Driver behavior and workload in an on-road automated vehicle

Stapel, Jork; Mullakkal Babu, Freddy; Happee, Riender

Publication date
2017

Document Version
Final published version

Published in
Proceedings Road Safety & Simulation International Conference 2017

Driver Behavior and Workload in an On-road Automated Vehicle

Jork Stapel, Freddy Antony Mullakkal-Babu, Riender Happee

Cognitive Robotics, Faculty of Mechanical Engineering, Delft University of Technology, J.C.J.Stapel@tudelft.nl
Transportation and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, F.A.MullakkalBabu@tudelft.nl
Cognitive Robotics, Faculty of Mechanical Engineering, Delft University of Technology, R.Happee@tudelft.nl

Abstract

Driver mental underload is an important concern in the operational safety of automated driving. In this study, workload was evaluated subjectively (NASA RTLX) and objectively (auditory detection-response task) on Dutch public highways (~150km) in a Tesla Model S comparing manual and supervised automated driving with moderators automation experience and traffic complexity. Participants (N=16) were either automation-inexperienced drivers or automation-experienced Tesla owners. Complexity ranged from an engaging environment with a road geometry stimulating continuous traffic interaction, and a monotonic environment with lower traffic density and a simple road geometry. Perceived and objective workload increased with traffic complexity. Automation use reduced perceived workload in both environments for automation-experienced drivers, but not for inexperienced drivers. However, the DRT did not reveal a reduced attentional demand with automation. This suggests that attentive monitoring requires a similar attentional demand as manual driving. The findings highlight the relevance of using system-experienced participants and the relevance of on-road testing for behavioral validity.

Keywords

Automated Driving; On-road; Workload; Experience; Underload

1. Introduction

Driver mental underload is an important concern in the operational safety of conditional automation. When automation relieves the driver from the continuous control tasks, mental underload can occur. [1] Over time, this can lead to a state of drowsiness, inattention and slower reactions. [2] This reduces the driver's ability to monitor the conditional automation and impairs his/her performance to intervene when necessary, leading to potential safety critical situations.

In order to address these effects, it is important to know how workload is affected by the use of automation, and how this effect changes between various conditions. Since workload is a broad construct with many influencing variables, this study focuses on two main moderating variables: the complexity of the driving environment and the driver's experience with the automation. Understanding the effect of these moderators can help to predict in which conditions underload is most likely to occur. Experience with driving automation for instance can lead to task execution at a lower cognitive level, or reduce the perceived complexity of the traffic situation. [3,4] It can also lead to better monitoring and improved cognitive readiness for familiar driving situations, resulting in higher control transition performance [3,4,5,6,7]. Moreover, automation experience may influence workload differently in high and low traffic complexity.

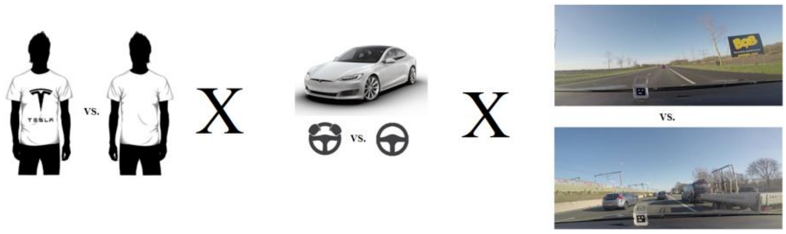

Figure 1: Illustration of the independent variables: automation experience, automation use, and complexity of the environment.

This study aims to establish how the workload under automated driving is moderated by traffic complexity and by the driver's prior experience with automated driving. Towards this, we conducted an on-road experiment which was carried out on Dutch public highways in a Tesla Model S. The change in workload was assessed subjectively as well as objectively. Traffic complexity was moderated by driving in a monotonic, low workload highway environment and on a complex, engaging highway. To moderate automation experience, participants were either automation-inexperienced drivers or automation-experienced Tesla owners. The conditions were driven both manually and with automation. This resulted in a 2(automation: on vs. off) x 2(environment:

monotonic vs. engaging) x 2(experience: experienced vs. inexperienced) factorial design as illustrated in Figure 1.

1.1. Preceding work

A considerable amount of studies have reported on the effects of automation on driver workload. In 2014, an empirical review summarized workload findings from 32 studies comparing different levels of automation. [8] Studies were mainly performed in simulators, and indicated a workload reduction of 21% on average from manual to automated driving. Ratings ranged from 23% to 66% for manual and from 11% to 40% for automated driving.

The influence of traffic complexity on workload has been studied extensively. (e.g. [1,2,9,10,11,12,13]). It can be as large as the use of driving automation, with a 35% workload increase from low to high traffic complexity in manual driving. [9] During conditional automated driving, traffic increases the demands for the monitoring task. This in turn can reduce the development of drowsiness. A significant drowsiness increase and attention decrease were found during conditional automated driving, [10] but the effect was smaller for higher traffic densities.

While task complexity increases workload, experience with automation may reduce it. Young [4,14] performed several studies comparing the effect of automation levels on workload for drivers with different levels of driving skill. Although both factors (automation and manual driving experience) significantly improved secondary task performance, the effect of automation was stronger than that of driving experience.

Until recently, the influence of experience with automation could hardly be investigated due to the unavailability of automation-experienced drivers. Simulator studies on workload in automation often include a familiarization period, but the exposure times are too short for the development of experience. [15,7] Some studies have approximated automated driving experience by using ACC experienced drivers [16,5,17] or developed special procedures to create experience through training, [7] but no studies were found that moderated conditional automation experience with actual users of automated driving. Some effects of experience, such as the perceived risk and trust, may also be hard to study in simulators since they pose limitations on the behavioral validity. [18,19,20,21,22]

Some recent studies involving on-road evaluation of mental workload in automated driving can serve as a basis to hypothesize results. Here, we will briefly elaborate on these studies.

Banks and Stanton [23] studied the workload of automation-inexperienced drivers during a short but engaging trip in a prototype conditionally automated vehicle. In contrast to findings from simulators, the perceived workload was higher during automated driving compared to manual driving. The participants' lack of prior training with the system, the high number of tasks and reported issues with the automation's behavior may all have contributed to the perceived workload increase.

Heikoop [24] performed a field test with professional drivers familiar to supercars, but with no prior experience with lateral automation in a Tesla Model S on the highway, following a lead vehicle after 30 minutes of test-track training. A simple secondary task (counting bridges) was performed during part of the trip. The perceived workload during automated driving was rated very low overall (average of 19%), which is even below findings from simulator literature and reduced over time, suggesting that accustomization occurred during the trip. Accordingly, signs of disengagement were reported from the stress-state questionnaire.

Eriksson [25] investigated the transition time in non-critical control transitions on the road in a Tesla Model S and compared it to a simulator study. Participants of the on-road experiment had prior experience with driving automation while the participants of the simulator study did not. Drivers in the on-road experiment regained control 32% (1.5 seconds) faster on average compared to the simulator drivers. The workload was perceived as low in both studies and no significant difference was found between the two studies, suggesting simulator validity for workload assessment in engaged driving with moderately automation-experienced drivers.

Naujoks [17] performed a field study measuring secondary task uptake, secondary task workload and compensatory behavior in congested traffic for driving manually, with ACC and ACC+ steer assist in a Mercedes-Benz E-Class. They explored the effect of automation experience by comparing drivers with and without prior ACC experience. They found that ACC experienced drivers perform more secondary tasks in automated driving than in manual driving, and when driving at lower speeds, suggesting reduced workload with automation at lower driving speeds. [4] The effect however was not present for ACC inexperienced drivers, suggesting that automation experience is a prerequisite for freeing cognitive resources for secondary tasks.

From the observations of preceding work the following hypothesis have been formulated for supervised automation:

H1. Automation will reduce workload (H1a), and this effect will be larger in the engaging condition (complex traffic) than in the monotonic condition (H1b).
H2. Workload will be higher in the engaging condition than in the monotonic condition for both manual and automated driving.
H3. Workload will be higher for automation-inexperienced drivers compared to the automation-experienced drivers in all automated driving conditions.

2. Methodology

2.1. Participants

Two groups (N=8 each) of participants took part in this experiment and were selected through convenience sampling. Automation-experienced Tesla owners, referred to as the experienced group, were recruited through the Dutch/Belgium section of the Tesla Motors forum. [26] Seven reported using a Tesla and its autopilot on a daily basis. One was an irregular user but reported 10,000 km traveled using autopilot. One of the experts was the safety instructor, who had observed 8 participants prior to taking part himself.

The automation-inexperienced participants were invited through the universities' employee mailing list and through a list of drivers who had indicated their interest to participate in research regarding automated driving. Inexperienced drivers were required to not have experienced driving automation before. Users of adaptive cruise control were excluded but users of regular cruise control were included. The demographics of both groups are summarized in table 1.

Table 1: Demographics of the two participant groups, with mean μ, standard deviation σ and [interval].

	Experienced group	Inexperienced group
Age	μ=43 σ=14 [27-69]	μ=41 σ=14 [21-61]
Years licensed	μ=22 σ=15 [4-51]	μ=21 σ=15 [3-43]
km driven past 12 months	μ=26.500 σ=21.500 [7.500-75.000]	μ=15.000 σ=13.000 [3.000-42.500]
gender	7 male, 1 female	8 male

2.2. Vehicle and instrumentation

The driving was performed with a rented Tesla model S 75D equipped with autopilot (hardware version 1; update 8.0) and the driver's seat on the left side. An overview of the instrumentation can be seen in Figure . Video was recorded with 3 GoPro cameras observing the traffic in front and behind of the car, as well as the driver. A webcam observed the instrument panel. Traffic flow and speed were logged off-line from the NDW open data server. [27]

An auditory detection response task (DRT) was performed as an objective measure of attentional demand of the driver task. The DRT was implemented in Python 3 on a Raspberry PI 3B running Raspbian Jessie. The DRT implementation and analysis were in line with ISO 17488:2016 [28], apart from the following notable exceptions:

- An auditory stimulus was provided randomly with an on-set interval of 3-5 seconds with a 2.3 kHz tone lasting 1s, irrespective of response time.
- Stimuli were presented over 5 minutes at a time (amounting to 72 stimuli per participant per condition), or until the participant left the considered road section.
- The button used to respond to the stimuli was strapped to the participant's right index finger, as the right hand had no driving related tasks other than steering during the DRT.
- Responses during special events like lane changes were not omitted from the analysis, as these events were considered to be part of the condition under observation.
- The DRT instruction was phrased as "Press the button as soon as you hear the signal, but keep your attention on the road".

Heart activity was recorded using an optical sensor mounted to the participant's right middle finger, powered by an Atmel AtMega328P embedded processor board as used in [29]. Vehicle motion was recorded using a MPU6050 IMU and GTPA013 GPS sensor connected to a second Atmel processor, all of which was kindly provided by the first author of [29].

Eye tracking was recorded with a head-mounted monocular eye tracker from pupil labs at 30 Hz. Markers were placed in the vehicle to track gazes relative to the vehicle's interior and to define regions of interest.

A safety instructor sat next to the participant and was proficient in the use of the autopilot and experienced in introducing new drivers to the vehicle. During the drive, his tasks were to inform or warn the driver when needed, to help with the navigation and vehicle settings from the center console and to provide answers to technical questions. He was also allowed to engage in idle conversations except when instructions were given by the experimenter or during the DRT. The participant was allowed to initiate a conversation at any time.

This paper focusses on the DRT and the subjective evaluation described below. Video footage, heart rate, vehicle motion, eye tracking and traffic information can be part of future publications.

a: Eye tracker
b: DRT button + heart activity sensor
c: Webcam
d: GPS antenna
e: Experimenters
f: IMU + DRT

Figure 2: Overview of the instrumentation

2.3. Subjective measures

A pre-drive questionnaire included questions regarding the participant's preparations for the experiment (e.g. read user manual, related articles, videos) and regarding physical state (amount of sleep, coffee consumption).

Three on-line questions were asked while driving, to which the participant responded with a verbal rating on a scale from 1 to 10. [26] The first question covered mental demand. The second regarded alertness and adopted the descriptions of the Karolinska sleepiness scale (KSS). Each time, the participant was reminded of the description of the given response and was permitted to revise the response accordingly. The third question was on trust and reflected the driver's trust in the automation, as well as his/her own driving abilities. In all 4 conditions the three questions were asked twice (before the DRT and after the DRT).

After the familiarization and after all 4 conditions, the vehicle was driven back to the starting location where the NASA Raw Task Load Index (RTLX) was filled out for each condition as well as for the familiarization. This questionnaire further included questions regarding confidence during manual and automated driving, which were answered for the two environments; and questions regarding trust in the automation (based on [30] and [15]). Open-ended questions were asked while driving between the test conditions. The questions covered usage and preferences of the automation.

2.4. Environment

Two highway sections were selected to conduct measurements in two levels of driving complexity; an engaging environment with a road geometry stimulating continuous traffic interaction, and a monotonic environment with lower traffic density and a simple road geometry and a low chance for high-attention scenarios to occur. (Figure 2)

For the engaging environment, the A10 (ring-East of Amsterdam) was selected for its high traffic density throughout the day and the 10-13 on/off-ramps (depending on direction traveled). To maximize the traffic interaction, the driver was instructed to drive in the right lane as much as possible and was allowed to overtake slow moving traffic. The shoulder lane was not used to avoid unpredictable behavior of the autopilot. The A10 was entered from the A1 and followed down till exit Oud Zuid, either driving manually or using the automation. The highway was then followed in the opposite direction until exit Zeeburg, during which the DRT was performed. The route was then repeated but with the remaining mode of automation use.

For the monotonic environment the A6 between Almere (exit 7) and Lelystad (exit 10) was selected, which is a straight two-lane highway with low traffic density and no on/off ramps between the two cities. Here the driver was instructed to remain in the right most lane, to not overtake slow traffic and to drive as fast as traffic permits, but not faster than 110 km/h. The drivers got stuck behind a truck or trailer driving ~90 km/h in 80% of the monotonic scenarios. The automation was either used on the way towards Lelystad or back towards Almere. The DRT was performed in the second half of each condition.

The two driving environments were located 15 minutes away from one another. The A1 connects the two locations and was used for the familiarization. The A1 was entered from the A9 and first traveled in eastern direction. When the familiarization was to be followed by the engaging environment in Amsterdam, the first available exit was taken before practicing the automated lane change, but no later than Naarden. When the familiarization was to be followed by the monotonic condition, the road was simply continued towards the A6. The order of visiting monotonic/engaging after the familiarization was counter balanced. The driven route is illustrated in Figure 2. For the inexperienced driver, the total trip lasted for 1.5h when first driving to the monotonic condition or 1.75h when first driving to the engaging condition. The experienced driver needed 1.5h for either route due to the shorter familiarization.

The McDonald's Amsterdam Zuidoost was selected as the start/end point of the route, as it was logistically located between the highway entrance and the Tesla supercharger, and because it provided the facilities needed for welcoming the participants.

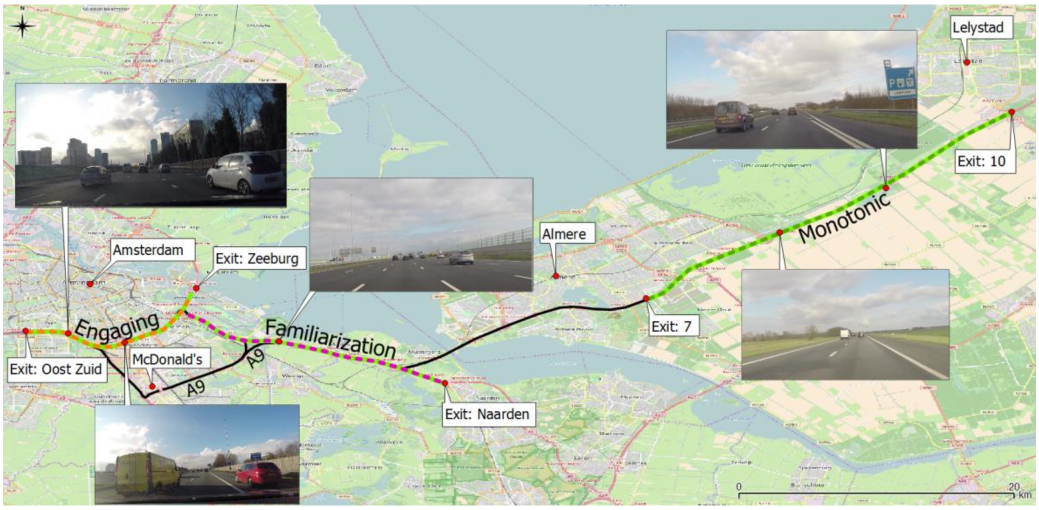

Figure 2: The driven route. Images were recorded during the drive of participant 5.

2.5. Procedure

Upon arrival, the participants were informed of the tasks and risks of the experiment. A pre-drive questionnaire was filled out and the procedures were explained. Prior to departure, the safety instructor informed the participants on the operation and limitations of the vehicle and the automation while the experimenter positioned the eye tracker, heart rate sensor, and DRT button. The participants were further instructed to remain attentive drivers at all times.

Once on the highway, a familiarization drive was performed, during which the participants were introduced to the general operation of the vehicle and the basic behavior of the automation. The performed tasks covered the different methods of activation and deactivation of the automation, the adjustment of the cruise speed setting and the automated lane change. Questions were asked to make sure that the driver understood the instrument panel and operation of the vehicle. The familiarization lasted as long as necessary to let the participant perform each task successfully once. The inexperienced drivers needed 20 minutes while the experienced drivers required 8 minutes on average. After this the three on-line questions were asked and the DRT was performed while using the automation. The participant then drove manually to the engaging or monotonic environments, where they performed 4 rides in manual/automated and engaging/monotonic conditions in a randomized order.

Each condition started with the instructions regarding driving behavior, followed by 5 minutes of driving. The 3 questions regarding mental demand, alertness and trust were asked, after which participants drove another 5 minutes while performing the DRT. The 3 questions were asked once more before driving to the next condition.

3. Analysis and Results

The drives were performed on workdays between the 3rd and 10th of March 2017. All drives were performed between 9:00 and 16:45 to avoid rush hours. No incidents occurred, but one of the inexperienced drivers merged very close ahead of a truck in the engaging condition while driving manually. All tests took place in normal weather conditions, except for two automation-experienced drivers, who drove in heavy rain. The automation operated reliably during these drives (road markings remained clear on the pervious concrete), but maintained a larger following distance despite using the same headway setting (setting 3 for all participants and all conditions).

Two experienced drivers did not follow all instructions during the monotonic condition. One occasionally overtook trucks despite being instructed not to do so. The other refused to turn off ACC during manual driving.

3.1. NASA RTLX

Subjective workload ratings were collected using the NASA RTLX. Due to missing values, the experienced and inexperienced group were represented by seven and eight participants respectively. For better comparison with literature, the responses have been converted from the 21-point scale to a percentage, with 1-21 mapped to 0%-100%. The descriptive statistics and some of the effects are summarized in Table 2 and are visualized in Figure 3, including the six contributing variables.

Table 2: means, standard deviations and effects of the overall perceived workload, RTLX converted to percent.

Overall workload (%)		Experienced N=7 μ	σ	*p*	Inexperienced N=8 μ	σ	*p*	*Inexperienced-experienced* *μ*	*p*
	Familiarization	18.2%	16.8		44.1%	16.3		*20.7%*	*.053*
Monotonic	Automated	10.2%	7.1		24.7%	16.2		*14.5%**	*.048*
	Manual	32.0%	19.6		29.6%	15.4		*-2.4%*	*.792*
	Man-auto	*21.8%***		*.008*	*4.9%*		*.463*		
Engaging	Automated	24.3%	20.3		42.9%	20.1		*18.6%*	*.098*
	Manual	48.3%	19.6		42.6%	17.1		*-5.7%*	*.556*
	Man-auto	*24.0%***		*.003*	*0.3%*		*.958*		
Engaging-monotonic	*Automated*	*14.1%*		*.066*	*18.2%**		*.015*		
	Manual	*16.3%**		*.015*	*13.0%**		*.033*		

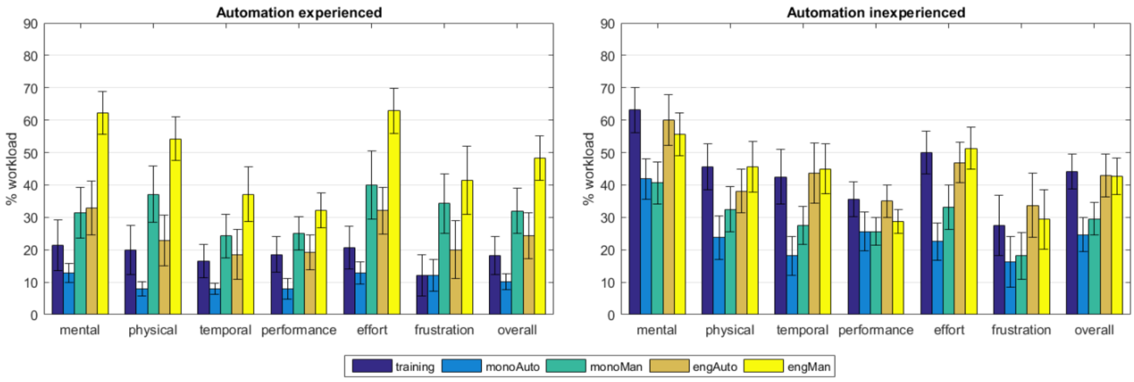

Figure 3: Results for the NASA RTLX for the experienced (left) and inexperienced (right) group, converted to percentage. Whiskers indicate standard errors.

Since the individual dimensions of workload were mostly showing a similar trend, only the overall workload ratings are further discussed here. The main effects are analyzed using a 2(automation: on vs. off) x 2(experience: experienced vs. inexperienced) x 2(environment: monotonic vs. engaging) factorial mixed ANOVA. The familiarization condition was excluded from this comparison, as it differs in road type.

The overall effect of automation showed that automated driving resulted in a 12.6% lower workload compared to manual driving ($F_{1,13}$=8.87, p=0.011), which supports the hypothesized main effect of automation (H1a). The hypothesized (H1b) interaction automation*environment was however not found in the overall workload ratings. ($F_{1,13}$=0.18, p=0.682).

The overall effect of environment showed a workload decrease of 15.4% with the monotonic versus the engaging condition ($F_{1,13}$=14.58, p=0.002). Hypothesis 2 is fully supported, as the workload in the monotonic environment was 16.1% lower for automated driving (SE=4.78%, p=0.005) and 14.7% lower for manual driving (SE=3.98%, p=0.003) when compared to the engaging environment. There were no significant interactions with the factor environment.

The overall effect of experience was not significant. ($F_{1,13}$=0.885, p=0.364) Experience did not reveal a significant effect in the manual driving conditions (p=0.633) indicating that both groups experience a similar workload during manual driving. The interaction automation*experience however was significant ($F_{1,13}$=5.95, p=0.030), with automation reducing workload by 22.9% compared to manual driving for the experienced drivers (SE=6.18%, p=0.003) and with no significant effect of automation in the inexperienced group (p=0.699). A post-hoc analysis shows that the hypothesized workload reduction with experience under automated driving (H3) was significant for the monotonic environment (mean difference=14.5%, SE=6.62%, p=0.048), but not for the engaging environment (mean difference=18.6%, SE=10.4%, p=0.098).

3.2. Detection response task (DRT)

The auditory DRT was performed as an objective measure of changes in mental workload. Due to missing values, the experienced and inexperienced group are represented by six and eight participants respectively. The descriptive statistics are summarized in Table 3 and are depicted in Figure 4.

Table 3: means, standard deviations and effects of the DRT reaction times and miss rates.

		Experienced N=6			Inexperienced N=8			*Inexperienced-experienced*	
Reaction time (ms)		µ	σ	*p*	µ	σ	*p*	*µ*	*p*
Monotonic	Automated	380	108		454	181		*74*	*.396*
	Manual	372	91		417	163		*46*	*.549*
	Man-auto	*8*		*.695*	*36*		*.068*		
Engaging	Automated	566	226		565	178		*0*	*.997*
	Manual	524	243		486	181		*38*	*.740*
	Man-auto	*41*		*.200*	*79**		*.011*		
Engaging-monotonic	*Automated*	*186***		*.005*	*112**		*.034*		
	Manual	*153***		*.005*	*68*		*.103*		
Miss rate (%)									
Monotonic	Automated	0.68%	1.68		1.05%	1.43		*0.4%*	*.664*
	Manual	0.23%	0.56		0.73%	1.12		*0.5%*	*.333*
	Man-auto	*0.5%*		*.531*	*0.3%*		*.611*		
Engaging	Automated	3.47%	2.66		7.81%	5.66		*4.3%*	*.109*
	Manual	1.84%	2.69		3.58%	2.51		*1.7%*	*.237*
	Man-auto	*1.6%*		*.372*	*4.2%**		*.016*		
Engaging-monotonic	*Automated*	*2.8%*		*.163*	*6.8%****		*.001*		
	Manual	*1.6%*		*.191*	*2.8% **		*.015*		

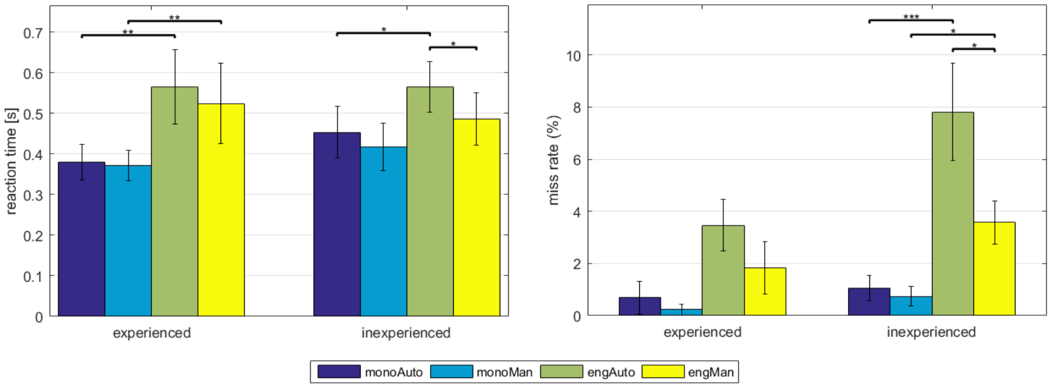

Figure 4: DRT reaction time (left) and miss rate (right) for the two groups. Whiskers indicate standard error and stars indicate significant differences at confidence P<0.05 (*), P<0.01 (**) and P<0.001 (***).

A factorial mixed MANOVA was performed on the variables reaction time and miss rate. Significant multivariate effects were found for environment (Pillai's trace=0.655, $F_{2,11}$=10.45, *p*=0.003), automation (Pillais' Trace=0.512, $F_{2,11}$=5.78, *p*=0.019) and for the interaction environment*experience (Pillais' Trace=0.467, $F_{2,11}$=4.81, *p*=0.032). No significant main effect was found for experience (Pillars' trace=0.233, $F_{2,11}$=1.67, *p*=0.23). Post-hoc analysis revealed that the effect of automation use was not significant in any of the conditions, except for the inexperienced group in the engaging environment, where automation use increased the miss rate (Mean difference = 4.2%, SE=1.5, *p*=0.016), as well as the reaction time (mean difference=79ms, SE=26, *p*=0.011).

The effect of environment was significant on the reaction times for the experienced group (automation: mean difference=186ms, SE=54, *p*=0.006. Manual: mean difference=153ms, SE=46, *p*=0.007) as well as for the inexperienced automated condition (mean difference=112ms, SE=47, *p*=0.034). The environment did affect the miss rate for the inexperienced group (automation: mean difference=6.8%, SE=1.6, *p*=0.001. Manual: mean difference=2.8%, SE=1.0, *p*=0.015)

4. Discussion

In this study, perceived workload (NASA RTLX) and objective workload (DRT) for manual and automated driving were assessed in a Tesla Model S on public roads. Effects of the moderators automation experience and traffic complexity were evaluated during attentive driving.

A strong effect of driving environment complexity was found both in perceived and objective workload, for both driver groups, and both in manual as well as in automated driving. This confirms hypotheses H2, but also shows that both groups remained sensitive to changes in driving complexity when using automation. This indicates that the monitoring task is demanding and that the demand changes with driving complexity.

For automation-experienced drivers, perceived workload reduced with automation use. For this group, the effect size of automation on workload (mean difference=22.9%, SE=6.18, p=0.003) was even larger than that of environment (mean difference=16.2%, SE=5.89, p=0.023 overall; mean difference=16.3%, SE=5.82, p=0.015 in manual driving), which is not in line with literature (comparing mean difference of 21% on factor automation from [8] with mean difference of ~36% on factor environment (manual driving) from [9]). In the automated monotonic condition with experienced drivers, perceived workload was even lower than values reported for on-road platooning after a 30 minute training (μ=19%; σ=13%). [24] This highlights the importance of using drivers with long-term experience in automated driving.

For automation-inexperienced drivers, automation use did not lower the workload compared to manual driving (neither for perceived nor objective load), meaning that experience is a prerequisite for workload reduction with automation. A similar effect of experience as prerequisite was found in a recent on-road automation study comparing manual drivers with ACC experienced drivers on secondary task uptake. [17] It should be emphasized that this finding is in conflict with results reviewed in [8], where automation-inexperienced drivers reported workload reduction due to automation of a magnitude similar to those found in this study for the experienced drivers. There is however one critical property for the studies reviewed in [8]: they mainly consist of simulator studies. This difference between real-world and simulator findings may relate to the poor validity of risk perception in driving simulators.

In contrast to the expectations, a dissociation between perceived workload (TLX) and objective attentional demand (DRT) was found for the effects of automation. The DRT did not show a reduction in attentional demand for automation. For the automation-inexperienced drivers, automation use even resulted in slower reaction times and a higher miss rate for the engaging condition, though when comparing to the experienced group, the reaction time difference seems to be caused by smaller variances and faster manual reactions rather than by slower reactions during automation. Therefore, the DRT results indicate that the drivers did not lower their attention when using the automation. Although this may indicate that the 10 minute periods of monitoring require as much attention as manual driving, some confounding effects have to be considered. The drivers have been encouraged to remain fully attentive due to the given instructions and the presence of a safety instructor and the two experimenters. A flooring effect on the sensitivity of the DRT may also be present. The sensitivity of the DRT to environment can possibly be attributed to the presence of lane changes in the engaging condition.

Limitations: The number of participants was low, which constrains how representative our findings may be for the general population. Since the variance is relatively large, increasing the sample size could test for significance in the smaller effect sizes. The instructions and presence of a safety instructor and experimenters motivated attentive supervision of the automation. Furthermore, the tesla users were sampled from a forum which actively discusses the limitations and abilities of the vehicle, making them more likely to demonstrate normative behavior. The two driver groups may also have differed in skill, since the experienced drivers had 77% more recent mileage than the inexperienced drivers. However the small between-group differences in the manual conditions on workload suggest that both groups experienced the manual driving similarly. Nonetheless, a more general population of Tesla owners would be beneficial for follow-up studies.

5. Conclusions

The following conclusions can be summarized:

- Automation lowered the perceived workload but in contrast to simulator studies, experience with the automation was a prerequisite for this effect, which highlights the importance of real-world testing.
- A strong effect of driving environment complexity was found under all conditions for both in perceived and objective workload. Both the perceived workload and attentional demand remained sensitive to changes in driving environment while using the automation.
- The DRT did not reveal an effect of automation use, which indicates that attentive monitoring can have an attentional demand similar to manual driving.

Acknowledgment

This work is supported by the NWO-TTW Foundation, the Netherlands, under the project “From Individual Automated Vehicles to Cooperative Traffic Management - Predicting the benefits of automated driving through on-road human behavior assessment and traffic flow models (IAVTRM)”-STW#13712.

References

[1] Dick de Waard, "The measurement of driver's mental workload," Traffic research centre, Rijksuniversiteit Groningen, 1996.

[2] Toshiya Hirose, Dai Kitabayashi, and Hidenobu Kubota, "Driving Characteristics of Drivers in a State of Low Alertness when an Autonomous System Changes from Autonomous Driving to Manual Driving," *SAE Technical Paper*, 2015-01-1407.

[3] Julie Paxion, Edith Galy, and Ctherine Berthelon, "Mental workload and driving," *frontiers in Psychology*, vol. 5, no. 1344, 2014.

[4] M. S. Young and N. A. Stanton, "What's skill got to do with it? Vehicle automation and driver mental workload," *Ergonomics*, vol. 50, no. 8, pp. 1324–1339, 2007.

[5] Annika F.L. Larsson, Katja Kircher, and Jonas Andersson Hultgren, "Learning from experience: Familiarity with ACC and responding to a cut-in situation in automated driving," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 27, pp. 229–237, 2014.

[6] T. J. Wright, S. Samuel, A. Borowsky, S. Zilberstein, and D. L. Fisher, "Experienced drivers are quicker to achieve situation awareness than inexperienced drivers in situations of transfer of control within a Level 3 autonomous environment," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 60, pp. 270-273, 2016.

[7] Maarten A. H. Krampell, "Accelerated Behavioural Adaptation through Targeted Training Programs – the Case of Highly Automated Driving," Department of Computer and Information Science (IDA) at Linköping University and the Swedish National Road and Transport Research Institute (VTI), Linköping , master's Thesis 2016.

[8] Joost C.F. de Winter, Riender Happee, Marieke H. Martens, and Neville A. Stanton, "Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence," *Transportation Research Part F: Traffic Psychology and Behaviour, Volume 27, Part B*, pp. 196-217, November 2014.

[9] Evona Teh, Samantha Jamson, Oliver Carsten, and Hamish Jamson, "Temporal fluctuations in driving demand:The effect of traffic complexity on subjective measures of workload and driving performance," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 22, pp. 207–217, 2013.

[10] A. Hamish Jamson, Natasha Merat, and Oliver Carsten, "Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions," *Transportation Research Part C: Emerging Technologies*, vol. 30, pp. 116-125, 2013.

[11] Christian Gold, Moritz Korber, David Lechner, and Klaus Bengler, "Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations: The Role of Traffic Density," *Human factors*, vol. 58, no. 4, pp. 642–652, 2016.

[12] Dick de Waard, Monique van der Hulst, Marika Hoedemaeker, and Karel A. Brookhuis, "Driver Behavior in an Emergency Situation in the Automated Highway System," *Transportation Human Factors*, pp. 67-82, 1999.

[13] Dyani J. Saxby, Gerald Matthews, Joel S. Warm, and Edward M. Hitchcock, "Active and Passive Fatigue in Simulated Driving: Discriminating Styles of Workload Regulation and Their Safety Impacts," *JOURNAL OF EXPERIMENTAL PSYCHOLOGY. APPLIED*, 2013.

[14] Mark Stuart Young, *Attention, automaticity, and automation: Perspectives on mental underload and perfrmance*.: University of Southampton, 2000.

[15] Matthias Beggiato, Marta Pereira, Tibor Petzoldt, and Josef Krems, "Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 35, pp. 75–84, 2015.

[16] Giulio Francesco Bianchi Piccinini, Carlos Manuel Rodrigues, Miguel Leitão, and Anabela Simões, "Reaction to a critical situation during driving with Adaptive Cruise Control for users and non-users of the system," *Safety Science*, vol. 72, pp. 116-126, 2015.

[17] Frederik Naujoks, Christian Purucker, and Alexandra Neukum, "Secondary task engagement and vehicle automation - Comapring the effects of different automation levels in an on-road experiment," *Transportation Research Part F*, vol. 38, pp. 67-82, 2016.

[18] Joost C. F. de Winter, P. M. van Leeuwen, and Riender Happee, "Advantages and disadvantages of driving simulators: a discussion," *Proceedings of Measuring Behavior*, p. 47050, 2012.

[19] Erwin R. Boer, Ahna Girshick, Tomohiro Yamamura, and Nobuyuki Kuge, "Experiencing the Same Road Twice: A Driver Centered Comparison between Simulation and Reality," *Proc. Driving Simulation Conference*, pp. 33-55, 2000.

[20] Oliver Carsten, Katja Kircher, and Samantha Jamson, "Vehicle-based studies of driving in the real workl: the hard truth?," *Accident Analysis and Prevention*, no. 58, pp. 126-174, 2013.

[21] Andreas Riener, "Assessment of Simulator Fidelity and Validity in Simulator and On-the-road Studies," *International Journal on Advances in Systems and Measurements*, vol. 3, pp. 110–124, 2010.

[22] David Hallvig et al., "Sleepy driving on the real road and in the simulator-A comparison," *Accident Analysis and Prevention*, vol. 50, pp. 44-50, 2013.

[23] Victoria A Banks and Neville A Stanton, "Keep the driver in control: Automating automobiles of the future," *Applied Ergonomics*, vol. 53, no. Part B, pp. 389-395, 2016.

[24] Daniël D. Heikoop, Joost C. F. de Winter, Bart van Arem, and Neville A. Stanton, "Acclimatizing to automation: driver workload and stress during partially automated car following in real traffic," *Manuscript submitted*.

[25] Alexander Eriksson, Victoria A Banks, and Neville A Stanton, "Transition to Manual: comparing simulator with on-road control transitions," *Accident Analysis and Prevention*, pp. 102C: 227-234, 2017.

[26] Tesla Motors Club LLC. (2017, March) TMC forum Belgium-and-the-Netherlands. [Online]. https://teslamotorsclub.com/tmc/forums/belgium-and-the-netherlands.118/

[27] National Data Warehouse for Traffic Information. (2017, March) NDW Open Actuele Dataservice. [Online]. http://opendata.ndw.nu/

[28] Nederlands Normalisatie-instituut, *NEN-ISO 17488:2016,IDT: Road vehicles -- Transport information and control systems -- Detection-response task (DRT) for assessing attentional effects of cognitive load in driving*. Delft, the Netherlands, 2016. [Online]. https://www.iso.org/standard/59887.html

[29] P. van Gent, H. Harah, N. van Nes, and B. van Arem, "Real-time, Nonintrusive Estimation of Driver Workload in a Simulated Environment," in *Manuscript submitted*.

[30] Jiun-Yin Jian, Ann M. Bisantz, Colin G. Drury, and James Llinas, "Foundations for an Empyrically Determined Scale of Trust in Automated systems," *International Journal of Cognitive Ergonomics*, 2000.

[31] Myra Blanco et al., "Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts," NHTSA, 2015.

[32] Ying Wang, Bryan Reimer, Jonathan Dobres, and Bruce Mehler, "The sensitivity of different methodologies for characterizing drivers’ gaze concentration under increased cognitive demand," *Transportation Research Part F: Traffic Psychology* , vol. 26, pp. 227–237, 2014.

[33] Sebastian Hergeth, Lutz Lorenz, and Roman Vilimek, "Keep your Scanners Peeled: Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving," *Human Factors*, vol. 58, no. 3, pp. 509-519, 2016.

[34] J. Navarro, M. François, and F. Mars, "Obstacle avoidance under automated steering: Impact on driving and gaze behaviours," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 43, pp. 315-324, 2016.