

Driving and tiredness: Results of the behaviour observation of a simulator study with special focus on automated driving

Clemens Kaufmann^a, Matthias Frühwirth^b, Dietmar Messerschmidt^c, Maximilian Moser^d, Arno Eichberger^e, Sadegh Arefnezhad^f

^aApptec Venture GmbH Factum, clemens.kaufmann@factum.at

^bHuman Research Institute of Health Technology and Prevention Research GmbH, matthias.fruehwirth@humanresearch.at

^cHuman Research Institute of Health Technology and Prevention Research GmbH, dietmar.messerschmidt@humanresearch.at

^dHuman Research Institute of Health Technology and Prevention Research GmbH, Medical University Graz, Austria, maximilian.moser@humanresearch.at

^eInstitute of Automotive Engineering, Graz University of Technology, arno.eichberger@tugraz.at

^fInstitute of Automotive Engineering, Graz University of Technology, s.arefnezhad@tugraz.at

ABSTRACT: *The development of automated driving is an ongoing process; nonetheless, certain problems remain unresolved. One of them is the question when the automated vehicle control system should hand over the control to a human driver and whether this can be done in a safe way. What happens if a driver is not ready to take over? Can the system somehow estimate the status of the driver? The WACHsens simulator study was designed with the aim to gain more knowledge about when and how drivers are getting sleepy with special focus on automated driving. The overall goal of the project was to merge data from vegetative vigilance, camera observation and driving behaviour. This article describes the process of the driving behaviour observation and the evaluation of the data collected during the observation.*

An enhanced observation scheme made it possible to determine, at any point in time of the 30 minutes drives, in which posture the test person is and in what degree of drowsiness the test person is. It is based on the variables and scales which have been used in other studies such as ORD (Observer Rating of Drowsiness) and ORS (Observer Rated Sleepiness). They were linked to the observation method of the Vienna driving test to allow continuous observation.

197 simulator test drives from 50 test persons were analyzed by the observers. Four different scenarios were evaluated for all test subjects: tired/manual, tired/automatic, rested/manual, and rested/automatic. The aim of the observation analysis was to investigate differences in body movements and activities according to

personal characteristics (age, gender, driving experience, experience with assistance systems) and regarding the different scenarios.

The categorization of the drowsiness level of the test persons by the observers corresponds very well with the subjective assessment of the test subjects (measured by the Karolinska sleepiness scale KSS). A comparison of the different scenarios shows that most of the signs of sleepiness or situations in which the test subjects fell asleep were observed during the tired/automated trips. But even during the rested/automated drive over 40% of the test persons showed signs of tiredness, roughly the same number actually fell asleep as in the tired/manual drive. No significant differences between the personal characteristics (gender, age, and experience with assistance systems) regarding the number of body movements (change of position and activities) or sleepiness levels could be found.

A significant difference was found between the different scenarios and the comparisons between the tired/rested trips and the manual/automated trips regarding the moment in which the test persons showed first signs of tiredness. During the automated trips and/or if the test subjects showed signs of progressing weariness, the first signs of tiredness were registered significantly earlier than during the trips in which the test subjects drove manually and/or were rested.

The results show that the mode of operation – manual or automated driving – impacts the time course and level of sleepiness while driving. This sheds light on the importance to carefully evaluate driving automation

systems that assume a driver as emergency fallback. Further research is recommended to investigate safe modes of control hand over in automated driving.

KEYWORDS: *Automated driving, tiredness, simulator, observation*

1. INTRODUCTION

An important milestone is the introduction of systems according to SAE driving automation Level 3 where people are allowed to carry out secondary activities instead of driving, but must be able to act as the fallback and take over control on request of the system. In this phase, which will also be permitted by the new version of the Vienna Convention, the human-machine interface, more precisely the transfer of control, is of crucial importance. As described in literature (Bainbridge, 1983), automation is intended to avoid errors but also risks to create and include new ones. The stronger the automation, the less the people can understand these automated systems and predict its behaviour. Automation Surprises are the new failures that arise when the drivers act as mere “operators”. Such adverse side effects are the loss of the “awareness mode”, complacency (excessive trust in technology), threat to the individual competence, operating errors and disregard of (frequent) alarms. Drivers who generally accept an integrated system tend to do so e.g. by the virtue of increased feeling of comfort or of misunderstanding of the actual functional limits, by decreasing level of attention while driving and taking responsibility for handing over the safe operation to the vehicle (Chaloupka et al. 1998).

Another problem in this regard is tiredness. A comparison of studies from different countries estimated (Royal Society for the Prevention of Accidents, 2001) that up to 50% of traffic accidents where trucks were involved and between 15 and 35% car accidents happened due to drowsiness of the driver. It is an undisputed fact that drowsiness drastically reduces responsiveness or, in the case of microsleep, can lead to accidents also without external danger. Driving in a state of sleepiness is often done by drivers intentionally (be it out of economic pressure from professional drivers or simply from the urge to arrive at their destination sooner and

without many interruptions); on the other hand, the subjective assessment of individual driving ability can be deceptive: In the course of the day, the probability of falling asleep at 4 p.m. is more than twice as high as at 10 a.m. or 12 a.m., but the subjective assessment of sleepiness is the same (Moller et al., 2006). With regard to automated driving, a study by Varhelyi & Kaufmann (2020) showed that test persons who were driving with an automated vehicle on a highway reported that they were feeling more tired than when driving (manually) by themselves. This means that automation renders the takeover process between the system and the driver even more difficult by inducing sleepiness the driver is, which results in delayed reaction times.

Therefore, important questions are: Who must take over the control of single functions or of the entire system at which time – the human or the vehicle? Should the system take over, it nevertheless needs to know about the current fitness of the driver at any given time, in order to carry out a controlled handover to the driver or to initiate emergency measures if needed. It is essential to assess whether a person is inattentive, occupied with other things, sleepy or even sleeping, or if there is some other reason why he/she is not able to take over the control of the function.

Currently, there are already in-vehicle systems that warn of drowsiness. They are based on different technical measurement approaches, e.g. the observation of the steering movements, tracking or eye blinking. Nevertheless, research is required to develop the system’s handover strategy to manual vehicle operation in a way which is timely feasible and acceptable for a smooth and complete operation of the vehicle. Therefore it is necessary to detect and identify the vigilance of human drivers without error. It also seems important that drivers understand a concrete situation to be able to react adequately - which in turn is associated with a certain cognitive lead time.

The data and results described in this article were gathered from the WACHsens project (Assessment of driver fitness for partial automated driving by physiological, behavioural, and camera-based sensors), a joint project of the Human Research, Graz University of Technology, AVL Powertrain UK, and Factum apttec Ventures. The overall aim of this simulator study was to merge data from vegetative vigilance, camera observation and driving behaviour in order to provide a basis for the development of drowsiness detection systems with special focus on automated driving. In this context, driving behaviour observation is

an important source of ground truth data for training and evaluation of detection systems.

This article describes the evaluation of the video data by using a suitably enhanced procedure of driving behaviour monitoring. The results are presented with respect personal characteristics (age, gender, driving experience, experience with assistance systems), subjective assessment of sleepiness, different states (rested/tired) and operating modes (manual/automated driving).

The findings of the study provide additional information about when drivers should be warned of their incapacity to drive, and how detection of weariness and decreasing attention could be controlled in automated vehicles.

2. METHODOLOGY OF THE OBSERVATION PROCEDURE

2.1 General considerations

In the foreground of the evaluation of the video data was the assessment of degree of tiredness of the test persons during simulated tests. There are already several studies that use different methods to distinguish whether a person is tired or not. An easy way is to let the test persons themselves assess how tired they feel at a certain point in time using subjective ratings such as the KSS (Karolinska Sleepiness Scale, see Akerstedt and Gillberg, 1990, Fors (2018)). Another widely used and less subjective method is the observer's assessment of tiredness status. Different rating scales and different time periods have been used for the assessment (see Naujoks et al. (2018), Wierwille and Ellsworth (1994), Ahlstrom et al. (2015)). The most frequently used scales are the Observer Rating of Drowsiness (ORD) developed by Wierwille & Ellsworth (1994) or the Observer Rated Sleepiness (ORS - Anund et al. 2013). Most scales which are assessing the tiredness level of a person focus in particular on the eye and eyelid movements, arm movements, upper body, arms movements and changes in facial expression.

In the WACHsens project, the tiredness has been rated by the test subjects themselves with KSS before and after each drive, and, continuously throughout the drive, by evaluating the tiredness by the observers. The observations scheme was developed on the basis of both scales (ORD and ORS) and was linked to the observation method of the Vienna driving test (Risser and Brandstatter 1985). Therefore, three ob-

servers were first separately coded test rides and afterwards discussed their observation results. This was repeated several times so that in the end the observers all agreed on the definition of each observation variable and the procedure of the observation itself. In contrast to other studies, the observation period or the period of assessment regarding the tiredness status was not limited to a short range e.g. before the occurrence of an event (trigger event) (see Naujoks et al. (2018), Wierwille & Ellsworth (1994)). Extending those procedures, the observation period lasted continuously over the entire driving period (approx. 30 minutes). At any time during the test drive it was possible to determine -on the basis of the observation protocol- in which posture the test person was and what their tiredness status was. When collecting data, it was necessary that all events (activities, changes in posture, changes in the state of tiredness) were annotated with a code so that they could be subsequently used in the next step in order to link and synchronize the observation data with data collected by other methods (physiological measurement such as EEG, eye-tracking and vehicle signals).

2.2 Observer variables

As mentioned, the observation sheet was developed based on the variables of ORD (Wiegand et al., 2009) and ORS (Anund et al., 2013). Different body parts were defined and variables for the following areas were defined as well:

- Positions and movements of the arms
- Hand position and movements of the hands
- Position of the upper body and upper body movements
- Position of the head and head movements
- Facial movements
- Tiredness level
- Control of the vehicle

In the following, a description of each variable:

2.2.1 Positions and movements of the arms. Positions of the hands

In principle, a distinction was made between the left and the right arm when entering the variables. The position of the arms was coded primarily according to a clock. If the test person had positioned his arm at the top of the steering wheel, a 12 was assigned in the observation sheet. Other annotated positions were

when the drivers placed their arms in their lap or took up another position (for example: arm rested on the side window frame). Activities of the right arm were also recorded when the arm was placed on the gear shift or when the hand was touching the display.

Other activities of the arms were annotated when the subjects scratched their face, rubbed their eyes or covered their mouths with their hand (yawning).

2.2.2 Hand position and movements of the hands
Again, a distinction was made between left and right hand. If the test driver enclosed the steering wheel with their hand, the variable “fixed” was annotated. The movements of fingers were monitored too. Also, when the test person shortly let off of the steering wheel but immediately grabbed it again without changing the position of the arms. Finally, “drumming” was registered as well when the subject drummed with his fingers on the steering wheel or on their arms.

2.2.3 Position of the upper body and upper body movements
The upper body posture mainly differentiated between sitting straight, bending forward and leaning to the side. Other observed activities were: short shaking movements of the upper body, stretching or raising or lowering of shoulders. The deep intake of breath of the test subject recognizable by a single strong lifting and lowering of the ribcage was defined as a sighing.

2.2.4 Position of the head and head movements
The basic position of the head posture was defined as “straight” when the subject held his head straight and looked forward. “Leaning”, “Leaning backwards” and “Leaning forward” has been assigned when the subject’s head rested for more than two seconds either on the side (to the left or right), backwards or forwards.

The activities were distinguished when the person made a brief glance (sideways or in the rear-view mirror, on the dashboard display or out of the window). However, this was only noted when the head really moved in this direction. If the test person glanced sideways (left or right with moving his/her head) for more than two seconds, it was coded “looking to the side”. Mere eye movements were not annotated with a code. “Short shaking movement” was assigned when the test subject shook his head briefly (less than two seconds). In contrast, “wiggle” was assigned

when the head moved from left to right for more than two seconds.

2.2.5 Facial movements

If there were no recognizable facial movements of the test person, the variable “expressionless” was assigned. “Yawning” was noted each time the subject opened his mouth and blew air out. “Lip movements”, “Tongue movements” were also annotated as well as short opening of the mouth. Other movements like “talking”, and “coughing” were assigned to other movements of the face.

2.2.6 Degree of tiredness

Based on the ORD and ORS scales, a four-part scale on the tiredness condition of the subject has been developed. The scale was defined as follows:

- 1 = no signs of tiredness: the subject has full control of the vehicle and is fully focused or shows no signs of tiredness
- 2 = signs of tiredness: face or eyes rubbing, restlessness, yawning, change of facial expressions and blinking, or rigid look forward with hardly any movements for a short period of time
- 3 = just before falling asleep: clear signs that the subject is struggling to keep awake. Change in eyelid movements (faster, heavier), subjects can hardly keep their eyes open
- 4 = falls asleep: subjects keep their eyes closed for two seconds or longer, with or without nodding of the head

2.2.7 Other

In addition, the following variables were registered for each separate subject:

- Subject: code number of the subject
- Date: date of test drive
- Ride: coded according to the four possible subject states and driving modes - rested/manual, awake/automated, tired/manual, tired/automated
- Hour/minute/second the exact time of the start of the ride
- Timeframe: video frame when a variable was recorded
- Video_Minute/Video_second: Time in minutes and seconds when a variable was recorded

2.3 Observation procedure

The observation was done while using four different videos which were synchronised and displayed at the same time on one video screen. The observers thus had a comprehensive view of the subject from three different angles. The camera position showed the face (head camera), the upper body, the arms and hands (two side cameras), a fourth camera showed the road from the driver's perspective (front camera). The observers used an observation guide where all variables were listed and described in detail. For the coding, a Microsoft Excel sheet containing all variables has been used. In the Excel sheet every change in posture and other activities was noted. At the beginning of the evaluation, the general data such as serial number, code of the subject, date and time were entered at the head of the sheet.

Once the subject began to drive, a base line for all variables was determined and entered in the first line. As soon as the test person changed his/her body position and performed an activity, the relevant variable was annotated in the second line together with the timeframe and the video minute and second. The third line was filled in when the test person performed another activity or changed his/her body position (again). From the completed sheet it is possible to determine and read each and every position and movement for a given time of the drive.

In the tiredness variable, the observers subjectively assessed the tiredness status of the subject on a four-part scale. When the signs of tiredness became more frequent, the timeframe and the video minute/second were noted in a separate line and the variable tiredness was given the corresponding variable (f.i. 2 = signs of tiredness). If there were no signs of tiredness within three minutes, this was entered in a new line along with the time(min/sec). However, if the test subject's tiredness symptoms have accumulated and they were about to fall asleep, this and the time was annotated (3 = is about to fall asleep). The observers also monitored whether the subject was falling asleep repeatedly in a three-minute period. If this was not the case despite visible signs of tiredness, this was noted in the observation sheet. If the test person fell asleep, no matter if it was a microsleep or a longer sleep phase, this was also noted together with the timeframe and the video minute/ second in a new line. This sometimes happened abruptly without the test subject showing visible signs of tiredness or shortly before falling asleep. As soon as the test person reopened their eyes, the times were entered and, by default, the variable for "is about to fall asleep " was assigned (this was useful

for test persons who fell asleep several times in succession). However if the test persons showed no signs of falling asleep over the next period of time the variable "shortly before falling asleep" was corrected to "signs of tiredness"; this was especially the case with test persons who fell asleep for one second but tried to keep awake through the driving activities.

At the end of the drive, the observers wrote a brief summary of the drive in a separate document describing specific incidents (such as vehicle handling problems, questions to the study supervisor in case of ambiguity, etc.) and, in particular, assessing the subject's tiredness status (when signs of tiredness were noticed, on what basis was the tiredness status noted in the data etc.). Finally, tiredness status was again assessed, based on the whole observation data by the observer on a three-part scale throughout the ride (no or few signs of tiredness - signs of tiredness but no sleep - falls asleep while driving).

3. TEST PROCEDURE

The test persons completed the rides in the simulator at two different days, one time in a rested state and the other time in a tired state. Each time the test person drove one time manually and one time automated. To exclude any test artifacts and effects by the array of the rides, random mixing of the order of subject states and driving modes has been used. 44% of the test persons drove first automated and then manually while 56% of the subjects first drove manually and then automated. In the second test, 42% drove first automated and then manually, while 58% drove first manually and automated in the second part. For the first test drive, 29 subjects were asked to come in a rested state and 21 test persons in a tired state. The tired state was defined as having been awake for at least 16 h continuously and doing the test at usual bed-time, or as sleep deprivation of at least half of their sleeping time the previous night). For the second test the same subjects were asked to come with the exact opposite conditions. Hence, each test person completed four trips. Two rides (automated and manual) in a rested state and two rides (automated and manual) in a tired state.

4. SAMPLE

In total, rides of 50 test persons were evaluated for the present video observation study. For techni-

cal reasons, three videos were defect and could not be played, therefore a total of 197 videos of about 30 minutes were analyzed.

The sample consists of 25 men and 25 women. Approximately the same number of younger (under 40 years, n = 19), middle aged (between 40 and 60 years, n = 15) and older subjects (60 years and older, n = 16) were monitored.

The majority of the test persons (60.9%) had a lot of driving experience with more than 100,000 km driven in total. 14.2% of the test subjects said they had between 50,000 and 100,000 km driven and 17.3% between 10,000 and 50,000 km. Approximately 10% of the test persons had only driven less than 10,000 km in total.

Almost two-thirds of the test subjects already had experience with assistance systems in the car.

5. RESULTS

First and foremost, it was examined whether there were differences in the observational variables according to gender, age and driving experience. In addition, an analysis was made to which extent observer assessment of the subjects' state of tiredness is connected with the self-assessment of the test subjects (KSS NASA TLX). Finally, it was examined whether there are differences between different scenarios (rested or tired state of driving or between manual and automated rides).

5.1 Overall observations

5.1.1 Body movements and activities

Head movements were the most frequent activity of all test persons. This was observed more than 22 times on average during the approximately 30-minute ride. Almost the same number of position changes and activities of the right and left arm were registered (right arm 20,6 times, left arm 18,2 times).

The activities of the left hand (7,7 times) and right hand (7,1 times) were also observed almost equally often. The upper body was moved 8,3 times on average by the test persons and almost 20 times the facial expression changed during the ride as recorded by the observers.

5.1.2 Changes in the tiredness level during the test ride

The observers noted an actual change in the tiredness level as soon as the first signs of tiredness were noticed, the subject was about to fall asleep or fell asleep while driving.

For one third of the rides, no signs of tiredness were registered at all, for about one third at least once it was noted that the subject appeared tired and in a little more than a third this happened more than once. In two-thirds of the rides it was not detected that the subject was about to fall asleep. In five percent of cases, this was noticed once, and in more than a quarter of the trips, subjects were more than once likely to fall asleep.

In almost 70 percent of the rides, the test subjects did not fall asleep. In 10% of the cases it was observed one time that the subject had either a microsleep or a longer sleep phase, and in about one-fifth of the rides test persons were falling asleep more than once (there were extreme cases where the subjects fall asleep more than 50 times within the 30 minute ride).

5.1.3 Assessment of the tiredness level after the test ride

After the ride, the observers assessed once again the tiredness levels of the subjects throughout the ride. A distinction was made as to whether or not there were any signs of tiredness being observed, if more signs of tiredness were observed (if the subject was tired for a long time but did not fall asleep) or if the subject fell asleep while driving. In approximately one-quarter of the rides the subjects were tired but did not fall asleep or the subjects fell asleep at least once.

Table 1: Overall tiredness level during the test ride

	Signs of tiredness		Shortly before falling asleep		Falling asleep	
	Count	%	Count	%	Count	%
Not observed	66	33,5	131	66,5	136	69,0
Observed once	61	31,0	11	5,6	21	10,7
Observed more than once	70	35,5	55	27,9	40	20,3
	197	100	197	100	197	100

Table 2: Assessment by the observers of the tiredness level after the test ride

	Count	%
None or hardly any signs of tiredness	99	50,3
Signs of tiredness but no sleep	46	23,4
Fell asleep	52	26,4
Total	197	100

The comparison of the observers' assessment with the different scenarios (rested – automated/manual, tired – automated/manual), showed that, especially in the case of manual driving, “no or only a few signs of tiredness” were observed frequently. This was the case for around three quarters of the rested/manual and over 50% for the tired / manual rides. On the other hand, this was registered “only” in one-third for the rested/automated and in quarter of the tired/automated drives.

The observers categorized 40% of the rides as having “signs of tiredness but no sleep” if the scenario was rested/automated. As for the other scenarios, this occurred at about one fifth of the rides. The test persons most often fell asleep when they were driving in the tired/automated scenario. Similarly, in a quarter of the trips, the subjects fell asleep in the scenario rested/automated and tired/manual. Sleeping test subjects could only be observed in 6% of tired/manual rides.

Comparing the settings of rested and tired scenarios it appears that in more than half of the rested rides “no or few signs of tiredness” were observed. However, nearly 30% of the participants “showed signs of tiredness but did not fall asleep”, and 13.8% of the test persons were falling asleep. During the

tired rides, the same number of rides (40%) were categorized as “none or hardly any signs of tiredness” or “asleep while driving”, while a fifth of the subjects showed “signs of tiredness but did not fall asleep”.

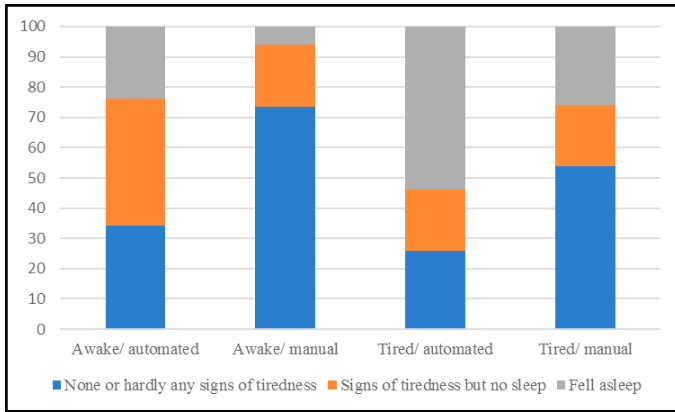
When comparing the automated and manual drives, “none or hardly any signs of tiredness” (63.6%) were seen during the manual scenario. In one-fifth of manual rides, subjects showed “signs of tiredness,” and in 16% of cases, subjects fell asleep during a manual ride. In comparison, 40.9% of the test person fell asleep while driving automated and 29.5% each showed either “none or hardly any signs of tiredness” or “signs of tiredness”.

5.2 Comparison between observer and test person's tiredness assessment

In order to compare the assessment of the tiredness status in the course of the entire test ride by the observers with the subjective tiredness assessment by the test persons after the trip, the assigned tiredness annotation was compared with the data of the KSS. It appears that the assessments between the respective categories, which the observers coded, differ significantly with regard to the KSS score

Table 3: Assessment by the observers of the tiredness level after the test ride for different scenarios in percentages

	Rested/automated	Rested/manual	Tired/automated	Tired/manual
None or hardly any signs of tiredness	34,2	73,5	26,0	54,0
Signs of tiredness but no sleep	42,1	20,4	20,0	20,0
Fell asleep	23,7	6,1	54,0	26,0
Total	100	100	100	100
	Rested	Tired	Automated	Manual
None or hardly any signs of tiredness	56,3	40,0	29,5	63,6
Signs of tiredness but no sleep	29,9	20,0	29,5	20,2
Fell asleep	13,8	40,0	40,9	16,2
Total	100	100	100	100



Graphic 1: Assessment of the tiredness status after the test ride for different scenarios

of the test subjects. Consequently, those who fell asleep while driving also reported a significantly higher degree of tiredness level after the drive than those who showed only signs of tiredness but did not fall asleep or who had none or hardly any signs of tiredness. Also, subjects who showed signs of tiredness self-assessed themselves as being tired significantly more often after the drive compared to the test persons who showed none or hardly any signs of tiredness.

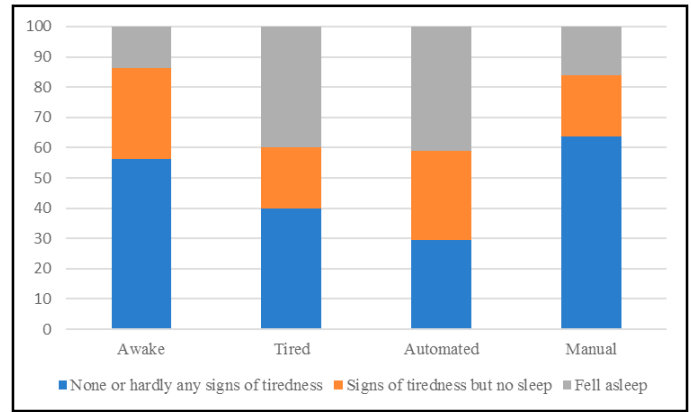
Table 4: Descriptive statistic of the assessment by the observers of the tiredness level after the test ride / KSS-Scores by the test persons and results from ONEWAYANOVA

	N	average	significant
None or hardly any signs of tiredness	99	5,23	
Signs of tiredness but no sleep	46	6,43	0,000
Fell asleep	51	8,06	
Total	196	6,25	

5.3 Comparison of changes in body position and activities

Subsequently it was calculated if there were any significant differences between gender, age, driving experience and experiences with assistance systems regarding the changes of the position of the body or body activities.

There were no significant differences in body posture or activity between the 25 men and 25 women. Also, the activities did not differ significantly from each other.



Graphic 2: Assessment of the tiredness status after the test ride with different preconditions

Between the age groups (under 40 years, between 40 and 60 years and 60 years and older), the only a significant difference was found in the facial expressions, with significantly more facial movements in the group of over 60 years in comparison to the other age groups.

There were several significant differences between the groups with different driving experience. It appears that the test persons with the shortest driving experience changed significantly more often the position or activities of the left arm, the left and right hand, the upper body and the head and also in the sum of all position changes and activities in comparison to the test persons with more driving experiences.

There were no significant differences in body posture and activity between the subjects who already had experience with assistance systems and those who had no previous experience.

In a further step, a comparison was made between the change of body position and activities with the observer’s assessment of tiredness status while driving to see if there were any differences between subjects, e.g. while falling asleep and those who showed no signs of tiredness.

Significant differences that appeared between the groups with different degrees of tiredness were expressed in the number of position changes and activities of the left and right arm, upper body, head and facial movements, and in the sum of all activities. Here, in all the characteristics, the group of those subjects who showed “none or hardly any signs of tiredness” differed from those with “signs of tiredness”, while those who “fell asleep while driving” in the sense that the group, with “none or hardly any signs of tiredness” showed significantly less activities than the other two groups.

5.4 Comparison of the different tiredness levels during and after the test ride

Regarding the changes in tiredness level during the test ride as well as the general assessment of the state of tiredness after the entire journey, no significant differences could be found between men and women. As a result, men and women showed symptoms of tiredness and were similarly close to falling asleep or falling asleep.

The observed changes in the tiredness level and the overall tiredness assessment after the test ride did not reveal any significant differences between age groups, between the groups with different driving experience and the groups with and without experience with assistance systems.

5.5 First signs of tiredness

In a further evaluation step, it was examined whether there were differences between the personality traits

of the test persons or between the different test scenarios regarding the first time at which the observers noticed signs of tiredness.

There were no significant differences between men and women, different age groups, groups with different driving experience, and experience with assistance systems between these times when the observers registered the first signs of tiredness. All three groups show these signs between the eleventh and thirteenth minute of driving (with the exception of the subjects with the shortest driving experience who showed signs of tiredness after just under nine minutes).

Comparing the scenarios in which the test persons drove rested or tired, it can be seen that in automated drives the test persons showed signs of tiredness significantly earlier (between five and six minutes) than when they were driving manually.

There is another significant difference between the different preconditions - rested and tired re-

Table 5: First signs of tiredness and results from ONEWAY ANOVA

		N	Average in minutes and seconds	Significant
Gender	Men	66	11:25	0,302
	Women	67	12:54	
Age	Under 40 years	54	12:30	0,693
	Between 40 and 60 years	44	12:36	
	60 years and older	35	11:08	
Driving experience in km	0-10k	10	08:43	0,582
	10-50k	26	12:01	
	50-100k	16	12:42	
	over 100k	81	12:32	
Experience with assistant systems	No	47	11:31	0,507
	Yes	86	12:31	
Scenarios	Rested/automated	28	12:15	0,000
	Rested/manual	21	18:23	
	Tired/automated	43	08:17	
	Tired/ manual	41	13:00	
Rested/tired drive	Tired	84	10:35	0,004
	Rested	49	14:53	
automated/manual drive	Manual	62	14:49	0,000
	Automated	71	09:51	
Assessment of the tiredness level after the test ride	None or hardly any signs of tiredness	35	20:04	0,000
	Signs of tiredness but no sleep	46	10:36	
	Fell asleep	52	08:15	
	Total	133	12:10	

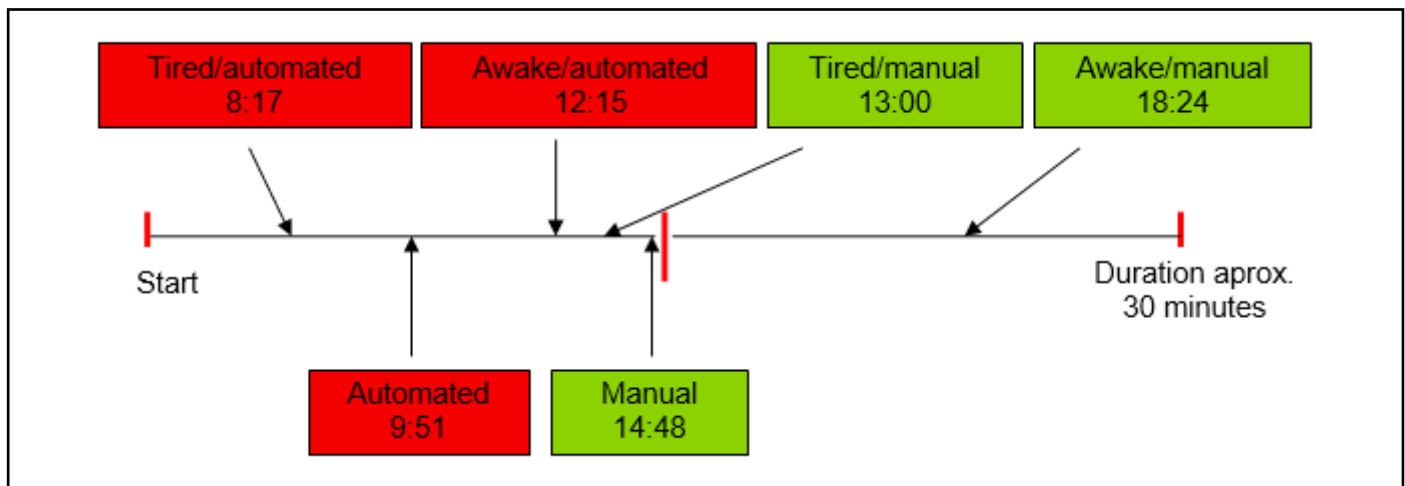


Figure 1: Average time when first signs of tiredness were noted in different scenarios

spectively manual and automated drives. While the first signs of tiredness during the rested rides were registered after about 15 minutes, this was already the case during the tired drives after just under ten and a half minutes. Even though the subjects were driving automatically they were significantly more likely to show signs of tiredness (at about minute 10) than if they were driving manually (at about minute 15).

Test subjects who fell asleep while driving showed significant early signs of tiredness (already after eight minutes and 15 seconds) compared to subjects who were assessed to be tired (first signs after ten and a half minutes) and those with hardly any or no signs of tiredness (if any signs then after more than 20 minutes).

6. SUMMARY AND DISCUSSION

197 simulator test drives under different test conditions were analyzed by observers. Four different scenarios were evaluated for all test persons. Tired/manual, tired/automated, rested/manual, and rested/automated.

The first aim of the observation analysis was to see if differences in body movements and activities exist:

- between the four different scenarios,
- according to personality traits (age, gender),
- between people with little and much driving experience and
- between persons with and without experience with assistance systems

The second aim was the classification of the subjects' state of drowsiness. To achieve this goal, a comprehensive observational scheme was developed to register every single of body, head and facial movements and activities as well as a continuous rating of the subject's state of drowsiness. The categorization of the state of drowsiness was carried out in two ways: on the one hand continuously during the entire drive and on the other hand by means of an overall assessment of the behaviour of the test persons after each individual test drive. Using the detailed observation data, it was possible to determine for each time point, which posture the subject was currently adopting, which activities were performed and in what state of drowsiness the subject was at that time. These results are useful for a better understanding of driving behaviour, but have also been used as training input for the development of drowsiness detection systems (Arefnezhad 2020a, Arefnezhad 2020b).

The sample of test persons is balanced by gender and age, with most of the subjects having more than 100,000 km of driving experience and the experience with assistance systems. It was found in the evaluation of all rides that in a quarter of the trips, the subjects fell asleep and in another quarter signs of tiredness were observed. Given that the level of sleep deprivation for the tired state tests was only moderate, this number seems quite high. In this simulator test, the test persons might have felt safe to allow themselves to relax more than they would have done in road tests. It remains an open question, how this numbers translate to real road driving tests.

The analysis of the data further shows that the observers' observations and categorizations regarding the subjects' state of tiredness are in line with

the subjective assessment of the subjects' state of tiredness. Those who gave a high value to the self-assessment using the KSS test were also identified by the observers as those who had significant problems staying awake or even falling asleep. Those who had least reported signs of tiredness were also rated by the observers as showing some signs of tiredness but still had no problems with falling asleep. And finally, those subjects who had no problems with tiredness were also recognized by the observer as such and rated accordingly.

Regarding the different scenarios, it can be said that most signs of tiredness or situations in which the subjects fell asleep were observed during the tired/automated ride (in over three quarters of these rides). But even during the rested/automated drive, more than 40% of the subject showed signs of tiredness, and in less than a fifth of the cases, subjects also fell asleep. This percentage is even higher than in the rides where the subjects were tired but controlled the vehicle manually (signs of tiredness and falling asleep in almost 50% of the cases). During the rested/manual rides, however, no signs of tiredness were observed in nearly three quarters of the cases. Thus, it turns out that when driving in automated mode drivers were more sleepy than in manual mode, either by a higher level of boredom or by a tendency to effectively permit to be tired (starting to sink away/to fall asleep) which seems logical as there was no urgent need to stay awake resp. to perform the driving task. In comparison, if the test person were driving manually, even if the test person started to drive already tired, they showed less often signs of tiredness.

Body movements and activities have been recorded for different body parts. Movements and activities of the arms, hands, upper body, head and facial expressions were analysed. Basically, on average, most movements and activities of the head and face were noted. This is followed by a change of position of the arms (more often movements of the right than of the left arm, which could be related to the fact that most people are right-handed). Far fewer movements, and only about half as often, were movements and activities of the upper body and hands.

There were no significant differences between the personality traits (gender, age) regarding the number of body movements (position changes and activities). As a result, men and women, and younger and older subjects moved or operated similarly during the different rides. Also, no significant differences were

found between the subjects who already had experience with assistance systems and those who had not. From this it can be concluded that the automated scenario or the handling of the display for activating the automated system alone has no influence on the subject when it comes to body movements.

Differences in the number of body movements were detected according to the driving experience. The test persons with the shortest driving experience had significantly more left-arm activities, left- and right-hand movements, upper-body and head movements and activities and activities in total than the test persons with more driving experience. There might be two possible explanations. First explanation would be that the test persons with more driving experience were also better able to cope with the test situation while test persons with less driving experience were more uneasy about the situation in the driving simulator. The other explanation would be that drivers with more body movements (thus: the drivers with less driving experience in this study) are trying to cope with their condition of tiredness and are struggling to stay awake with the help of body movements. Nevertheless, there is the question if less body movement is a real advantage while driving automatically. In cases where the driver must take over from the system - and especially in moments when the driver's attention and immediate action is requested without previous warning (e.g. in case of urgency) - a person with more body movements might have stayed more alert and cautious and therefore could easier take over from the system. On the other hand, a driver without body movements might not be aware of the situation and therefore would have more problems to take over from the system. In any case this would need further and more in-depth research.

There were also no significant differences between the personality traits (gender, age) or the groups with different driving experience, and the groups with and without experience with assistance systems with regard to the change in the state of tiredness during the drive. It turns out, that these features had no effect on the change in the state of tiredness.

This also applies to the point in time when the first signs of tiredness were registered by the observers. Men and women, younger and older, those with more and less driving experience and those with and without experience with assistance systems show first signs of tiredness at about the same time which is between the ninth and twelfth minute on average.

On the other hand, there is a significant difference between the different scenarios and the comparisons between the tired/rested rides and the manual/automated rides, each in the direction that the automated rides and/or when the test persons had the precondition of driving while tired, first signs of tiredness were registered significantly earlier than in the rides where the subjects were driving manually and/or rested. This shows once again that especially automated driving either induces more sleepiness by being boring, or to induce the tendency to permit the feeling of tiredness and to accept/admit the loss of control.

This study shows that there are still a lot of questions to be answered and research to be carried out before introducing automated driving in real life. The most interesting result was that automated driving, regardless of the tiredness status, leads to the fact that drivers were showing earlier signs of drowsiness when compared to the drives where the test persons were driving manually. This might be crucial for the system on a SAE3 level where the driver acts as a fallback level. In this level of automation, the driver is expected to act more like a pilot who must check the system periodically if it is working correctly as he/she must assume that a takeover request could come from the system anytime. That necessity somehow contradicts one of the proposed advantages of automated driving, namely that the driver would be free to do other things. But this continuous monitoring of the system might also induce and increase the feeling of tiredness, because it is a generally boring task. It is also proven that tiredness leads to a longer reaction time. This might not be of importance in situations where the system can warn the driver several seconds before a planned take over of control (end of the drive, low fuel, exit from the highway etc.), but it surely will make the takeover in a critical situation more difficult (system failure when for instance the system cannot detect the lane, misinterprets the behaviour of other road users etc.). In such situations the driver should -as quick as possible- be able to correctly read and assess the situation and understand that he/she must take over from the system. This task could take up to some seconds which might be too long in a crucial situation. So, the question still remains, if a driver who is tired or preoccupied with secondary tasks is really the best response for a fallback level and whether he/she would be able to react correctly in critical situation. Another question remains - who bears the legal responsibility in a situation in which the driver should takeover but is not

mentally or physically capable to react correctly. Will drivers take the risk and trust both the system and themselves that they will always do the right thing? This is also related to another issue, namely personal attitudes, which this study did not deal with (but the issue has been already examined for example in: Paris & Van den Broucke 2007). No distinction between persons with high safety attitudes compared to persons with low safety attitudes could be found in the study and also, effects such as an overall trust in the system (“the system will react correctly for me”) could not be detected. Further research is needed to address these critical issues of system-driver interaction in the near future.

Acknowledgements: The WACHsens project was carried out by Human Research Institute of Health Technology and Prevention Research GmbH, Graz University of Technology, AVL Powertrain UK Limited, and Factum apptec Ventures GmbH. It was co-funded by the Austrian ministry of Mobility, Innovation and Technology via the program mobility of the future (Austrian Research Promotion Agency FFG grant no. 860875). The authors are indebted to all drivers who voluntarily participated in the experiment, to the experimenters, and the many people who helped to set up the tests.

LITERATURE

- Ahlstrom, C., Fors, C., Anund, A. et al. *Eur. Transp. Res. Rev.* (2015) 7: 38. <https://doi.org/10.1007/s12544-015-0188-y>
- Akerstedt, T. and Gillberg, M. Subjective and objective sleepiness in the active individual. *Int. J. Neurosci.*, 1990, 52: 29–37.
- Anund A, Fors C, Hallvig D, Akerstedt T, Kecklund G (2013) Observer rated sleepiness and real road driving: an explorative study. *PLoS One* 8(5):e64782
- Arefnezhad S, Samiee S., Eichberger A., Frühwirth M., Klotz E. (2020a) Applying Deep Neural Networks for Multi-level Classification of Driver Drowsiness Using Vehicle-based Measures. *Expert Systems with Applications*, 162, [113778]. <https://doi.org/10.1016/j.eswa.2020.113778>
- Arefnezhad S., Eichberger A., Frühwirth M., Kaufmann C., & Moser M. (Accepted/In press). Driver Drowsiness Classification Using Data Fusion of Vehicle-based Measures and ECG Signals. Paper presented at IEEE SMC 2020; IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN, AND CYBERNETICS, Toronto, Canada.
- Bainbridge, L. (1983) Ironies of automation. *Automatica* 1983, 19(6):775-779.

Chaloupka C, Risser R, Antoniadou A, Lehner U, Praschl M (1998) Abschätzung von reaktiver Anpassung an fahrzeugtechnische Veränderungen. In: BASt, Heft M 84. Bergisch Gladbach; 1998.

Fors C., Ahlstrom C. & Anund A. (2018) A comparison of driver sleepiness in the simulator and on the real road, *Journal of Transportation Safety & Security*, 10:1-2,72-87, DOI: [10.1080/19439962.2016.1228092](https://doi.org/10.1080/19439962.2016.1228092)

Moller HJ, Kayumov L, Bulmash EL, Nhan J, Shapiro CM (2006) Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness. *J Psychosom Res* 2006, 61(3):335-342.

Naujoks, F., Höfling, S., Purucker, C., & Zeeb, K. (2018). From partial and high automation to manual driving: Relationship between non-driving related tasks, drowsiness and take-over performance. *Accident; Analysis and Prevention*, 121, 28-42.

Paris H., Van den Broucke S. (2008). Measuring cognitive determinants of speeding: An application of the theory of planned behaviour. *Transportation Research Part F* 11 (2008) 168–180

Risser, R., Brandstatter, C., (1985). *Die Wiener Fahrprobe. Freie Beobachtung.* Literas Universitätsverlag, Wien.

Royal Society for the Prevention of Accidents (RoSPA). (2001). *Driver fatigue and road accidents: A literature review and position paper.* Royal Society for the Prevention of Accidents

Várhelyi A., Kaufmann C., Johnsson C. & Almqvist S. (2020) Driving with and without automation on the motorway – an observational study, *Journal of Intelligent Transportation Systems*, DOI: [10.1080/15472450.2020.1738230](https://doi.org/10.1080/15472450.2020.1738230)

Wierwille, W.W. & Ellsworth, L. A. (1994). Evaluation of driver drowsiness by trained observers. *Accident Analysis and Prevention*, 26(5), 571-581.

Wiegand, D. M. , Hanowski, R. J., McDonald, S. E., Virginia Polytechnic Institute and State University., & National Surface Transportation Safety Center for Excellence. (2009). *Development and Evaluation of a Naturalistic Observer Rating of Drowsiness Protocol Final report.* Blacksburg, Va: Virginia Tech Transportation Institute.