Driver’s Experience and Mode Awareness in between and during Transitions of different Levels of Car Automation

Paula Lassmann1, Ina Othersen2, Matthias Sebastian Fischer1, Florian Reichelt1, Marcus Jenke1, Gregory-Jamie Tüzün1, Kassandra Bauerfeind2, Lisa Mührmann2, & Thomas Maier1

1University of Stuttgart, IKTD, Germany
2Volkswagen AG, Wolfsburg, Germany

Abstract

Highly automated driving will have a significant impact on our future mobility. When a driver uses a system that comprises different SAE levels (L0, L2 and L3) the Human Machine Interface (HMI) needs to support the mode awareness of the driver at all times. While in L2 the driver has to monitor constantly, in L3 he can spend time on non-driving-related-tasks. The publicly funded project TANGO (Technology for automated driving, optimized to the benefit of the user) enables the design of an “attention and activity assistant” for automated truck driving in L2 and L3. The HMI of the project provides information about the automation level through different modalities: visually (instrument cluster & LED strip), auditory (sounds and voice announcements) and haptically via a tactile seat matrix. By conducting a driving simulation study, the usability of the HMI was investigated. The goal was to determine the ability of the driver to differentiate cognitively three SAE levels with the support of the TANGO HMI.

Introduction

The vision of automated driving stands for an increase in road safety and efficiency, a fatigue-free and stress-free driving experience as well as a safe use of built-in information and communication systems while driving. The fact that such a need exists among car drivers has already been sufficiently demonstrated in various studies (cf. Petermann-Stock et al., 2013; Wulf et al., 2012). However, the benefits of automation can also be beneficial for another group of users - professional drivers. They could be supported in their daily work routine by hours of monotonous journeys. The altered human-vehicle interaction has not yet been sufficiently examined in the field of trucks.

Within the project of TANGO, the research project concentrates on SAE Level 2 (L2) and SAE Level 3 (L3) systems (SAE International, 2018) and their transitions. L2 provides the driver with combined support in longitudinal and lateral guidance by means of an automated function. However, the driver must constantly monitor the driving situation and be prepared to intervene immediately. This monitoring function

is omitted in L3, where the system takes over the complete vehicle guidance in certain conditions (e.g. traffic jam or on motorways). In these situations, the driver no longer has to be “in the loop” and can therefore potentially turn to activities not related to driving. The driver only has to be ready for manual vehicle guidance within a period of several seconds when the system requests the driver to take over the vehicle guidance.

Especially in these quite similar modes of automation, an adequate awareness of the situation and the system is mandatory in order to avoid errors and to achieve an ideal human machine interaction (Sarter & Woods, 1995; Kolbig & Müller, 2013). Situation awareness is defined as “the perception of the elements in the environment within a span of time and space, the comprehension of their meaning and the projection of their status in the near future” (Endsley, 1988a, p. 97). It is understood as a dynamic process in which errors but also corrections can take place at all three levels (Endsley, 1988a, 1995a, 1995b). In the course of automated systems, situation awareness must also be enriched by system awareness. It can be understood as part of situation awareness and thus includes the same processes: the knowledge and understanding of system information and system-relevant environmental information as well as the anticipation of the information (Sarter & Woods, 1995; cf. Othersen, 2016). If this system awareness is incomplete, the probability of mode confusion raises. Thereby, the system reacts differently than expected by the user. The user may behave inappropriately (e.g. monitoring activities in L3) or miss actions (e.g. missing monitoring activities in L2; Bredereke & Lankennau, 2002). Studies could identify the missing supervision and higher attention to side task or longer viewing distances from the road in L2 (cf. Buld et al., 2002). Petermann-Stock (2015) also identified uncertainties regarding the system status and the required action through increased focus on relevant displays. Above all, an over-confidence in low automation levels, where a lack of monitoring with the potential oversight of system errors occurs, should be prevented.

The human machine interaction changes significantly through the use of automation, so that earlier actions are replaced by supervision or withdrawal from the driving task. The aim of efficient HMI should therefore be to provide important information for adequate awareness of the situation and the system as well as to prevent mode confusion as far as possible. The multimodal HMI developed in the TANGO project will be evaluated for the first time in a driving simulator study with professional drivers. The research questions of the study are as follows:

- Do people know which SAE Level they are in?
- Do people know their tasks according to the SAE level?
- How efficient, effective and satisfyingly is the level change supported by the TANGO HMI?
- How do people react to a critical driving situation?
Methodology

Experimental setup

The study took place at the vehicle ergonomics test facility at the research and teaching area Industrial Design Engineering of the University of Stuttgart. This fully variable model of a vehicle interior is based on a static, electrically adjustable seat box with driver and passenger seats. The driving simulation is shown on five monitors (Samsung, 1920 x 1080 pixels, 59”) with a 210° field of vision covering, two side mirrors and a rear-view mirror. For simulation, the software SILAB (WIVW, Version 5.0) was in use. The rides took place on a two-lane motorway with an emergency lane. During all automated rides, the vehicle’s speed was set to 100 km/h on the right lane. The vehicle was occasionally overtaken by other road users.

The study design included two independent variables: One variable was automation level (L0, L2, L3), respectively transitions, as a within factor. Each participant experienced each automation level and possible transition. However, the order was counterbalanced in between two groups. The second independent variable was the arousal of a critical event (same situation, either in L2 or L3) as a between factor, which occurred at the end of test run two. The difference between the take-over situations before the critical situation was the fact, that in L2 no warning was given, whereas in L3 the system gave a take-over request (TOR) (for an overview of the ride see fig. 2 below).

User centred HMI

For promoting the mode awareness, the automation level state was supported by different HMI elements. A schematic layout of the vehicle cockpit is given in fig. 1. In this HMI, the L2-mode was called “Assistance Plus” (colour code: blue) and L3-mode “Autopilot” (green). Orange and red were used for warnings. The HMI included an instrument cluster, a head unit display above the centre console and a detachable tablet placed on a holding to the right. A colour-coded LED strip showed the level colour and was positioned at the bottom edge of the windshield. For (de-)activating L3, two push-buttons (that lit up in green when the L3 was available) on the steering wheel had to be pushed simultaneously. One push-button on the centre console, which lit up in blue, (de-)activated L2. The push-button in the middle of the steering wheel was used for the Sign Detection Task (SDT; see chapter Data and analysis). A tactile seat matrix (TSM; Schwalk et al., 2015), was used for tactile feedback during the TOR and after activation of the automation. It consisted of a 4x4-matrix in the backrest and a 3x5-matrix inside the seating area. All processes were supported by voice announcements and icons within the instrument cluster. As soon as a change from L2 to L0 was recommended, the driver had 2.2 s to deactivate Assistance Plus either by pressing the above-mentioned button or by driving related intervention (using the accelerator or brake pedal or oversteering). However, Intervention always resulted in a level change to L0. If the driver neither pressed the button nor intervened, the automation switched off. In L3 the driver had 15 s to react to a level change. If the driver did not react the system did a safe stop.
Procedure

At the beginning, participants received an introduction to the research topic, the detailed description of the system functionality and the different tasks they have to perform according to the SAE levels. Afterwards they answered demographic questions and went through an acclimatisation ride in which all SAE levels and all tasks could be experienced.

One test ride consisted of three consecutive runs with automation and four transitions, with a total run time of approximately 90 min. Transitions were announced by the system according to the study setup. After each transition, questions were asked about the transition itself in terms of effort, mode awareness as well as the HMI without pausing the ride (see fig. 2). Participants answered all questions throughout the study verbally, which the investigator documented. In addition, the simulation was paused in the middle of each test drive (system freeze) for answering questions on the mode awareness. At the end of the first drive, the participants took a break during which they had to complete another questionnaire regarding the overall driving experience in this first half of the simulation. Subsequently, the second test drive took place – similar to the first one. However, at the end of the test drive, a critical event (system error) occurred, in which the driver had to take over. This critical event consisted of a traffic jam that suddenly occurred after a hill and therefore could not have been noticed early. In L2 the system did not brake on time. Therefore the drivers needed to initiate a transition and brake themselves to avoid a crash. However, in L3 the system announced a take-over within 15 s. After finishing the second ride, in addition to the same questions that were answered at the end of the first drive, another questionnaire referred to the perception of the critical event.
Figure 2. Schematic study design and procedure (two rides per person with two different test conditions as in between factor).

Data and analysis

A non-driving-related-task as well as a driving-related task assessed the mode awareness. The driving related task was supposed to measure the monitoring performance of the drivers and at the same time, assess the mode awareness. The drivers had to detect speed limit signs, which is called the Sign Detection Task (Lassmann et al. 2019). The design of the SDT is based on detection tasks of driving relevant stimuli that have been used for assessing vigilance and therefore monitoring performance (cf. Greenlee et al. 2017; Heikoop et al. 2017). In L2, the participant was supposed to press the SDT-button located on the steering wheel when a specified speed limit sign (100 km/h) could be seen on the roadside for 3.5 s. If the driver reacted to another road sign, this was considered as error. The SDT requires visual attention and could be compared to monitoring activity in terms of suddenly appearing obstacles in L2 (Lassmann et al. 2019). The hit rate and the response time to the stimulus were recorded. In case of pressing the SDT-button in L3 it was considered as a mode confusion since monitoring is not required in this level. In addition to the SDT, a secondary task (quiz, based on Petermann-Stock et al., 2013) was offered on the tablet to the right. The drivers could decide themselves if they interacted with the quiz (either in the fixed position or hand-held). It consisted of 262 questions in German with four answer options each. The quiz has been proofed as an engaging task for truck drivers in several studies within the TANGO project (e.g. Bieg et al. 2019).
Another objective measure of mode awareness is a variation of the SAGAT (Situation Awareness Global Assessment Technique; Endsley, 1988b) survey method. The SAGAT records the respective situation-specific knowledge of the person via questions concerning perception, understanding and anticipation of the situation after freezing the simulation. In this study, the HMI was hidden or frozen after a transition instead of the entire simulation and questions were asked about the respective automation level and the distribution of responsibilities. Besides that, the Mode Awareness questionnaire from Benecke (2014) was used as subjective data. This includes the areas of perception, understanding and anticipation of the system status. In addition, individual items were asked for the critical event with regard to effort (Subjective Experienced Strain Scale; Eilers et al., 1986), subjective reaction quality and subjective criticality. All questions were implemented using a five-level Likert scale.

For this purpose, mean value differences were calculated using the Wilcoxon rank-sum test as well as variance statistical methods with and without repeated measurements for the factors measurement time and automation level at a significance level of \( \alpha = .05 \).

**Participants**

The driving simulation was performed with 30 participants (aged 22 to 60 years, \( M = 41.6, SD = 10.8 \)). The group consisted of twenty truck drivers, three bus drivers and seven other frequent drivers with an average annual kilometrage of 85,500 ± 36,667 km (range 20,000 to 200,000 km). For 2.5 hours of simulated driving and questioning the participants received an incentive of 100 Euro. Due to measurement failures, motions sickness or language problems during the study, eight participants were excluded from the analysis, which leaves 22 subjects.

**Results**

**Mode Awareness**

**Sign Detection Task (SDT)**

Only in L2, people should perform the SDT. However, two participants hit the button continuously, three subjects once, while being in L3. The rest (77.0 %) performed correctly by not hitting the button. In L2 hit rates reached a mean of 77.6 % (16.9) with values ranging from 50 to 100 % with no change over time (\( F[2,42] = 1.900; \ p = .162; \ \eta^2 = .083 \)). Seven persons had a mean under 70.0 %, whereas the hit rate of three of them increased throughout the study. 15 had a hit rate of over 88 %. The mean of the reaction times of the hits was 1.95 s (.35). No change over time occurred either (\( F[2,42] = 1.042; \ p = .362; \ \eta^2 = .047 \)).

**Secondary Task – Quiz**

The results showed a shift of attention in higher automation level (\( F[2,20] = 103.33; \ p \leq .001; \ \eta^2 = .91 \)) (see fig. 3). The drivers did less quiz questions during L0 (\( M_{L0}=1.33 \)) than in L2 (\( M_{L2}=34.77 \)) and L3 (\( M_{L3}=56.39 \)). In addition, there was an effect of time (First, Second and Third Time in either L0, L2 or L3) for all three
automation levels ($F[2,20] = 5.14; p \leq .05; \eta^2 = .34$). All test persons performed less quiz questions over time ($M_{M1}=36.01; M_{M2}=31.02; M_{M3}=25.41$).

The correlation of the performance frequencies of the quiz and the hit rates of L2 are negatively correlated (Spearman: $r = -.488; p \leq .05)$. Looking at the correlations of each measurement time, only a correlation of M3 exists: $r = -.566 (p \leq .01)$.

**Subjective Mode Awareness**

The analysis of the subjective evaluation of mode awareness showed that participants in the two groups did not differ in terms of mode awareness ($F[1,20] = 0.004; p = .952; \eta^2 = .000$). However, there was an effect on the time of measurement ($F[1,20] = 10.664; p \leq .01; \eta^2 = .348$). The mode awareness improved during the ride ($M_{before} = 4.47; M_{after} = 4.68$).

During the freezing situation, 21 of 22 participants were able to reproduce the current automation level they were in, as indicated by the correct labelling of the automation mode (i.e. ‘Assistance Plus’/ ‘L2’), or the matching colour (i.e. ‘blue mode’). Participants in L2 and L3 were equally aware of the system mode during the freezing situation ($Mdn_{L2}=5.00, Mdn_{L3}=4.83, W=70, p=.209$). However, 31% of the participants in L2 and 22% in L3 were unsure about the correct tasks they had to perform.

Concerning the system error that had to be detected either by the participants themselves in L2 or by a TOR by the system in L3, no significant differences could be found in the degree of effort ($Mdn_{L2}=2.17, Mdn_{L3}=1.60; W=38, p=.125$), the speed while overtaking ($Mdn_{L2}=4.90, Mdn_{L3}=4.67, W=74, p=.220$) and their confidence with the quality of their reaction ($Mdn_{L2}=4.83, Mdn_{L3}=4.80, W=58, p=.882$). Finally, both
groups evaluated the driving task as well adopted ($Mdn_{L2}=4.83$, $Mdn_{L3}=4.90$, $W=64$, $p=.689$), and quickly surveyed ($Mdn_{L2}=4.83$, $Mdn_{L3}=4.80$, $W=58$, $p=.882$).

Transition and Critical TOR

In the following section, only the transitions into and between automation levels (see fig. 4) are addressed. The change into L0 (critically and uncritically) is focused in fig. 5. The analysis of the transitions shows that a few participants had problems changing levels. 22.73% did not make the safe transition from L0 to L2 at the first attempt. They switched at least once to L2 and back to L0 due to another button, pedal or steering wheel operation. 18.19% had the same problems transitioning from L0 to L3. The same number of participants did not change directly from L3 to L2. They changed into L0 before reaching the right level. There were no problems when changing from L2 to L3.

The transition times of three participants were considered as outliers (3σ) and therefore excluded from statistical analysis. Each participant changed from L0 to L2 and L3 once. They changed twice from L2 to L3 and L3 to L2 due to the study setup, without an effect of time of measurement (first or second time; L2 to L3: $t(18)=0.498$; $p = .624$; L3 to L2: $t(18)=1.895$; $p = .074$). Therefore the mean of both values was used for the following analysis. Reaction times differed in terms of transitions ($F[2,253]=8.480; p = .001; \eta^2 = .320$). The transition from L3 to L2 ($M_{L3\rightarrow L2}=6.90$) took more time than the changes from L0 to L2 ($M_{L0\rightarrow L2}=4.28; p = .005$) and L2 to L3 ($M_{L2\rightarrow L3}=4.12; p \leq .001$).

The following results are displayed in fig. 5. Due to small in between group sizes (9 and 10 participants), the data was not analysed statistically. After the request to change from L2 to L0, the automation was switched off after 2.2 s. No driver reacted via intervention or button press within this time. At this point, no statement can be made about the drivers’ handling of the situation. In the critical event, all drivers acted within 6.90 s with a mean of 4.54 s (onset: fastest transition minus 2 s reaction time).
by using the pedal or steering wheel intervention during the critical event. None of the participants collided with another vehicle.

The uncritical transition from L3 to L0 was made within 3.20 and 11.32 s with a mean of 6.77 s, whereas in the critical situation, the reaction time decreased slightly to a mean of 6.22 s (range from 3.31 to 8.2 s). Without a critical event, 50 % used the buttons to change level. All other participants used pedal or steering wheel intervention. The buttons were used less in the case of the critical event (33 %). None of the participants provoked a safe stop (15 s after the announcement).

<table>
<thead>
<tr>
<th>Transition</th>
<th>L2 → L0</th>
<th>L2 → L0 critical event</th>
<th>L3 → L0</th>
<th>L3 → L0 critical event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants in group</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Time [s] from announcement to completed transition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collision during critical event</td>
<td>N/A</td>
<td>0 %</td>
<td>N/A</td>
<td>0 %</td>
</tr>
<tr>
<td>Transition initiation</td>
<td>Button(s)</td>
<td>0 %</td>
<td>0 %</td>
<td>41.67 %</td>
</tr>
<tr>
<td>Interventions</td>
<td>0 %</td>
<td>100 %</td>
<td>58.33 %</td>
<td>70 %</td>
</tr>
<tr>
<td>No reaction *</td>
<td>100 %</td>
<td>0 %</td>
<td>0 %</td>
<td>0 %</td>
</tr>
</tbody>
</table>

* L2: no reaction within 2.2 s
L3: no reaction within 15.0 s (safe stop)

Figure 5. Transition times for changing into L0 in an uncritical and critical situation (between factor).

Discussion

The study identified the influences of the various automation levels on mode awareness. According to the SDT, two subjects had a continuous task confusion. Seven people failed the monitoring performance, since they would not have detected important signs and reacted to them within 3.5 s in over 30 % of the cases in L2. There might be three factors which had influence on the SDT performance: distraction, task confusion due to insufficient knowledge about the tasks or mode confusion. The quiz was a visual distracting non-driving related task, which was available throughout the levels. Visual distraction leads to a bad performance for detecting obstacles (Lorenz et al. 2015) as well as in the SDT (Lassmann et al. 2019). This thesis is supported by the negative correlation of the SDT hit rates and the performance frequency of the quiz. Subjects who were rather involved in the quiz, missed road signs. Regarding the factors task and mode confusion, the objective data gathered during the freezing situation might help to explain the results. Overall 21 of 22 participants were able to
correctly reproduce the current automation level they were in, but 31% of the participants in L2 and 22% in L3 were unsure about the respective tasks. These results could actually be an indicator of mode confusion, which refers to a discrepancy of the participants’ belief about which aspects of vehicle performance are controlled by themselves and which are controlled by the automation at a particular instance (Cummings & Ryan, 2014).

In terms of the transition, participants did fairly well. After having tried the transition once during the acclimatisation ride, most succeeded in transitioning at the first attempt within a few seconds. Most problems that occurred were due to the fact that people either still pressed a pedal, pressed the button for too long or did not trust the trajectory of the simulation. For most people this happened only once during a whole test ride. In addition, subjects did not change levels faster while doing it the second time which speaks for good usability at the first place. In summary, according to the results, the HMI supported the user during transitions well. Nevertheless, a quote of 100% transitions at first attempt would be desirable.

For the changes from L3 into a lower level, people took more time, which is in line with the findings of Gold et al. (2013): the longer the possible time frame for take-over, the longer the take-over takes. Even during the critical situation, take-over times did not change much, which supports the thesis, that people were rather trustful of the system. A timeframe of 2.2 s for the transition from L2 to L0 was not enough for drivers to react to the change and the readiness of the driver for take-over was not checked. This shows the danger of a L2-system: undertaking a transition from L2 to L0 without an explicit driver interaction, monitoring or a fallback action might lead to a situation of an unsupervised car in motion. For this reason a driver monitoring will be implemented in the TANGO system to check the driver’s readiness. However, all drivers became aware of the critical situation in L2 and reacted in time to prevent an accident, which leads to the assumption that drivers were aware of the monitoring task and also the mode. In terms of take-over from automation to manual driving, the intervention seems to be more intuitive for drivers than pressing a button.

In summary, the results of the study seem divers. According to the findings of Lee and See (2004), the misbelief about the vehicle’s operation is a result of overtrust or undertrust in the automated system. In this context, it could be assumed that participants in L2 had overtrust in the automated system, as they incorrectly thought that they could fill in the quiz, despite being supposed to watch the traffic. Overtrust can lead to misuse of the automated system, where the driver applies the automation to a roadway environment that is outside of the automation’s operational scenarios. In the critical take-over situation, however, the drivers with L2 were able to update their situational awareness quickly enough so that there were no problems with out-of-the-loop performance. On the contrary, participants in L3 had possibly distrust in the system, as they thought that they had to watch the traffic, despite the automated system taking over this task completely. Participants believed that the automation performance was less than it actually was, which leads to a disuse of the automated system and thus removing the possible benefits of the automation (Lee & See, 2004).
Conclusion

This study on mode awareness with regard to different automation levels was able to show that the test persons could subjectively indicate the correct automation level, but made mistakes in indicating the tasks which they had to perform. This corresponds to the objective performance in the secondary task. On the one hand these results show, that the HMI succeeded to convey the information of different automation modes that were obvious to the driver. On the other hand the results could actually be an indicator for mode confusion which refers to a discrepancy between how the participants believed the vehicle to be operating and how the vehicle was actually operating during L2 and L3 – e.g. that monitoring the system could be achieved while performing a visual non-driving related activity. Therefore the tasks during automation should be emphasised more clearly – either by instruction or by the system - and internalised by the driver. However, in conclusion the TANGO HMI supports the driver well, especially in regard to transitions, but can be improved regarding assistance of mode and task awareness.

References


SAE International (2018). Taxonomy and definitions for terms related to on-road automated motor vehicles (J 3016 Aufl.).
