Highly automated driving: How to get the driver drowsy and how does drowsiness influence various take-over aspects?

Veronika Weinbeer[1,2,*], Christoph Baur[2], Jonas Radlmayr[2], Julian-Sebastian Bill[3], Tobias Muhr[1], and Klaus Bengler[2]

[1]AUDI AG, Germany
[2]Technical University of Munich, Germany
[3]Otto-von-Guericke University Magdeburg, Germany

* e-mail address: veronika.weinbeer@audi.de

***Abstract*— In the context of highly automated driving (HAD) the driver state drowsiness is becoming increasingly important. For assessing the usefulness of different strategies to manage drowsiness during HAD, appropriate test methods are needed. To determine whether a right-hand-drive vehicle (RHDV) is a suitable method, a study with 31 participants was conducted on the motorway A9 in Germany. Two investigators evaluated the drowsiness level (DL) of the participants during the test drive. Depending on the participant's DL Requests to Intervene (RtI) were triggered. There was no statistically significant influence of drowsiness on take-over-time aspects. Our results indicate that extremely drowsy drivers are still able to perceive and to react to a RtI. However, it should be considered that the take-over scenario used in this study was rather simple and quality aspects could not be assessed by the RHDV-setting. This study demonstrates that it is possible to induce and enhance drowsiness by controlling several influencing factors such as caffeine and atmosphere. Thus, this approach can be helpful for future studies when evaluating the effectiveness of different reactivation or transition strategies to manage drowsiness during HAD.**

Keywords*— highly automated driving; driver state; drowsiness; sleepiness; expert rating; observer-rated sleepiness, human-machine interaction; right-hand-drive vehicle; wizard-of-oz;

I. INTRODUCTION

Various investigations of accident reports as well as driving simulator studies show that there is a link between an increase of drowsiness and a decrement of driving performance (e.g., [1–3]). These findings suggest that drowsiness might crucially influence take-over performance after an automated drive. Nevertheless, this effect has hardly been studied in the context of HAD. However, we need knowledge about the development of drowsiness during HAD in order to investigate the influence of drowsiness on take-over performance. Further, it is necessary to design reactivation strategies and to derive requirements for driver-monitoring systems.

II. STATE OF THE ART

A. *Drowsiness and Driving Performance*

In the field of sleep research the distinction between the terms sleepiness, drowsiness, and fatigue is inconclusive as various researchers pointed out [4–7]. In this context Johns describes that drowsiness "*is a transitional state between wakefulness and sleep"* [8]. Also, no golden standard of the definition and measurement of sleep propensity [8] and fatigue exists [9]. Fatigue, for example, is described as "*a biological drive for recuperative rest*" [10]. Furthermore, this article presents a relation between fatigue and impaired performance capabilities. Another model describes that task-related-active and task-related-passive fatigue as well as sleep-related fatigue are influencing driving performance. Furthermore, these two different types of task-related fatigue can affect sleep-related fatigue [11]. Contrary to these models, other scientists state that fatigue and sleepiness are two different phenomena and therefore should be differentiated [7]. In another study the authors conclude that the reasons for drowsiness, fatigue or sleepiness can differ. However, they assume that for these constructs the reduced driver performance and its negative influence on traffic safety are similar [13]. The effect of sleepiness on the probability of incidents and accidents was evaluated in a driving simulator study. Results suggest that the probability of incidents and accidents increases extremely as level 7 on the Karolinska-Sleepiness Scale is exceeded [14]. Also, various driving simulator studies found a significant decrement of different driving performance metrics such as the standard deviation of the lateral position as drowsiness evolved [2, 15, 16]. Furthermore, slowed down reaction times were observed when comparing the results of reaction-time tasks with and without sleep deprivation [17, 18].

To sum up, there is not yet any consensus about the definitions and the distinction of the terms drowsiness, fatigue, and sleepiness. However, the link between a negative influence on driving performance and an increase of sleepiness, fatigue or drowsiness seems to be common to these constructs.

B. *Factors Influencing Take-Over Performance*

Automation aims to enhance safety, comfort and transportation efficiency [19]. To achieve these goals various challenges must be overcome. Various authors pointed out that human factors like skill degradation, the ability to monitor and to detect system failures as well as overreliance need to be considered in the context of automation (e.g., [20–23]). During

HAD drivers do not constantly need to monitor the system. Though, drivers still must be able to take back control appropriately when there is a request to intervene [24]**.** In this context researches examined whether and to what extend take-over time in critical (e.g., [25–28]) and non-critical situations [29, 30], non-driving-related tasks (e.g., [31, 32]), traffic density (e.g., [32, 33]), age [34], trust (e.g., [35–37]), and take-over designs (e.g., [38]) influence take-over performance. Take-over performance can be divided in time and quality aspects and can be assessed using different metrics like for example hands-on time and driver-intervention time, lateral and longitudinal acceleration as well as the minimal time to collision (TTC_{min}) [39]. However, so far little is known about the influence of drowsiness on take-over performance. Therefore, the following section provides an overview of studies investigating the influence of drowsiness or longer automation duration in the context of HAD.

C. *Influence of Drowsiness on Take-Over Performance*

One driving simulator study found no effect of automation duration (5 and 20 min) on take-over performance [40]**.** Another approach was used to investigate the influence of HAD on the development of drivers' drowsiness. Therefore, four DLs were estimated in real time depending on eyelid-movements. It was found that a quiz as a non-driving-related activity has the potential to keep the drowsiness level constant, whereas HAD without task-engagement led to an increase of drowsiness. The latter was also observed in the manual driving condition. Additionally, participants showed large differences in the development of drowsiness. In this study, it was not possible to analyze take-over performance statistically. However, based on descriptive results the authors assume that a take-over time of 12s might be sufficient for taking over control even when experiencing high levels of drowsiness [41]. Depending on the subjective assessment of drowsiness on the Stanford-Sleepiness Scale take-over requests were triggered in another driving simulator study. No effects of drowsiness on intervention time, TTC_{min}, and longitudinal acceleration were observed. However, drowsiness significantly influenced the lateral acceleration [42]. To the authors' knowledge, this is the only study proving for at least one parameter that drowsiness might be an important factor in the context of HAD.

It is apparent, that these approaches differ strongly and hence are difficult to compare. Therefore, the aim of the study at hand is to investigate the development of drowsiness during simulated HAD. As knowledge about the drowsiness process can be helpful in deciding about the most suitable approach (drowsiness level or automation duration) regarding the research question in the context of HAD. As the influence of drowsiness was only assessed in driving simulator studies there is a need for real driving studies [43]. Therefore, we modified a vehicle (similar to the RRADS [44]) to generate a more realistic setting in real driving environment.

III. Method

A. *Right-Hand-Drive Vehicle (RHDV) and Test Track*

The usage of the HAD-system considered here is limited to motorways. To evaluate the development of drivers' drowsiness during HAD and its influence on take-over aspects (TOAs) we adopted an AUDI Q7 right-hand-drive vehicle (RHDV) (see fig.1). For this, pedal dummies and a steering wheel dummy were integrated at the passenger's seat. The steering wheel can be turned to the left and right before reaching the stop position (see fig.1 bottom-right). Springs were integrated to simulate the resistance of the break and gas pedal. Also driving school wing mirrors and a rearview mirror were integrated so that participants had the opportunity to monitor the traffic. A curtain was attached between the drive wizard (investigator) and the participant. Furthermore, four cameras were integrated. The video images were displayed in real time on a screen on the second row. This made it possible for the two raters (investigators sitting at the back seat) to observe and rate participants' DL and to analyze the hands-on time afterwards.

Fig. 1 Modified vehicle (RHDV) to simulate HAD. Top: steering wheel (left) and pedal dummies (right); Bottom: abstract take-over scenario (left) and steering impulse (right).

A button on the center console displayed the pilot status by an LED: pilot not available, pilot available, pilot active and please take over. Additionally, a LCD (system status display) was attached to the middle of the steering wheel. A static ego-vehicle, the current pilot state, speed, direction indicators, and the arrival time (121 minutes at the beginning of the examination) were displayed on this screen.

Request to Intervene (RtI) and continous-reaction task

Depending on the driver's DL the RtI and the take-over-scenario were triggered simultaneously. The RtI was presented visually "please take over" (on the system status display and on a 12" tablet on the center console) and acoustically. For this, the take-over sound of the series Audi Adaptive Cruise Control system was used. Three displays attached to the dashboard (see fig. 1 bottom-left) were supposed to abstractly represent three traffic lanes. In case of a RtI stylized brake lights were presented on the right and middle display representing the end of a traffic jam. Consequently, the task of the participants was to perform a lane change maneuver to the left as realistically as possible by using the steering wheel dummy as well as the pedal dummies. During the take-over scenario, the participants did at no time intervene in the real driving process. The brake lights disappeared when participants pressed the brake pedal or gave a steering impulse. In case of no reaction the brake lights were presented for a maximum of 7s.

After the take-over situation, a continuous visual-motoric-cognitive task was conducted for five minutes in order to investigate the development of drivers' drowsiness in a continuous task, as it might happen in a following less automated drive. Participant's task was to observe a specific area on the LCD-screen mounted on the steering wheel. If an object entered this ego area its color changed depending on the entrance area from grey to red, blue, yellow or purple. Then, participants had to press the corresponding colored steering wheel button. After pressing the correct button, the color of the object changed back to grey. If the participant failed to press the correct button the color changed back to grey when the object left the ego area.

B. *Drowsiness Assessment*

The method of observer-rated sleepiness (ORS) was used as it allows to assess the drowsiness process in real time, to avoid interruptions by interviewing the participants and to ensure that results are independent from a specific driver-monitoring system. According to ORS, the drowsiness level is rated on the basis of various indicators, such as duration of eyelid closure that are assigned to different DLs [45]. A regression analysis revealed a significant association between ORS and the subjective rating of sleepiness on the Karolinska-Sleepiness-Scale (KSS) and between ORS and the occurrence of various mannerisms. However, the model fit was low [46]. Besides, another examination found a moderate correlation between ORS and blink duration and performance observations [47]. Furthermore, Wierwille and Ellsworth (1994) proved a high interrater reliability ($R = .81$, $p < .001$). Our evaluation is based on the first scale developed to assess the DL of participants by raters by Wierwille and Ellsworth [45]. Two adjustments were implemented to this scale based on the findings of Karrer-Gauß [48] and Wiegand and colleagues [49]. Karrer-Gauß modified the original scale by changing the eyelid closure time of the item "very drowsy" from 2-3 seconds to 1-2 seconds. In her examination this level was found to be highly relevant [48]. Therefore, this level was added to the original scale of ORS [45] and named "drowsy". DL2 was supplemented by some indicators reported by Wiegand and colleagues [49] (see. table 1). According to an expert estimate the reliability and validity of this kind of rating is higher if extreme drowsiness states are assed. Also the experts point out that the reliability of this method depends on the raters' expertise [50]. To increase the reliability of our results two investigators (sitting on the back seat of the RHDV) evaluated the DL of the participants.

Two investigators rated the DL of the participants on an application on a tablet. The ratings as well as the corresponding timestamps were recorded. There was no evaluation for one minute followed by a further one minute interval after which the DL was assessed by the investigators. This process was repeated during the entire test drive as this procedure allows to observe the development of drowsiness as a function of time. A one-minute interval was found to be appropriate to observe changes in eyelid parameters [51]. Based on pretests the one minute interval appeared to be too short in some cases, as even during that interval participants reached a higher DL. To make sure that the RtI was triggered immediately when participants experienced specified drowsiness levels, investigators could communicate by nodding in case of uncertainty. Therefore, an analysis of the interrater reliability afterwards would be inconclusive. However, this limitation was accepted as the properness of the drowsiness level is a key factor in order to analyze the influence of specific drowsiness levels on take-over aspects.

TABLE 1 LEVELS OF DROWSINESS AND INDICATORS

Level of Drowsiness	**Indicators**
1 *not drowsy*	appearance of alertness present; normal facial tone; normal fast eye blinks; short ordinary glances; occasional body movements / gestures; [45]
2 *slightly drowsy*	still sufficiently alert; less sharp / alert looks; longer glances; slower eye blinks; first mannerisms as: rubbing face / eyes, scratching, facial contortions, moving restlessly in the seat; [45] and [49]
3 *moderately drowsy*	mannerisms; slower eye lid closures; decreasing facial tone; glassy eyes; staring at fixed position; [45]
4 *drowsy*	eyelid closures (1-2s); eyes rolling sideways; rarer blinks; no proper focused eyes; decreased facial tone; lack of apparent activity; large isolated or punctuating movements; [48]
5 *very drowsy*	eyelid closures (2-3s); eyes rolling upward /sideways; no proper focused eyes; decreased facial tone; lack of apparent activity; large isolated or punctuating movements; [45]
6 *extremely drowsy*	eyelid closures (4s or more); falling asleep; longer periods of lack of activity; movements when transition in and out of dozing; [45]

IV. EXPERIMENTAL DESIGN AND PROCEDURE

A. *Experimental Design*

In order to understand the development of drowsiness during HAD, the time required to reach DL4 or DL6 for the first time and the number of participants reaching DL6 are determined. Table 2 presents the experimental design used in this study.

TABLE 2 EXPERIMENTAL DESIGN

	Levels of Drowsiness		
	DL1	**DL4**	**DL6**
Group A n = 16	Group A (DL1)	Group A (DL4)	Group A (DL6)
Group B n = 15	x	Group B (DL4)	Group B (DL6)

The effect of drowsiness on various take-over-time aspects is analyzed in a within-subjects comparison of the three different DL of Group A (DL1, DL4, and DL6). Thus, participants, who had not reached the highest DL were excluded from the TOA analysis. To control for training effects Group B served as a control group. Consequently, TOAs of Group A (DL4) and Group B (DL4) are compared.

B. *Procedure and Drowsiness Manipulation*

The study lasted for a maximum of 3 hours per participant and was conducted from 9 a.m. to 12 a.m. or from 1 p.m. to 4 p.m.. In an invitation letter participants were asked to avoid drinking caffeinated beverages 5 hours before participating in the study and to make sure that they do not get hungry during the study. Participants were informed that no passenger and steering wheel airbag are available due to the vehicle adaption.

They were asked to turn their mobile phone off, not to drink or eat and to avoid chewing gum during the entire study. Participants were aware, that they can cancel the study at any time. Before the ride, participants got a short introduction regarding the handling of the simulated autopilot. They were also informed about the drive wizard. After the activation participants should relax as best as possible. Therefore, discrete relaxing background music was played.[1] The volume could not be adjusted. Participants were instructed to avoid closing their eyes and to fall asleep during the test drive. They should also avoid talking to the investigators.

Thereafter, the take-over scenario, the RtI as well as the visual-motoric-cognitive task were explained and presented to the participants for the first time. In the case of a take-over scenario the participants should react as realistically as possible. This time a steering impulse to the left was necessary. Also, participants were requested to follow the road traffic regulations by using mirrors and indicators. Furthermore, it was explained that a steering impulse is unavoidable to handle this situation successfully.

The take-over scenario and the visual-motoric-cognitive task were exercised one more time on the way to the motorway entrance. This time participants had to steer to the right in order to prevent trainings effects. At the beginning of the motorway entrance participants were asked to draw the curtain and to enter the final destination into an application. When entering the motorway the pilot status changed to available. After activating the autopilot the relaxation phase began. Thereafter, RtI followed by the continuous-visual-motoric task were triggered in different DLs depending on the participant's group (see table 2).

The measurement started at the motorway entrance of Lenting. HAD was simulated from Lenting to the interchange Nürnberg-Ost and back again. This track was selected because the motorway must not be left during the test drive. Automatic cruise control, lane-keeping assist as well as the active speed limiter were deactivated during this study. The maximum speed was 130km/h. Lane changes were performed very cautiously.

[1] Various tracks of the album Ambiente No.1 (Dr. Stein) were used and repeated during the relaxation phases.

V. RESULTS

A total of 31 employees of the AUDI AG participated in this study. The sample consisted of n = 12 females and n = 19 males. The mean age of the sample was 30.61 years (SD = 8.16). Participants had owned their driving license on average for 13.45 years (SD = 8.20). One participant had to be excluded from the evaluation as constantly narrowed eyes did not allow to assess the DL by the ORS-method.

A. *Development of Drowsiness*

Table 3 presents the cumulative percentage of participants reaching DL4 or DL6 for the first time as a function of time - regardless of the study group. In total 76.67% of the participants reached DL4 and 63.33% reached DL6. On average, participants reached DL4 after 30.49 minutes (SD = 18.71) and DL6 after 42.04 minutes (SD = 15.44) for the first time. The number of participants that reached DL6 in the afternoon (n = 8) was greater than in mid-morning (n = 3).

TABLE 3 CUMULVATIVE PERCENTAGE UNTIL PARTICIPANTS REACHED DL4 OR DL6 FOR THE FIRST TIME AS A FUNCTION OF TIME (SAMPLE SIZE: N = 30)

time (minutes)	**DL4 (cumulative percentage)**	**DL6 (cumulative percentage)**
0	0.00 %	0.00 %
5	3.33 %	0.00 %
10	10.00 %	0.00 %
15	20.00 %	0.00 %
20	23.33 %	3.33 %
25	30.00 %	10.00 %
30	46.67 %	16.67 %
45	60.00 %	40.00 %
60	73.33 %	56.67 %
75	76.67 %	60.00 %
>75	76.67 %	63.33 %
	never reached DL4: 23.33%	never reached DL6: 36.67%

B. *Take-Over Aspects*

Hands-on time and driver-intervention time were assessed in this study. Driver-intervention time was defined as the time from the beginning of the RtI until reaching the mechanical stop position of the steering wheel dummy.

One participant did not steer to the left. However, the participant took both hands on the steering wheel and asked whether steering is now necessary. Hence, only the hands-on time of this participant could be assessed. Furthermore, only take-over aspects of the participants that reached the specified drowsiness levels were analyzed.

In total 12 data sets were available for the hands-on and driver-intervention time in Group A (DL4) and 10 data sets for the same metrics in Group B (DL4). The TOAs of Group A (DL4) and Group B (DL4) were normally distributed (Shapiro-Wilks-test: hands-on time Group A (DL4) p = .071, hands-on time Group B (DL4) p = .118, driver-intervention time Group A (DL4) p = .557, and driver-intervention time Group B (DL4) p = .820). Also, the assumption of variance homogeneity was not violated (Levene's test p > .05). No training effects were observed for the hands-on as well as for the driver-intervention times, as assessed by a t-test for independent samples (hands-on time: t(20) = -0.286, p = .778, and driver-intervention time: t(20) = -0.471, p = .643).

In total 9 participants of Group A experienced DL1, DL4, and DL6. Hands-on times for Group A (DL4) were not normally distributed (Shaprio Wilk's test: p = .021). Thus, the influence of drowsiness on hands-on time was analyzed using a Friedman test. There was no influence of drowsiness on hands-on time ($\chi 2(2) = 2.00$, p = .368). The average hands-on time was 1.74s (SD = 0.27) when participants were not drowsy and became slightly shorter when participants were drowsy (M = 1.49s, SD = 0.29) or extremely drowsy (M = 1.49s, SD = 0.35).

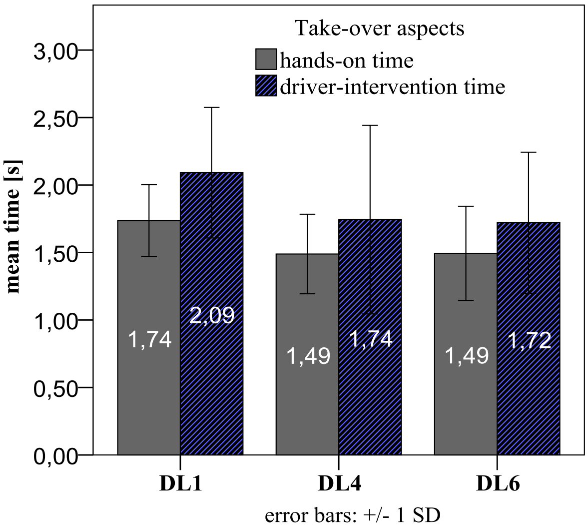

Fig. 2 Hands-on time (n = 9) and driver-intervention time (n = 8) of Group A depending on the drowsiness level (DL1, DL4, and DL6)

Similar results were observed for the driver-intervention time (n = 8) (see fig. 2). Driver-intervention times of this sample were normally distributed, as assessed by Shapiro-Wilk's test (DL1: p = .252, DL4: p = .343, and DL6: p = .942). An ANOVA with repeated measures was conducted to analyze the influence of drowsiness on driver-intervention time. The assumption of sphericity was not violated (Mauchly's test: $\chi 2(2) = 0.749$, p = .688). The mean driver-intervention time of Group A (DL1) was 2.09s (SD = 0.48) when participants were not drowsy. Driver-intervention times became slightly faster with an increase of drowsiness (Group A (DL4): M = 1.74s, SD = 0.70), and Group A (DL6): M = 1.72s, SD = 0.52) (see fig.2). However, drowsiness also did not significantly influence driver-intervention time (F(2, 14) = 2.64, p = .107).

Apart from the quantitative analysis of the influence of drowsiness, it was observed that some participants gave a startled sound in case of a RtI when experiencing higher levels of drowsiness, whereas such a behavior was not observed in the not drowsy condition.

C. *Right-hand-drive vehicle to simulate HAD*

At the end of the study participants were asked to rate whether they had felt to be driven highly automated on a scale ranging from 1 (no impression of HAD) to 10 (strong impression of HAD). The average rating was 6.89 (SD = 1.52). Furthermore, participants were asked to describe how their feeling of HAD could be optimized in the RHDV-setting. The main findings of this open question are presented in the following. Two participants said that a rotating steering wheel during HAD would be helpful for total immersion. Five participants stated that the drive wizard should be more hidden (e.g., complete separation of driver and participant, smaller driver). Two participants mentioned that a real cluster instrument could improve their feeling of HAD. In addition, two participants mentioned that a computer voice might be helpful.

VI. Discussion and Limitations

Results of the drowsiness process indicate that drowsiness evolves very individually and can be experienced even after a short period of HAD. One participant already reached DL4 within the first 5 minutes, whereas others did not reach this drowsiness level during the entire test drive. Those inter-individual differences are in line with existing studies in the context of HAD [41, 42]. In the current study optimal conditions to generate drowsiness were realized (for example relaxing music, no communication, no caffeinated beverages 5h prior to the trial). Therefore, drowsiness might occur later in reality, e.g. when caffeinated beverages are consumed or conversations are allowed. As 76.67% of the participants reached DL4 and 63.33% reached DL6 during normal working hours and without sleep deprivation it is likely that drowsiness occurs when drivers want to use the time provided by a motorway pilot to relax.

Also, our results indicate that a manipulation of automation duration is less suitable to investigate drowsiness effects. Large individual differences were observed by large standard deviations when participants had initially reached DL4 (SD = 18.71min) or DL6 (SD = 15.44min) while some participants reached neither of these drowsiness levels. Analyzing the influence of drowsiness afterwards by comparing groups of shorter and longer automation duration could therefore lead to an underestimation of possible drowsiness effects. The knowledge about the cumulative percentage of participants reaching a certain drowsiness level as a function of time can be helpful for future studies. In order to investigate the influence of drowsiness, the cumulative percentages allows for an a priori estimation of the time needed to reach a certain DL. According to table 3, a period of about 60 minutes is an appropriate time frame for a test drive. In this study an extra time above 60 minutes did not lead to a considerable increase of the number of participants reaching DL4 or DL6. In fact, only one further participant reached DL4 or DL6 between 60 and 75 minutes. After 75 minutes, only one more participant reached DL6.

The assessed TOAs hands-on and driver-intervention time did not change significantly with an increase of drowsiness. Contrary to other studies that observed slowed down reactions when executing simple reaction time tasks in case of drowsiness [17, 18, 52], take-over-time aspects even became slightly faster with an increase of drowsiness in the current study. Though we tried to control for training effects, these effects cannot be completely excluded. All participants reacted to an RtI. All steering impulses were performed to the left correctly – except for one participant of the non-drowsy condition who asked whether steering is necessary. Although no statistically significant interaction between drowsiness and take-over-time aspects were found there are indicators that drowsiness might influence take-over-quality aspects. For instance when experiencing higher levels of drowsiness some participants gave a startled sound in case of a RtI and reaction times became slightly shorter. Thus, drowsiness might lead to startled or surprised reactions. Therefore, the RHDV-method should be optimized for future studies, so that take-over-quality aspects (e.g., the existence of mirror checks when performing a lane change maneuver) can be assessed. Also, it needs to be considered that the sample size was small in this study. Therefore, future studies might use larger samples.

The RHDV is yet limited to take-over-time aspects. For time aspects it was found that the mean hands-on time of the not drowsy group ($M = 1.79$s) was very similar compared to the mean hands-on time ($M = 1.74$s) of an existing driving simulator study using also a take-over time of 7s [25]. This indicates that at least hands-on times can be assessed using the RHDV-setting. In the driving simulator study the take-over situation was a broken-down vehicle on the right lane of a motorway. Participants' task was to stop or to execute a lane change maneuver. Therefore, the situations of the driving simulator study and of the current study are comparable to a certain extent.

However, there is still a need for real-driving studies to investigate take-over performance in reality. Therefore, it could be helpful to improve and validate the RHDV-setting by comparing the results of the RHDV-method to the results of real-driving studies in non-critical situations. This could be helpful to enhance the transferability of the results obtained by the RHDV-setting when investigating take-over performance in critical driver states.

In order to ensure that transitions are executed in a controlled manner an adaption of the RtI might be useful for drowsy drivers during HAD. As loud requests to intervene might lead to fast and in some cases startled reactions instead of controlled and safe transitions, future research is needed in order to investigate the impact of RtI in case of severe drowsiness.

VII. Conclusion

In this study a modified RHDV was used for assessing the influence of drowsiness on take-over-time aspects. Therefore, two investigators evaluated participants' drowsiness level during the test drive using the ORS-method. Requests to intervene were triggered when participants reached a certain drowsiness level. There was no significant influence of the drowsiness level on take-over-time aspects. However, in contrast to the participants of the non-drowsy condition some participants experiencing higher levels of drowsiness showed surprise when a RtI happened. Thus, further research is needed to ensure that transitions are executed in controlled manner, taking into account that faster reactions might not always be safe reactions. Therefore, an adaption of the RtI or an implementation of reactivation strategies according to the drowsiness state might be helpful. Finally, the current study demonstrates that it is possible to induce and enhance drowsiness during normal working hours and without sleep deprivation by controlling several influencing factors such as caffeine and atmosphere. This knowledge can be helpful for future studies evaluating the effectiveness of different strategies to manage drowsiness in the context of HAD.

Acknowledgement

This work results from the joint project Ko-HAF - Cooperative Highly Automated Driving and has been funded by the Federal Ministry for Economic Affairs and Energy based on a resolution of the German Bundestag.

References

[1] A. J. Filtness, K. A. Armstrong, A. Watson, and S. S. Smith, "Sleep-related crash characteristics: Implications for applying a fatigue definition to crash reports," (eng), *Accident; analysis and prevention*, vol. 99, no. Pt B, pp. 440–444, 2017.

[2] R. Rossi, M. Gastaldi, and G. Gecchele, "Analysis of driver task-related fatigue using driving simulator experiments," *Procedia - Social and Behavioral Sciences*, vol. 20, pp. 666–675, 2011.

[3] F. Sagberg, "Road accidents caused by drivers falling asleep," *Accident Analysis & Prevention*, vol. 31, no. 6, pp. 639–649, 1999.

[4] W. R. Pigeon, M. J. Sateia, R. F. Ferguson "Distinguishing between excessive daytime sleepiness and fatigue: Toward improved detection and treatment," *Journal of Psychosomatic Research*, vol. 54, no. 1, pp. 61–69, 2003.

[5] E. Ream and A. Richardson, "Fatigue: A concept analysis," *International Journal of Nursing Studies*, vol. 33, no. 5, pp. 519–529, 1996.

[6] A. Shahid, J. Shen, and C. M. Shapiro, "Measurements of sleepiness and fatigue," (eng), *Journal of psychosomatic research*, vol. 69, no. 1, pp. 81–89, 2010.

[7] J. Shen, J. Barbera, and C. M. Shapiro, "Distinguishing sleepiness and fatigue: focus on definition and measurement," (eng), *Sleep Medicine Reviews*, vol. 10, no. 1, pp. 63–76, 2006.

[8] M. Johns, "Rethinking the assessment of sleepiness," *Sleep Medicine Reviews*, vol. 2, no. 1, pp. 3–15, 1998.

[9] R. O. Phillips, *What is fatigue and how does it affect the safety performance of human transport operators?: Fatigue in Transport Report I*, 2014.

[10] A. Williamson *et al.*, "The link between fatigue and safety," (eng), *Accident; analysis and prevention*, vol. 43, no. 2, pp. 498–515, 2011.

[11] J. F. May and C. L. Baldwin, "Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, no. 3, pp. 218–224, 2009.

[12] J. L. Hossain *et al.*, “Subjective fatigue and subjective sleepiness: two independent consequences of sleep disorders?,” (eng), *Journal of sleep research*, vol. 14, no. 3, pp. 245–253, 2005.
[13] W. Vanlaar, H. Simpson, D. Mayhew, and R. Robertson, “Fatigued and drowsy driving: a survey of attitudes, opinions and behaviors,” (eng), *Journal of safety research*, vol. 39, no. 3, pp. 303–309, 2008.
[14] M. Ingre *et al.*, “Subjective sleepiness and accident risk avoiding the ecological fallacy,” (eng), *Journal of sleep research*, vol. 15, no. 2, pp. 142–148, 2006.
[15] T. Akerstedt, B. Peters, A. Anund, and G. Kecklund, “Impaired alertness and performance driving home from the night shift: a driving simulator study,” (eng), *Journal of sleep research*, vol. 14, no. 1, pp. 17–20, 2005.
[16] A. Anund *et al.*, “Driver impairment at night and its relation to physiological sleepiness,” *Scand J Work Environ Health*, vol. 34, no. 2, pp. 142–150, 2008.
[17] M. W. Johns, R. Chapman, K. Crowley, and A. Tucker, “A new method for assessing the risks of drowsiness while driving,” *Somnologie*, vol. 12, no. 1, pp. 66–74, 2008.
[18] P. Philip *et al.*, “Fatigue, Sleepiness, and Performance in Simulated Versus Real Driving Conditions,” *J Sleep Res*, vol. 28, no. 12, pp. 1511–1516, 2005.
[19] V. d. A. VDA, “Automation: From Driver Assistance Systems to Automated Driving,” 2015.
[20] E. L. Wiener and R. E. Curry, “Flight-Deck Automation: Promises and Problems,” *National Aeronautic and Space Administration*, 1980.
[21] M. Saffarian, de Winter, J. C. F., and R. Happee, “Automated Driving: Human-Factors Issues and Design Solutions,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, no. 1, pp. 2296–2300, 2012.
[22] R. Parasuraman and D. H. Manzey, “Complacency and bias in human use of automation: an attentional integration,” (eng), *Human factors*, vol. 52, no. 3, pp. 381–410, 2010.
[23] L. Bainbridge, “Ironies of Automation,” *Automatica*, vol. 19, no. 6, pp. 775–779, 1983.
[24] *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, SAE International J3016, 2016.
[25] C. Gold, D. Damböck, L. Lorenz, and K. Bengler, ““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*, vol. 57, no. 1, pp. 1938–1942, 2013.
[26] D. Damböck, “Automationseffekte im Fahrzeug - von der Reaktion zur Übernahme,” Dissertation, Fakultät für Maschinenwesen, Technische Universität München, München, 2013.
[27] B. Mok *et al.*, “Emergency, Automation Off: Unstructured Transition Timing for Distracted Drivers of Automated Vehicles,” in *2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC 2015): [Las Palmas], Gran Canaria, Spain, 15 - 18 September 2015*, Piscataway, NJ: IEEE, 2015, pp. 2458–2464.
[28] K. Zeeb, A. Buchner, and M. Schrauf, “What determines the take-over time? An integrated model approach of driver take-over after automated driving,” (eng), *Accident; analysis and prevention*, vol. 78, pp. 212–221, 2015.
[29] N. Merat, A. H. Jamson, F. C. Lai, M. Daly, and O. M. Carsten, “Transition to manual: Driver behaviour when resuming control from a highly automated vehicle,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 27, pp. 274–282, 2014.
[30] A. Eriksson and N. A. Stanton, “Takeover Time in Highly Automated Vehicles: Noncritical Transitions to and From Manual Control,” (eng), *Human factors*, vol. 59, no. 4, pp. 689–705, 2017.
[31] C. Gold, I. Berisha, and K. Bengler, “Utilization of Drivetime – Performing Non-Driving Related Tasks While Driving Highly Automated,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 59, no. 1, pp. 1666–1670, 2015.
[32] J. Radlmayr, C. Gold, L. Lorenz, M. Farid, and K. Bengler, “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*, vol. 58, no. 1, pp. 2063–2067, 2014.
[33] C. Gold, M. Körber, D. Lechner, and K. Bengler, “Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations: The Role of Traffic Density,” (eng), *Human factors*, vol. 58, no. 4, pp. 642–652, 2016.
[34] M. Körber, C. Gold, D. Lechner, and K. Bengler, “The influence of age on the take-over of vehicle control in highly automated driving,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 39, pp. 19–32, 2016.
[35] S. Hergeth, L. Lorenz, and J. F. Krems, “Prior Familiarization With Takeover Requests Affects Drivers?: Takeover Performance and Automation Trust,” *Human factors*, vol. 59, no. 3, pp. 457–470, 2016.
[36] A. Feldhütter, C. Gold, A. Hüger, and K. Bengler, “Trust in Automation as a Matter of Media Influence and Experience of Automated Vehicles,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 60, no. 1, pp. 2024–2028, 2016.
[37] C. Gold, M. Körber, C. Hohenberger, D. Lechner, and K. Bengler, “Trust in Automation? Before and After the Experience of Take-over Scenarios in a Highly Automated Vehicle,” *Procedia Manufacturing*, vol. 3, pp. 3025–3032, 2015.
[38] P. Kerschbaum, L. Lorenz, and K. Bengler, “Highly automated driving with a decoupled steering wheel,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1, pp. 1686–1690, 2014.
[39] C. Gold, “Modeling of Take-Over Performance in Highly Automated Vehicle Guidance,” Dissertation, Fakultät für Maschinenwesen, Technische Universität München, München, 2016.
[40] A. Feldhütter, C. Gold, S. Schneider, and K. Bengler, “How the Duration of Automated Driving Influences

Take-Over Performance and Gaze Behavior," in *Advances in Ergonomic Design of Systems, Products and Processes*, C. M. Schlick et al., Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2017.

[41] N. Schömig, V. Hargutt, A. Neukum, I. Petermann-Stock, and I. Othersen, "The Interaction Between Highly Automated Driving and the Development of Drowsiness," *Procedia Manufacturing*, vol. 3, pp. 6652–6659, 2015.

[42] J. Goncalves, R. Happee, and K. Bengler, "Drowsiness in Conditional Automation: Proneness, diagnosis and driving performance effects," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC): 1-4 Nov. 2016*, [Piscataway, New Jersey]: IEEE, 2016, pp. 873–878.

[43] M. Körber and K. Bengler, "Potential Individual Differences Regarding Automation Effects in Automated Driving," in *INTERACCIÓN 2014: Proceedings of the XV International Conference on Human Computer Interaction : 2014, Puerto de la Cruz, Tenerife, Spain*, New York: ACM, 2007, pp. 1–7.

[44] S. Baltodano, S. Sibi, N. Martelaro, N. Gowda, and W. Ju, "The RRADS platform". Retrieved from: http://wendyju.com/publications/RRADS_FINAL.pdf. (checked on: 13.11.2017)

[45] W. W. Wierwille and L. A. Ellsworth, "Evaluation of driver drowsiness by trained raters," *Accident Analysis & Prevention*, vol. 26, no. 5, pp. 571–581, 1994.

[46] S. M. Belz, G. S. Robinson, and J. G. Casali, "Temporal separation and self-rating of alertness as indicators of driver fatigue in commercial motor vehicle operators," (eng), *Human factors*, vol. 46, no. 1, pp. 154–169, 2004.

[47] A. Anund, C. Fors, D. Hallvig, T. Åkerstedt, and G. Kecklund, "Observer rated sleepiness and real road driving: an explorative study," (eng), *PloS one*, vol. 8, no. 5, pp. e64782, 2013.

[48] K. Karrer-Gauß, "Prospektive Bewertung von Systemen zur Müdigkeitserkennung: Ableitung von Gestaltungsempfehlungen zur Vermeidung von Risikokompensation aus empirischen Untersuchungen," Dissertation, Verkehrs- und Maschinensysteme, Technische Universität Berlin, Berlin, 2011.

[49] D. M. Wiegang, J. McClafferty, S. E. McDonald, and R. J. Hanowski, "Development and Evaluation of a Naturalistic Observer Rating of Drowsiness Protocol," *The National Surface Transportation Safety Center for Excellence*, 2009.

[50] C. Platho, A. Pietrek, and H. Kolrep-Rometsch, *Erfassung der Fahrermüdigkeit: Bericht zum Forschungsprojekt FE 82.513/2010*. Bremen: Fachverl. NW in der Carl Schünemann Verl. GmbH, 2013.

[51] D. Sandberg, "The performance of driver sleepiness indicators as a function of interval length," in *2011 14th International IEEE Conference on Intelligent Transportation Systems - (ITSC 2011)*, 2011, pp. 1735–1740.

[52] P. Philip, J. Tillard, M. A. Quera-Salva, B. Bioulac, and T. Åkerstedt, "Simple reaction time, duration of driving and sleep deprivation in young versus old automobile drivers," *J Sleep Res*, vol. 8, no. 1, pp. 9–14, 1999.