Impact of the driver's visual engagement on situation awareness and takeover quality

Paul Marti, Christophe Jallais, Arnaud Koustanaï, Anne Guillaume, Franck Mars

Impact of the driver's visual engagement on situation awareness and takeover quality

Paul Marti[1,3], Christophe Jallais[2], Arnaud Koustanaï[1], Anne Guillaume[1], Franck Mars[3]

[1] Renault SAS, Laboratoire d'accidentologie, de biomécanique et d'étude du comportement humain, Nanterre, France
(E-mail : paul.marti@orange.fr, arnaud.koustanai@renault.com, anneivyguillaume@gmail.com)

[2] Université Gustave Eiffel, IFSTTAR, TS2-LESCOT, Lyon, France
(E-mail: Christophe.jallais@univ-eiffel.fr)

[3] Nantes Université, CNRS, Laboratoire des sciences du numérique de Nantes, Nantes, France
(E-mail : franck.mars@ls2n.fr)

Corresponding author:
Dr Franck Mars
UMR CNRS 6004 – Laboratoire des Sciences du Numérique de Nantes (LS2N)
1, rue de la Noë – B.P. 92101
44321 Nantes Cedex 03
E-Mail : franck.mars@ls2n.fr

ABSTRACT

During automated driving (SAE Level 3), drivers can delegate control of the vehicle and monitoring of the road to an automated system. They may then devote themselves to tasks other than driving and gradually lose situational awareness (SA). This could result in difficulty in regaining control of the vehicle when the automated system requires it. In this simulator study, the level of SA was manipulated through the time spent performing a non-driving task (NDRT), which alternated with phases where the driver could monitor the driving scene, prior to a critical takeover request (TOR). The SA at the time of TOR, the visual behaviour after TOR, and the takeover quality were analysed. The results showed that monitoring the road just before the TOR allowed the development of limited perception of the driving situation, which only partially compensated for the lack of a consolidated mental model of the situation. The quality of the recovery, assessed through the number of collisions, was consistent with the level of development of SA. The analysis of visual behaviour showed that engagement in the non-driving task at the time of TOR induced a form of perseverance in consulting the interface where the task was displayed, to the detriment of checking the mirrors. These results underline the importance of helping the driver to restore good SA well in advance of a TOR.

Keywords: Automated driving; Taker-over request; Human factors; Non-driving task; Situation awareness.

1. Introduction

Developments in the automotive industry have contributed to the creation of vehicles in which increasing numbers of functions are delegated to an automated system. The Society of Automotive Engineers International (SAE) has defined six levels of automation, from the lowest level (Level 0) to the highest level of automation (Level 5; SAE International, 2016). At Level 3, the driver can delegate vehicle control and road monitoring to the automated system on roads that meet certain conditions. This delegation allows drivers to engage in non-driving tasks while the driving is automated. However, the driver must be receptive at all times to a request to intervene from the automation and ready to serve as a "fallback ready user". The driver must take over manual control of driving or, if that is not possible, must achieve a minimal risk condition. The takeover request (TOR) can occur for two reasons: first, the system may anticipate that the vehicle is approaching the end of its operational domain, in which case it will be a planned and non-critical takeover. The second reason is that the system may be not able to handle a situation, or a problem may occur that prevents the system from operating properly.

Unplanned critical takeovers are sometimes characterised by a high level of hazard with a short time to regain control of the vehicle. Numerous scientific works have studied the risk factors during the takeover. Zhang et al. (2019) synthesised the results of 129 studies analysing the takeover time according to various criteria, such as time budget (from 2 s to 15 s), traffic complexity, and driver age. They concluded that performing a non-driving task, even without a handheld device, increased the takeover time compared to not performing such a task. Task modality, classified as either auditory or cognitive, did not influence the takeover time. Other studies have focused on situational awareness (SA) and takeover (e.g. Endsley, 2018; Ma & Kaber, 2005; White et al., 2019; Winter et al., 2014). The present study contributes to these efforts by studying how a non-driving related task (NDRT) can affect the driver's SA and the takeover quality in a critical situation.

A major concern in the delegation of vehicle control and road supervision is that automated systems can cause the driver to be "out-of-the-loop". To formalise this problem, Merat et al. (2019) distinguished three states. When the driver acts on the vehicle's controls and supervises the road, they are "in-the-loop". If they delegate the operational control of the vehicle to the automated system but continue to monitor the road scene, they are "on-the-loop". When the driver no longer monitors the situation, they are considered to be "out-of-the-loop". In Level 3 automation, drivers may be on-the-loop if they supervise the road or out-the-loop if they are engaged in non-driving activities. Potential drivers of autonomous vehicles have been shown to want to perform new activities that require attentional resources such as reading, writing messages, eating, drinking, browsing the internet, or making phone calls (Pfleging et al. 2016; Shi & Frey, 2021). Drivers can thus be expected to engage in non-driving activities that may divert their attention from the road (Naujoks et al., 2016). In turn, that engagement can lead to impaired SA.

Situational awareness is a dynamic process that can be defined as the perception and understanding of a context, which enables the individual to anticipate an impending scenario. Endsley & Kiris (1995) proposed a three-level model. The first level of SA is the perception of elements relevant to the task in the current situation. At the second level, the operator integrates the perceived elements to understand their impact on the current objectives. At the last level, the operator anticipates the dynamics of the elements to predict how they will affect the environment and the operator's goals. In other words, SA displays different levels of information elaboration. The perception of the context (Level 1) is based on immediate information, which is processed to give meaning to the current situation (Level 2). Because the construction of a mental model allowing projection into the future (Level 3) requires the integration of information over a longer period, monitoring of the road scene determines the quality of SA. Merat et al. (2019) illustrated how SA fits into the driving activity (see Figure 1).

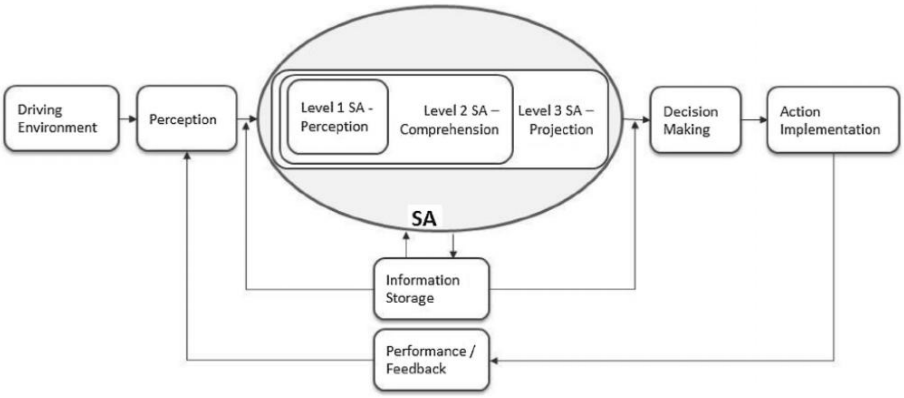

Figure 1. SA in driving information processing; reproduced from Merat et al (2019)

During manual driving, performing a secondary task or talking with another person can impair SA (Gugerty et al., 2003; Heikoop et al., 2018; Ma & Kaber, 2005). For example, Green (1999) showed that self-paced glances at in-vehicle controls and displays typically did not exceed 1.2 s or 1.5 s. Rockwell (1988) also showed that drivers were loath to wait more than 2 s to obtain road information. This 2-s rule can be considered the critical value for distraction in cars. It appears to be the interval beyond which the absence of road scene monitoring critically impairs the perceptual determinants of SA.

Because drivers are prone to be distracted and no longer control the vehicle, the 2-second rule ceases to be valid in highly automated vehicles. For example, Zeeb et al. (2015) performed a simulator study in which participants had to drive for 26 min in highly automated driving. They had to perform "Texting" and "Internet Search" on the multimedia system. By the end of the driving task, some drivers spent up to 55 s without looking at the road. Drivers who spent considerable time without looking at the road showed longer braking reaction times and higher collision rates than those who did not. Winter et al. (2014) also presented a meta-analysis of 32 studies about the impact of adaptive cruise control and highly automated driving on workload and SA. Highly automated driving impaired SA when the driver chose to perform an NDRT at the expense of monitoring the driving scene (Carsten et al., 2012; Merat et al., 2019). Marti et al. (2021) reported that the difficulty of a NDRT performed before a TOR did not influence the success of a critical takeover. However, looking at the task display at the time of TOR increased the risk of collision.

Studies in various fields have investigated the relationship between situation awareness and visual strategies. In the air traffic control domain, Moore & Gugerty (2010) found that controllers' SA score depended on visual attention paid not only to the most important aircrafts, but also to surrounding aircrafts. Too much visual focus on important aircrafts can create attention tunnelling and degrade SA. Van de Merwe et al. (2012) studied the gaze behaviour of pilots when a malfunction occurred during a flight. Gaze entropy (visual scanning activities around the cockpit) increased as SA decreased. Gartenberg et al. (2014) tried to characterize SA recovery after a task interruption. Shorter fixation durations, increased the number of objects scanned, longer resumption lags, and a greater likelihood of refixating objects that were previously looked at were identified as indicators of SA recovery. More closely related to the autonomous vehicle domain, Kunze et al. (2019) showed that drivers who performed shorter fixations in a peripheral search task during the 40 seconds before a TOR exhibited higher SA scores. In another study, Liang et al. (2021) found that greater gaze dispersion and more time looking at the road scene were positively correlated with SA scores. They also showed that previous engagement in a NDRT impairs SA after the TOR. Thus, SA depends on an appropriate distribution of visual attention.

It is therefore reasonable to assume that the quality of SA depends on the driver's visual behaviour and determines the driver's ability to safely regain control of the vehicle. This is especially true in complex situations, such as unexpected obstacle avoidance. A question is how long it takes to reconstruct a sufficiently elaborated SA. It is hypothesised that if the driver’s perception of the immediate context is sufficient (Level 1 SA), looking at the road 2 s before resuming vehicle control would be effective to obtain a good quality of takeover. This outcome would be independent of the driver's visual behaviour during the preceding automated driving phase. By contrast, if there is an advantage to having more elaborate SA (i.e., an appropriate mental model of the situation built over time), takeover quality should improve when the driver has not been distracted for long before resuming control.

The present study examines the impact of visual distraction on the quality of vehicular control. Two temporal windows of visual engagement in a NDRT preceding the TOR are distinguished. This gives rise to four experimental conditions in which the gaze behaviour and the quality of the takeover will be analysed.

- The first group of participants will not have a NDRT to perform. They will have every opportunity to build a good Level 3 SA over the course of autonomous driving and they will be attentive to the road scene when the TOR is delivered.
- The second group of participants will be placed in identical conditions to the previous one, except that they will be distracted by the NDRT for the two seconds before the TOR. This condition will therefore allow to evaluate the influence of impairing the immediate perception of the road scene in drivers who have had time to build up a good SA beforehand.
- The third group will have to perform the NDRT during an extended period of 5 minutes but 2 seconds before the TOR. The NDRT interruption will allow the driver to reacquire a vision of the immediate driving environment right before the TOR intervenes. These participants will therefore not be able to build and update SA during the NDRT, but will not be distracted from the road scene at the time of the TOR.
- The last group of participants will have to perform the NDRT until the TOR. SA will be the most severely impacted and the TOR will intervene when the driver’s gaze is not on the road. The driver will only have the time between him and the obstacle to analyse the situation and decide on the manoeuvres to be executed.

2. Materials and Methods

2.1. Participants

There were 88 participants (33 women, 55 men) aged between 18 and 56 years (M = 24.01, SD = 7.67) in the study. They all held a valid driving licence and had average driving experience of 7,704 km/year (SD = 8500, Min = 200, Max = 40000). They had normal vision or vision corrected with contact lenses. All participants gave written informed consent in accordance with the Declaration of Helsinki. The experiment was approved by the non-interventional research ethics committee of Nantes University (CERNI, IRB #IORG0011023; approbation #08072021-3).

2.2. Experimental setup

Participants drove on a fixed-base simulator (see Figure 2) consisting of an adjustable seat, a steering wheel with force feedback, a gear lever, clutch, accelerator and brake pedals. The software SCANeR Studio (v1.8) displayed the driving scene on three screens at a field of view of about 120°. An 11" tablet, positioned about 35° to the right and 15° below the direction straight ahead of the driver's head, served as the centre console where the NDRT was displayed.

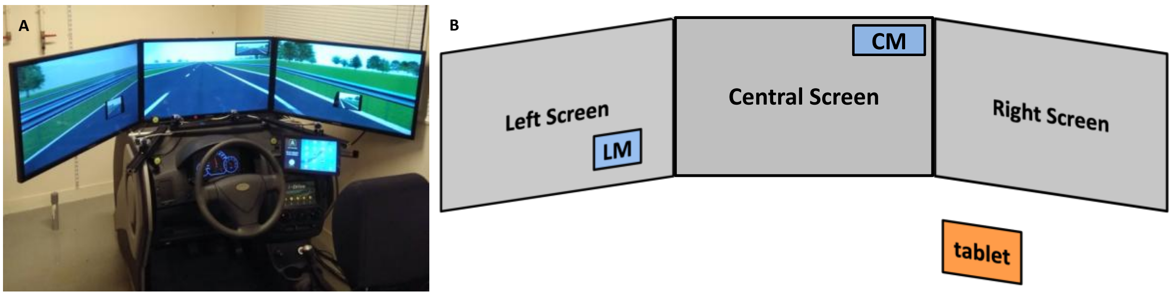

Figure 2. A. The LS2N driving simulator. B. Areas of interest considered for the analysis of gaze behaviour: the road scene in grey, the left (LM) and central (CM) mirrors in blue and the tablet in orange.

Gaze behaviour was recorded via a Smart Eye Pro eye tracker (version 5.9), which included four cameras: two below the central screen and one below each lateral screen. Gaze and vehicle data were recorded and synchronised at 60 Hz.

2.3. Procedure

2.3.1. Installation and instruction

First, participants were invited to settle in the simulator seat and the eyetracker calibration was performed. Then, participants were informed about the operation of an automated SAE Level-3 vehicle. They were asked to switch from manual to automated driving when the automated system required it, using a touchscreen button on the tablet. A confirmation pictogram was displayed (see Figure 3B). Then, participants could release the control of the vehicle by removing their hands from the steering wheel and their feet from the pedals. In the case of an unplanned TOR, a red pictogram associated with an auditory warning was displayed. The drivers had 8 s to resume control of the vehicle, either by pressing a touchscreen button on the tablet or by acting on the pedals (accelerator pedal threshold at 1% of total possible depression, brake pedal threshold at 1 N) or steering wheel (torque threshold at 1 N.m). During automated driving, participants were instructed to perform an NDRT. As Shi & Frey (2021) indicated that participants primarily wanted to engage in reading, the task consisted of reading aloud a text that scrolled automatically on the side tablet. The text was an excerpt of the story of Tom Sawyer. Participants were asked to read the story as they would read it for another person. After pretests, the scrolling speed was chosen to make the task demanding enough that participants did not have time to look at the road.

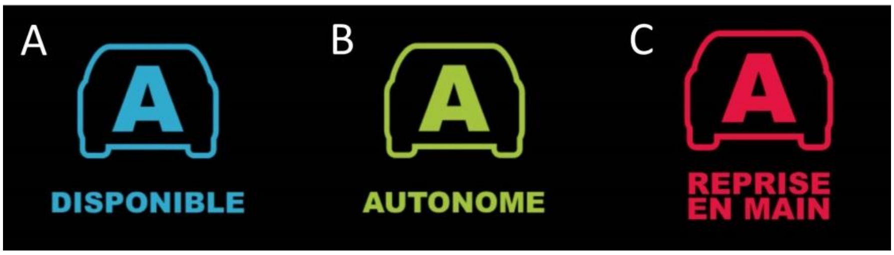

Figure 3. Pictograms displayed on the HMI: A: autonomous driving available; B: autonomous driving activated; C: take-over request (8s).

2.3.2. Scenario

Prior to conducting the experimental trial, participants completed two familiarization trials. The first allowed participants to test the transition between manual and autonomous driving 2 times. In the second training trial, they were asked to perform the NDRT in the autonomous driving phase.

Then, the experimental trial was carried out. Participants drove on a 3-lane highway at 110 km/h, with moderate traffic, for 8 min before a critical unplanned TOR happened. They started on a highway insertion ramp in manual mode. One minute after entering on the highway, they were invited by the system to switch to automated mode. After 2 minutes of automated driving, 5 minutes remained before the TOR. During that time, the task of the participants (performing the NDRT or not) depended on the experimental condition (see below). As soon as the scenario ended, all screens were switched off and the participants reported their SA. This was performed on a touch tablet. Participants were asked to report the vehicles they were aware of at the time of the TOR on a scheme representing the road scene in top view. Their own vehicle and the obstacle were already placed, so they only had to place other vehicles they had seen before the TOR.

2.3.3. Unplanned critical takeover

At the end of the experimental trial, participants had to regain control of the vehicle in a complex critical situation (see Figure 4) that unfolded as follows:

(1) 5 min before the TOR, the participant’s vehicle started to follow a lead vehicle with a time headway of 3 s.

(2) 2 min before the TOR, the participant’s vehicle and the lead vehicle began to move into the centre lane to overtake two slower vehicles.

(3) 1 min before the TOR, two rapidly moving vehicles – travelling 2 s apart – approached from behind in the left lane. They travelled at 120 km/h and started to overtake the participant’s vehicle. The first of these fast vehicles was called FV1 and the second FV2. They were visible in the central and left mirrors 20 s before the TOR.

(4) At 1 s before TOR, the lead vehicle changed lanes to avoid a vehicle that had stopped across the right and centre lanes.
(5) At the time of TOR, participants had only a partial direct view of the obstacle vehicle, because the lead vehicle was still changing lanes. The emergency TOR was delivered. If the NDRT was in progress, it was deactivated.

Participants had 8 seconds to resume control of the vehicle before reaching the obstacle (time to collision). To successfully intervene, participants could either brake and try to move between the two fast overtaking vehicles in the left lane or could change lanes after they had both passed. Alternatively, they could stop in the centre lane before reaching the obstacle.

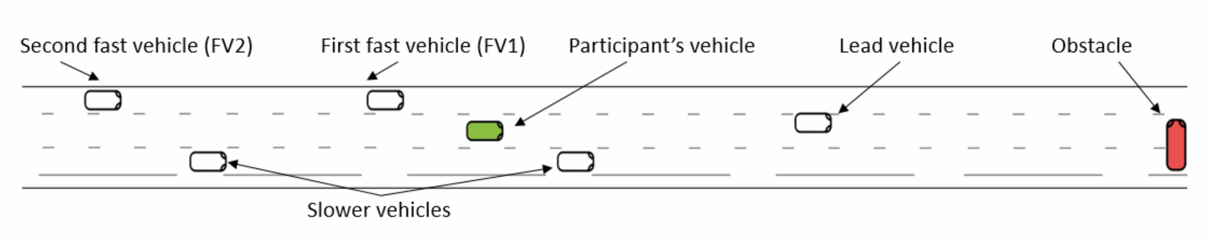

Figure 4: Critical case

Although the time to collision was 8 s at the moment of the TOR, drivers had less time to regain SA due to the movements of other vehicles. Indeed, FV1 disappeared in the blind spot 3.2 s (SD = 0.1) after the TOR. If participants did not detect the vehicle while it was visible in the mirrors, they were likely to initiate a lane change without being aware of the vehicle's presence; this would result in a collision or a late abortion of the manoeuvre. For the analysis of gaze behaviour after the TOR, we refer to this crucial 3.2-s period as the "critical phase".

2.3.4. Experimental conditions

Participants were instructed to continuously monitor the road scene, except when they were asked to perform the NDRT. The session started with 1 min of manual driving, followed by 2 min of automated driving without any NDRT. The last 5 min of the trial depended on the experimental condition the participant had been assigned to (see Figure 5):

- Full SA: No NDRT was required.
- SA_NDT: The NDRT was required only during the last 2 s before the TOR.
- NDT_SA: Participants had to perform the NDRT during the 5 min, except for during the last 2 s before the TOR.
- Full NDT: The participant had to perform the NDRT continuously up to the TOR.

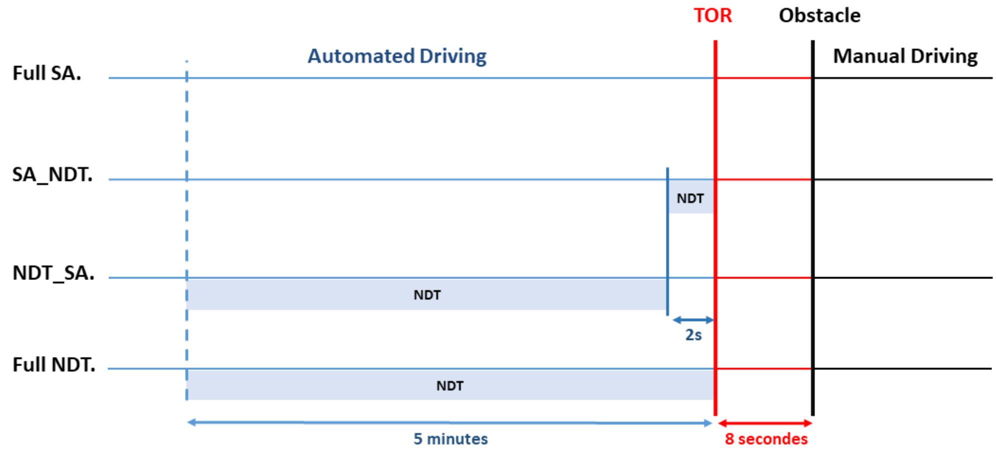

Figure 5. Timeline of the experimental conditions

Participants were randomly assigned to one of the experimental conditions.

3. Data Analysis

The effects of the experimental conditions on the number of collisions and the number of times FV1 was reportedly perceived were analysed using chi-square tests. Because the number of times FV2 was reported was too low, Fisher's exact test was used.

Gaze behaviour during the critical phase (post-TOR) was analysed by considering three areas of interest (AOIs): the tablet, the left and centre mirrors, and the road scene (see Figure 2B). For each participant, we analysed gaze data to extract fixations and saccades.

First, the time between the TOR and the first fixation toward the road or the mirrors was calculated for each participant. Shapiro's test showed that the dataset was not normally distributed. Hence, a non-parametric method (Kruskal-Wallis) was used to assess the effect of the experimental conditions.

Second, for each AOI considered independently, the numbers of participants who made at least one fixation at that AOI after the TOR were compared using the chi-square test.

The proportion of fixation time spent on the different AOIs was also compared. For each participant, this period corresponded to the cumulative fixation time on each AOI, divided by the total fixation time during the critical phase. Since the dataset was not normally distributed, the Kruskal-Wallis test was used. Saccade time was not taken into account.

Finally, the evolution of gaze distribution during the 3 seconds after the TOR was scrutinized. For each second during this period, the percentage of time spent on each AOI was calculated. As the dataset was not normally distributed, we used the Kruskal-Wallis test to compare the effect of the conditions on the mean percentage of time for each AOI at each second.

4. Results

Eighty-eight trials were performed in total, one for each participant. Of these, 45 resulted in a collision with another vehicle: 23 with FV1, 10 with FV2, and 12 with the obstacle. No collision occurred with the vehicles travelling in the right-hand lane.

4.1. Effect of conditions on the occurrence of collisions

When considering the occurrence of collisions, a main effect of the experimental conditions was found ($chi^2 = 8.50$, p = <.05). Figure 6 shows that the percentage of collisions was higher in the Full_NDT condition (73% or 16/22 participants) than in the NDT_SA condition (59% or 13/22 participants). In turn, the NDT_SA condition resulted in more collisions than either the Full_SA or SA_NDT conditions (36% or 8/22 participants for each of the latter groups).

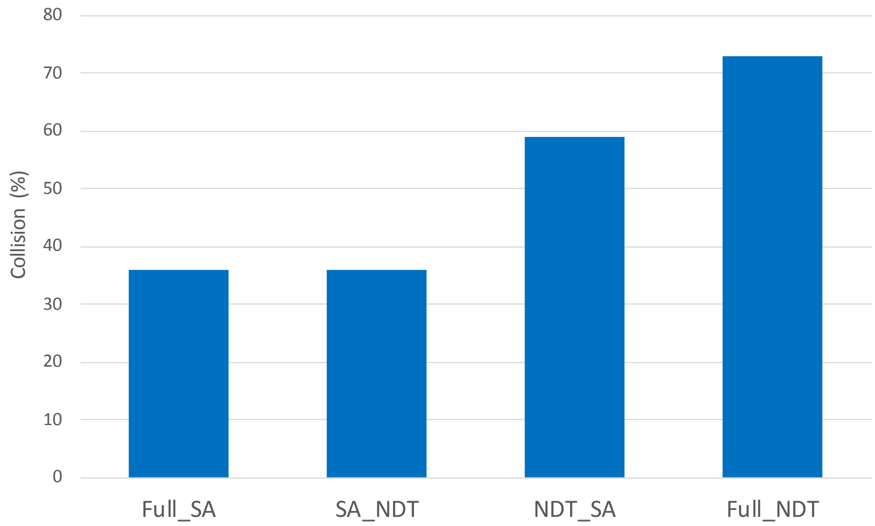

Figure 6. Percentage of participants who collided in each condition

4.2. Effect of conditions on the awareness of the first vehicle

Since none of the participants attempted to avoid the obstacle from the right, only the fast vehicles in the left lane were relevant. FV2 was rarely reported by participants, as the distance made it relatively inconspicuous in the mirrors (4/22 participants for Full_SA, 6/22 participants for SA_NDT, 3/22 participants for NDT_SA, and 0/22 participants for Full_NDT). The results of the Fisher exact test (p >.05) did not show any effect of the conditions on drivers' awareness of FV2. By contrast, as shown in Figure 7, FV1 was often reported in the Full_SA and SA_NDT conditions (82% or 18/22 participants in both cases). It was reported less often in the NDT_SA condition (45% or 10/22 participants) and rarely (9% or 2/22 participants) in the Full_NDT condition (chi2 = 32.267, p <.05).

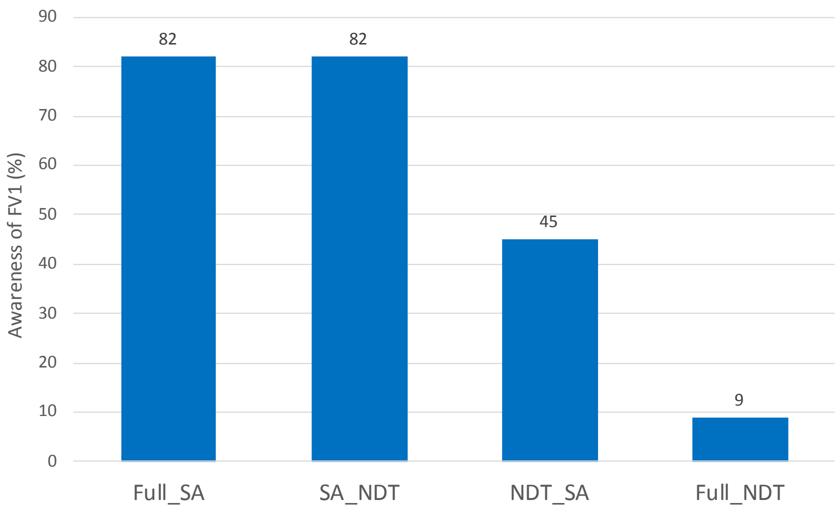

Figure 7. Percentage of participants in each experimental condition who were aware of FV1 at the moment of TOR

4.3. Effect of conditions on fixations during the critical phase after TOR

4.3.1. First fixation after the TOR on the road or the mirrors

Figure 8 shows a significant effect of the experimental conditions on the time of the first fixation to the road or the mirrors (chi^2 = 47,835, $p < 0.05$). While almost all participants in the Full_NDT and SA_NDT groups looked at these AOIs at the time of the TOR or shortly thereafter, those in the SA_NDT and Full_NDT groups took more than a second on average to do so.

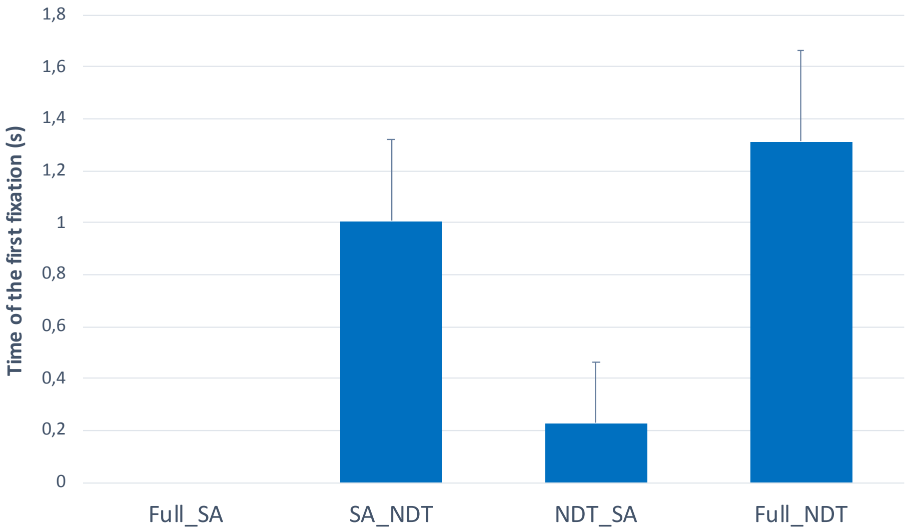

Figure 8. Mean time of the first fixation to the road or one of the mirrors after the TOR

4.3.2. Number of participants who looked at AOIs

More participants made fixations to the left and/or central mirrors in the Full_SA and NDT_SA conditions than in the Full_NDT and SA_NDT conditions (chi^2 = 13.149, $p < 0.05$).

No significant effect was found for fixations towards the road or the tablet for the various experimental conditions (see Figure 9).

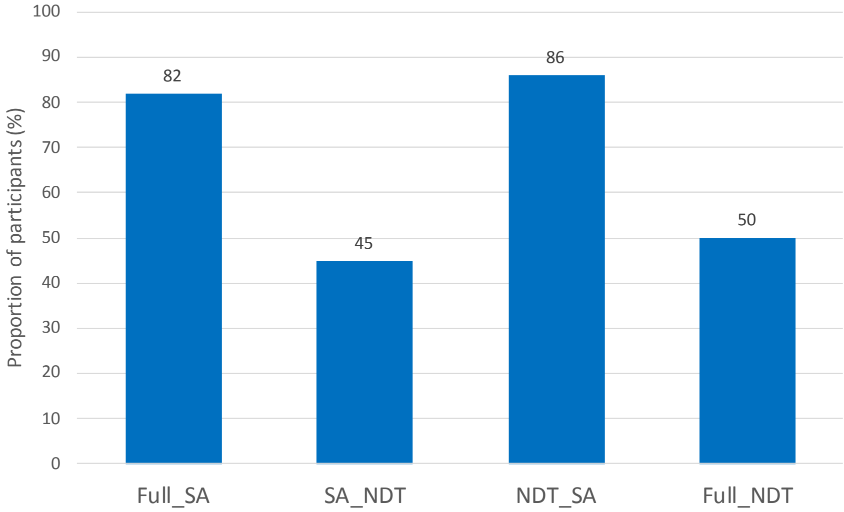

Figure 9. Percentage of participants in each condition who made fixations at the left and centre mirrors at least once

4.3.3. Proportion of fixation time spent on AOIs

Figure 10 shows the proportion of fixation time spent looking at the three AOIs. Participants spent more time to look at the left and central mirrors in the Full_SA and NDT_SA conditions than in the Full_NDT and SA_NDT conditions ($chi^2 = 17.35$, $p < 0.05$). Conversely, they spent more time looking at the tablet in the Full_NDT and SA_NDT conditions than in the Full_SA and NDT_SA conditions ($chi^2 = 14.83$, $p < 0.05$). Time spent to make fixations at the road was not significantly different between the different conditions ($chi^2 = 1.17$, $p = 0.328$).

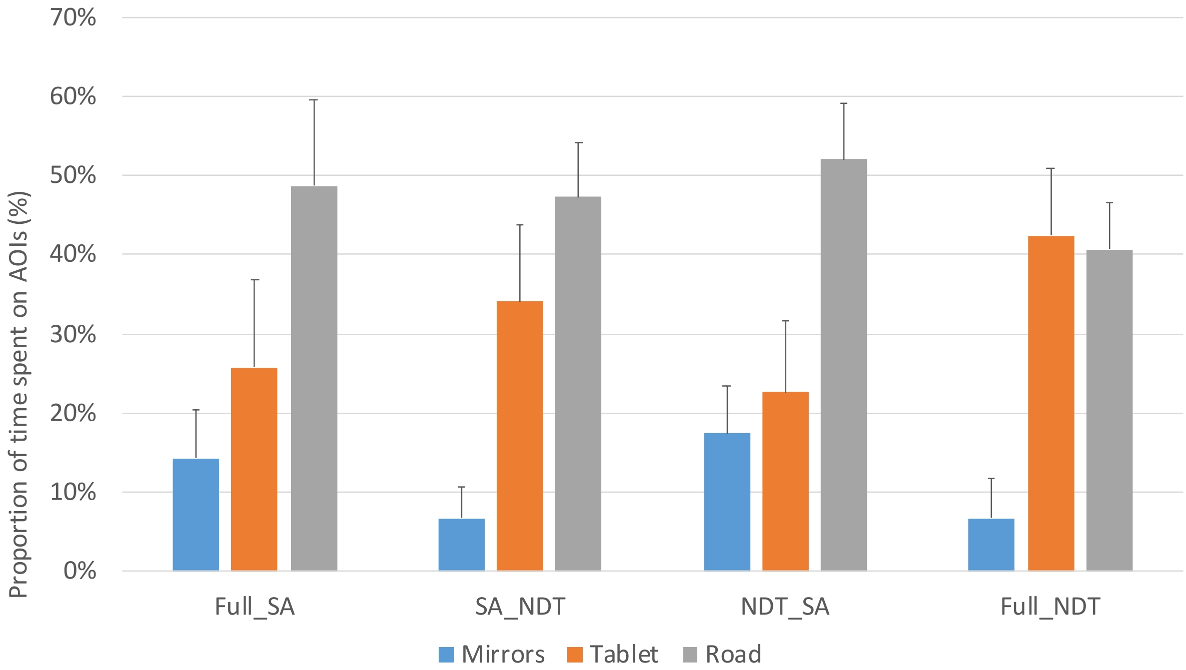

Figure 10. Influence of experimental condition on distribution of fixation during critical phase after TOR. (The three AOIs are depicted in Figure 2.B)

4.3.4. *Effect of conditions on the time spent each second on each AOI for seconds*

Figure 11.A represents the percentage of time spent looking at the road during each of the first 3 seconds following the TOR. An effect of the experimental condition is observed only for the first second ($chi^2 = 32.077$, $p < 0.05$). The Full_NDT group spent significantly less time on the road compared to NDT_SA and SA groups. The SA_NDT group only differed from the Full_SA group. The same analysis was performed on the time spent looking at the tablet (see Figure 11.B). We also found an effect of conditions only for the first second ($chi^2 = 42.824$, $p < 0.05$). The Full_SA and NDT_SA groups were significantly different from the Full_NDT and SA_NDT groups. Considering now the time spent on the centre and left mirrors (see Figure 11.C), again, the effect of the conditions was found only for the first second ($chi^2 = 23{,}388$, $p < 0.05$), with the SA_NDT group showing significantly less time spent on the mirrors than the Full_SA and NDT_SA groups.

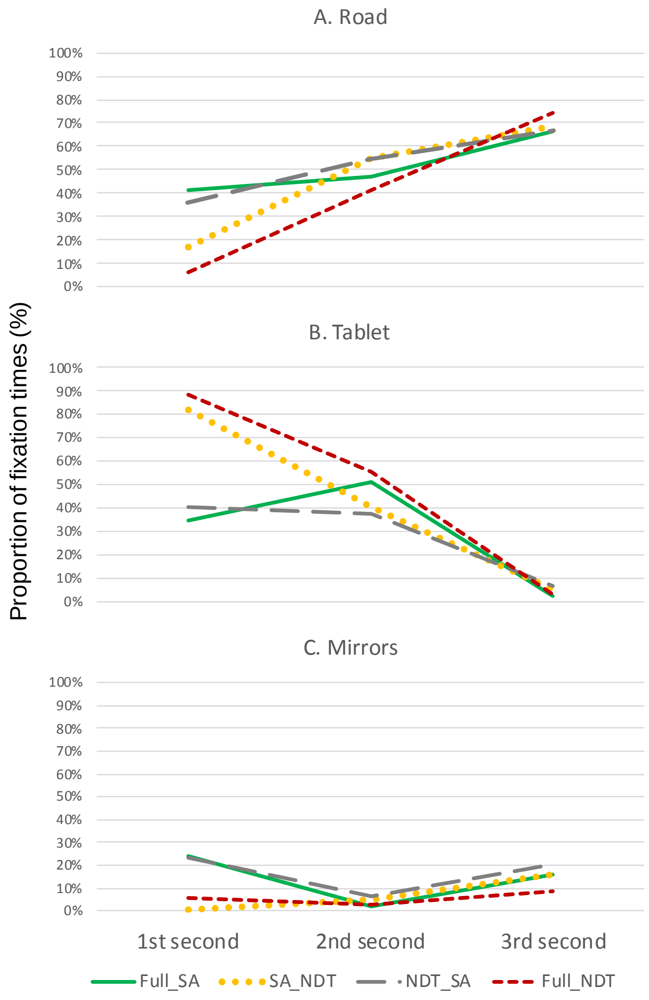

Figure 11. Percentage of time spent during the 3s-period following the TOR on A. the road, B. the tablet and C. the mirrors, depending on the experimental condition.

5. Discussion

In this study, we manipulated the driver's ability to monitor the driving scene over a long period or only during the 2 s preceding a TOR. By doing so, we were able to assess the importance of being on-the-loop at the critical moment and thus being able to rely on relatively elaborate SA. In other words, we investigated whether giving drivers 2 s before they resumed control of the vehicle was sufficient to restore SA. To assess this, we asked drivers to indicate whether they had been aware of vehicles in the driving environment just before the TOR. Our results support the idea that 2 s of distraction did not impact SA if it was previously adequate. In both the SA_NDT condition and the Full_SA condition, most drivers reported being aware of the vehicle coming fast in the left lane. However, recovering the ability to monitor the driving scene just prior to the TOR was useful for drivers who had been distracted in the preceding minutes. Drivers in the NDT_SA condition reported the presence of FV1 more often than those in the Full_NDT condition. This result suggests that stopping the NDRT 2 s before the TOR allowed the development of a degree of SA, which partially compensated for the lack of a consolidated mental model of the situation.

Having good SA does not necessarily imply a smooth takeover. Indeed, the critical situation set up in this study was really difficult to negotiate, which explains the high number of collisions. The role of the NDRT in the occurrence of those collisions has been demonstrated. However, the complexity and kinematics of the driving scenario have been identified in the literature as one of the factors responsible for the reduction of take-over quality (Gold et al., 2016; Louw, Markkula, et al., 2017; Radlmayr et al., 2014; Scharfe et al., 2020; Zhang et al., 2019). Another essential point is the taker-over time budget in critical situations. Damböck & Bengler (2012) tested time budget of 4 s, 5 s, 6 s and 8 s. They found that drivers crashed more often in all time budgets except for the 8 s condition compared to manual driving. Gold et al., (2013) also found a better take-over performance for a 7s time budget than for a 5s time budget. In Mok et al. (2017) drivers could handle a critical situation with a take-over time between 5s and 8s whereas they were practicing an active NDRT. Our results showed that even with an 8s time budget and a good SA, drivers could have difficulty regaining vehicle control in a complex situation. Indeed, 1/3 of drivers had a collision in the Full_SA and SA_NDT groups, even though they reported the presence of the most dangerous vehicle in the left lane. In situations similar to the critical case we used, returning control of the vehicle to the driver would not be the best solution. It would probably be better for the automated system to decide, based on the assessment of the criticality of the situation, to perform emergency braking rather than to issue a TOR.

However, the quality of the takeover was consistent with the level of elaboration of SA. Drivers in the Full_SA and SA_NDT groups passed the critical case with the same level of success (64%), which was notably better than in the NDT_SA group (41%) – which in turn was better than the performance of the Full_NDT group (27%). Thus, just as for SA, drivers benefited from having been able to supervise the driving scene well before the TOR. The final 2 s were not sufficient to obtain a good quality of takeover if SA was low to begin with.

Previous studies have shown that the level of mental load associated with an NDRT performed at the time of a TOR was not predictive of the takeover quality (Bueno et al., 2016; Marti et al., 2021). This point does not fit well with the idea that the more distracting an NDRT is, the more difficult takeover will be. The main difference between those studies and ours is that we manipulated the duration of engagement in the NDRT before takeover, rather than its difficulty level. In addition, Marti et al. (2021) showed that the only determinant of successful critical control was whether the driver was looking at the road rather than a peripheral NDRT at the time of the TOR, regardless of the difficulty of the NDRT. This finding might have reflected in the outcomes of our experiment if participants who looked at the tablet at the time of TOR (SA_NDT and Full_NDT) had shown greater difficulty in taking over than did undistracted drivers (Full_SA and NDT_SA). This was not the case. Instead, the analysis of gaze behaviour after the TOR showed a perseverance effect, with drivers in the SA_NDT and Full_NDT conditions spending more time looking at the tablet and less time checking their mirrors. This is consistent with the observation that the first fixation to the road of the mirrors was performed about 1 s later in those conditions. The drivers who were not performing the NDRT at the time of the TOR, even if they had spent about 5 min doing it before (NDT_SA group), did not take significantly more time to look at the road or the mirror. Additional analyses were conducted to examine in more detail the evolution of gaze distribution during the seconds after the TOR. They showed that the perseverance effect lasted only about 1 s. This is consistent with Louw, Madigan, et al. (2017) who observed that out-of-the-loop drivers spent less time looking at the road centre during the first second after an unexpected alert TOR. In our case, it appears that participants performing the NDRT kept on looking at the tablet even though only a pictogram requesting to takeover was displayed on it. The warning on the centre console could be a source of distraction, whereas the driver should pay full and immediate attention to the reconstruction of SA to ensure the success of the takeover. This suggests that avoiding the display of information on the centre console during a TOR (e.g., blanking the screen) in order to discourage drivers from looking at information on the device would lead to better performance. This hypothesis could be tested in future studies.

That said, the perseverance at looking at the tablet did not result in increased collisions in the SA_NDT condition, probably because drivers were aware of overtaking vehicles before performing the NDRT. Taken together, the results suggest that continuous – or at least very frequent – monitoring of the road scene is essential for building Level 3 SA, according to Endsley's terminology. Level 3 requires developing a mental model of the situation that allows anticipating its future state. The late disengagement from an NDRT may be sufficient to correctly perceive the immediate environment (SA Level 1); however, it may be insufficient to consider the dynamics of the situation. If the TOR intervenes while the driver is engaged in a task on the central console, this can lead to a form of perseverance in looking at the display. However, this perseverance was less critical to the success of the takeover than having a sufficiently elaborated SA in our use case.

Some studies have explored solution to help the driver to rebuild SA during TOR. For instance, Yousfi (2018) tested blind spot warning devices during takeover. She found that the use of the blind spot detector reduced the collision rate with a vehicle in the left lane. White et al. (2019) proposed a "top-down" guidance check after the emergency TOR; drivers who had this system glanced more at their mirrors than drivers without. In the same vein, Carsten & Martens (2019) advised a set of design principles to improve the human-machine interface for automated cars.

6. Conclusions

In our study, we manipulated the drivers' engagement in a reading NDRT to prevent or allow them to build and maintain SA. We have shown that as long as SA was adequate, drivers were able to successfully manage the TOR even if they have been distracted for a short period just before the TOR. Two seconds of NDRT did not degrade SA enough to impair take-over quality. In future work, it would be interesting to further manipulate the duration of distraction periods prior to the TOR to better understand from what time point distraction impacts Level 3 SA. Conversely, it would be interesting to reproduce the experiment by increasing the time given to drivers to rebuild SA in the NDT_SA condition.

Participants engaged in the NDRT at the time of the TOR exhibited different visual patterns during the first second of the takeover, with more time spent on the display and less time spent at recovering SA (i.e., looking at the road and checking the mirrors). This raises the question of the best design strategy to inform the driver of a need to regain control without delivering a visual alert that may keep the driver's eyes away from the road scene. Driving aids guiding the gaze towards the important elements to be taken into account could even be useful.

The concept of SA appears central in current issues about vehicle automation. The question of how to help drivers to maintain or restore sufficient SA has not yet been definitely answered. Our results confirm the idea that particular effort must be made by designers to restore a good level of SA quickly when the driver is out-of-the-loop. Further work is needed to better identify the behavioural markers of SA, for example in terms of mirror-checking routines or sequences of actions using the vehicle’s controls. To conclude, we refer to Louw et al. (2015), who stated that “until there is an effective strategy to help drivers regain situation awareness during the resumption of control from Highly Automated Driving, they should be encouraged to remain in the driving loop”.

Conflict-of-Interest Statement:

The authors declare that they have no conflict of interest.

Acknowledgements:

This research was supported the French National Research Agency (Agence Nationale de la Recherche, AUTOCONDUCT project, grant n° ANR-16-CE22-0007-05) and by Association Nationale de la Recherche et de la Technologie (ANRT). The authors warmly thank Denis Creusot for his technical support and the participants for their cooperation.

1. References

Bueno, M., Dogan, E., Hadj Selem, F., Monacelli, E., Boverie, S., & Guillaume, A. (2016). How different mental workload levels affect the take-over control after automated driving. *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2040‑ 2045. https://doi.org/10.1109/ITSC.2016.7795886

Carsten, O., Lai, F., Barnard, Y., Jamson, A., & Merat, N. (2012). Control Task Substitution in Semiautomated Driving : Does It Matter What Aspects Are Automated? *Human factors*, *54*, 747‑ 761. https://doi.org/10.1177/0018720812460246

Carsten, O., & Martens, M. (2019). How can humans understand their automated cars? HMI principles, problems and solutions. *Cognition, Technology & Work*, *21*. https://doi.org/10.1007/s10111-018-0484-0

Damböck, D., & Bengler, K. (2012). Übernahmezeiten beim hochautomatisierten Fahren. *5. Tag. Fahr*, 15,16.

Endsley, M. (2018). Situation Awareness in Future Autonomous Vehicles : Beware of the Unexpected. *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*, 303- 309.

Endsley, M. R., & Kiris, E. O. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(2), 381- 394. https://doi.org/10.1518/001872095779064555

Gartenberg, D., Breslow, L., McCurry, J., & Trafton, J. (2014). Situation Awareness Recovery. *Human factors*, *56*, 710- 727. https://doi.org/10.1177/0018720813506223

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). “Take over!” How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1938- 1942. https://doi.org/10.1177/1541931213571433

Gold, C., Körber, M., Lechner, D., & Bengler, K. (2016). Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations The Role of Traffic Density. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*, 642–652. https://doi.org/10.1177/0018720816634226

Green, P. (1999). *Visual and Task Demands of Driver Information Systems* (Technical Report UMTRI-98-16). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Gugerty, L., Rando, C., Rakauskas, M., Brooks, J., & Olson, H. (2003). Differences in Remote versus in-Person Communications While Performing a Driving Task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *47*, 1855- 1859. https://doi.org/10.1177/154193120304701605

Heikoop, D. D., de Winter, J. C. F., van Arem, B., & Stanton, N. A. (2018). Effects of mental demands on situation awareness during platooning : A driving simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, 193- 209. https://doi.org/10.1016/j.trf.2018.04.015

Kunze, A., Summerskill, S. J., Marshall, R., & Filtness, A. J. (2019). Automation transparency : Implications of uncertainty communication for human-automation interaction and interfaces. *Ergonomics*, *62*(3), 345- 360. https://doi.org/10.1080/00140139.2018.1547842

Liang, N., Yang, J., Yu, D., Prakah-Asante, K. O., Curry, R., Blommer, M., Swaminathan, R., & Pitts, B. J. (2021). Using eye-tracking to investigate the effects of pre-takeover visual engagement on situation awareness during automated driving. *Accident Analysis & Prevention*, *157*, 106143. https://doi.org/10.1016/j.aap.2021.106143

Louw, T., Madigan, R., Carsten, O., & Merat, N. (2017). Were they in the loop during automated driving? Links between visual attention and crash potential. *Injury Prevention: Journal of the International Society for Child and Adolescent Injury Prevention*, *23*(4), 281- 286. https://doi.org/10.1136/injuryprev-2016-042155

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 & Prevention*, *108*, 9- 18. https://doi.org/10.1016/j.aap.2017.08.011

Louw, T., Merat, N., & Jamson, A. (2015). Engaging With Highly Automated Driving : To Be Or Not To Be In The Loop? *Driving Assessment Conference*, 190- 196. https://doi.org/10.13140/RG.2.1.2788.9760

Ma, R., & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, *35*(10), 939- 953. https://doi.org/10.1016/j.ergon.2005.04.002

Marti, P., Jallais, C., Koustanaï, A., Guillaume, A., & Mars, F. (2021). Effect of driver’s distraction on gaze behaviour, mental workload and takeover quality. *Driving Simulation Conference Europe 2021 VR*, 99-106. https://hal.archives-ouvertes.fr/hal-03343285

Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J. D., Johansson, E., Green, C. A., Katazaki, S., Monk, C., Itoh, M., McGehee, D., Sunda, T., Unoura, K., Victor, T., Schieben, A., & Keinath, A. (2019). The "Out-of-the-Loop" concept in automated driving : Proposed definition, measures and implications. *Cognition, Technology & Work*, *21*(1), 87-98. https://doi.org/10.1007/s10111-018-0525-8

Mok, B., Johns, M., Miller, D., & Ju, W. (2017). Tunneled In : Drivers with Active Secondary Tasks Need More Time to Transition from Automation. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2840–2844. https://doi.org/10.1145/3025453.3025713

Moore, K., & Gugerty, L. (2010). Development of a Novel Measure of Situation Awareness : The Case for Eye Movement Analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*(19), 1650-1654. https://doi.org/10.1177/154193121005401961

Naujoks, F., Purucker, C., & Neukum, A. (2016). Secondary task engagement and vehicle automation – Comparing the effects of different automation levels in an on-road experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *38*, 67-82. https://doi.org/10.1016/j.trf.2016.01.011

Pfleging, B., Rang, M., & Broy, N. (2016). Investigating user needs for non-driving-related activities during automated driving. *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, 91–99. https://doi.org/10.1145/3012709.3012735

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How Traffic Situations and Non-Driving Related Tasks Affect the Take-Over Quality in Highly Automated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 2063-2067. https://doi.org/10.1177/1541931214581434

Rockwell, T. H. (1988). Spare Visual Capacity in Driving-Revisited : New Empirical Results for an Old Idea. *Vision in Vehicles II*, 317-324.

SAE International. (2016). *Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems* (Technical Report J3016_201609). SAE International.

Scharfe, M. S. L., Zeeb, K., & Russwinkel, N. (2020). The Impact of Situational Complexity and Familiarity on Takeover Quality in Uncritical Highly Automated Driving Scenarios. *Information*, *11*(2), 115. https://doi.org/10.3390/info11020115

Shi, E., & Frey, A. T. (2021). Non-Driving-Related Tasks During Level 3 Automated Driving Phases—Measuring What Users Will Be Likely to Do. *Technology, Mind, and Behavior*, *2*(2). https://doi.org/10.1037/tmb0000006

van de Merwe, K., Dijk, H., & Zon, R. (2012). Eye Movements as an Indicator of Situation Awareness in a Flight Simulator Experiment. *The International Journal of Aviation Psychology*, *22*, 78-95. https://doi.org/10.1080/10508414.2012.635129

White, H., Large, D., Salanitri, D., Burnett, G., Lawson, A., & Box, E. (2019). Rebuilding Drivers' Situation Awareness During Take-Over Requests in Level 3 Automated Cars. *In Proceedings of the Contemporary Ergonomics & Human Factors, Stratford-uponAvon, UK.*

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

https://doi.org/10.1016/j.trf.2014.06.016

Yousfi, E. (2018). Conscience de la situation et performance de conduite lors de la reprise en main en conduite automobile déléguée. *PhD thesis, Bibliothèque numérique paris 8*. https://octaviana.fr/document/2018PA080098.

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accident Analysis & Prevention*, *78*, 212‐ 221. https://doi.org/10.1016/j.aap.2015.02.023

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