Predicting self-assessment of the out-of-the-loop phenomenon from visual strategies during highly automated driving

Damien Schnebelen[1], Camilo Charron[1,2], & Franck Mars[1]
[1]Centrale Nantes, CNRS, LS2N, Nantes, [2]Université Rennes 2
France

Abstract

During highly automated driving, drivers do not physically control the vehicle anymore, but they still have to monitor the driving scene. This is particularly true for SAE level 3 (SAE International, 2016), as they need to be able to react quickly and safely to a take-over request. Without such an (even partial) monitoring, drivers are considered out-of-the-loop (OOTL) and safety may be compromised. This OOTL phenomenon may be particularly important for long automated driving periods. The current study aimed at scrutinizing driver's visual behaviour for a long period of highly automated driving (18 minutes). Intersections between gaze and 13 areas of interest (AOI) were analysed, considering both static (percentage of time gaze spent in one single AOI) and dynamic (transitions from one AOI to another) patterns. Then, a prediction of the self-reported OOTL level (subjective assessment) from gaze behaviour was performed using Partial Least Squares (PLS) regression models. The outputs of the PLS regressions allowed defining visual strategies associated with good monitoring of the driving scene and paved the way for an online estimation of the OOTL phenomenon based on driver's spontaneous visual behaviour.

Introduction

In manual driving, drivers must gather information about the driving scene and the vehicle (perceptual process), interpret this information (cognitive process) and act appropriately (motor process), which in turn generate information. However, with the imminent deployment of highly automated vehicles on the roads (between 2020 and 2030 depending on the organization (Chan, 2017)), where the operational driving task is performed by automation, drivers are likely to become supervisors of the driving scene. In this case, the perceptual-motor loop is neutralized, which has consequences on perception and cognition (Mole et al., 2019). This is referred to the out-of-the-loop (OOTL) phenomenon.

In automated driving, the OOTL phenomenon was investigated by comparing the driver's behaviour during automated and manual driving. In terms of gaze behaviour, automated driving leads to greater horizontal dispersion (Louw & Merat, 2017; Mackenzie & Harris, 2015), and a decrease of the percentage of glances to the road centre (Louw et al., 2015; Mackenzie & Harris, 2015). Similarly, in curve driving,

In D. de Waard, A. Toffetti, L. Pietrantoni, T. Franke, J-F. Petiot, C. Dumas, A. Botzer, L. Onnasch, I. Milleville, and F. Mars (2020). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2019 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

automated driving has been shown to enhance long-term anticipation (through look-ahead fixations) to the detriment of the short-term anticipation used to guide the vehicle (Mars & Navarro, 2012; Schnebelen et al., 2019).

The consequences of the OOTL phenomenon were also observed during level 3 automated driving , where drivers had to take control of the vehicle when automation required it. Indeed, in response to a critical case, drivers had longer reaction times in automated driving than in manual driving (Feldhütter et al., 2017; Neubauer et al., 2012; Saxby et al., 2013; Zeeb et al., 2015; Zeeb et al., 2017). Such changes in driver behaviour during takeover have been attributed to drivers being more OOTL during automated driving.

Drivers' performance during takeover is also affected by the duration of automation, with higher reactions times after a prolonged period of automation than after a short drive (Bourrelly et al., 2019; Feldhütter et al., 2017). Feldhütter et al. (2017) have shown, for instance, that a 20-minutes' drive in automated mode is sufficient to increase the reaction time to a takeover request. Drivers experienced mind wandering, distracting themselves from the supervision task, which impaired the perceptual and cognitive processing of information.

Recently, Merat et al. (2019) proposed an operational definition of the OOTL concept. It relies on two aspects: To be out-of-the-loop, drivers must not have physical control of the vehicle (no motor process), and must not monitor the driving scene (perception/cognition process). When the driver is in manual control, he is considered to be in-the-loop. An intermediate state, the on-the-loop (OTL) level, has been introduced to designate situations in which the driver correctly monitors the driving situation during autonomous driving. Thus, estimating the driver's ability to manage imminent takeover situations is a matter of determining whether the driver is OOTL or OTL based on the observation of his/her monitoring of the situation. However, the question of how to model and quantify what constitutes proper monitoring of the driving scene remains open.

Two principal issues were addressed in the present study:

- What is a good monitoring of the driving situation? In other words, can we identify the gaze behaviour characteristic of OOTL drivers?
- Is it possible to predict the driver's OOTL state from the observation of spontaneous gaze strategies?

In the current study, participants experienced an 18-min drive of automated driving (similar to Fleurette et al., 2017) without any non-driving activities to perform. The assessment of the OOTL state was based on the self-reported time of mind wandering during the drive. The driver's gaze behaviour was analysed considering 13 areas of interest, using static (percent of time on each AOI) and dynamic (transitions matrix from and to each AOI) patterns.

Material and method

Participants

This study involved 12 participants (N = 12; 3 females; 9 males), with a mean age of 21.4 years (SD = 5.34). Most of them were students from Centrale Nantes. They held a valid driver's licence (average driving experience: 9950 km/year, SD = 5500) and signed written informed consent to participate in this study.

Experimental device

The experiment took place on a driving simulator (Figure 1), consisting in 3 screens (120° Field of View), with one additional screen for the HMI. The eye tracker (SmartEye Pro v5.9) computed gaze intersections with the screens at 20 Hz.

Most of the road was a 40 km two-lane dual carriageway, with a speed limit of 130 km/h in accordance with French regulations. Occasional changes in road geometry (temporary 3-lane traffic flow; highway exits; slope variation) and speed limits (130 km/h to 110 km/h) have been included to make driving less monotonous. In both directions on the highway, traffic was fluid, with 8 overtaking situations.

Figure 1. Driving Simulator Setup.

Procedure

After a presentation of the driving simulator and a short drive in manual driving mode, participants were trained to activate (pressing a button) and disactivate (pressing the button, pedals or steering wheel) the automated mode. Instructions corresponding to a level 3 (SAE) automated driving were given: Automated driving was available only for a portion of road, and drivers had to take over the system when required (auditory + visual signals). Then, they experienced 4 takeover situations, with relatively long

(45 s; 2 situations) or short (8 s, 2 situations) time-to-collision. No collision occurred during the training session.

Then, the experiment proper started. Participants activated the automated driving mode just before entering the highway. Gaze data were recorded as soon as the vehicle was correctly inserted in the lane and reached 130 km/h. No major driving events appeared for the first 15 min on the highway to let the driver enough time to become out-of-the-loop. The driver did not perform any secondary task during that time. A critical case occurred at the 18th minute, and the scenario ended thirty seconds after. Participants were then asked to report on a continuous Likert scale the proportion of time spent thinking at something else than the driving task throughout the trial. Since this paper focuses on the link between gaze behaviour and the OOTL scores, the results on the critical case will not be presented here.

Data structure and annotations

Definition of the OOTL score Y
The evaluation of the percentage of time spent thinking about something else than the driving task may be considered as a self-assessment of the OOTL phenomenon. In that sense, the higher the percentage was, the more drivers estimated they were out-of-the-loop. Percentages for all participants were stored in a vector with 12 elements, named OOTL score and denoted Y.

Definition of the matrix of gaze behaviour X
The driving scene was divided into 13 areas of interest (AOI) (see figure 2):

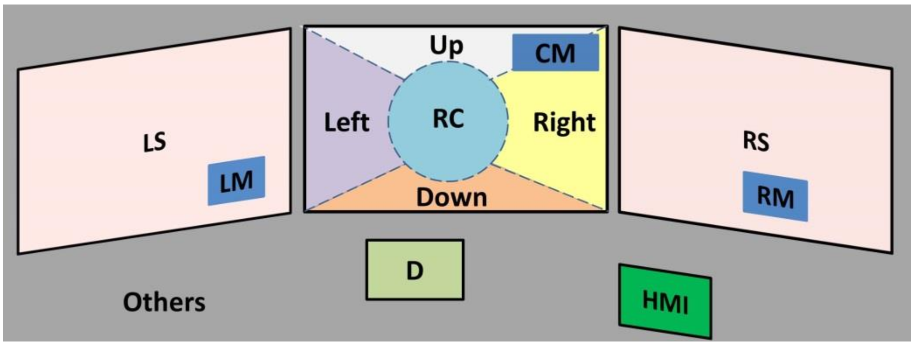

Figure 2. Division of the driving environment into 13 areas of interest.

- The central screen contained six areas: The central mirror (area CM), the road centre (RC), defined as a circular area of 8° radius in front of the driver, and 4 additional areas defined relatively to the road centre (Up, Left, Down, Right). The Percentage Road Centre (PRC) defined as the proportion of time spent in RC has been introduced by Victor (2005). A decrease of the PRC was found to be a good indicator of distraction during driving, as drivers reduced this time when visually or auditory distracted (Victor et al., 2005)
- Each peripheral screen contained two areas: The lateral mirror (LM, RM) and the remaining peripheral scene (LS, RS)

predicting OOTL phenomenon during automated driving

- The dashboard (D) and the HMI (HMI) All gaze data directed outside of all the previous areas were regrouped in area Others.

Drivers gaze behaviours for each participant were considered in this study as the combination of static (percentage of time in one AOI) and dynamic (transitions matrix between AOIs) patterns. Thus, a vector of 182 numerical indicators (= 13x13 transitions + 13 percentage of time on each AOI) summarizes gaze behaviour for one participant. When considering all participants, the matrix of gaze behaviour was named X and its size was 12 (participants) x 182 (visual indicators).

Due to the small number of observations (12) compared to the number of visual indicators (182), we used the PLS regression to predict the OOTL score from gaze behaviour. This method performed a decomposition of X and Y in orthogonal components in order to explain the maximum of the variance of Y. The components actually reflect the underlying structure of the prediction model.

Data analysis

Two sequential stages composed the analysis (Figure 3):

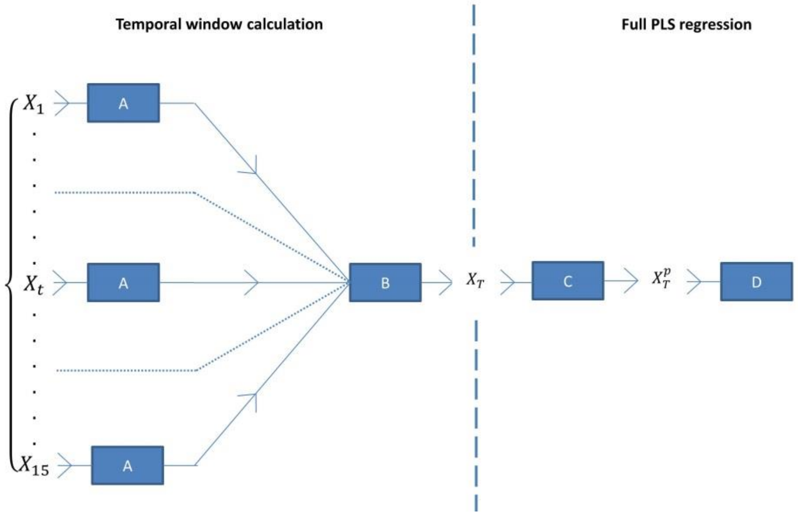

Figure 3. Multi-step approach for data analysis.

- The first one (steps A and B) focused on selecting the best time window (T) to predict the OOTL score. To do so, 15 matrixes of gaze behaviour were computed and labelled X_t. It differed by the time on which visual indicators were computed, that varied from 1 to 15 minutes.
- The selection was then based on the most stable (over time) and accurate (in terms of percent of variance explained) model of prediction. The second one (steps C & D) consisted of predicting the OOTL score using X_T as predictors and the PLS regression model. After reducing the dimension of X to increase

prediction power (step C), the model was tested using the training and the validation data set (step D).

The details of data analysis are presented in the results section.

Results

OOTL Score

The OOTL scores (Figure 4) showed large variations between participants (range ~= 75%). The median score was 43%. Even in the absence of a secondary task, some participants (9 to 12) declared that they spent 80% of the time thinking at something else than the driving task.

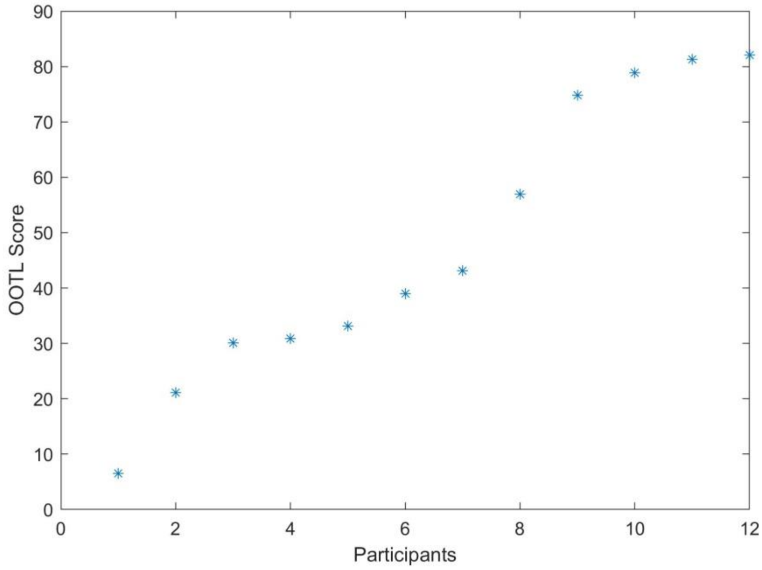

Figure 4. OOTL scores reported by the participants.

Time window selection

On the first step (A), the optimal number of components for each matrix X_t was obtained by minimizing the mean square error of prediction. This number of components, reflecting the structure of the prediction model, actually changed depending on the integration window (figure 5), but reached a stability level for time windows higher than 9 minutes. The most appropriate temporal window, labelled T, was selected (step B) as the one maximizing the variance of the OOTL score explained, among the stable models. All subsequent analysis referred to the matrix of gaze behaviour computed over T = 11 minutes of automated driving.

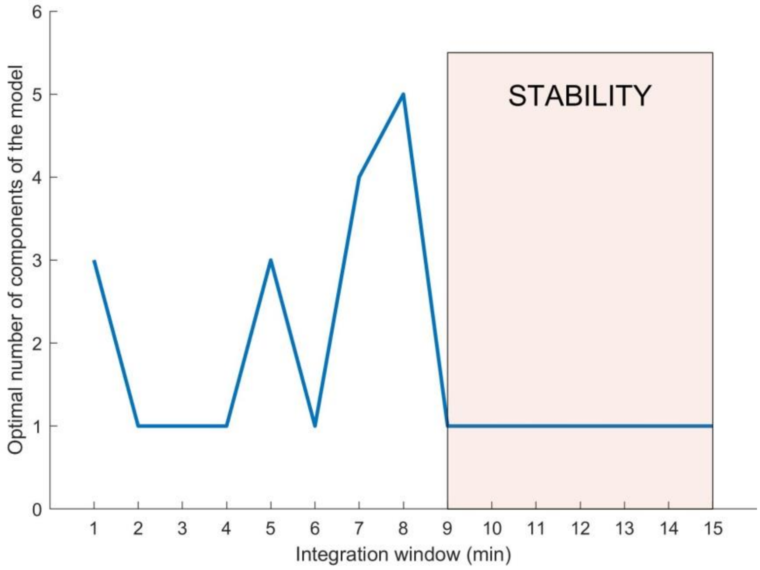

Figure 5. Optimal number of components of the PLS regression as a function of the integration window. The figure shows that model stability was achieved from the 9th minute.

Reduction of the number of visual indicators

After selecting the most appropriate time window, the prediction model explained 62.53% of the variance of Y, using the 182 visual indicators. Then, the aim was to reduce the number of visual indicators by selecting only the most relevant visual indicators.

The PLS regression is a linear model: the variable to be estimated ($\hat{Y}$) and the predictor (X_T) are linked by a matrix of coefficients C: $\hat{Y} = C * X_T$. The relevant indicators were determined by the absolute magnitude of their coefficient: If the magnitude was close to zero, the contribution to the prediction was negligible. On the contrary, a high magnitude indicated a very important indicator for the prediction.

In practice, the coefficients magnitudes were compared with an increasing threshold value. A new regression model was computed for each partial matrix (i.e. a matrix comprising only those indicators whose coefficient amplitude exceeded the threshold value). The threshold was increased by step of 0.005 until the percentage of variance of Y explained by the partial model stopped increasing. With our data, the maximum of explained variance was 85.64%, with only 8 visual indicators (Figure 6).

On these 8 indicators, 5 contributed to an increase of the OOTL score (in red on Figure 6): Taking the eyes off the central mirror to look away from the driving scene, taking the eyes off the road centre area to look down or away from the driving scene, spending too much time in the down area. By contrast, 3 indicators contributed to a reduction of the OOTL score (green arrows on Figure 6): Redirecting the gaze to the road centre or to the left side of the driving scene from any area outside the driving scene, take your eyes off the road centre to check the left rear-view mirror.

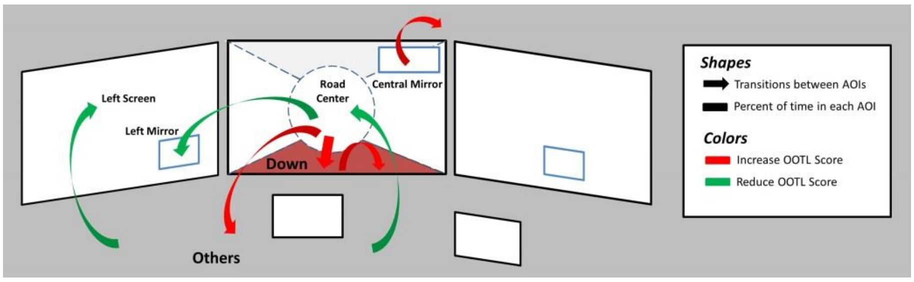

Figure 6. Visual indicators relevant for OOTL score prediction.

Final prediction of the OOTL score

A final PLS model (step D) was computed to predict the OOTL score from the best partial matrix (containing the 8 visual indicators relevant for the prediction). The prediction of the model compared to real values of the OOTL score is presented on Figure 7.

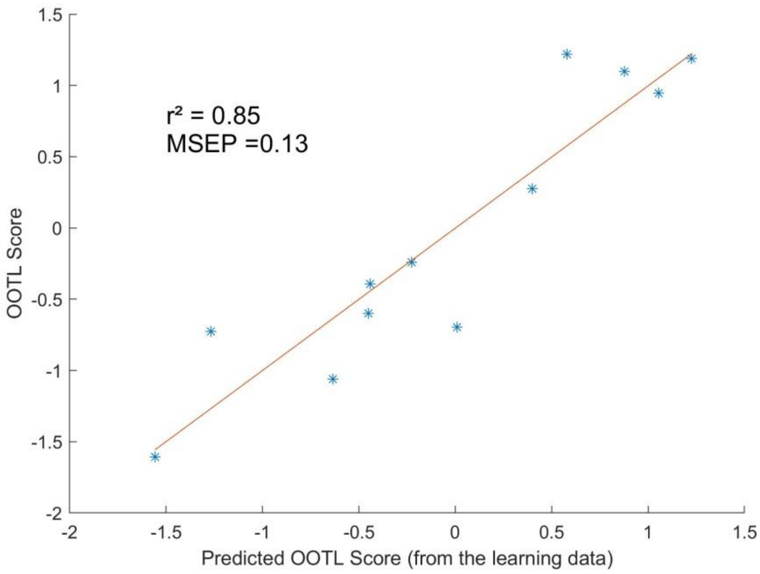

Figure 7. Correlation plot between the OOTL score and the prediction of the OOTL score by PLS regression.

The PLS regression performed a good estimation of the OOTL score, with a low mean square error of prediction (0.13) and a significant positive correlation between the estimated and real values (r = 0.92, p<0.01).

Discussion

During automated driving, the OOTL phenomenon results from an incorrect monitoring of the driving situation (Merat et al., 2019). The alternative state, namely being OTL (on-the-loop), corresponds to passive drivers who satisfactorily monitor their driving environment. However, a more precise definition of what constitutes a great monitoring of the environment is still needed to distinguish OTL from OOTL drivers. This study investigated this issue in a highway driving context with the analysis of drivers gaze behaviour, with both static (percent of time in AOI) and dynamic (transitions between AOI) patterns. The methods consisted in using PLS regressions to identify the most characteristic elements of the gaze behaviour of OTL and OOTL drivers. The multi-step approach began with 182 visual indicators as an input matrix, and retained in the end only 8 relevant elements to predict an accurate OOTL score.

The results revealed that drivers with a lower OOTL score made more transitions from the road centre to the left mirror. After spending time looking at area unrelated to driving (“others” area), they returned more frequently to the road (road centre area) or to the left screen where they could monitor traffic. Conversely, drivers with higher OOTL scores made more transitions from the road centre to areas irrelevant to driving. They spent more time and made multiple fixations in the lower part of the front screen.

These findings may be interpreted in terms of the adequacy of the driver’s gaze strategy to maintain good situation awareness (Endsleigh, 1995) in autonomous mode. Situation Awareness (SA) during automated driving actually involved three levels: Perception, Comprehension and Projection (Merat et al, 2019). In the current study, OTL drivers remained dynamically aware of their surrounding by regularly checking the left lane and mirror. This certainly have helped to anticipate future hazards. They also remained attentive to the road well-ahead in time. In other words, these gaze strategies allowed to perceive, comprehend and project on the future state of the driving situation in an appropriate way, i.e. to have a good enough SA. On the other hand, the OOTL drivers’ gaze was more strongly attracted by irrelevant information inside or outside the simulator. Even when looking at the driving scene, the driver favoured the road immediately in front of them (down area), suggesting a lack of visual anticipation.

In the current study, PLS regressions appear to be a relevant approach to predict the driver’s state from spontaneous gaze behaviour. Indeed, PLS regressions allowed finding one optimal temporal window, reducing the dimensions of the matrix of gaze behaviour from 182 to 8 relevant elements, but also indicated whether they contributed to increase or decrease the OOTL score. Then, the prediction of the OOTL score given by the model was accurate with a strong correlation between the predicted and the real values. However, a validation step (i.e. testing the model with another set of gaze behaviour data) is required to confirm the results presented here.

In the current study, the OOTL score could be predicted from the driver’s spontaneous strategies over 11 minutes of automated driving. For further research, it may be interesting to apply this model on shorter durations of automated driving, and to apply similar methods to other driving contexts.

Conclusion

The current study used PLS regression to satisfactory predict driver's state from their visual monitoring of the driving situation. The analysis of gaze behaviour proved that an appropriate gaze strategy for being on the loop requires to get information on the oncoming traffic as well as interleaving glances on the road centre. To provide a more accurate detection of the OOTL phenomenon during automated driving, the analysis of gaze behaviour might be coupled with other approaches, for example by incorporating physiological measurements or the analysis of the driver's posture in the diagnosis.

References

Bourrelly, A., de Naurois, C.J., Zran, A., Rampillon, F., Vercher, J.-L., & Bourdin, C. (2019). Long automated driving phase affects take-over performance. *IET Intelligent Transport Systems, 13,* 1249 - 1255.

Chan, C.-Y. (2017). Advancements, prospects, and impacts of automated driving systems. *International Journal of Transportation Science and Technology, 6*, 208-216. https://doi.org/10.1016/j.ijtst.2017.07.008.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64.

Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017). How the duration of automated driving influences take-over performance and gaze behaviour. In *Advances in Ergonomic Design of Systems, Products and Processes* (p. 309–318). Springer.

Louw, T, Kountouriotis, G., Carsten, O., & Merat, N. (2015). Driver Inattention During Vehicle Automation : How Does Driver Engagement Affect Resumption Of Control? *New South Wales*, 16.

Louw, T. & Merat, N. (2017). Are you in the loop ? Using gaze dispersion to understand driver visual attention during vehicle automation. *Transportation Research Part C: Emerging Technologies, 76*, 35–50.

Mackenzie, A.K., & Harris, J.M. (2015). Eye movements and hazard perception in active and passive driving. *Visual Cognition, 23*, 736–757.

Mars, F., & Navarro, J. (2012). Where we look when we drive with or without active steering wheel control. *PLoS One, 7*(8), e43858.

Merat, N., Seppelt, B., Louw, T. et al. (2019). The "out-of-the-loop" concept in automated driving : Proposed definition, measures and implications. *Cognition, Technology & Work 21*, 87–98. https://doi.org/10.1007/s10111-018-0525-8.

Mole, C.D., Lappi, O., Giles, O., Markkula, G., Mars, F., & Wilkie, R.M. (2019). Getting Back Into the Loop : The Perceptual-Motor Determinants of Successful Transitions out of Automated Driving. *Human Factors, 61*, 1037-1065. https://doi.org/10.1177/0018720819829594.

Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors, 54*, 734–746.

SAE International. (2016). *Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems (J3016)*.

Saxby, D.J., Matthews, G., Warm, J.S., Hitchcock, E.M., & Neubauer, C. (2013). Active and Passive Fatigue in Simulated Driving : Discriminating Styles of Workload Regulation and Their Safety Impacts. *Journal of Experimental Psychology. Applied*, *19*, 287-300. https://doi.org/10.1037/a0034386.

Schnebelen, D., Lappi, O., Mole, C., Pekkanen, J., & Mars, F. (2019). Looking at the Road When Driving Around Bends : Influence of Vehicle Automation and Speed. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01699.

Victor, T. (2005). *Keeping eye and mind on the road* (PhD Thesis). Acta Universitatis Upsaliensis.

Victor, T.W., Harbluk, J.L., & Engström, J.A. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F*, *8*, 167-190. https://doi.org/10.1016/j.trf.2005.04.014.

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.

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