Automation aftereffects: the influence of automation duration, test track and timings

Citation for the original published paper (version of record):
Pipkorn, L., Victor, T., Dozza, M. et al (2021)
Automation aftereffects: the influence of automation duration, test track and timings
IEEE Transactions on Intelligent Transportation Systems, In press
http://dx.doi.org/10.1109/TITS.2020.3048355

(article starts on next page)

Automation Aftereffects: The Influence of Automation Duration, Test Track and Timings

Linda Pipkorn, Trent Victor, Marco Dozza, and Emma Tivesten

***Abstract*— Automation aftereffects (i.e., degraded manual driving performance, delayed responses, and more aggressive avoidance maneuvers) have been found in driving simulator studies. In addition, longer automation duration seems to result in more severe aftereffects, compared to shorter duration. The extent to which these findings generalize to real-world driving is currently unknown. The present study investigated how automation duration affects drivers' take-over response quality and driving performance in a road-work zone. Seventeen participants followed a lead vehicle on test track. They encountered the road-work zone four times: two times while driving manually, and after a short and a long automation duration. The take-over request was issued before the lead vehicle changed lane to reveal the road-work zone. After both short and long automation durations, all drivers deactivated automation well ahead of the road-work zone. Compared to manual, drivers started their steering maneuvers earlier or at similar times after automation (independently of duration), and none of the drivers crashed. However, slight increases in vehicle speed and accelerations were observed after exposure to automation. In sum, the present study did not observe as large automation aftereffects on the test track as previously found in driving simulator studies. The extent to which these results are a consequence of a more realistic test environment, or due to the duration between the timings for the take-over request and the conflict appearance, is still unknown.**

***Index Terms*— Automated driving, driver response, driving performance, take-over request, driver behavior, automation.**

I. Introduction

VEHICLE automation that can relieve the driver from the driving task, is still under development. This type of unsupervised vehicle automation differs from the assisted vehicle automation currently present in on-market vehicles, which requires the driver to be responsible for the driving task at all times [1]. Unsupervised driving automation, on the other hand, enables the vehicle to take full responsibility of the driving task (longitudinal and lateral control, event detection and response). The driver is then allowed to disengage from driving and engage in non-driving related tasks (e.g., playing a game). However, the driver is required to appropriately resume manual control when notified by the system [2]. Such a notification, often referred to as a *take-over request* (TOR), takes place for situations when the limitations of the system are encountered.

The extent to which drivers can safely and in a controlled manner resume manual control after a period of automated driving, is subject to ongoing research. Previous research on human collaboration with automation gives us reason to be cautious [3], [4]. This is because human limitations exist in the collaboration with automated systems. Increased automation results in an altered task for the driver, which involves less active participation in controlling the system, and more monitoring of the system performance. Consequently, the drivers may enter an "out-of-the-loop" state [5], [6], which may limit their ability to safely resume manual control when needed [3].

In the context of vehicle automation, the ability to resume manual control has, to a great extent, been assessed through the *take-over time* (TOT) and to some extent by analyzing *driving performance* after the takeover [7], [8]. The TOT is defined as the time from the TOR until automation is deactivated by either steering, braking, or a button press. A recent review of 129 studies pointed out a wide variety of mean TOTs, ranging from 0.69 s up to 19.79 s, with an average mean TOT of 2.72 s [9]. Several factors have been shown to have an impact on the TOT including: the take-over time budget (Time to collision (TTC) when TOR is issued); if the take-over procedure is practiced beforehand, the presence of secondary tasks (especially hand-held tasks); and if a TOR is present or not. Whereas many studies have focused on the effect of automation on the TOT, *automation aftereffects* (i.e., the effect of automation on the driving performance following the takeover) have not received as much attention [7], [8]. It is especially important to also consider the driving performance after automation, since the TOT has not been shown to predict if the driving performance will be degraded. That is, a long TOT does not necessarily result in a delayed event response or a poorer driving performance, if a quick and appropriate response follows this TOT [10]. However, the longer the TOT, the shorter the available time for a driver to respond to a subsequent event if the TOR is issued at a specific time before reaching a conflict object. Thus, to understand the mechanisms behind an observed automation aftereffect (e.g., a less calibrated sensorimotor control as hypothesized by [8]), it is important to disentangle the influence of TOT on the driving performance from other possible influencing factors.

Manuscript received May 27, 2020; revised October 2, 2020 and November 13, 2020; accepted November 24, 2020. This work was supported by the European Project L3Pilot (data collection and data analyses) under Grant 723051. The Associate Editor for this article was C. Ahlstrom. *(Corresponding author: Linda Pipkorn.)*

Linda Pipkorn and Marco Dozza are with the Department of Mechanics and Maritime Sciences, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: linda.pipkorn@chalmers.se).

Trent Victor and Emma Tivesten are with the Volvo Cars Safety Center, Volvo Cars Corporation, 405 31 Gothenburg, Sweden.

This article has supplementary downloadable material available at https://doi.org/10.1109/TITS.2020.3048355, provided by the authors.

Digital Object Identifier 10.1109/TITS.2020.3048355

Some studies have already observed unsupervised automation aftereffects, when drivers need to resume manual control and respond to a subsequent event that requires driver intervention (e.g., a road-work zone, a stationary vehicle in lane). A period of automated driving may lead to a reduced driving performance compared to manual driving [7], [11], [12]. In addition, some evidence exists on that *automation duration* (i.e., the time drivers are exposed to automation) also reduces driving performance after a TOR [13], [14]. That is, longer automation duration seems to result in more degraded performance (e.g., longer response times, greater number of uncontrolled maneuvers, greater accelerations) than shorter automation duration [13], [14]. However, at least one study showed no effects of automation duration on driving performance [15]. A degraded driving performance after a longer automation duration, compared to a shorter duration, may be related to increased driver drowsiness or fatigue [13], or driver vigilance decrements [16].

Notably, the current evidence behind the existence of automation aftereffects stems mainly from studies performed in driving simulators, and with event-response designs in which the drivers first need to resume manual control in response to a TOR, and then shortly after respond to a critical event. A common scenario in these studies is a broken-down vehicle in the lane, which the driver encounters at high speeds typically above 100 kph [11]–[15]. Thus, it is currently unknown the extent to which these findings generalize to real-world driving (non-simulator), lower speeds, and less critical events. A logical next step is therefore to investigate if automation aftereffects are also present in a more realistic driving environment (e.g., on a test track) with a real vehicle, for an event that is encountered in low speeds (e.g., with a traffic jam pilot system). The use of a more realistic environment and a real vehicle is especially relevant when trying to understand and model steering and braking behaviors (i.e., the driving performance following a control transition) since kinematic cues are not present in most driving simulators [17]. Some studies (e.g., [18], [19]) have investigated automated driving take-overs in realistic environments (on public roads) for take-overs not followed by a conflict event. Whereas [18] focused on TOTs, [19] also investigated the driving performance. Specifically, [19] found that all drivers were able to safely resume manual control after automation, and stabilize their manual driving performance within 5 s.

Therefore, the aim of this study was to examine the effect of automation duration on the *driver take-over response* and *driving performance* after a TOR in a simulated road-work zone on test track. The road-work zone, which required the driver to act after the take-over was completed, was encountered multiple times at low speed (about 60-70 km/h). By comparing the present study's results with previous driving simulator studies, this study also aims to better understand the potential factors (e.g., test environment, experimental protocols) that contribute to automation aftereffects. The following research questions were addressed: 1) what is the effect of automation duration on the *driver take-over response* after a TOR?, 2) what is the aftereffect of automation duration on *driving performance?*, and 3) how do these test-track results compare to previous results from driving simulators?

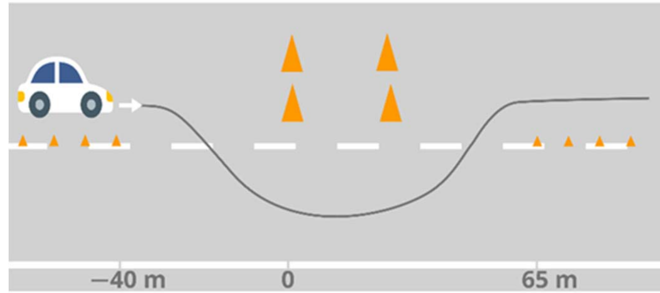

Fig. 1. The cone zone used in the test.

II. Methods

This test-track study investigates drivers' take-over response and driving performance when a driver encounters a road-work zone, after a long automation duration and a short automation duration. The road-work zone was simulated by placing cones on the test track in a way that invited the drivers to carefully maneuver the vehicle in order not to collide with any cones (see section *B*. for details); in the paper we refer to this surrogate road-work zone as the *cone zone.*

A. Participants

Eighteen Volvo Cars employees were recruited for the study. To minimize biases, the participants had no work duties associated with the development of automated driving, did not work as test drivers, and had not been part of a similar study before. All participants had driven at least 5000 km during the year prior to the study. One participant was excluded from the analysis due to missing data, resulting in a final sample size of 17 participants that was used for the analyses presented in this paper. Out of the participants, 11 (65%) were male and 6 (35%) were female, aged between 29 and 63 years (M = 43.9 years, SD = 9.7). All participants signed a consent form prior to participation. The study was reviewed and approved by the regional national ethical review board in Gothenburg, Sweden (Dnr:369-16).

B. Testing Environment and Equipment

The study was conducted on a two-lane rural-road test track, located outside of Gothenburg, Sweden [20]. To assess the manual driving performance after a period of manual driving or when the drivers had resumed manual control after automation (short or long duration), we designed a driving scenario based on the following criteria: 1) the driver shall be required to act to avoid obstacles, 2) the driver shall not be forced to perform a critical evasive maneuver, but still be required to perform fine lateral control, and 3) the situation shall resemble something that drivers may experience in real traffic (i.e., as part of a road-work zone). The cone zone used in the present study, designed to meet these criteria, was built up by four larger cones and eight smaller cones as shown in Fig. 1. The cones were placed on the test track in such a way that the driver had to follow a specific trajectory to avoid

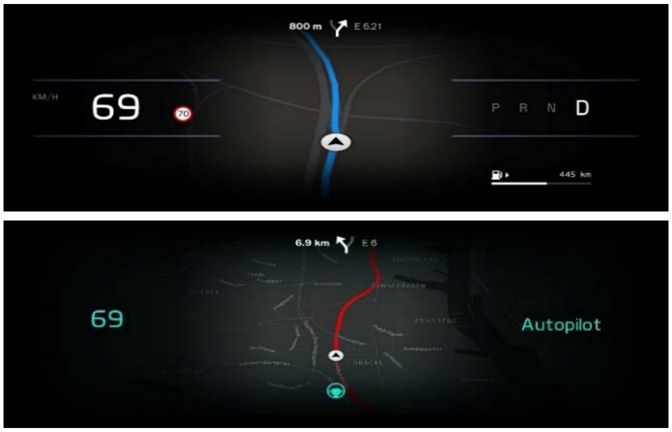

Fig. 2. Top: The HMI view in manual mode. Bottom: The HMI view in automated mode.

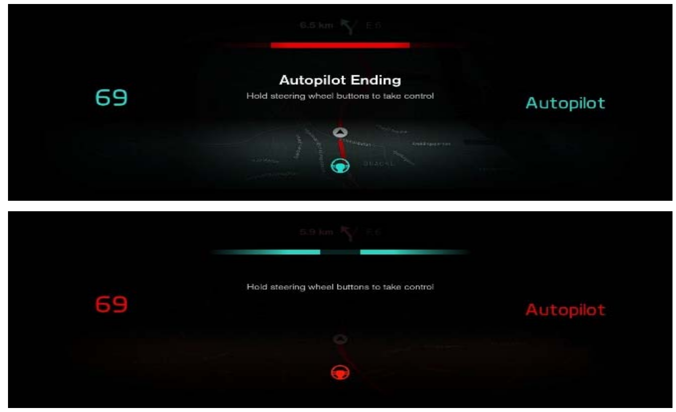

Fig. 3. Top: The HMI view for TJP deactivation (the TOR). Bottom: The HMI view when the two steering wheel buttons are being pressed and the turquoise bars move toward each other.

hitting any cones (Fig. 1). The larger cones were the same type of cones typically used in road-work zones on Swedish roadways.

1) Test Vehicles: The test vehicle (TV) used in the study was a Volvo XC90 (Model Year 2017). The TV was rebuilt to incorporate a W*izard-of-Oz* [21] experimental platform to simulate automation. Inside the TV, three cameras recorded the video of the drivers' face and upper body, as well as the forward roadway. To enable the participants to play a game while in automated driving mode, a dashboard-mounted tablet was installed in the TV. In addition, the TV was equipped with a custom human-machine interface (HMI) in the dashboard behind the steering wheel. The HMI provided the driver with information on driving mode (manual or automated) as shown in Fig. 2. Further, the TV was equipped with a DeweSoft data logger. The collected data included: vehicle controller-area-network signals and GPS data (recorded at 100 Hz), HMI signals (recorded at 0.4 Hz), and video data (recorded at 30 Hz). The only other vehicle present during the study was a robot-controlled XC60 (Model Year 2018), that we refer as lead vehicle (LV) because participants were instructed to follow it. The LV was programmed to follow a pre-defined path and speed profile.

2) The Automated Driving System: The automated driving system (ADS) was simulated by a w*izard driver* controlling the vehicle by using a steering wheel positioned in front of the middle backseat of the TV. The simulated ADS was an *unsupervised traffic jam pilot* (TJP): an ADS designed for low speeds, that allows the participant to disengage from the driving task during automation, although s/he must be prepared to resume manual driving when requested. The simulated TJP required that the speed was lower than 70 km/h and that the LV was present. When the TJP was active, and any of the requirements was no longer met, the TJP notified the driver about the need to resume manual driving. Visual feedback about the current driving mode was presented in the dashboard located behind the steering wheel (Fig. 2).

a) Activation of TJP: When TJP was available for activation, the system notified the driver by an audio tone and a message in the DIM reading "Autopilot available". The driver pressed two buttons on the steering wheel for 0.6 seconds to activate TJP. The driver received feedback when TJP was activated, through a voice sounding "Autopilot active" and the HMI (Fig. 2 bottom).

b) Deactivation of TJP: When the requirements for the TJP were no longer met, the system notified the driver through a take-over request (TOR). The TOR consisted of an audio tone, a seat-belt tensioning, and a message in the DIM reading "Autopilot ending" (Fig. 3 top). The participants had 6 s to deactivate automation (this time was visualized in the DIM with a red shrinking bar, see Fig. 3 top). If the participants did not deactivate TJP within 6 s, the DIM view changed to "Moving to a safe stop" but the participants could still deactivate automation (this only happened once in the experiment). Deactivation was performed by pressing for 0.6 s the same two buttons previously used to activate the system. The remaining time was visualized in the DIM with two turquoise bars (Fig. 3 bottom) approaching each other and meeting when the deactivation was completed. When TJP was deactivated, the HMI changed to the manual driving mode view (Fig. 2, top) and a voice sounded: "Drive the car".

C. Study Procedure

Prior to the test, the participants received information about the test. The purpose explained to the participants, was to evaluate driver experiences during automated driving in traffic jam conditions. In addition, the participants were given information about the automated driving system. This information included: 1) the specifications and requirements for the TJP, 2) that the vehicle was responsible for the driving task when TJP was active, and 3) the need for having a safety driver in the backseat. The participants were instructed to follow the LV closely to avoid other "imaginary" vehicles to cut in between the TV and the LV. Apart from following the LV closely, the participants were free to drive as they normally would as long as they obeyed the traffic rules. The participants were informed about the need to activate and deactivate the TJP when notified by the system. In addition, they were instructed to play a game [22] on the dashboard-mounted tablet, when in automated driving mode. The participants practiced activating

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

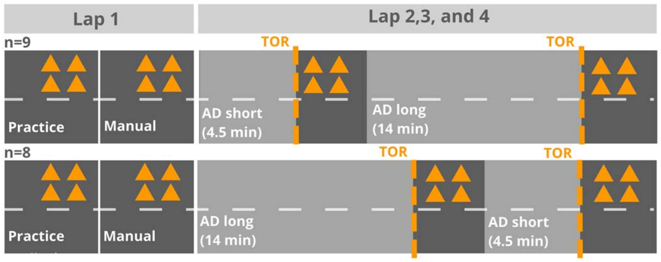

Fig. 4. The study design.

and deactivating the TJP and how to play the game, both at stand-still before the test and during a first training lap.

The total test duration (excluding the training lap) was 30 minutes. Each participant drove 4 laps around the test track and drove both manually and with automation. The TV always followed the LV: in manual mode the participants were free to decide the speed and the time headway, but in automated mode the time headway was kept at 2.5 s. Each participant started the test in manual driving mode. During the first lap, the participants encountered the cone zone twice: the first time, after 1.5 minutes, was used as a practice and the second time, after 4.5 minutes, was used as a manual baseline (Fig. 4). The purpose of the practice was to familiarize the drivers with the design of the cone zone. The purpose of the manual condition was to collect a baseline of driving performance in the cone zone before automation exposure.

To address the defined research questions, the three remaining laps (Laps 2-4) included one long automation duration (14 minutes) and one short (4.5 minutes). These two durations were counterbalanced among the participants, with 9 of the participants experiencing the short duration first followed by the long duration, and the remaining 8 participants experiencing the long duration first, followed by the short duration. When in automated driving mode, the participants received the TOR about 5-7 s before the cone zone was revealed (hereafter referred to as the time of *conflict appearance)* due to the LV speeding up. At conflict appearance, the LV speed was 70 km/h and the LV was 1.5 s from the first cone in the cone zone. Thus, the TOR was given when the TV was about 9-11 s from the first cone in the cone zone. In this paper, the three driving conditions will be referred to as: *manual* for the manual baseline, *AD long* for the long automation duration, and *AD short* for the short automation duration. The complete design is illustrated in Fig. 4.

D. Data Processing and Coding

The videos of all drives were analyzed to assess crash involvement (crash or no crash), take-over outcome (successful or failed), and response process. Crash outcome was originally included as a dependent variable, but no crashes occurred during the experiment. A take-over was coded as failed if the first button press did not succeed to deactivate the system, and the participant needed an additional try to deactivate the TJP.

1) Coding of Relevant Time Points for Take-Over Response: The driver take-over response process (i.e., the driver actions following the *TOR*) were coded using video views and HMI signals. The take-over response actions were: *hands on wheel*, *eyes forward*, *eyes on HMI*, *AD deactivated,* and *2nd try to deactivate TJP.* In addition, the moment when the LV started turning as observed from the video of the forward roadway (*LV start turn*) and the start of the driver steering maneuver (d*river steering start*) were coded. Table I summarizes how these variables were coded.

TABLE I
RESPONSE PROCESS VARIABLE DEFINITIONS

Variable	Variable Description
TOR	The time point when the TOR was issued—obtained from HMI signals.
TJP deactivated	The time point when the automation was deactivated.
eyes forward	The time point when the driver started redirecting the glance towards the forward path (transition included).
eyes on HMI	The time point when the driver started redirecting the glance towards the HMI (transition included).
hands on wheel	The time point when the driver touched the steering wheel with at least one hand or part of a hand.
2nd try to deactivate TJP	The time point when the second button press started. This was only coded if the first button press did not result in a successful deactivation of AD.
driver steering start	The time point when the driver started performing a voluntary steering maneuver (as observed from video but with the steering wheel angle signal as support) intended to avoid the cone zone.
LV start turn	The time point when the LV is observed (from the video view of the forward path) to start turning.

2) Vehicle Signals: The driving performance in the cone zone was assessed using vehicle signals including longitudinal vehicle speed, longitudinal and lateral accelerations, steering wheel angle, and longitudinal distance to the first cone in the cone zone (based on GPS position). All extracted signals were resampled (mean aggregation) to 20 Hz.

E. Data Analysis

To answer the defined research questions, driver take-over response and driving performance in the cone zone were assessed and compared across all conditions.

1) Driver Take-Over Response: The driver take-over response was assessed through: 1) the take-over outcome (i.e., if the drivers needed a second try to deactivate automation), 2) the individual take-over response process (Table I) and 3) the take-over time (TOT; i.e., the time from the TOR until AD deactivated). The individual take-over response process was assessed by visualizing the time points (i.e., hands on wheel, eyes forward, eyes on HMI, TJP deactivated and 2nd try to deactivate TJP; Table 1), for each individual participant, using scatterplots. The take-over response process was anchored at the *TOR*.

2) The Driving Performance in the Cone Zone: The driving performance in the cone zone was assessed through: 1) the conflict outcome, 2) the vehicle signals (vehicle speed, longitudinal and lateral accelerations, and the steering wheel angle) in the *interval-of-interest* (i.e., the recording between 100 m before and 100 m after the first cone in the cone zone) and 3) some *driving performance metrics*; namely, the time to collision (TTC) at driver steering start, maximum vehicle speed, maximum longitudinal acceleration, maximum lateral acceleration, and minimum steering wheel angle within the interval-of-interest. The TTC at driver steering start was calculated by dividing the distance at driver steering start by the longitudinal vehicle speed. The distance used in the TTC calculation was obtained from a GPS signal measuring the distance from the front of the vehicle until the first large cone in the cone zone [23]. The four remaining driving performance metrics were the maximum values within the interval-of-interest of the following vehicle signals: longitudinal vehicle speed, longitudinal and lateral acceleration, and steering wheel angle.

3) Statistical Analysis: Descriptive statistics (i.e., frequencies, boxplots, means and standard deviations) were used to understand how the driver take-over response, the conflict outcome, and the take-over outcome differed depending on automation duration. Bayesian hierarchical varying-intercept models were used to assess the effect of automation exposure and its duration on the TOTs and driving performance. One model was fit to each of the each of the driving performance metrics and the TOT. The general formula for the varying-intercepts models is represented in (1).

$$y_{ij} = \beta_0 + \beta_1 \cdot x_{1ij} + \beta_2 \cdot x_{2ij} + u_j + \varepsilon_{ij} \tag{1}$$

y_{ij} is the modelled response (e.g., TOT) for condition $i = 1, \ldots, 3$ (i.e., manual, short or long) and participant $j = 1, .., 17$, u_j is the random effect on the intercept with standard deviation τ^2 and ε_{ij} is the model error: $\varepsilon_{ij} \sim N(0, \sigma^2)$. β_0 is the global intercept which in this case corresponds to the manual mean (μ_{manual}) and β_1, β_2 represent the effect of AD long ($\mu_{ADlong} - \mu_{manual}$) and AD short ($\mu_{ADshort} - \mu_{manual}$). To quantify the effect of automation exposure and duration, posterior distributions were obtained for the difference in means for: 1) AD long vs. manual ($\mu_{ADlong} - \mu_{manual}$), 2) AD short vs. manual ($\mu_{ADshort} - \mu_{manual}$) and 3) AD long vs. AD short ($\mu_{ADlong} - \mu_{ADshort}$).

In Bayesian statistics, posterior distributions represent the most credible (likely) parameter value together with the uncertainty of this value [20]. The uncertainty can be represented with a 95% highest posterior density (HPD): the values inside the HPD have a total probability of 95%, meaning that these values are more credible than the values outside the interval. Thus, a 95% HPD obtained for differences between group means, as done in this study, offers a way to determine effect sizes, as well as the uncertainty for this effect size. Specifically, the HPDs can be compared to a so-called region of practical equivalence (ROPE) with user-specified limits, often centered around zero [24]. In line with the *new statistics* [25], this paper will present estimates of effect sizes (i.e., differences in means), but will leave further assessment of practical significance (e.g., setting ROPE limits) to the reader. In the present study, the effect sizes were obtained through the posterior distributions of the difference in means (including a 95% HPD), which were visualized together with a vertical reference line marking the zero. For an HPD that does include the zero, a ROPE centered around zero would always overlap with the HPD, and a difference of zero cannot be ruled out.

Prior to fitting the models to the data, TOT and TTC at driver steering start were transformed using the natural logarithm (i.e., $\ln(y_{ij})$). The reason was that the data distributions was slightly skewed (i.e., the transformation was applied to achieve distributions that were closer to normal). Prior to graphing the posterior distributions of the log transformed parameters, these were transformed back by taking the exponential (i.e., μ_{manual} in original units $= \exp(\beta_0 + (\sigma^2 + \tau^2)/2)$, $\mu_{ADlong}$and$\mu_{ADshort}$ in original units $= \exp(\beta_0 + \beta_1 + (\sigma^2 + \tau^2)/2)$and $\exp(\beta_0 + \beta_2 + (\sigma^2 + \tau^2)/2)$. The remaining metrics (maximum speed, maximum longitudinal and lateral acceleration, and minimum steering wheel angle) were modelled as normally distributed. This decision was based on observations of the data distributions.

Model-specification and fitting were performed with the Python (version 3.7.6) packages PyMC3 3.7 [26] and Bambi 0.1.5 [27]. Weakly informative priors were placed on the parameters using the default priors in Bambi [27]. To fit the models the Markov Chain Monte Carlo algorithm No-U-Turn Sampler (NUTS) was used [28]. For each model, the sampler was tuned with 1000 samples and then 4000 posterior samples per chain were drawn (2 chains). The model convergence was verified through: (a) visual inspection of the generated trace plots (after 1000 burn-in samples were discarded) and (b) the obtained Gelman Rubin R hat [29] statistic that should be close to 1. Finally, the generated predictive distributions were compared to the empirical data in order to understand how well the models could describe the data. Further information can be found in the Supplementary materials.

III. Results

A. Driver Take-Over Response

1) Individual Driver Take-Over Response: Fig. 5a shows the participants' individual take-over response process, for AD short and AD long. Most participants put their hands on the wheel and glanced towards the HMI or on the forward roadway shortly after the TOR. Typically, these actions (i.e., hands on wheel and glances towards the HMI or on the forward roadway) were performed within 2 seconds from the TOR, for both conditions. Most of the participants (11/17 in AD short and 14/17 in AD long) glanced towards the HMI before they glanced to the forward roadway. That is, when the participants noticed the TOR, they redirected their eyes from the dashboard-mounted tablet towards the HMI, and then

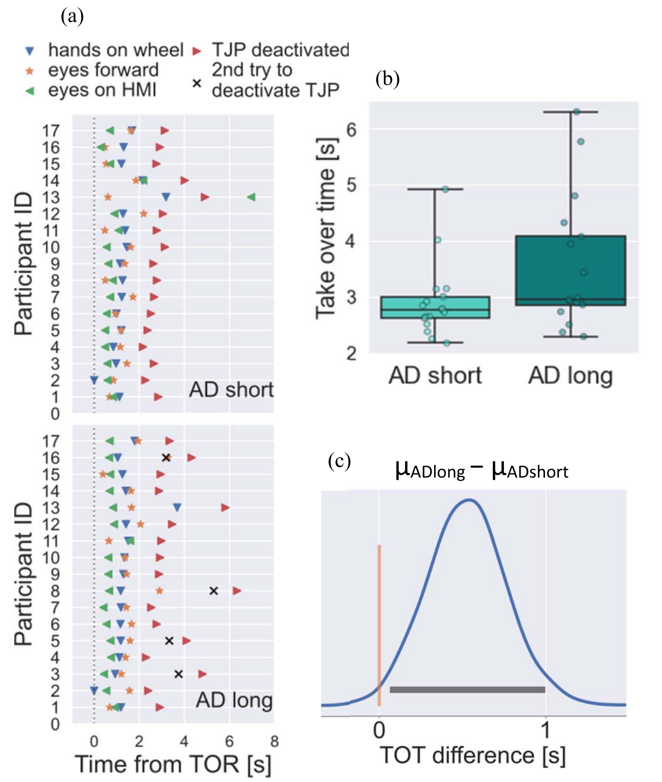

Fig. 5. (a) The individual take-over responses for each participant after a short (top graph) and a long automation (bottom graph) duration, (b) the take-over time for AD short and AD long, and (c) the posterior distribution of the difference in mean TOT between AD long and AD short (note that the orange vertical line in Fig. 5c marks the zero reference and the horizontal grey bar the 95% HPD).

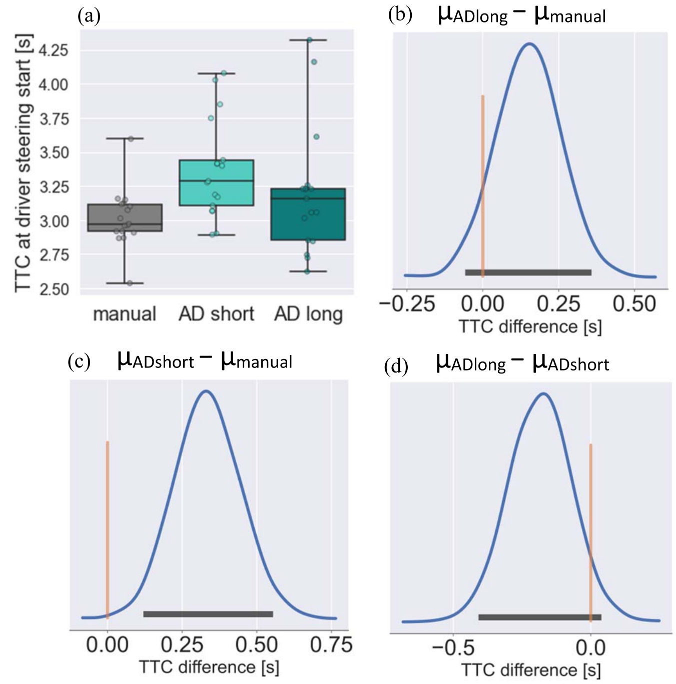

Fig. 6. (a) The TTC at driver steering start for manual, AD short and AD long, (b)-(d) posterior distributions and highest probability density intervals for the difference in mean TTC for: b) AD long vs. manual, c) AD short vs. manual, and d) AD long vs. AD short. Note that the orange vertical line in Fig. 6b-d marks the zero reference and the horizontal grey bar the 95% HPD.

on the forward roadway. In addition to the fast re-direction of gaze, most participants also had put their hands on the steering wheel within 2 s after the TOR. Only one of the participants (the same participant for AD short and AD long) had at least one hand on the steering wheel at the TOR both in AD long and AD short. Fig. 5a also reveals that a longer automation exposure resulted in four failed first attempts to take over control (i.e., four participants required a second additional button press to deactivate automation) whereas no failed takeovers was recorded after the participants had been exposed to a short automation duration.

Note that 75% (3/4) of these drivers were in the group that first experienced a short automation duration followed by the long duration, and 25% (1/4) of these drivers were in the group that first experienced the long automation duration followed by the short. The reason behind these failed attempts was that the drivers did not keep the steering wheel buttons pressed long enough (< 0.6 s). In general, most drivers obtained the necessary motor readiness (i.e., put hands on wheel and deactivated automation) within 3 s. Most drivers (all except three, all in AD long) had achieved this motor readiness before the conflict appearance.

2) Take-Over Times: On average, the long automation duration resulted in increased TOTs compared to the short duration. In Fig. 5b, the mean TOT for AD long (M = 3.53 s, SD = 1.18 s) is larger than for AD short (M = 2.91 s, SD = 0.67 s), and in Fig. 5c the whole 95% HPD for the mean TOT difference between AD long and AD short is above zero signifying that the effect size is credible. Further, the posterior distribution in Fig. 5c shows that, on average, a 9.5 min longer automation duration results in a 0.52 s longer TOT. Removing the four drivers that failed to deactivate automation at their first attempt resulted in a slightly increased mean TOT after AD long (M = 3.12 s, SD = 0.91 s) compared to AD short (M = 2.91, SD = 0.67 s).

B. The Driving Performance in the Cone Zone

Overall, the participants managed to maneuver through the cone zone successfully (i.e., no crashes; see Section D in Chapter II), both in manual as well as after AD long and AD short.

1) Time to Collision at Driver Steering Start: Automation exposure (including both durations) resulted in that participants, on average, started steering to pass the cone zone earlier (at higher TTC) compared to manual. The participants started steering at the largest mean TTC for AD short (M = 3.37 s, SD = 0.36 s), followed by AD long (M = 3.19 s, SD = 0.46 s) and then manual (M = 3.01 s, SD = 0.21 s) as shown in Fig. 6a. Further, the effect of automation on the TTC at driver steering start was largest for the short duration, compared to the long duration: in Fig. 6b-c, the whole 95% HPD for the mean TTC difference between AD short and manual is above zero, whereas 92% of the corresponding 95% HPD for AD long and manual is above zero. However, since in Fig. 6d, the 95% HPD still includes the zero, both durations

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

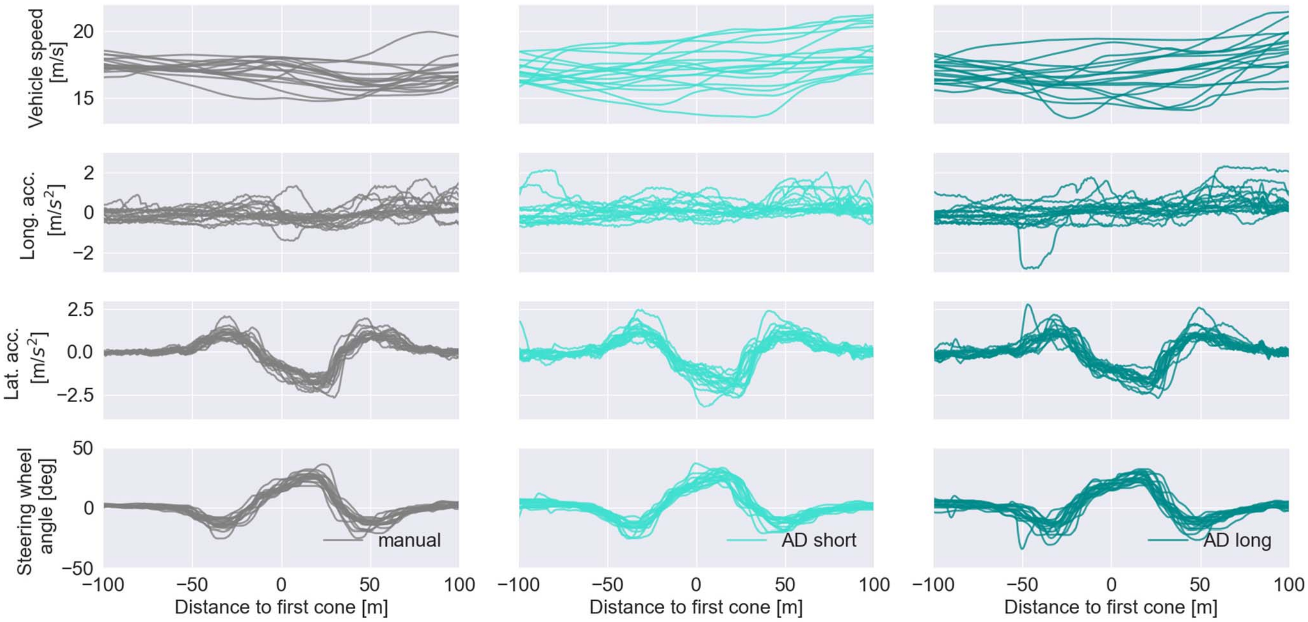

Fig. 7. Vehicle signals (vehicle speed, longitudinal and lateral acceleration, and steering wheel angle) for manual, AD short and AD long 100 m before and 100 m after the first cone in the cone zone.

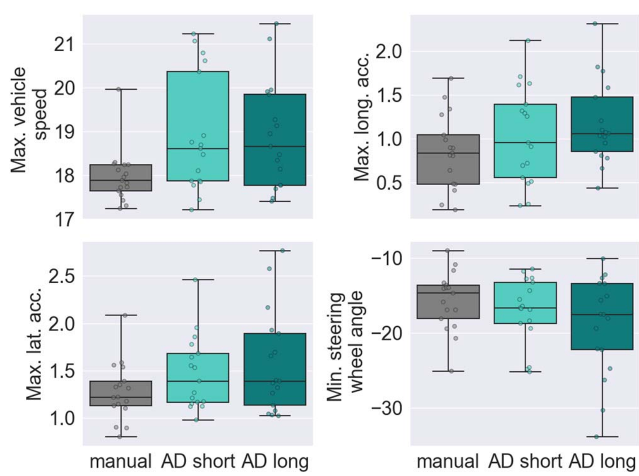

Fig. 8. Driving performance metrics (maximum vehicle speed, maximum longitudinal and lateral acceleration, and minimum steering wheel angle) for manual, AD short, and AD long.

may result in similar TTC values (i.e., no credible effect of automation duration).

2) *The Vehicle Speed, Accelerations and Steering Wheel Angle:*

a) Vehicle speed: Automation exposure (including both durations) resulted in increased vehicle speed within the interval-of-interest: in Fig. 7, an increase in vehicle speed can be observed towards the end of the interval-of-interest for AD long and AD short, whereas the vehicle speed for manual is almost constant within the interval. Further, Fig. 8 reveals that the mean maximum vehicle speed for AD long (M = 18.90 m/s) and AD short (M = 18.92 m/s) was slightly higher than for manual (M = 17.96 m/s).

Similar effect sizes for the effect of automation on the maximum vehicle speed was observed for the long and the short automation duration.

In Fig. 9 (column 1 and 2), both the 95% HPD for the mean maximum vehicle speed difference between AD long and manual and the corresponding 95% HPD for AD short compared to manual was above zero. Consequently, the automation duration was not found to affect the generated maximum vehicle speed: in Fig. 9 (column 3), the 95% HPD for the mean difference between AD long and AD short was neither fully above nor fully below zero, and the most credible difference was almost zero (–0.021 m/s).

b) Accelerations: In addition to an increase in vehicle speed, automation exposure (including both durations) also resulted in increased maximum accelerations: the mean maximum longitudinal acceleration was greater for AD long (M = 1.17 m/s^2, SD = 0.48 m/s^2) and AD short (M = 1.04 m/s^2, SD = 0.55 m/s^2) compared to manual (M = 0.85 m/s^2, SD = 0.42 m/s^2). The mean maximum lateral acceleration was also slightly larger for AD long (M = 1.57 m/s^2, SD = 0.54 m/s^2) and AD short (M = 1.47 m/s^2, SD = 0.39 m/s^2) compared to manual (M = 1.28 m/s^2, SD = 0.31 m/s^2). For both the longitudinal and the lateral acceleration, a larger effect size was observed for the long automation duration, compared to the short duration. In Fig. 9, both the 95% HPD for AD long compared to manual for the maximum longitudinal acceleration and the 95% HPD for AD long compared to manual for the maximum lateral acceleration were above zero. For AD short compared to manual, however, the effect size was not as large: 87.7% of the 95% HPD for the longitudinal acceleration and 95.7% of the 95% HPD for the lateral acceleration, were greater than zero, but both still included the zero. However, since in Fig. 9 (column 3), the 95% HPDs for both the maximum longitudinal and lateral accelerations

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

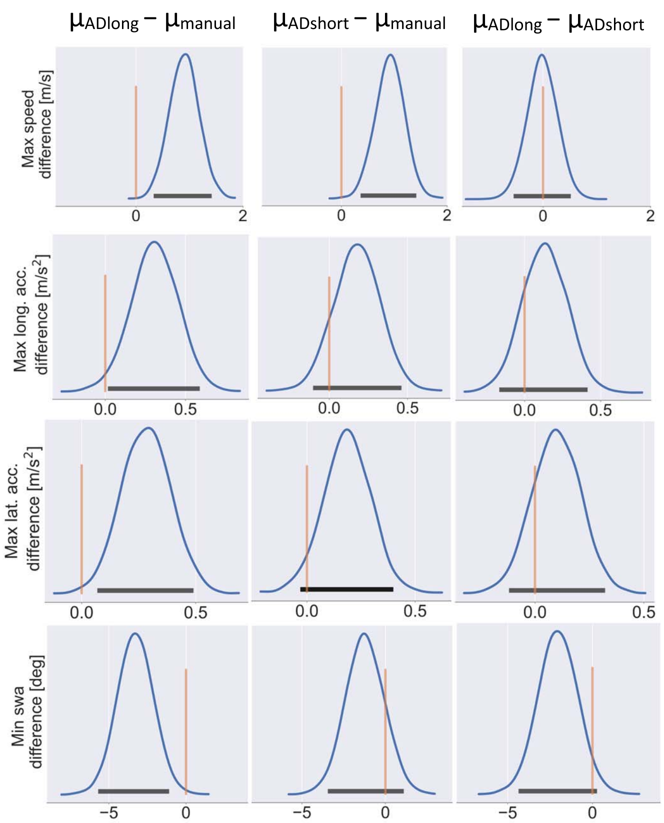

Fig. 9. Posterior distributions (including the highest-posterior density intervals) for the mean difference in maximum vehicle speed (first row from the top), longitudinal acceleration (second row from the top), lateral acceleration (third row from the top) and minimum steering wheel angle (swa; bottom row) for AD long vs. manual, AD short vs. manual, and AD long vs. AD short. Note that the vertical orange line marks the zero reference and the grey horizontal bar the 95% HPD.

still include the zero, we cannot rule out the possibility of that both durations may result in similar accelerations (i.e., no credible effect of automation duration).

c) Steering wheel angle: Automation exposure (including both durations) resulted in decreased minimum steering wheel angle: the mean minimum steering wheel angle was smaller for AD long ($M = -19.12$, $SD = 6.68$) and AD short ($M = -17.06$, $SD = 4.43$) deg. compared to manual ($M = -15.75$, $SD = 3.96$) deg. Further, a larger effect size was observed for the long automation duration compared to the short duration. In Fig. 9 (column 1 and 2), the whole 95% HPD for the difference in mean minimum steering wheel angle for AD long compared to manual was below zero, whereas 86.5% of the corresponding 95% HPD was below zero for AD short compared to manual. However, since in Fig. 9 (column 3), the 95% HPD still include the zero, we cannot rule out the possibility of that both durations may result in similar minimum steering wheel angles (i.e., no credible effect of automation duration).

IV. Discussion

A. The Effect of Automation Duration on Driver Take-Over Response and Driving Performance

The present study found surprisingly small effects of automation duration on driving performance, suggesting that the automation effect is well within acceptable limits for the studied driving scenario and durations. After a long and a short automation duration the drivers generated similar maximum longitudinal vehicle speed, accelerations, and minimum steering wheel angles within the interval-of-interest used in the present study. These findings are in line with [15], but contrasts to [13] and [14], who found that a longer duration resulted in a degraded driving performance (higher lateral accelerations and more collisions or loss of control). A reason behind this difference, may be that automation durations were not consistent across the studies. In fact, the present study and [15] compared about 5 minutes of exposure to automation to 15-20 minutes, whereas [13] and [14] compared 10 minutes with 1 hour, and 25 minutes with 50 minutes, respectively. Another possibility for the observed differences in driving performance across studies, may be that the different studies used events of different criticalities. That is, [15] and [14], were notified about the need to take over control at 6-7 s TTC, whereas the drivers in [13] and the present study were notified at TTC 9-11s. However, considering that [13] and [14] (i.e., TTC 7 s vs. TTC 10 s) observed greater effects compared to [15] and the present study (i.e., TTC 6 s vs. TTC 9-11 s), the different criticalities seems to be less likely to explain the observed differences across studies.

The greatest effect of automation duration found in the present study was on the driver take-over response. That is, a longer automation duration results in a lower probability for the drivers to successfully deactivate automation. Further, we observed longer TOTs after the longer automation duration, compared to the shorter. However, this result is likely due to the greater number of failed deactivations after a long duration compared to a shorter duration of automation. In fact, a driver who fails to press the buttons long enough the first attempt, will need additional time to succeed at the second try. Since removing the four failed attempts markedly reduces the mean TOT (from 3.53 s to 3.12 s) after a long automation duration, we propose that the increased TOT rather stems from drivers not mastering the HMI than from driver fatigue or vigilance decrements.

The reason behind why the drivers failed to deactivate automation at the first attempt could be that the drivers forgot the need to press the buttons. Assuming the likelihood of forgetting increases with time from the practice, this is supported by the fact that a greater number (3/4) of the participants that failed belonged to the group that experienced the long automation duration after the short automation duration (i.e., later during the test drive). However, since no driver failed to deactivate automation after the short automation duration, it seems that the automation duration also matters. One reason could be that the drivers had longer time to immerse into playing the game, and when it was time to deactivate automation they responded intuitively by "clicking" the buttons instead of pressing them. Their intuitive response is likely "clicking" on the buttons, since that is what is typically required for activating/deactivating existing systems such as for example adaptive cruise control (ACC).

A different HMI design or additional practice (which may naturally occur if the system would be used regularly in a commercial vehicle) may resolve the problem of failed deactivations, and consequently the TOTs after a longer exposure would decrease. For example, an HMI that do not require the drivers to press the buttons (but instead allow a short click on the buttons) will likely have solved the issues with failed first deactivation attempts observed in this study. In addition, an HMI letting the drivers deactivate automation by steering or braking may also have been beneficial. Such an HMI may also have shortened the time the drivers needed to look down on the HMI before they looked up on the forward path. This is because steering and braking typically do not require the drivers to look down and search for the controls (which may be the case for buttons on the steering wheel). A driver monitoring system (DMS) may be useful to identify drivers that are distracted or drowsy and at risk of not mastering the HMI and may therefore fail to deactivate automation at the first attempt. However, the present study did not manipulate any factors related to distraction or drowsiness. In fact, all drivers in the present study were instructed to engage in a visually demanding task while driving with automation (i.e., the drivers were almost only looking off path). Therefore, to inform future DMS systems, more research is needed to better understand potential factors (e.g., drowsiness, glance behaviors) that may correlate with the take-over outcome (i.e., failed or successful automation deactivation).

The present study found that, on average, a driver would need 3.22 seconds (for both automation durations) to deactivate automation in response to the TOR. This time is smaller than the predicted mean TOT of about 3.75 s obtained from [7, pg. 652], when assuming a take-over time-budget of 10 s. On the other hand, a mean TOT of 3.22 s is in line with the predicted mean TOT of about 3.20 s obtained from [9, pg. 294], assuming the same take-over time budget (referred to as the *Time Budget to Collision* in [9]). Further, the present study found that, on average, a driver would need 0.52 s additional TOT after a 9.5 min longer automation duration. This result of increased average TOTs (even if only slightly) after a longer duration is in line with [13] and [15]. [15] only observed a slight difference in mean TOT (i.e., 0.1 s longer for the 15 minutes longer duration), whereas [13] observed a difference in mean TOT similar as to the present study (i.e., 0.5 s longer for the 50 minutes longer duration).

B. The Automation Aftereffects on Driving Performance

Surprisingly, in the present study the drivers started their steering maneuver earlier after automation (both durations) compared to manual. Thus, our findings contradict the literature reporting a delay in the driver response after a period of automated driving (i.e., that drivers start steering closer to the conflict object after automation, compared to manual). One possible explanation to why this result contrasts to [7], [11], [12] is the use of different timings for *TOR triggering* and the *conflict appearance*. In the mentioned studies, these two timings happen simultaneously, whereas in the present study the TOR triggering took place about 5-6 s prior to conflict appearance. Consequently, the drivers in the present study had the chance to become *ready-to-act* (i.e., by putting their hands on the steering wheel and deactivating automation) before the conflict appearance and could start steering at similar times as when in manual driving mode. However, in the mentioned studies the drivers had not received a notification about the need to start preparing at the conflict appearance. Consequently, the drivers in the afore mentioned studies likely showed a delayed reaction due to the time needed for moving the hands to the steering wheel and the feet to the pedals before they could act. Another explanation for the observed lack of delayed driver response in the present study, compared to previous studies, could be the fact that the drivers in the present study had prior experience of the cone zone (i.e., they had practiced beforehand), which was not the case in [11]-[12].

The present study observed some effects of automation on the longitudinal vehicle speed and some influence of automation on the accelerations and steering wheel angle within the cone zone. That is, the participants increased the vehicle speed (both maximum and within the interval-of-interest) as well as the longitudinal and lateral accelerations, and decreased the steering wheel angle, after automation compared to manual. For the vehicle speed, the observed difference between the manual condition and the automation condition, indicates that drivers keep a constant speed when driving manually, whereas, after automation, the vehicle seems to slow down a little, potentially as a result of the drivers taking time to locate the accelerator pedal. Then, when they had located the pedals, the drivers accelerated and increased their speed significantly, potentially to keep up with the LV. Shortly after, when the drivers entered the cone zone, they showed a slightly different steering behavior (i.e., increased lateral accelerations, decreased steering wheel angle) compared to the manual condition. This could be explained by the interactions between the observed increase in vehicle speed and the generated accelerations. For the vehicle speed to increase, the longitudinal acceleration must have increased, but increased vehicle speed may also result in increased lateral accelerations and decreased steering wheel angles needed to avoid colliding with any cones. Thus, it may be that the steering behavior within the cone zone, after automation, is merely the result of the decreased speed while the drivers were busy finding the pedals. Therefore, the extent to which the observed difference in driving performance may also be due to a less calibrated perceptual-motor control loop remains unknown.

Overall, the increase in accelerations after automation compared to manual driving were lower than the accelerations reported in [11] and [12]. Our data suggests a maximum increase of accelerations (based on the maximum longitudinal and lateral acceleration metrics) in the range of 1.2-1.3 times higher for automation compared to manual. This difference is smaller compared to the accelerations reported in driving simulator experiments from [12] and [11]. [12] found the mean accelerations, calculated as mean(a_c) where $a_c = \mathrm{sqrt}(a_{long}^2 + a_{lat}^2)$, to be 2-3 times higher after automation compared to manual, whereas [11] found the average max lateral acceleration to be 1.7 times higher after automation compared to manual in the *distracted* condition.

Finally, when comparing manual driving with manual driving after automation, it is important to put the findings into a bigger picture. For example, in the present study automation was found to increase mean lateral accelerations compared to manual (i.e., from 0.19 m/s^2 to 0.28 m/s^2). However, this increase resulted in lateral accelerations, after automation, that were below 2.5 m/s^2, which is well below the thresholds for being considered an "evasive steering maneuver" (defined in [30] as a steering maneuver with a lateral acceleration of above approximately 4 m/s^2). Thus, in the present study, the effect of automation on driver behavior seems to be minor (i.e., automation did not result in any particularly critical situation for any of the drivers). By only focusing on the relative differences in driving performance for automation vs. manual, we may just observe the natural differences between manual driving and manual driving after automation. That is, an observed increase in vehicle speed may be because drivers need to calibrate their longitudinal control, which is not necessarily safety critical. Thus, the effects we found should be considered as mild.

C. The Influence of Test Environment and Test Protocol (TOR Timing)

Overall, the present test track study did not observe as large effects of automation as previously found in driving simulator studies. However, it is not clear in which way this difference may be the cause of different test environments (i.e., test track vs. driving simulators) or differences in test protocols (i.e., the difference in timings for the TOR and the conflict appearance). The present study collected data on a test track where real motion cues and force feedback are present, which is not the case in driving simulator studies were motion cues and force feedback are simulated or absent [31]. In addition, driving simulators have been shown to demonstrate relative validity, rather than absolute validity. That is, driving simulator studies are able to produce results in similar directions and when compared to other test environments, but the actual result magnitudes are not necessarily correct [31].

Therefore, comparisons across studies regarding the presence and absence of effects are likely to be trusted, but the comparisons of actual values (e.g., accelerations) may be affected by the differences in test environments. Consequently, previous studies that have investigated automation aftereffects in driving simulators seem therefore to overestimate (exaggerate) the effect of automation and automation duration on driving performance and response process to a TOR. To the knowledge of the authors, no previous studies have directly compared automation aftereffects across test environments. However, as a good starting point, Eriksson *et al.* [18] investigated the effect of test environment (driving simulator vs. on-road test) on the TOT. Eriksson *et al.* [18] confirmed that absolute validity could not be established and that driving simulators may exaggerate the effects of automation, since the observed mean TOT was 1.4 s longer in the driving simulator compared to the on-road study.

D. Limitations and Future Work

The results presented in this paper should be viewed in the lights of its limitations. To begin with, the experiment was performed on a test track and not in real traffic. Thus, the experiment lack realism since no surrounding traffic (only a lead vehicle) was present, the construction zone was simulated, the participants knew that they were part of an experiment and a test leader was present. However, it is difficult to perform controlled experiments with a degree of realism without encountering difficult ethical considerations associated with the risk of crashing. Further, the positive aspects of the test track (e.g., real kinematics) needs to be balanced against the limitations (e.g., the difficulties in obtaining exact timings across participants for the take-over request and the conflict appearance). Further, the participants were Volvo Cars employees who are not directly involved in the development of vehicle automation. Thus, the extent to which the results generalize to other populations remains unknown. The video reduction to obtain variables for the response process was performed by only one person. Thus, the present results depend on the judgement from one person and may have differed slightly if another person would have performed the video reduction.

First, in order to fully understand the impact of the test environment, two studies with the same experimental protocol could be performed in a driving simulator and on the test track. Second, independent of the test environments, a study which controls for the influence of the driver take-over response, on the driving performance should be performed. That is, we hypothesize that the test protocols used in previous driving simulator studies (i.e., when TOR and conflict appearance timings coincide), may have influenced the presence and size of automation aftereffects. Practically, a future study that would control for the driver take-over response, would provide the drivers with enough time after automation to become ready-to-act before presented with the conflict. Such a setup would enable us to understand if the observed automation aftereffect is merely a consequence of the time needed to become ready-to-act (i.e., the time needed for the driver take-over response process) or if there are some other underlying mechanism (e.g., sensorimotor, cognitive reaction times) at play.

V. Conclusion

To conclude, our findings demonstrate that after a duration of 4.5-14 minutes automated driving, drivers are able to safely deactivate automation and maneuver through a road-work zone. Nevertheless, the longer the exposure to automation (14 min) the more important it is for drivers to master the HMI to appropriately take over control. The present test-track study did not observe as large automation aftereffects as indicated by previous driving simulator studies. The extent to which this is due to the use of different test environments (test track vs. simulator) or different protocols needs to be investigated but is likely an overestimation due to driving simulator characteristics. However, a key point is that independent of test environment, the timing of when the TOR is triggered and when the conflict becomes visible for the driver is an important influential factor, because this relation may influence the presence and magnitude of an automation aftereffects (larger aftereffects when TOR closer to conflict). Previous studies which have observed automation effects, triggered the TOR at the same time as the conflict became visible. Consequently, when comparing driving performance in a manual condition with manual driving after automation, such an automation aftereffect is at least partly due to the additional time drivers need to become *physically ready-to-act* after automation (put hands on wheel etc.). Drivers in manual mode can act directly (e.g., without taking extra time to locate steering wheel and pedals) as they are already controlling the vehicle. For a fair comparison, future studies should control for the influence of the driver take-over response (e.g., the take-over time) on the driving performance, by giving the drivers enough time to become physically ready-to-act, before the conflict becomes visible.

ACKNOWLEDGMENT

The authors would like to thank Pär Gustavsson for contributions to the experimental protocol. The authors would also like to thank Elin Meltzer, Joel Johansson, and Regina Johansson for performing the data collection and Carol Flannagan for discussions about, and assistance with, the Bayesian Statistics.

REFERENCES

[1] Thatcham Research. (2018). *Assisted and Automated Driving*. [Online]. Available: https://www.abi.org.uk/globalassets/files/publications/public/motor/2018/06/thatcham-research-assisted-and-automated-driving-definitions-summary-june-2018.pdf

[2] *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, Standard J3016, SAE International, 2016.

[3] L. Bainbridge, "Ironies of automation," *Automatica*, vol. 19, no. 6, pp. 775–779, Nov. 1983.

[4] B. D. Seppelt and T. W. Victor, "Potential solutions to human factors challenges in road vehicle automation," in *Road Vehicle Automation*. Cham, Switzerland: Springer, 2016, pp. 131–148.

[5] M. R. Endsley and E. O. Kiris, "The out-of-the-loop performance problem and level of control in automation," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 37, no. 2, pp. 381–394, Jun. 1995.

[6] N. Merat *et al.*, "The 'out-of-the-loop' concept in automated driving: Proposed definition, measures and implications," *Cognition, Technol. Work*, vol. 21, no. 1, pp. 87–98, 2019.

[7] A. D. McDonald *et al.*, "Toward computational simulations of behavior during automated driving takeovers: A review of the empirical and modeling literatures," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 61, no. 4, pp. 642–688, Jun. 2019.

[8] C. D. Mole, O. Lappi, O. Giles, G. Markkula, F. Mars, and R. M. Wilkie, "Getting back into the loop: The perceptual-motor determinants of successful transitions out of automated driving," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 61, no. 7, pp. 1037–1065, Nov. 2019.

[9] B. Zhang, J. de Winter, S. Varotto, R. Happee, and M. Martens, "Determinants of take-over time from automated driving: A meta-analysis of 129 studies," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 64, pp. 285–307, Jul. 2019.

[10] T. Louw, G. Markkula, E. Boer, R. Madigan, O. Carsten, and N. Merat, "Coming back into the loop: Drivers' perceptual-motor performance in critical events after automated driving," *Accident Anal. Prevention*, vol. 108, pp. 9–18, Nov. 2017.

[11] T. Louw, N. Merat, and H. Jamson, "Engaging with highly automated driving: To be or not to be in the loop?" in *Proc. Driving Assessment Conf.*, 2015, pp. 1–8.

[12] C. Gold, D. Damböck, L. Lorenz, and K. Bengler "'Take over!' How long does it take to get the driver back into the loop?" in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 57, no. 1. Los Angeles, CA, USA: Sage Publications, 2013, pp. 1938–1942.

[13] A. Bourrelly, C. J. de Naurois, A. Zran, F. Rampillon, J.-L. Vercher, and C. Bourdin, "Long automated driving phase affects take-over performance," *IET Intell. Transp. Syst.*, vol. 13, no. 8, pp. 1249–1255, Aug. 2019.

[14] O. Jarosch and K. Bengler, "Is it the duration of the ride or the non-driving related task? What affects take-over performance in conditional automated driving?" in *Proc. Congr. Int. Ergonom. Assoc.* Cham, Switzerland: Springer, 2018, pp. 512–523.

[15] A. Feldhütter, C. Gold, S. Schneider, and K. Bengler "How the duration of automated driving influences take-over performance and gaze behavior," in *Advances in Ergonomic Design of Systems, Products and Processes*. Berlin, Germany: Springer, 2017, pp. 309–318.

[16] E. T. Greenlee, P. R. DeLucia, and D. C. Newton, "Driver vigilance in automated vehicles: Hazard detection failures are a matter of time," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 60, no. 4, pp. 465–476, Jun. 2018, doi: 10.1177/0018720818761711.

[17] C.-N. Boda, M. Dozza, K. Bohman, P. Thalya, A. Larsson, and N. Lubbe, "Modelling how drivers respond to a bicyclist crossing their path at an intersection: How do test track and driving simulator compare?" *Accident Anal. Prevention*, vol. 111, pp. 238–250, Feb. 2018.

[18] A. Eriksson, V. A. Banks, and N. A. Stanton, "Transition to manual: Comparing simulator with on-road control transitions," *Accident Anal. Prevention*, vol. 102, pp. 227–234, May 2017, doi: 10.1016/j.aap.2017.03.011.

[19] F. Naujoks, C. Purucker, K. Wiedemann, and C. Marberger, "Noncritical state transitions during conditionally automated driving on german freeways: Effects of non–driving related tasks on takeover time and takeover quality," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 61, no. 4, pp. 596–613, Jun. 2019, doi: 10.1177/0018720818824002.

[20] (2020). *AstaZero*. [Online]. Available: http://www.astazero.com/

[21] A. Habibovic, J. Andersson, M. Nilsson, V. M. Lundgren, and J. Nilsson "Evaluating interactions with non-existing automated vehicles: Three Wizard of Oz approaches," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2016, pp. 32–37.

[22] (2020). *DOTS*. [Online]. Available: https://www.dots.co/

[23] *Operational Definitions of Driving Performance Measures and Statistics*, Standard SAE J2944, 2015.

[24] J. K. Kruschke, *Doing Bayesian Data Analysis: A Tutorial With R, JAGS, and Stan*. New York, NY, USA: Academic, 2014.

[25] G. Cumming, "The new statistics: Why and how," *Psychol. Sci.*, vol. 25, no. 1, pp. 7–29, Jan. 2014.

[26] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, "Probabilistic programming in Python using PyMC3," *PeerJ Comput. Sci.*, vol. 2, p. e55, Apr. 2016.

[27] T. Yarkoni and J. Westfall, "Bambi: A simple interface for fitting Bayesian mixed effects models," *OSF Preprints*, 2016, doi: 10.31219/osf.io/rv7sn.

[28] M. D. Homan and A. Gelman, "The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo," *J. Mach. Learn. Res.*, vol. 15, no.1, pp. 1593–1623, Jan. 2014.

[29] A. Gelman and D. B. Rubin, "Inference from iterative simulation using multiple sequences," *Stat. Sci.*, vol. 7, no. 4, pp. 457–472, Nov. 1992.

[30] T. Victor, M. Dozza, J. Bärgman, C. N. Boda, J. Engström, C. Flannagan, J. D. Lee, and G. Markkula, "Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk," Transp. Res. Board Nat. Academies, Washington, DC, USA, Tech. Rep. S2-S08A-RW-1, 2015.

[31] D. L. Fisher, *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*. Boca Raton, FL, USA: CRC Press, 2011.

Linda Pipkorn received the M.Sc. degree in mechanical engineering from the Chalmers University of Technology in 2018, where she is currently pursuing the Ph.D. degree in human factors of automation. The aim of her research is to contribute to the understanding of the drivers' response process after a period of automated driving.

Trent Victor received the Ph.D. degree in psychology from Uppsala University, Uppsala, Sweden, in 2005. He is currently a Senior Technical Leader Crash Avoidance with the Volvo Cars Safety Centre, an Adjunct Professor with the Chalmers University of Technology (affiliated with the SAFER Vehicle and Traffic Safety Center), Gothenburg, Sweden, and an Adjunct Professor with the University of Iowa. At the Volvo Cars Safety Centre he provides leadership safety-related issues in the development of crash avoidance systems, autonomous vehicles, safety impact analyses, and human factors. He has over 22 years of work experience in design and evaluation in the field of intelligent vehicle systems but his deepest technical knowledge is within driver attention and visual behavior.

Marco Dozza received the Ph.D. degree in bioengineering from the University of Bologna, Italy, in collaboration with Oregon Health and Science University, Portland, OR, USA, in 2007. After graduation, he worked as a System Developer with Volvo Technology, a research and innovation company inside the Volvo Group. Since 2009, he has been with the Chalmers University of Technology, Gothenburg, Sweden, where he is currently a Professor and leads the unit for Crash Analysis and Prevention. He is also an Examiner for the course Active Safety in the Master's Programme for Automotive Engineering and collaborates with SAFER, the Vehicle and Traffic Safety Center with Chalmers.

Emma Tivesten received the Ph.D. degree in machine and vehicle systems from the Chalmers University of Technology in collaboration with Volvo Car Corporation, Gothenburg, Sweden, in 2014. She is currently working as a Senior Research Leader with the Volvo Cars Safety Centre. Her research focuses on driver performance in manual and assisted driving, including both naturalistic driving as well as test-track experiments.