IT'S OUT OF OUR HANDS NOW! EFFECTS OF NON-DRIVING RELATED TASKS DURING HIGHLY AUTOMATED DRIVING ON DRIVERS' FATIGUE

Oliver Jarosch[1,3], Matthias Kuhnt[2], Svenja Paradies[1], Klaus Bengler[3]
[1] BMW Group Research and Technology, Munich, Germany
[2] Chemnitz University of Technology, Chemnitz, Germany
[3] Technical University of Munich, Munich, Germany
E-Mail: oliver.oj.jarosch@bmw.de

Summary: With introduction of conditional automation in vehicles the driver can engage in non-driving related tasks (NDRTs) and only has to intervene in case of take-over requests (TOR). Therefore, active fatigue, which is the most frequent form of fatigue in manual driving, is assumed to be replaced by passive fatigue, intensified through monotony and monitoring tasks in conditional automated driving (CAD), SAE Level 3. To investigate effects of NDRTs on drivers' fatigue and take over capability a driving simulator study was conducted. In total, 56 participants experienced two rides on a highway with CAD. During the two rides, participants had to fulfill both a monotonous monitoring task and an activating task. As in CAD the system is executing longitudinal and lateral control, drowsiness detection referring to driving performance becomes inoperative. Non-invasive methods for drowsiness detection that are not related to driving performance have to be investigated. Therefore, fatigue was measured with percentage of eye-lid closure (PERCLOS), blink related eye-tracking parameters, and the self-report Karolinska Sleepiness Scale (KSS). Results suggest that fatigue can be caused through a monitoring task in highly automated driving. PERCLOS could be confirmed as a valid parameter for detecting fatigue in CAD. Further, passive task related fatigue caused by a 25 min monotonous monitoring task does not affect the drivers' take over capability negatively.

INTRODUCTION

Due to continuous progress in automation and automation technology, the human in his classic role as operator, is replaced more and more by automated systems. With introduction of conditional automated driving (CAD; SAE, 2014), the driver can engage in non-driving related tasks (NDRTs) and only hast to take over in case of a take-over request (TOR). Next to its benefits, automation also brings along problems that are already extensively discussed in the human factors literature. Bainbridge (1983) pointed out that automation may cause an expansion of problems rather than to eliminate them. The switched role from the active operating to system supervising, increases monotony due to monitoring tasks. With introduction of CAD, the role of the human as driver will change in a few years, too.

Fatigue and drowsiness are known to be relevant causations for road accidents and have been investigated in the past (i.e. Treat et al., 1979). In the fatigue model of May and Baldwin (2009), fatigue is related to an increased crash risk as fatigue causes poor reaction times and a decreased driving performance. In this model, it is distinguished between three different types of fatigue: *sleep related fatigue, active task related fatigue*, and *passive task related fatigue*.

Sleep related fatigue is linked to the circadian rhythm of a human being and intensifies with sleep deprivation and sleep restriction. Pack et al. (1995) suggested, that sleep related accidents occur more frequently when people are sleepier because of the circadian rhythm. Deteriorated reaction times (Jewett et al., 1999) and driving performance seemed to be negatively affected by sleep related fatigue (Lenne, Triggs, & Redman, 1997; Philip et al., 2005). *Active task related fatigue* in driving is caused by mental overload conditions like high traffic density, poor visibly conditions or the need to complete a secondary task next to the driving task (Gimeno et al., 2006). *Passive task related fatigue* instead is connected to mental underload conditions, monotony and monitoring tasks. It occured when there was little traffic and monotonous situations (Gimeno et al., 2006). Driving performance (Matthews et al., 2002) and reaction times (Saxby et al., 2013) were negatively affected through *active* and *passive task related fatigue*.

With introduction of CAD *active task related fatigue* was potentially reduced, due to lower mental workload when the automation was executing the driving task (Neubauer et al., 2012). However, researchers assumed that aspects of *passive task related fatigue* might be promoted in CADS (May & Baldwin, 2009; Neubauer et al., 2012), as automation might lead to a more monotonous situation for the driver as he only has to intervene in case of TORs.

In manual driving, driver drowsiness detection mostly referred to driving-performance parameters (Forsman et al., 2013). As in CAD guidance is executed by the system, non-invasive drowsiness detection methods have to be investigated. An appropriate drivers' state is vital in case of a TOR.

Summarized, consequences of fatigue in manual driving are increased crash risk due to deteriorated reaction times and reduced driving performance. Transferred to CAD, fatigue affects reaction times and quality upon TORs. It is also assumed, that CAD reinforces monotony whilst driving and that NDRTs can potentially influence drivers' fatigue.
Objective of this study was to find out (i) how passive task related fatigue can be measured in CAD, (ii) whether NDRTs are viable methods of manipulating passive task related fatigue in CAD, and (iii) whether passive task related fatigue affects reaction time and quality upon TOR.

METHOD

Participants

Fifty-six employees of the BMW Group voluntarily participated in the study. The sample consisted of 9 female and 47 male participants. Mean age was 30.10 years (*SD* = 9.00, min = 20, max = 56). The subjects were experienced drivers with mean driving experience of 12.29 years (*SD* = 9.36). The majority of the sample had experienced at least one driving assistance system (79.3 %). Here adaptive cruise control was the most experienced one (75%).

Experimental Design

All participants experienced two different NDRTs and two different take-over scenarios in a counterbalanced order during two experimental rides, resulting in a two-factor within design. Two types of NDRTs, presented on a tablet installed in the central console of the car, were used to affect drivers' fatigue. A monotonous monitoring task was used to induce *task related passive fatigue*. In reference to Warm et al. (2008), a task with a low event rate and a fixed event location was selected. Different letters ("P", "q", "p" and "d") were presented on the screen for a variable time between 10s – 15s in a mixed order. Subjects had to touch the screen every time the "p" was displayed. The other task instead had the purpose to keep drivers on an adequate level and prevent them from fatigue. Therefore, in reference to Schömig et al. (2015) a quiz task was used. Participants had to choose the right answer to a question out of four possibilities. To test effects of resulting fatigue on take-over performance two different take-over scenarios (accident on the ego-lane / a sensor failure in a bend) occurred in the end of the two rides. Both scenarios are not predictable and highly critical but differ in the complexity of the driver intervention.

Driver fatigue, affected through the two NDRTs, was assessed using the self-report Karolinska Sleepiness Scale (KSS), developed by Åkerstedt & Gillberg (1990). As objective method for measuring fatigue, percentage of eyelid closure over time (PERCLOS; Wierwille et al., 1994) was used. Körber et al. (2015), who investigated fatigue caused by monotony in automated driving, showed a significant increase on blink rate (blinks/min) and blink duration with increasing time on task. Therefore, these parameters were observed as well. Self-reported and objective sleepiness was measured repeatedly over the two experimental rides for each seven defined times (t1 – t7; see figure 1). KSS was assessed after eye-tracking to not affect data. Take-over performance was assessed using take-over time (braking > 10% of pedal position / steering input > 2^0) and driving related parameters (acceleration [longitudinal, lateral], steering [steering-angle, steering-angle velocity] and tracking [standard deviation of lateral position; SDLP]).

Data were collected in single 105-minute experiments. The virtual driving scenario for all sessions was a three-lane freeway with a hard shoulder. At the beginning of each experiment, participants were briefed on the driving simulator and the CAD which provided lateral and longitudinal control, including lane changes and overtaking. In a first 10 min training session, participants were familiarized with the simulator, the CAD and the TOR. A one minute baseline for eye-tracking data was recorded when the car drove with CAD. The following two experimental rides were identically concerning route (~ 30 min, 59 km), traffic (low to middle traffic) and weather conditions (cloudy, no rain). The TOR situation happened in min 25 (after 50 km). After the first experimental ride, participants had to leave the simulator for a break to regularize fatigue affected through the first NDRT. In the second ride, each the other NDRT was presented and the other take-over scenario occurred.

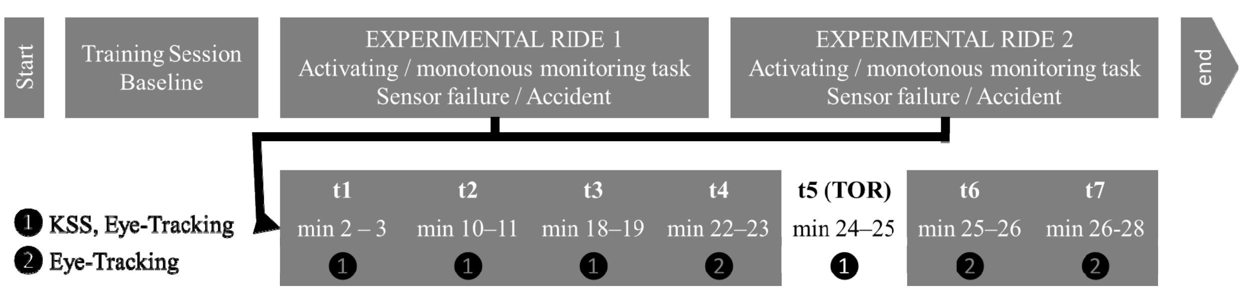

Figure 1: Procedure

Apparatus

The study was conducted in a driving simulator with a motion system. All visual channels were rendered at 60 frames/s, predominantly at a resolution of 1920 x 1200. Seven forward channels were front-projected providing a horizontal field of view of 240° x 45°. The two rear channels could be seen through the vehicle's side mirrors. One LCD panel with a resolution of 960 x 480 inside the mockup displays the rear-view. The simulator also incorporated a six degree-of-freedom hexapod motion system. As mockup, a BMW 5 Series Touring was used. A Dikablis 3.0 head-mounted eye-tracker was used for measuring PERCLOS and blink related parameters.

RESULTS

A significance level of $\alpha = .05$ was set for all hypothesis tests. All reported results were conducted using multivariate general mixed model (GLM). If Mauchly test for sphericity returned significant results, Greenhouse-Geisser correction was used.

Self-Report Measures

Karolinska Sleepiness Scale. Results showed a significant main effect of the NDRT on sleepiness over time, $F(1, 52) = 53.52$, $p < .001$, $r = .71$. The monotonous monitoring task induced a higher level of sleepiness compared to the activating task (Figure 2). There was also a significant main effect of time on KSS, $F(2.42, 125.91) = 37.24$, $p < .001$, $r = .65$, and a significant interaction effect between the NDRT and time, $F(2.46, 127.75) = 17.386$, $p < .001$, $r = .50$. Self-reported sleepiness increased significantly ($p < .001$) during the monotonous monitoring task. During the activating task sleepiness did not change significantly.

Eye Tracking Data

PERCLOS. Results showed a significant main effect of the NDRT on PERCLOS over time, $F(1, 53) = 43.92$, $p < .001$, $r = .67$. The monotonous monitoring task induced a higher level of PERCLOS compared to the activating task. There was also a significant main effect of time on PERCLOS, $F(1.70, 89.93) = 15.62$, $p < .001$, $r = .48$, and a significant interaction effect between the tasks and time, $F(1.73, 91.50) = 17.85$, $p < .001$, $r = .50$. As Figure 3 illustrates, PERCLOS increased with time spending on the NDRT only in the monotonous monitoring task condition.

Blink rate. Results showed a significant main effect of the NDRT on blink rate over time, $F(1, 53) = 95.12$, $p < .001$, $r = .80$. The monotonous monitoring task induced a higher level of blink rate compared to the activating task (see Figure 4). There was also a significant main effect of time on blink rate, $F(4.12, 218.34) = 8.91$, $p < .001$, $r = .38$, and a significant interaction effect between the NDRT and time, $F(4.44, 235.52) = 20.23$, $p < .001$, $r = .53$.

Blink duration. Results showed a significant main effect of the NDRT on blink duration over time, $F(1, 53) = 11.72$, $p = .001$, $r = .43$. The monotonous monitoring task induced a higher level of blink duration compared to the activating task. There was also a significant main effect of time, $F(1.61, 85.17) = 15.93$, $p < .001$, $r = .48$, and a significant interaction effect between the NDRT and time, $F(1.78, 94.37) = 9.69$, $p < .001$, $r = .39$. Blink duration increased during the monotonous monitoring task with increasing time and decreased after the TOR (see figure 5).

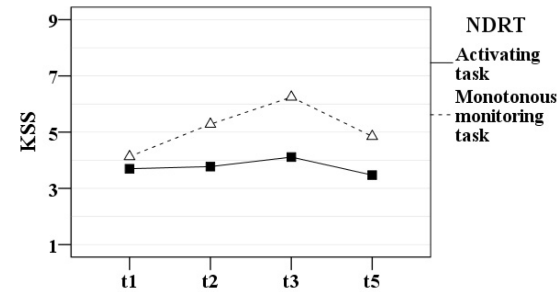

Figure 2. KSS Rating over time

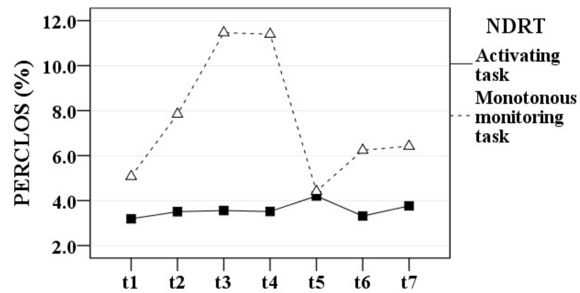

Figure 3. PERCLOS over time

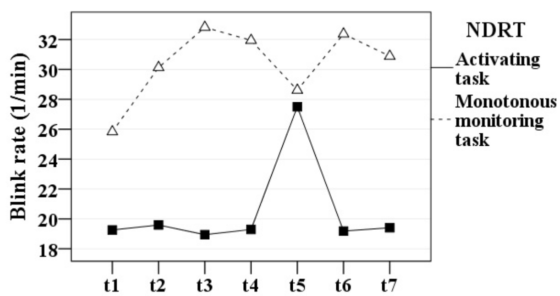

Figure 4. Blink Rate over time

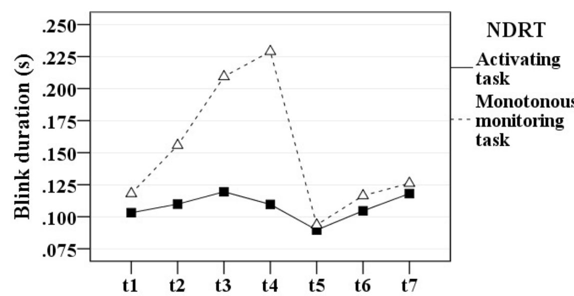

Figure 5. Blink duration over time

Take-over performance

Take-over time. No differences were found for take-over times between the different NDRTs. *Driving-related parameters*. No differences were found for the different NDRTs. Two accidents occurred after the TOR. One after the activating and one after the monotonous monitoring task.

Table 1. Mean values for Take-over Performance + (SD)

situation		accident		sensor failure	
task		Monitoring task	Quiz	Monitoring task	Quiz
Take-over-time (in s)		2.65 (1.68)	3.14 (1.6)	2.47 (0.8)	2.31 (0.79)
acceleration	Longitudinal (in m/s²)	7.31 (2.21)	7.99 (2.68)	1.8 (2.6)	0.8 (0.8)
	Lateral (m/s²)	1.8 (1.0)	2.27 (1.16)	2.47 (0.56)	2.52 (0.82)
steering	steering angle (in °)	25.52 (42.2)	29.4 (30.45)	11.3 (2.63)	11.14 (3.89)
	steering-angle-speed (°/s)	72.11 (77.8)	95.9 (81.02)	33.11 (17.1)	29.96 (18.4)
tracking	SDLP (in m)	0.35 (0.1)	0.35 (0.07)	0.24 (0.13)	0.24 (0.13)

DISCUSSION

The objectives of this study were to investigate effects of different NDRTs in CAD on task related fatigue and to investigate if resulting drowsiness affects take over performance.

As hypothesized NDRTs affected drivers' fatigue during CAD. During the ride with the monotonous monitoring task, an increase in the self-reported KSS could be found. Also eye-tracking data support this finding. PERCLOS increased with time-on-task whilst participants dealt with the monotonous monitoring task. When participants dealt with the quiz task, it stayed on a significant lower level. A more precise look at eye-tracking data provides an explanation, how the increase of PERCLOS comes to be. When participants had to deal with the monotonous monitoring task, at first blink frequency accumulates. With increasing time on task, next blink duration ascends. This is in line with findings of Hargutt (2003), who reports, that early stages of fatigue are connected to a rise in blink frequency. Subsequent stages of fatigue are rather connected to an increase in blink duration. A slight increase of PERCLOS directly after the TOR (see Figure 3) during the quiz task ride can be explained with an increase of blink frequency under stressful or frightening conditions (Harrigan & O´Connell, 1996).

The monotonous monitoring task that was used to induce task related passive fatigue, fulfilled its purpose. KSS and eye-tracking parameters indicate, that fatigue increased during this task. PERCLOS and related eye-tracking parameters could be confirmed as reliable drowsiness detection technologies as these parameters changed significantly over time during the ride with the monotonous monitoring task. Applicability of eye-tracking as drowsiness detection in CAD for different races and light conditions should be further investigated. Detrimental effects of task related passive fatigue on take-over performance could not be demonstrated in this study. There were no significant effects in take-over performance dependent on the monotonous monitoring task, used to induce drowsiness, compared to the quiz task and its activating effect (Schömig et al., 2015).

LIMITATIONS

All participants were employees of the BMW Group and familiar with driver's assistance systems. Further, a majority of participants were male. Therefore, a transfer of these findings on basic population is not possible without restrictions. Our data further suggests that a longer exposure time may yield more promising results. This is supported by findings in manual driving (i.e. Schmidt et al., 2009) who show performance decrements in manual driving. Therefore we suggest to prolong the length of the rides in further studies to investigate effects of NDRTs in CAD on take-over performance.

CONCLUSION

Summarized, the monotonous monitoring task induced task related fatigue after a time-on task of 25 min, which could be demonstrated by a rise of subjective KSS ratings, PERCLOS and blink related parameters. However, task related passive fatigue evoked with a tiring monitoring task, did not negatively affect the take-over performance of the drivers. However, our data further suggests that effects may be found in rides with longer time on task.

REFERENCES

Åkerstedt, T., & Gillberg, M. (1990). Subjective and Objective Sleepiness in the Active Individual. *International Journal of Neuroscience*, *52*(1-2), 29–37.

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775–779.

Forsman, P. M., Vila, B. J., Short, R. A., Mott, C. G., & van Dongen, H. P. A. (2013). Efficient driver drowsiness detection at moderate levels of drowsiness. *Accident; analysis and prevention*, *50*, 341–350.

Hargutt, V. (2003). *Das lidschlussverhalten als Indikator für Aufmerksamkeits-und Müdigkeitsprozesse bei Arbeitshandlungen* (Vol. 233).

Harrigan, J. A., & O'Connell, D. M. (1996). How do you look when feeling anxious? Facial displays of anxiety. *Personality and Individual Differences*, *21*(2), 205–212.

Jewett, M. E., Dijk, D.-J., Kronauer, R. E., & Dinges, D. F. (1999). Dose-response relationship between sleep duration and human psychomotor vigilance and subjective alertness. *Sleep: Journal of Sleep Research & Sleep Medicine*.

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. *Procedia Manufacturing*, *3*, 2403–2409.

Lenné, M. G., Triggs, T. J., & Redman, J. R. (1997). Time of day variations in driving performance. *Accident Analysis & Prevention*, *29*(4), 431–437.

Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(2), 659–686.

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(3), 218–224.

Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(5), 734–746.

Pack, A. I., Pack, A. M., Rodgman, E., Cucchiara, A., Dinges, D. F., & Schwab, C. W. (1995). Characteristics of crashes attributed to the driver having fallen asleep. *Accident Analysis & Prevention*, *27*(6), 769–775.

SAE International (2014, January 16). *Taxonomy and Definitions for Terms Related to 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*(4), 287.

Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., & Othersen, I. (2015). The Interaction Between Highly Automated Driving and the Development of Drowsiness. *Procedia Manufacturing*, *3*, 6652–6659.

Gimeno, P., Pastor Cerezuela, G., & Choliz Montanes, M. (2006). On the concept and measurement of driver drowsiness, fatigue and inattention: implications for countermeasures. *International journal of vehicle design*, *42*(1-2), 67–86.

Treat, JR, Tumbas, N. S., McDonald, S. T., Shinar, D., & Hume, R. D. (1979). *Tri-level study of the causes of traffic accidents. executive summary.*

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance Requires Hard Mental Work and Is Stressful. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 433–441.

Wierwille, W. W., Wreggit, S. S., Kirn, C. L., La Ellsworth, & Fairbanks, R. J. (1994). *Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness. final report.*