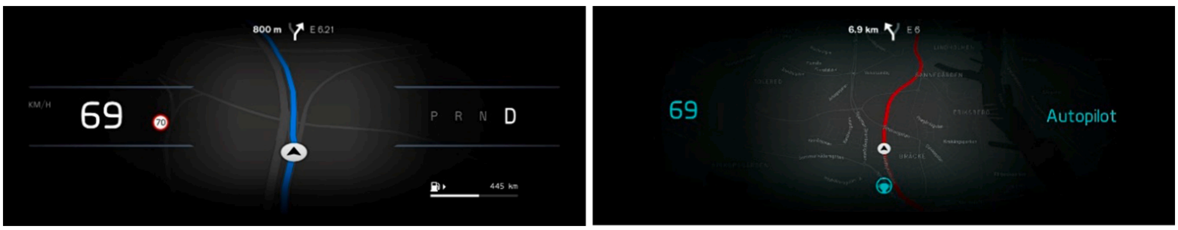

Fig. 1. The HMI display in ACC and manual mode (left) and in AD mode (right).

TV was rebuilt to simulate AD using a W*izard-of-Oz* protocol (Green & Wei-Haas, 1985; Wang, Sibi, Mok, & Ju, 2017). The TV included an extra steering wheel and set of pedals, in the middle of the back. Both the steering wheel and the pedals were hidden from the participant in the driver seat. This setup enabled the *Wizard* to control the vehicle from the backseat when the simulated AD was engaged, giving the participant the impression that AD controlled the vehicle. The role of the wizard was to perform all parts of the driving task when AD was active. The Wizard was assisted in both lateral and longitudinal control by an in-production system called *Pilot Assist*.

A tablet was mounted on top of the center stack in the TV which the drivers could use when AD was activated. The TV was also equipped with a custom human–machine interface (HMI) in the instrument cluster, which provided the driver with information on the driving mode, as shown in Fig. 1. Fig. 1 (left) shows the HMI view when the driving mode changed from AD to manual. When ACC was activated, the HMI view was the same, except a small icon was visible in the bottom left corner and the set ACC speed was shown next to the current vehicle speed. The TV was equipped with a DeweSoft data logger that recorded vehicle signals (100–200 Hz), GPS data (100 Hz), HMI signals specific to the Wizard-of-Oz setup, and video data. The video data were collected using three cameras that recorded video (20 Hz) of the drivers' face and upper body and the forward roadway.

2.2.1. *Activating and deactivating AD*

When AD was available, the system notified the driver with an audio tone and a message in the instrument cluster reading, "Autopilot available". The driver could then press two buttons on the steering wheel for 0.6 s to activate the AD system (i.e., giving control to the Wizard). The driver received feedback when AD was activated by a voice saying, "Autopilot active" and an updated HMI view (Fig. 1, right).

The AD system notified the driver about the need to resume manual control with a TOR, consisting of an audio tone and a message on the instrument cluster reading, "Autopilot ending" (Fig. 2, left). The participants had 6 s to deactivate AD (this time was visible in the instrument cluster as a shrinking red bar; see Fig. 2, left). Deactivation was performed by pressing the same two buttons previously used to activate the system for 0.6 s. Once the buttons were pressed, the time remaining was visible on the instrument cluster as two turquoise bars approaching each other and meeting when the deactivation was completed (Fig. 2, right). When AD was deactivated, the control shifted from the Wizard to the participant, the HMI changed to the manual driving mode view (Fig. 1, right), and a voice said, "Drive the car".

2.3. *Study design*

The study had a between-subject design with three conditions: driving with ACC (the *ACC baseline*), driving with AD, and receiving a late TOR at 9 s TTC (the *TOR9* condition), and AD with an early TOR issued at 18 s TTC (the *TOR18* condition). The selection of TTC (calculated as distance divided by speed) was based on two criteria. The most critical condition (9 s time budget) was selected to give drivers enough time to finish their preparatory actions before the conflict onset. Given a previous study (reported in Pipkorn et al., 2021a) it took up to 6 s for drivers, with a similar Wizard of Oz vehicle and TOR procedure, to deactivate AD. Therefore, to make sure the participants had enough time to deactivate AD before conflict onset, the TOR was issued 6 s before the conflict onset which occured at 3 s TTC. Together, this corresponds to a TOR issued at 9 s TCC. The less critical condition (18 s) was selected to give participants the chance to deactivate AD well before the conflict onset. The independent variables were driving mode (ACC vs. AD) and TOR timings (9 s vs. 18 s TTC). On a high level we were interested in the influence of these independent variables on the drivers' response process within a conflict scenario. The response process consisted of drivers' response to TOR (applicable for AD) and drivers' response to the conflict scenario (applicable for AD and ACC). The dependent variables are defined in detail in Section 2.5.2.

The conflict scenario used in the present study was replicated from a previous study reported in Victor et al. (2018) and Pipkorn et al. (2021b): a longitudinal cut-out scenario occurred after 30 min of driving, with a stationary balloon car as the conflict object (see Fig. 3). The object was first visible approximately 11 s before reaching it, when the TV was going through a curve and over a crest. The object then became visually obstructed again by the LV when the road straightened out and the TV was about 8 s from the conflict object. Finally, the object became visible again when the LV performed the cut-out maneuver (i.e., the conflict onset) and the TV was about 3 s away from the conflict object. The AD system did not steer or brake, so the driver had to steer or brake to avoid a crash.

Fig. 2. The HMI view for the TOR (left). The HMI view when the two steering wheel buttons are being pressed; the turquoise bars are moving toward each other (right).

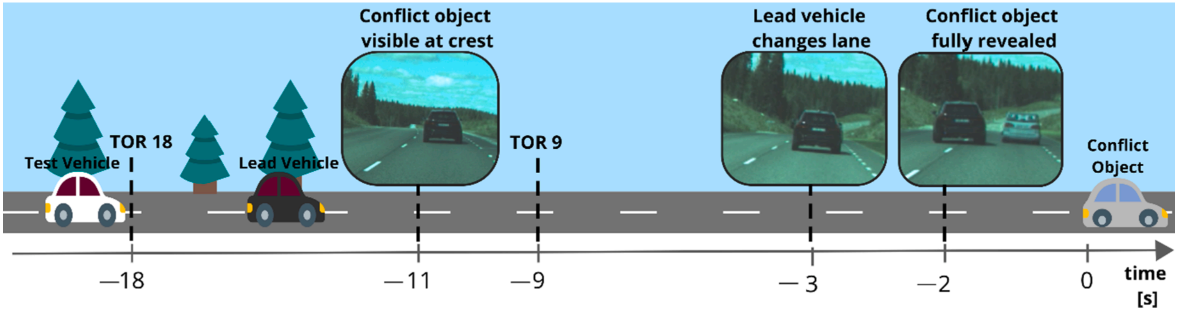

Fig. 3. Timeline from 20 s before TV reached the conflict object.

2.4. Study procedure

Upon arrival, all participants were asked to read an information sheet about the study. They were informed that the study's purpose was to understand the user's experience during AD. The TOR9 and TOR18 participants were informed about their right to: (a) bring items (e.g., magazine, notebook, or phone) of their choice to use during AD, and (b) terminate the test at any time. Importantly, the participants were also instructed to be prepared to respond to TORs and resume manual driving at any time during the drive. After receiving the information, all participants signed an informed consent form and then sat in the TV driver's seat. Two other persons were seated in the back during the test: the test leader and the Wizard. The test leader instructed the participants and administered a short questionnaire after the participants had experienced the conflict. In fact, two questions, previously asked in Victor et al. (2018), were asked to gain insight into the participants' cognitive process (e.g., perception, understanding of their responsibilities) in the conflict scenario. The first question was: "Did you perceive the conflict object prior to the lead vehicle performing the cut-out?" The second was: "Did you realize that you had to act to avoid crashing with the conflict object?" Both questions were answered verbally (yes/no/uncertain) by the participant, and the responses were noted by the test leader.

Before the actual test started, the participants in the TOR9 and TOR18 conditions practised activating and deactivating AD several times, both at a standstill and during a short drive. The only instruction given to the ACC participants was to select the specified time headway and then drive as they normally would. Independent of condition, the TV followed the LV, which was programmed to follow a pre-defined path and speed profile (70 km/h on straight road segments and 50 km/h on curves). The headway between the TV and the LV was set to 2 s for all three conditions, using the in-production driver assistance systems; consequently, the TV maintained the same speeds. The in-production ACC system was used in ACC mode. While in AD mode, the participants in the TOR9 and TOR18 conditions were free to engage in secondary tasks of their choice, including using the center-stack-mounted tablet.

Table 1
Coding scheme for the response process and conflict scenario.

Coded variables	*Time point or glance location*
Take-over request (TOR)	When the TOR was issued. This was assessed using a binary signal that switched from 0 to 1 when the TOR was issued.
Automation deactivated	When the AD was deactivated. This was assessed using a binary signal that switched from 1 to 0 when AD was deactivated.
2nd try to deactivate automation	When the second button press started; coded only if the first button press did not result in a successful deactivation of AD.
Glance location at TOR	The location of the glance when the TOR was issued: FWD if the driver was glancing forward, TASK if the driver was glancing towards a secondary task device/object, and OTHER if the glance was directed elsewhere. The glance location was assessed manually by determining drivers' gaze direction from video.
First glance to HMI	The onset of the first glance towards the HMI after the TOR. The time was assessed manually by determining drivers' gaze direction from video.
First glance forward	The onset of the first glance forward (onto the forward path or the conflict object) after the TOR. The time was assessed manually by determining drivers' gaze direction from video.
Hands on wheel	When the driver touched the steering wheel with at least one hand or part of a hand. Contact time was determined based on manual assessment of the video view of the driver.
End of secondary task engagement	The end of secondary task engagement (e.g., driver moved glance away from the tablet or let go of the mobile phone). The time was determined based on manual assessment of the video view of the driver.
Onset of last on-path glance	When the driver kept the eyes forward or on the threat constantly until reaching the conflict object. The time was determined based on manual assessment of the video view of the driver.
First brake	When the driver pressed the brake pedal the first time. The time was determined by a binary signal that switched from 0 to 1 when the brake pedal was pressed.
Start of steering intervention	When the driver started performing a voluntary steering maneuver to avoid the conflict object. The time was determined by manual assessment of the video, with corroboration from the steering wheel angle signal.
LV cut-out	When the lead vehicle showed the first sign of the cut-out maneuver. The time was determined based on manual assessment of the video view of the forward roadway.
Reaching conflict object	When the test vehicle reached the conflict object. The time was determined based on a vehicle signal measuring the longitudinal distance between the front of the test vehicle and the closest part of the conflict object (i.e., the signal was 0 when the TV reached the conflict object).

2.5. Data processing and coding

Video segments from 30 s prior to the conflict to 10 s after the object was reached were extracted to examine the driver response process. The segments were then observed for each participant, to assess crash involvement (i.e., crash or no crash) and code the times of different actions in the driver response process as well as times relevant for the conflict (Table 1). Further, vehicle data (speed, lateral and longitudinal accelerations) and GPS data (longitudinal distance to conflict object and lateral distance to road centre) were collected for a 300-meter interval from 200 m before the conflict object to 100 m after.

2.6. Data analysis

The coded variables (see Table 1) represent important events, used to examine the driver response process. The response process was examined in different ways by anchoring different actions to three different variables: TOR, LV cut-out, and Reaching conflict object.

2.6.1. The drivers' response to take-over request

To understand how TOR timings influence the driver response process, response times were calculated by anchoring the following actions to the TOR: First glance to HMI, First glance forward, Hands on wheel, Automation deactivated, 2nd try to deactivate automation, End of secondary task engagement, Onset of last on-path glance, First brake, and Start of steering intervention. The response times were then visualized in scatter plots for each participant, with vertical lines marking TOR, LV cut-out, and Reaching conflict object. To demonstrate the timing variability across subjects (e.g., due to braking), shaded areas marking the range (i.e., maximum and minimum time points) were created for the LV cut-out and the Reaching conflict object timings. The glance location at the TOR was also included as a string, coded as TASK if the glance was related to the secondary task, FWD if the participant was looking forward, and OTHER if the participant was looking somewhere else (e.g., down, out the side window).

2.6.2. The drivers' conflict response

The overall driving performance when passing the conflict object was visualized using vehicle data and GPS data: vehicle speed, longitudinal acceleration, lateral acceleration, and the lateral distance between the right side of the test vehicle and the road centre were plotted against the longitudinal distance to the conflict object. In addition, the maximum absolute lateral and longitudinal acceleration were computed for each condition within the 300-m distance interval (200 m before and 100 m after the conflict object). Each maximum absolute acceleration was summarized with a mean and standard deviation. Descriptive statistics (i.e., counts) were used to summarize the drivers' responses to the two post-drive questions. To establish how far away from the conflict object the participants were when they started to steer, the start of the driver steering intervention was anchored to Reaching conflict object (i.e., the front of the TV reached the rear of the balloon vehicle at 0 s). We refer to this interval as the *steering response time to the conflict object*. Then, to establish how long it took for drivers to react to the LV cut-out, the start of the driver steering intervention was anchored to LV cut-out (i.e., the LV initiated the cut-out at 0 s). We refer to this interval as the *steering response time to LV cut-out*.

2.6.3. Statistical analysis

To assess the influence of take-over timing and driving mode on the driver response process, a set of *Bayesian generalized linear models* was fit to the response times previously defined in Sections 2.5.1 and 2.5.2. A generalized linear model is a type of linear regression that does not require likelihood functions to be Gaussian (i.e., lognormal; McElreath, 2016). The Bayesian framework was chosen over the more traditional frequentist paradigm; it enables a more informative estimation of model parameters and the contrasts between them, since it includes the uncertainty of the estimated parameters and contrasts (Kruschke, 2014).

The general formula for all the models is represented in (1).

$$y \sim \log N(\mu, \sigma^2), \mu = X\beta \tag{1}$$

In Eq. (1), y is the *dependent variable*, X is the predictor (or *independent variable)*,β is the parameter vector, including the estimated parameters. All response times were modelled as lognormal due to their right-tailed nature (Eriksson & Stanton, 2017). Further, the mean μ represents the lognormal distribution's mean, represented by the linear predictor $X\beta$, with a variance σ^2.

2.6.3.1. The driver response time to take-over request models.

One model was fit to each of the following response times to the TOR: first glance to HMI, first glance forward, hands on wheel, automation deactivated, and onset of last on-path glance. For these models, $\beta = [\beta_0, \beta_{TOR18}]$ and $X = [1, X_{TOR18}]$ where X_{TOR18} is a dummy-coded vector (0 = TOR9, 1 = TOR18). β_0 is the global intercept which, in this case, corresponds to the late TOR mean (μ_{TOR9}), and β_{TOR18} represents the deviation from the intercept of TOR18 (μ_{TOR18}). The following priors were placed on the parameters: $\beta_0 \sim N(0.5, 1), \beta_{TOR18} \sim N(0, 1), \sigma \sim \mathit{halfN}(1)$. The priors were selected using *prior predictive checks*: we generated 2000 samples from the prior joint distribution to ensure that the prior was appropriate (i.e., could generate reasonable samples).

2.6.3.2. The steering response time models.

One model was fit to each of the following response times: steering response time to the conflict object and steering response time to LV cut-out. For these models, $\beta = [\beta_0, \beta_{TOR9}, \beta_{TOR18}]$ and $X = [1, X_{TOR9}, X_{TOR18}]$ where X_{TOR9} was a dummy-coded vector (0 = ACC, 1 = TOR9, 0 = TOR18) and X_{TOR18} was also a dummy-coded vector (0 = ACC, 0 = TOR9, 1 =

TOR18). β_0 is the global intercept which, in this case, corresponds to the ACC mean (μ_{ACC}). β_{TOR9} and β_{TOR18} represent the deviations from the intercepts of TOR9 ($\mu_{TOR9} - \mu_{ACC}$) and TOR18 ($\mu_{TOR18} - \mu_{ACC}$), respectively. The following priors were placed on the parameters: $\beta_0 \sim N(0.5, 1), \beta_{TOR9}, \beta_{TOR18} \sim N(0, 1), \sigma \sim halfN(1)$.

2.6.3.3. Bayesian inference and group comparison. The analyses were performed using Python ver. 3.7.6 and the probabilistic programming library PyMC3 ver. 3.9.3 (Salvatier et al., 2016). The Markov Chain Monte Carlo algorithm No-U-Turn Sampler (NUTS) was used to fit the models (Hoffman & Gelman, 2014). All models were fitted with four Markov Chain Monte Carlo chains. 5000 samples were drawn from the posterior distribution for each chain (after 2000 samples had been used for tuning the sampler and then discarded). The model convergence was verified through (a) visual inspection of the generated trace plots and (b) the obtained Gelman-Rubin R hat (Gelman & Rubin, 1992) statistic that should be close to 1. Finally, each model's goodness-of-fit was assessed by comparing the posterior predictive distribution against empirical data (i.e., a posterior predictive check; results of this procedure together with the R hat statistic can be found in the Supplementary material). The posterior predictive distribution is the distribution of new values predicted by the model, given the old data.

The models were summarized with a mean, standard deviation, and 95% Highest Posterior Density (HPD) interval for each of the model parameters' posterior distributions (the tables can be found in the Supplementary material). To assess the influence of take-over timing and driving mode on the driver responses to the TOR and the LV cut-out, we computed the posterior distributions for mean response times and the difference in means in line with Kruschke (2013). The mean response time distribution on the original scale (i.e., in seconds) was computed as $exp\left(\mu + \frac{\sigma^2}{2}\right)$, where μ is the mean of the lognormal distribution and σ the standard deviation. The distribution of the difference in means was then formed by subtracting the mean of one condition (e.g., TOR9) from the mean of the other condition (e.g., TOR18). The posterior distributions indicate the most likely (or credible) parameter values (e.g., differences in mean response times), and their uncertainty is represented by the width of the distribution. This width is represented by the 95% highest posterior density (HPD) interval, which includes the 95% most probable values (Kruschke, 2018). Therefore, each posterior distribution of mean response times and difference in means were summarized with a mean (most likely parameter value) and a 95% HPD interval.

The most important output of a Bayesian analysis is the complete posterior distribution, as it enables readers and decision makers to

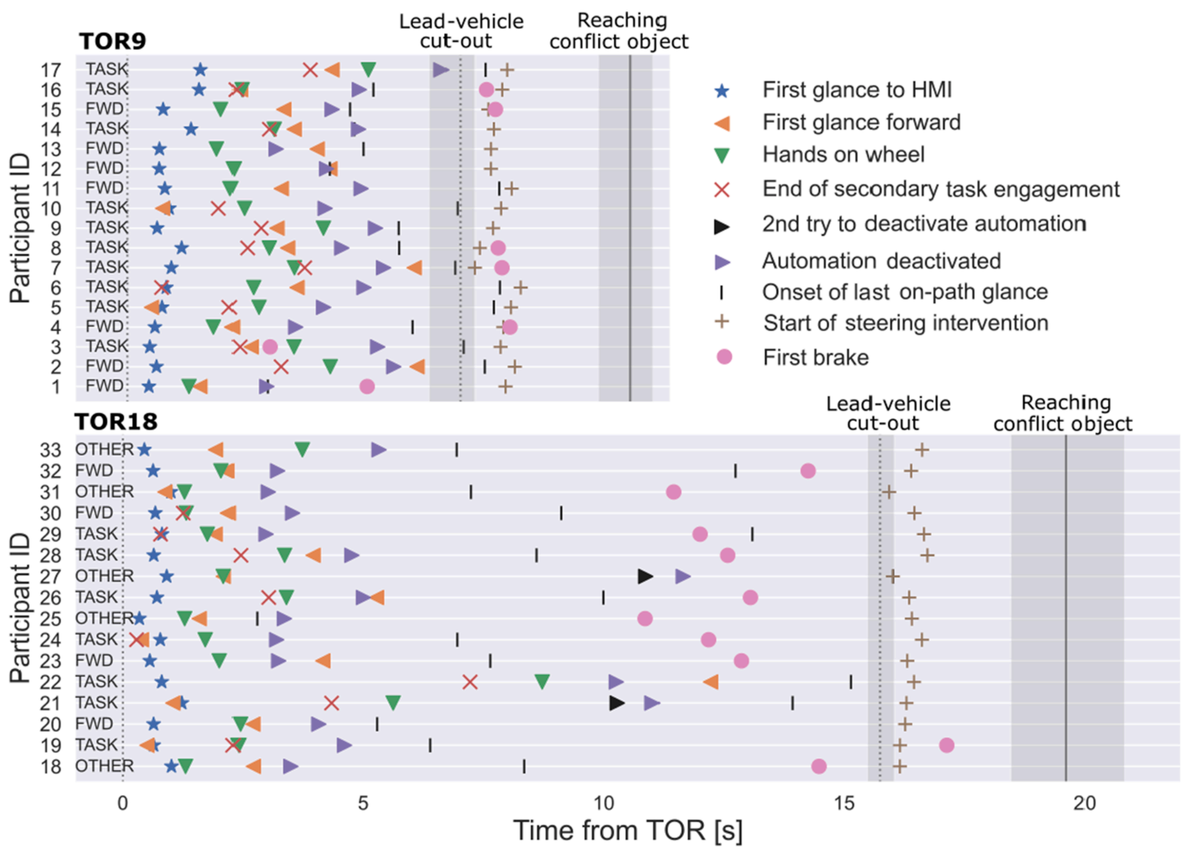

Fig. 4. The driver response process for TOR9 (top panel) and TOR18 conditions (bottom panel). The text along y axis marks the glance position at the TOR. The various markers represent the driver's actions between TOR and Reaching conflict object (see Legend).

evaluate the information in the context of their specific practical application. In general, within the Bayesian framework it is suggested that one avoid making discrete decisions (such as accepting or rejecting a parameter value), as such decisions may lead people to ignore the actual magnitude of the effects as well as the uncertainty (e.g., Kruschke & Liddell, 2018). The posterior distribution for the difference in mean response times represents a measure of the actual effect size, given our data and the model. The results should be interpreted as follows: a difference of zero indicates no difference in response times (and thus *no effect*), a positive difference represents an increase in response times (i.e., a tendency for an effect in the positive direction) and a negative difference represents a decrease in response time (i.e., a tendency for an effect in the negative direction). Although we encourage the readers to use the actual values (and effect sizes), we will indicate when the 95% HPD interval does not include zero, as this result indicates that it is unlikely that there is a real difference between the two distributions.

3. Results

3.1. The drivers' response to take-over request

Fig. 4 displays the driver response process for the participants in the two TOR timing conditions. At the TOR, 17 participants had an ongoing glance related to the secondary task (TASK), 11 were looking forward (FWD), and five were looking somewhere else (OTHER). The figure also shows that, independent of the TOR timing, the drivers glanced towards the HMI and the forward road, put their hands on the wheel, ended their secondary task engagement, and deactivated AD within about six seconds of the TOR. The order of the actions varied, but some trends could be observed. For example, most participants first seemed to glance towards the HMI (26 out of 33), followed by either ending the secondary task if present (11 out of 26) or putting their hands on the steering wheel (14 out of 26). Further, all 33 drivers had put their hands on the steering wheel before finally deactivating AD. Typically, the participants (27 out of 33) glanced forward before deactivating AD.

Three participants in the TOR18 condition needed a longer time to deactivate AD. Two of these participants (Participant IDs 21 and 27 in Fig. 4) did not manage to deactivate AD on their first attempt, since they either pressed the buttons for less than 0.6 s or they pressed next to the buttons instead of on them. The third participant with a long take-over time was engaged in two secondary tasks (mobile phone and notebook) and sat in a relaxed position with both feet up on the car seat at the time of the TOR. Consequently, she needed time to reposition her legs and put away the objects before her hands were free to deactivate AD and drive manually.

Differences in brake onset and onset of the last on-path glance can be observed for the two TOR timings. For TOR9, nine out of the ten participants who braked started to do so before the LV cut-out. In contrast, for TOR18, five out of the seven who braked started braking after the LV cut-out. Further, when the TOR was given earlier, the onset of the last on-path glance generally occurred a longer time after the TOR. For comparison, only four of the 15 ACC participants braked during the conflict scenario: two before the LV cut-out and two after.

3.1.1. Influence of take-over request timings on the driver response process

3.1.1.1. Hands-on-wheel response time. When the TOR was issued, none of the participants in the TOR9 or TOR18 conditions had their hands on the steering wheel. The participants (independent of TOR timing) needed 2.7 s on average (Table 2) to place at least part of a hand on the steering wheel. The most credible difference in mean hands-on-wheel response times for the two conditions was estimated to be –0.04 s (mean) with 95% HPD interval [–0.96, 0.85]. Since zero is included in this 95% HPD, a difference of zero is among the credible values. Thus, according to our data and model, there is no evidence of a consistent difference in mean hands-on-wheel response time, whether the TOR is issued at 9 or 18 s TTC.

3.1.1.2. Visual response times. On average, the participants (independent of TOR timing) took about 0.8 s (Table 2) to redirect their eyes to the HMI in response to the TOR. The most credible difference in means between the two conditions was estimated to be –0.02 s (mean) with 95% HPD interval [–0.21, 0.17]. Thus, similar to the hands-on-wheel response time, there is no evidence of any difference in the time needed for drivers to direct their first glance to the HMI for TOR9 and TOR18. Further, on average, the participants needed 3.4 s to direct their eyes to the forward road in response to the TOR when the TOR was issued late (9 s TTC) and 2.9 s when the TOR was early (18 s TTC). The most likely difference in means between the two conditions was estimated to be –0.56 s (mean) with 95% HPD

Table 2
The posterior distributions (summarized with mean and 95% HPD) for mean response times for each condition (TOR9 and TOR18) and the differences in means between the conditions for the response time to TOR models. The 95% HPD interval that does not include a difference of zero among the credible values is bolded.

Dependent variable	Mean TOR9	Mean TOR18	Difference in means: TOR18-TOR9
Hands on wheel [s]	2.7 [2.1, 3.4]	2.7 [2.0, 3.3]	–0.04 [–0.96, 0.85]
First glance to HMI [s]	0.78 [0.65, 0.92]	0.76 [0.63, 0.91]	–0.02 [–0.21, 0.17]
First glance forward [s]	3.4 [2.2, 4.9]	2.9 [1.7, 4.1]	–0.56 [–2.3, 1.2]
Automation deactivated [s]	4.4 [3.6, 5.2]	5.0 [4.1, 5.9]	0.58 [–0.63, 1.8]
Onset of last on-path glance [s]	5.7 [4.7, 6.8]	9.3 [7.5, 11.0]	3.6 [1.6, 5.8]

interval [–2.3, 1.2]. Thus, our data and the model suggest a slight increase in the time needed for drivers to look forward in response to the later TOR. Finally, on average, the participants needed 5.7 s until the onset of the last on-path glance when the TOR was issued late (9 s) and 9.3 s when the TOR was issued early (18 s). The difference in means for the two conditions was estimated to be 3.6 s (mean) with 95% HPD interval [1.6, 5.8]. Thus, when the TOR was issued early, on average, the participants took 3.6 s longer to keep their eyes constantly on the path. Since zero is not in the 95% HPD, a difference of zero is not among the credible values. Thus, according to our data and model, there is evidence of an increase in the time needed for drivers to begin their last on-path glance before reaching the conflict object for TOR18 compared to TOR9.

3.1.1.3. Deactivation response time. On average, the participants needed 4.4 s to deactivate AD in response to the TOR issued late and 5.0 s in response to the TOR issued early. The most likely difference in means between the two conditions was estimated to be 0.58 s (mean) with 95% HPD interval [–0.63, 1.8], and 83% of the HPD interval is above zero. Thus, the earlier TOR resulted in a slight increase in the participants' deactivation response time. However, a difference of zero is still among the credible values since the 95% HPD interval includes zero. Thus, according to our data and model, there is an 83% chance that the TOR issued at 18 s will result in a longer deactivation response time than the TOR issued at 9 s.

3.2. Drivers' response to the conflict scenario

Fig. 4 (Section 3.1) indicates that all participants started their steering intervention after the *Lead-vehicle cut-out* (vertical dotted line) but before *Reaching conflict object* (vertical solid line). Fig. 5 (top panel) further reveals that all participants responded with a timely steering manoeuvre and easily avoided the conflict object. The average speed for the TOR9 and ACC participants remained similar for the complete 300-m interval, while the TOR18 participants slowed down slightly. However, as shown in Fig. 5 (second panel from top), none of the drivers braked to a complete stop, and the lowest speed was about 50 km/h. Thus, no drivers performed any harsh braking in response to the conflict object. ACC participants generated 1.4 times higher mean maximum lateral accelerations than TOR18 participants, but TOR18 participants generated 1.1 times higher mean maximum longitudinal accelerations. Further, TOR9 participants generated 1.2 times higher mean maximum longitudinal and lateral accelerations than ACC participants.

3.2.1. Participants reported experience about the conflict scenario

All 15 ACC participants reported that they had observed the stationary conflict object before the LV cut-out. They also responded

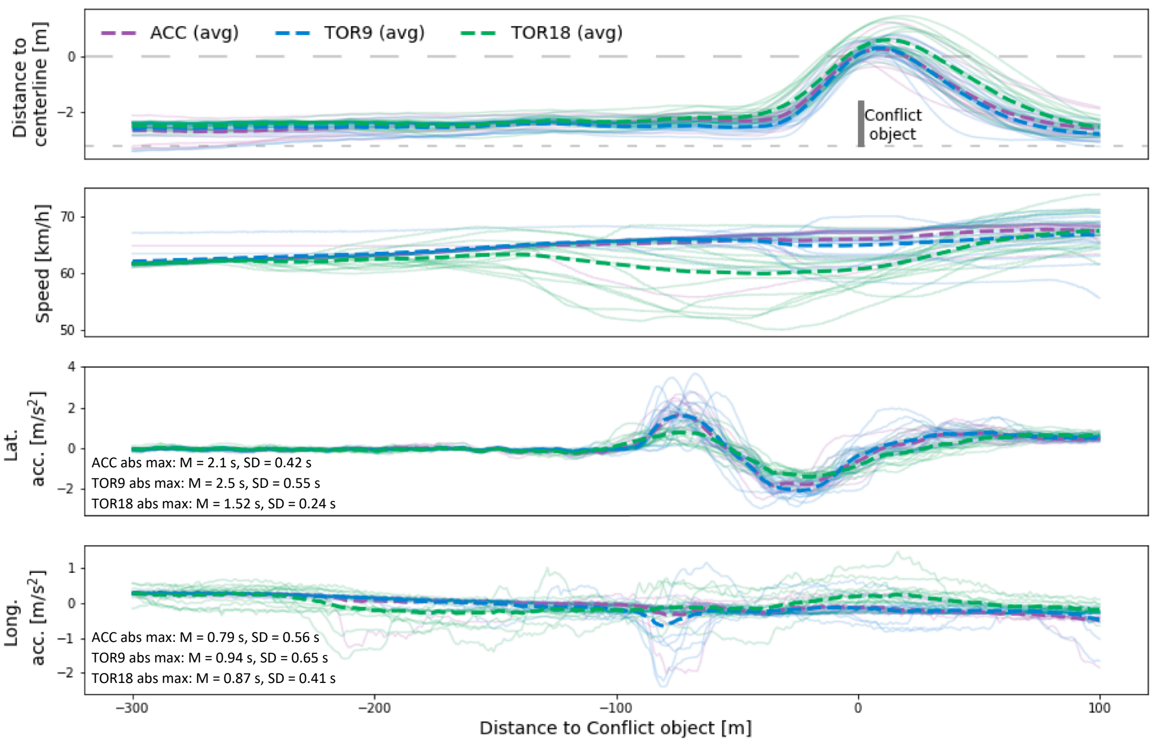

Fig. 5. The manual driving performance (distance to centerline, speed, lateral and longitudinal accelerations) when passing the conflict object. For lateral and longitudinal accelerations, the maximum absolute values within the 300-m interval from 200 m before to 100 m after the conflict object are displayed in the lower left corners.

that they understood the need to act to avoid the conflict object. Although only six out of 17 of the TOR9 participants observed the conflict object before the LV cut-out manoeuvre, all 17 reported that they understood their responsibility to act to avoid crashing. For the TOR18 participants, on the other hand, 12 out of the 16 participants reported that they observed the conflict object before the LV cut-out manoeuvre (i.e., when it was visible at the crest). Most of the TOR18 participants (15 out of 16) answered "yes" when asked if they understood their responsibility to act to avoid a crash (the remaining participant answered "no").

3.2.2. Influence of driving mode on drivers' conflict response

3.2.2.1. Steering response time to lead-vehicle cut-out. Fig. 6 (left) shows the steering response time to lead-vehicle cut-out. The ACC participants had the shortest mean steering response times, followed by the TOR18 participants and then the TOR9 participants. In Table 3, the average mean response times for the ACC, TOR9, and TOR18 participants were 0.47 s, 0.75 s, and 0.65 s, respectively. The greatest effect was observed for TOR9 compared to ACC: the 95% HPD interval [0.08, 0.49] for the difference in means for TOR9 and ACC does not include zero. Thus, according to our data and model, there is evidence of an increase in steering response time to lead-vehicle cut-out for AD with a TOR issued at 9 s TTC compared to ACC.

3.2.2.2. Steering response time to the conflict object. Fig. 6 (right) shows the steering response time to the conflict object. On average, the TOR18 participants started steering away from the conflict object earlier than participants in the other conditions. That is, Table 3 indicates an increase in steering response time to the conflict object for the TOR18 condition compared to driving with ACC and driving with AD and receiving the late TOR (0.43 s [0.09, 0.79] and 0.72 s [0.40, 1.1], respectively). Since zero is not in any of these 95% HPDs, a difference of zero is not among the credible values. Thus, according to our data and model, there is evidence of an increase in steering response time to the conflict object for AD with a TOR issued at 18 s TTC compared to the other two conditions.

4. Discussion

4.1. The influence of take-over timing on the driver response process

Our findings suggest that TOR timing (early or late), in general, influences some actions of the response process more than others. The results shown in Fig. 4 suggest that the TOR timing has little influence on the time drivers need to redirect their glance to the HMI, redirect their first glance forward, put their hands on the steering wheel, end their secondary task, and deactivate AD. These actions seem to be clustered in a six-second interval after the TOR. On the other hand, TOR timing seems to influence the actions taking place later than 6 s after the TOR. Compared to the later TOR, in response to the early TOR drivers took 3.6 s longer, on average, to begin their last on-path glance. These drivers had more time to make sure the automation was successfully deactivated and that any device (e.g., mobile phone) was stored away properly before the LV cut-out directed their attention to the forward road. In other words, the early TOR elicited the same preparation time as the late TOR, while leaving more time for detecting the object and planning the avoidance action. For example, when drivers received the early TOR, they could have looked forward early enough to identify the upcoming conflict object when it was briefly visible from the crest of the hill. The early detection of the conflict object likely explains why more participants braked before the LV cut-out when the TOR was issued early. This observation is supported by the post-drive interview responses: 75% of the participants in the TOR18 condition identified the conflict object before the LV cut-out. In contrast, only about 35% of those in the TOR9 condition did. In other words, since preparing to act (i.e., putting hands on wheel, deactivating AD) is independent of TOR timing, an early TOR gives drivers more time to visually assess the situation and encourages the driver to engage in precautionary braking. Therefore, our findings do not support the current understanding in the literature (e.g., see McDonald et al., 2019; Zhang et al., 2019) that a longer take-over time budget results in longer take-over time as a general rule. The possible reasons behind this discrepancy will be discussed below.

4.1.1. Take-over times

Although the model in McDonald et al.'s review (2019) predicts a 2.43-second increase in take-over time when a TOR is anticipated by 9 s, our model suggests an average increase of only 0.58 s. These contrasting findings may be explained in the following ways: 1) their model is mainly valid for short take-over time budgets (critical situations) and 2) our study uses a different TOR and deactivation

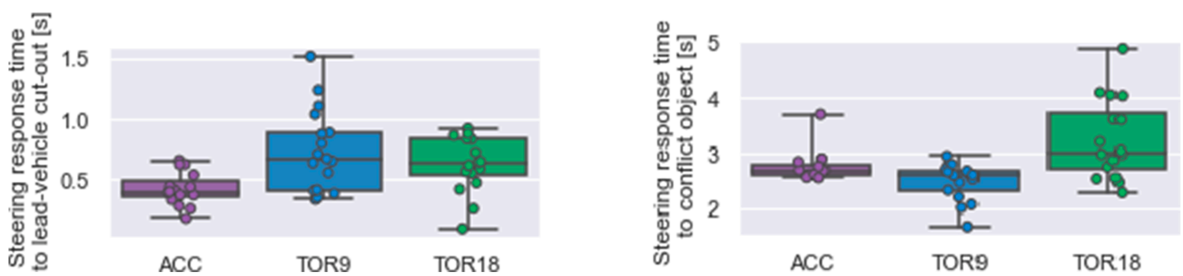

Fig. 6. The drivers' steering response times to the lead-vehicle cut-out (left) and to the conflict object (right).

Table 3
The posterior distributions (summarized with means and 95% HPD intervals) for mean response times for each condition (ACC, TOR9, and TOR18) and the difference in means between the conditions for the conflict intervention models. The 95% HPD intervals that do not include a difference of zero among the credible values are bolded.

Dependent variable	Mean ACC	Mean TOR9	Mean TOR18	Difference in means: TOR18–ACC	Difference in means: TOR9–ACC	Difference in means: TOR18–TOR9
Steering response time to conflict object [s]	2.8 [2.6, 3.0]	2.5 [2.3, 2.7]	3.2 [3.0, 3.5]	0.43 [0.09, 0.79]	–0.28 [–0.6, 0.02]	0.72 [0.40, 1.1]
Steering response time to lead-vehicle cut-out [s]	0.47 [0.36, 0.59]	0.75 [0.58, 0.93]	0.65 [0.50, 0.80]	0.18 [–0.02, 0.36]	0.28 [0.08, 0.49]	–0.11 [–0.33, 0.12]

strategy. Most studies included in the work by McDonald et al. (2019) issued a TOR when a situation (e.g., looming) required an immediate response; therefore, the driver may have responded to the situation rather than the TOR. Also, these studies allowed drivers to deactivate AD by simply steering or braking. A longer take-over time budget in these settings likely resulted in a longer take-over time because drivers waited for the situation to become critical before starting the avoidance manoeuvre, since they didn't need to perform specific actions to deactivate AD. However, in the present study, the drivers had to deactivate AD by pressing steering wheel buttons before they could respond to the conflict (as opposed to just braking or steering to deactivate AD). The deactivation strategy in the present study, selected to avoid the risk of accidental deactivation and mode confusion, required more deliberate actions and was more complex than in previous studies. Moreover, the TOR strategy in the present study gave drivers the chance to deactivate AD before being presented with the conflict. Therefore, the situation kinematics when the TOR was issued were not the same as in previous driving-simulator studies. In other words, when a TOR is issued before a situation requires an immediate response, drivers may respond to the TOR rather than to the threat itself.

Finally, three participants in the TOR18 condition seemed to deviate from the other drivers in the same condition. They all needed ten seconds or more to deactivate AD. If these drivers had been in the TOR9 condition, they would likely have responded too late or even crashed. Their longer take-over times may stem from both an increased hands-on-wheel response time and altered seating position (Participants 21 and 22). The longer hands-on-wheel times are likely influenced by the performance of a secondary task with a handheld item. This result is in line with previous research indicating that a handheld item usually prolongs take-over times more than an item that is not handheld, such as a mounted tablet (McDonald et al., 2019; Zeeb, Härtel, Buchner, & Schrauf, 2017). Moreover, Participants 21 and 27 had difficulty deactivating AD and needed an additional attempt. According to United Nations Economic Comission for Europe (2021), a TOR should be escalated after 4 s (e.g., with an increasingly frequent audio tone), and a minimum-risk manoeuvre should be started by the vehicle after 10 s at the earliest. The average driver in the present study would have deactivated AD in the four seconds before the escalated TOR, but some drivers would still have received an escalated TOR. A minimum-risk manoeuvre starting after 10 s would have started before all drivers in the present study had deactivated AD.

4.2. Automation effects: The influence of driving mode on drivers' response to conflict

Our results show that drivers' response to the conflict scenario depends on both driving mode (i.e., ACC, AD) and the TOR timing (i.e., 9-s or 18-s time budget). In fact, for the given TOR strategy and the tested TOR timings, we observed the earliest response for the TOR18 drivers who started to steer about 0.43 s earlier (95% HPD interval [0.09, 0.79]) than the ACC drivers, while the TOR9 drivers started to steer 0.28 s (95% HPD interval [−0.60, 0.02]) later than the ACC drivers. The early TOR likely resulted in the earliest response because of the precautionary braking that occurred for most drivers within the TOR18 condition and not in the ACC or TOR9 condition. As the TOR18 participants braked prior to the conflict onset, they were farther away from the conflict object when the LV started the cut-out and could therefore also both detect and respond to the event earlier compared to the other participants, who typically maintained their speed while approaching the conflict object.

In contrast to results reported by Gold et al. (2013), we did not observe severe automation effects on drivers' responses to the LV cut-out scenario for AD (compared to the ACC baseline). Recall that we used an ACC baseline as a substitute for the manual baseline used in Gold et al. (2013), since a fixed time headway between the test vehicle and the lead vehicle was needed to achieve the same conflict criticality within each condition. Firstly, we did not observe the delayed response for AD compared to ACC that was observed for AD compared to manual driving in Gold et al.'s results (2013): TOR18 participants generally started steering earlier than ACC participants, and TOR9 participants responded slightly later than ACC participants (the 95% HPD for the difference in mean steering response times is shifted towards negative values). In addition, the observed difference in mean steering response times from the conflict object observed by Gold et al. (2013) (1.47 s for the 7-s TOR timing, and 0.86 s for the 5-s TOR timing compared to the manual baseline) are not included in the 95% HPD for TOR9 compared to ACC. Finally, we did not observe as large an increase in accelerations after AD compared to ACC as Gold et al. did for accelerations after AD compared to manual driving (2013). In the present study, the largest increase in accelerations was about 1.2 times, for AD with the late TOR compared to ACC. This is markedly lower than the 2- to 3-fold increase in accelerations observed by Gold et al. (2013).

As hypothesized, one of the explanations behind these differing findings may be the fact that the TOR was issued before the LV cut-out; unlike in previous driving-simulator studies, the TOR9 and TOR18 drivers were physically ready to act and had fully transitioned to manual driving before the conflict object was revealed in the LV cut-out. In fact, Fig. 4 shows that all drivers had put their hands on the steering wheel and deactivated AD before the LV cut-out. Consequently, the drivers all had the opportunity to start to steer at the same time as the ACC drivers. However, as the present study did not include time budgets as short as those reported by Gold et al. (2013), we are not able to completely disentangle the influence of the TOR time budget from the influence of the TOR timing in relation to the conflict onset. Overall, our findings suggest that responding to TORs in AD does not always need to end up in a safety–critical situation if the TOR is issued early enough and/or prior to conflict onset (e.g., using vehicle-to-vehicle communication). It seems that not only does the take-over time budget matter, but the relation between the conflict onset and the TOR timing is also important for understanding automation effects.

4.2.1. The influence of the visual component of the response process on drivers' conflict response

Even though our findings do not completely support a group-level effect of AD and the late TOR on the drivers' response to the LV cut-out scenario, a slight delay in response could still be observed for some of the TOR9 drivers compared to the ACC drivers (see Fig. 6 right). This slight delay may indicate mechanisms underlying the automation effects, beyond the physical response time needed for drivers to become ready to act. The response process's visual component may explain the delay in the response of the TOR9 drivers compared to that of the ACC drivers. Fig. 4 (top panel) shows that for many drivers, the onset of the last on-path glance was very close to the time of the LV cut-out (or even after). They may have been looking away when the LV started to change lanes; consequently, they may have noticed the LV cut-out later and therefore acted later. Thus, the response process to a TOR may include off-path glances which delay drivers' glances toward the road, so they might not detect a conflict onset right away. In fact, Pipkorn, Dozza & Tivesten (2021) observed a significant increase in glances towards the instrument cluster seconds after the TOR. In other words, when nothing in the environment calls for drivers' attention, a TOR can trigger off-road glances as part of the response process. Such off-road glances may be safety-relevant and could delay drivers' detection and response to events. For example, so-called "Perfect mismatches" between the chance timing of off-road glances and a simultaneously occurring critical situation (e.g., decreasing TTC) have previously been observed to contribute to near-crashes and crashes in real traffic (Victor et al., 2015). Thus, it is important to design a TOR that takes into consideration the possibility that drivers may not look immediately on-path in response to the TOR. In fact, some drivers may not look on-path until they have deactivated AD. However, off-path glances during the response process to a TOR may become less frequent with practice. Therefore, a future study should investigate whether a driver's visual behaviour after a TOR depends on the number of take-overs experienced by that driver.

4.2.2. Driver response to conflicts for ACC, AD, and a near-perfect assisted driving system

Interestingly, none of the drivers in the present study crashed in the LV cut-out scenario. In contrast, about one-third of the drivers in the study by Victor et al. (2018) crashed in the same conflict scenario after having supervised a near-perfect assisted driving system. The participants who crashed started their steering intervention when they were 1 s or less away from the conflict object—and some drivers did not even act at all (Pipkorn et al., 2021b). In the present study, the average steering response time was 2.5 s from the conflict object for TOR9, 2.8 s for ACC, and 3.2 s for TOR18. Importantly, since none of the 95% HPD intervals corresponding to these average mean response times includes a response time of 1 s (see Table 3), we can be certain that the drivers in the present study started to steer away at a safer distance from the conflict object. Thus, it seems that both ACC and AD (different driving modes) can result in a safer response than was observed in a near-perfect assisted driving system (Victor et al., 2018). The likely explanation behind the reduced crash rate in the present study is that the TOR helped the drivers understand their responsibilities to act in the conflict scenario. In fact, the participants who crashed in Victor et al.'s study reported that they did not realize the need to act—despite being specifically informed before the drive about the system limitations and the conflict scenarios which would require them to act (Gustavsson et al., 2018). In contrast, in the present study, all except one participant driving with AD reported that they had understood the need to act in the LV cut-out scenario. In summary, a TOR can be designed to prime drivers to take the appropriate actions in response to an upcoming conflict. When drivers received a TOR and resumed manual driving, there seemed to be no confusion about the current driving mode or the responsibility to act in conflicts. It was clear to the drivers that they were driving manually and therefore needed to perform the avoidance manoeuvre. However, since all ACC participants in the present study also managed to resolve the conflict despite receiving no information from the vehicle about the upcoming conflict, explanations other than the presence of a TOR are called for. One difference between driving with ACC (the present study) and an assisted driving system (Victor et al., 2018) is the capability of the system to assist with steering. While an ACC only supports the driver with longitudinal control, an assisted driving system supports drivers with both longitudinal and lateral control. Consequently, drivers of an assisted driving system need only provide occasional steering inputs, while the ACC participants were responsible for steering before the conflict. Thus, it is unlikely that these participants would expect the system to handle the conflict by steering because it was obviously outside the system's operational domain.

4.3. Limitations and future work

The findings in this paper are based on a test-track experiment with a real vehicle, which provides a higher degree of realism than the driving-simulator studies reported in the literature. However, a test track is still not the same as driving on a real road with real traffic. Thus, the results may be influenced by the absence of real traffic, as well as by the fact that the conflict object was a balloon car (the realism was compromised for the sake of safety). However, it would not have been possible to perform a similar study that includes a LV cut-out scenario on a public road because of the ethical concerns. Further, the presence of a test leader and a Wizard in the vehicle may have had an effect. Although none of the participants (Volvo car employees in the Gothenburg area of Sweden) were directly involved in vehicle automation product development, they may not be representative of an international (or even a Swedish) population. To understand how the present results generalize to more realistic settings and a wider population, a future study should investigate the response process using naturalistic driving data. Finally, the present study did not control for the engagement in secondary tasks. The reason was to keep the AD condition as realistic as possible; strictly controlling the engagement in secondary tasks may not allow us to get a realistic understanding of drivers' response to conflict scenarios when they are able to choose to engage in secondary tasks. However, as a similar number of participants engaged in secondary tasks at the time of the TOR within the two AD conditions, the influence of this engagement on the drivers' response to the TOR should be distributed evenly across them. Thus, the present study traded a higher degree of realism for some controllability. Our findings may not apply to AD studies with forced engagement in secondary tasks or supervised automation.

5. Conclusions

This study shows that an earlier TOR (18 s before a conflict object used in a conflict scenario) may make AD safer. We have shown that the time needed for drivers' preparatory actions (placing hands on wheel, deactivating AD) is independent of TOR timing, so an earlier TOR may result in drivers' earlier detection of upcoming conflicts and earlier braking in preparation for a conflict. With the TOR timings of 18 s and 9 s used in this study, we could not confirm the severe automation effects previously observed in driving-simulator studies. In fact, this study shows that when a TOR is issued early (18 s before a conflict object) drivers respond earlier than those driving with ACC. Further, when a TOR is issued late (9 s before a conflict object) drivers respond only slightly later than those driving with ACC. This slight delay in response was not as severe as previously observed in driving simulators. In addition, we could not confirm previously observed lane excursions, harsh braking and increase in accelerations for AD compared to a manual baseline in driving-simulator studies. The deviating findings may stem from the different TOR strategy and time budgets used in the simulator studies: they issued the TOR at higher criticality and before conflict onset, providing drivers with less time for their preparatory actions before presenting them with the conflict scenario. Overall, the present study shows that AD does not need to end up in a highly critical situation if the TOR is issued early enough. Further, this study could not confirm the crash rates observed in the previous study using the same conflict scenario (but a near-perfect assisted driving system). One reason may be that AD deactivation and driving with ACC (as in this study) communicate more clearly to drivers when they are responsible for handling conflicts than a near-perfect assisted driving system does. To understand how these findings generalize to public roads with real traffic, a future study should investigate the safety of the driver response process in naturalistic driving.

CRediT authorship contribution statement

Linda Pipkorn: Conceptualization, Methodology, Data curation, Formal analysis, Visualization, Writing – original draft. **Emma Tivesten:** Methodology, Supervision, Writing – review & editing. **Marco Dozza:** Supervision, Funding acquisition, Writing – review & editing.

Acknowledgements

This research was supported by the European Project L3Pilot, grant number 723051. The authors would like to thank Elin Meltzer, Kristina Svahnström, Pär Gustavsson, and Regina Johansson for performing the data collection, Carol Flannagan for discussions about (and assistance with) the Bayesian Statistics, and Dr. K. Mayberry for language revisions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2022.02.014.

References

AstaZero. (2020). *AstaZero*. http://www.astazero.com/.

Eriksson, A., Petermeijer, S. M., Zimmermann, M., de Winter, J. C. F., Bengler, K. J., & Stanton, N. A. (2018). Rolling Out the Red (and Green) Carpet: Supporting Driver Decision Making in Automation-to-Manual Transitions. *IEEE Transactions on Human-Machine Systems, 49*(1), 20–31. https://doi.org/10.1109/thms.2018.2883862

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. https://doi.org/10.1177/0018720816685832

Euro NCAP - Cut-out scenario. (2021). https://euroncap.newsmarket.com/images-and-videos/video/euro-ncap—cut-out-scenario/a/a587a765-9996-44d4-8e22-b61c5dd9b973.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

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

Green, P., & Wei-Haas, L. (1985). The Rapid Development of User Interfaces: Experience with the Wizard of OZ Method. *Proceedings of the Human Factors Society Annual Meeting, 29*(5), 470–474. https://doi.org/10.1177/154193128502900515

Gustavsson, P., Victor, T. W., Johansson, J., Tivesten, E., Johansson, R., & Aust, L. (2018). What were they thinking ? Subjective experiences associated with automation expectation mismatch. In *Proceedings of the 6th Driver Distraction and Inattention Conference* (pp. 1–12).

Happee, R., Gold, C., Radlmayr, J., Hergeth, S., & Bengler, K. (2017). Take-over performance in evasive manoeuvres. *Accident Analysis & Prevention, 106*, 211–222. https://doi.org/10.1016/J.AAP.2017.04.017

Hoffman, M. D., & Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research, 15*, 1593–1623.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology, 142*(2), 573–603. https://doi.org/10.1037/A0029146

Kruschke, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan, second edition* (2nd ed.). Academic Press. https://doi.org/10.1016/B978-0-12-405888-0.09999-2.

Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation: *Advances in Methods and Practices in Psychological. Science, 1*(2), 270–280. https://doi.org/10.1177/2515245918771304

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin and Review, 25*(1), 178–206. https://doi.org/10.3758/S13423-016-1221-4/FIGURES/12

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. *Accident Analysis and Prevention, 108*(September), 9–18. https://doi.org/10.1016/j.aap.2017.08.011

McDonald, A. D., Alambeigi, H., Engström, J., Markkula, G., Vogelpohl, T., Dunne, J., & Yuma, N. (2019). Toward Computational Simulations of Behavior During Automated Driving Takeovers: A Review of the Empirical and Modeling Literatures. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 61* (4), 642–688. https://doi.org/10.1177/0018720819829572

McElreath, R. (2016). *Statistical rethinking: A bayesian course with examples in R and stan* (1st ed.). CRC Press. https://doi.org/10.1201/9781315372495.

Pipkorn, L., & Dozza, M. Tivesten, E. (2021). Driver visual attention before and after take-over requests in automated driving on public roads. *[Manuscript Submitted for Publication]*.

Pipkorn, L., Victor, T., Dozza, M., & Tivesten, E. (2021a). Automation Aftereffects: The Influence of Automation Duration, Test Track and Timings. *IEEE Transactions on Intelligent Transportation Systems*. https://doi.org/10.1109/TITS.2020.3048355.

Pipkorn, L., Victor, T., Dozza, M., & Tivesten, E. (2021b). Driver conflict response during supervised automation: Do hands on wheel matter? *Transportation Research Part F: Traffic Psychology and Behaviour, 76*, 14–25. https://doi.org/10.1016/j.trf.2020.10.001

SAE International. (2018). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016)*.

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science, 2016*(4), 1–24. https://doi.org/10.7717/peerj-cs.55

Thatcham Research. (2019). *Defining Safe Automated Driving*. https://www.thatcham.org_wp-content_uploads_2020_10_Defining-Safe-Automation-technical-document-September-2019.pdf.

United Nations Economic Comission for Europe. (2021). *UN Regulation No. 157 - Automated Lane Keeping Systems (ALKS)*. https://unece.org/sites/default/files/2021-03/R157e.pdf.

Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., Lee, J. D., & Markkula, G. (2015). *Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk*. https://doi.org/10.17226/22297.

Seppelt, B. D., & Victor, T. W. (2016). Potential Solutions to Human Factors Challenges in Road Vehicle Automation. In G. Meyer, & S. Beiker (Eds.), *Road Vehicle Automation 3* (pp. 131–148). Cham: Springer. https://doi.org/10.1007/978-3-319-40503-2.

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 60*(8), 1095–1116. https://doi.org/10.1177/0018720818788164

Vogelpohl, T., Kühn, M., Hummel, T., Gehlert, T., & Vollrath, M. (2018). Transitioning to manual driving requires additional time after automation deactivation. *Transportation Research Part F: Traffic Psychology and Behaviour, 55*. https://doi.org/10.1016/j.trf.2018.03.019

Wandtner, B., Schömig, N., & Schmidt, G. (2018). Effects of Non-Driving Related Task Modalities on Takeover Performance in Highly Automated Driving. *Human Factors, 60*(6), 870–881. https://doi.org/10.1177/0018720818768199

Wang, P., Sibi, S., Mok, B., & Ju, W. (2017). Marionette: Enabling On-Road Wizard-of-Oz Autonomous Driving Studies. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI '17*, 234–243. https://doi.org/10.1145/2909824.3020256.

Zeeb, K., Härtel, M., Buchner, A., & Schrauf, M. (2017). Why is steering not the same as braking? The impact of non-driving related tasks on lateral and longitudinal driver interventions during conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 50*, 65–79. https://doi.org/10.1016/j.trf.2017.07.008

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. https://doi.org/10.1016/j.trf.2019.04.020