Workload evaluation of effects of a lane keeping assistance system with physiological and performance measures

Yu-Jeng Kuo, Corinna Seidler, Bernhard Schick, & Dirk Nissing
[1]Kempten University of Applied Science, [2]Rhein-Waal University of Applied Science
Germany

Abstract

The present study investigated the mental workload associated with driving a vehicle equipped with Lane Keeping Assistance System (LKAS). Specifically, an experiment was carried out with 16 participants driving with LKAS in four real-world scenarios. Effects on mental workload were evaluated with psychophysiological measures such as heart rate and skin conductance response (SCR). The driving performance, which is also a measure of evaluating mental workload, was assessed by measures such as steering reversal rate, variation of lateral position and steering effort. The result suggested that LKAS has reduced physical workload in the steering task. However, the lane keeping performance was not improved. Moreover, the NASA-TLX showed that participants perceived higher mental workload while driving with LKAS. This effect was mirrored in the SCR. The objective data showed that LKAS was associated with higher steering reversal rate, which might explain the reason of participants perceiving higher mental workload. Overall, it was suggested that the mental workload was higher with the tested LKAS.

Introduction

The development of the Advanced Driver Assistance Systems (ADAS) has advanced a lot since the late 90s. From the passive Anti-lock Braking System (ABS) in 1987 (Bosch), to the introduction of Adaptive Cruise Control (ACC) in 1999, various driving tasks in modern cars have been gradually delegated to automated control system (Bengler et al., 2014). Few years after the introduction of ACC, the Lane Keeping Assistance System (LKAS) was introduced by Honda in 2004 (Ishida & Gayko, 2004). In contrast to longitudinal motion managed by ACC, the LKAS is designed specifically for lateral control. The idea behind LKAS is simple: to support staying in a lane. The system constantly measures the distance to the lane marking via one or more camera, and applies steering torque to keep vehicle from leaving the lane.

Discussions regarding the effect of vehicle automation on human can be found widely in the literatures. Stanton and Marsden (1996) stated a number of arguments favouring automation of the driver role, such as the automation could improve well-being, improve road safety, and enhance product sales. Ultimately, it may relieve the driver of excessive and complex driving activities. Brookhuis et al. (2001), however, pointed out that the ADAS may introduce these benefits, but the consequences (e.g. increasing

In D. de Waard, A. Toffetti, L. Pietrantoni, T. Franke, J-F. Petiot, C. Dumas, A. Botzer, L. Onnasch, I. Milleville, and F. Mars (2020). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2018 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

complexities of the cockpit, decreasing alertness and attention from the driving task, and negative effect on skills) should also be identified. Although in the context of Intelligent Vehicle-Highway System, Hancock and Parasuraman (1992) also commented that such assistive function might intend to mitigate the mechanical effort from the driver, yet it could also hypothetically increase driver's cognitive workload if monitoring the system is required.

When a driver assistant system functions as expected, it has been reported in a review paper that the averaged self-reported workload (0% = minimum, 100 = maximum) decreased from 43.5% in manual driving to 38.6% in ACC driving (De Winter et al., 2014). Furthermore, some evidences even suggested that lateral support relieves mental workload to a greater extent than ACC (Young et al. 2002; Carsten et al. 2012). In contrast, if automation does not behave as one anticipates, it could result in increasing driver’s mental workload. For instance, Banks and Stanton (2015) showed in a field study that participants reported higher subjective mental workload and lower trust when driving with automated vehicle (with longitudinal, lateral support and auto-overtake system) in comparison to manual driving. The results indicated that the unexpected lane changes and unsafe auto-overtake offerings were possibly part of system’s weaknesses.

As argued by Sarter et al., (1997), when a new automation is introduced into a system, new coordination demands between human and machine often come along. Moreover, it is particularly difficult for human to coordinate activities, when the intentions of machine agents are not clear. This observation is similar to the findings in our previous pilot study, in which the participants subjectively reported overall higher workload levels while driving with LKAS than driving without it. The paper concluded that the unexpected system failure, inconsistent feedback and lack of transparency were the main reasons of having this outcome (Schick et al., 2019).

Similar to our pilot study, the primary purpose of this study is also to investigate the mental workload associated with LKAS. However, it differs in two ways. Firstly, only drivers who have had experience with LKAS were selected as participants. The experience with automation, as suggested in Stapel et al., 2017, is a prerequisite of reducing perceived workload. Secondly, the objective data (i.e. driving performance) were presented, which should reveal the driving behaviour when driving with LKAS. In total, four distinct real-world scenarios were designed, which consisted of various curviness of the route and driving velocity. The participants were asked to drive through all scenarios two times (with and without LKAS). The mental workload was assessed with objective and subjective measures.

Based on the work mentioned above, two hypothesis have been formulated:

H1: Drivers’ perceived mental workload would be higher with LKAS than without it.

H2: Lane keeping performance would be better when driving with LKAS.

workload evaluation of effects of a lane keeping assistance system

Experimental Design

To elicit different levels of workload with LKAS, four driving scenarios were designed based on the combination of cruising velocity and road geometry. The cruising speed was either at 120 km/h (Low-speed, L) or 160 km/h (High-speed, H), whereas the road geometry was either curvy (c) or straight (s). Hence, the four scenarios were abbreviated as Lc (Low-speed-curvy), Ls (Low-speed-straight), Hc (High-speed-curvy) and Hs (High-speed-straight).

The scenario Lc was a 7 km rural road (B19, Waltenhofen – Oberdorf) that consisted of a number of minor curvy sections. The scenario Ls and Hs were each 5 km straight motorway sections (A980, Waltenhofen – Dreieck Allgäu). Essentially, these two scenarios shared the same motorway, but in opposite direction. Finally, the scenario Hc was a 10 km motorway (A7, Dreieck Allgäu – Oy-Mittelberg) which consisted of two high radius curves (each with a radius of approximately one km). The order of the scenarios was predefined, as driving through all scenarios in a randomized order would have taken too much time travelling between each scenario.

For one complete lap, the participant first started with scenario Lc (go and back), followed by a single scenario Ls (go), then through scenario Hc (go and back), and finally finish in scenario Hs (back). Unless explicit speed limit encountered, the driver tried to maintain the speed at 120 km/h in scenario Lc and Ls, and at 160 km/h in scenario Hc and Hs. It took in total about 25 minutes to finish one lap. In order to investigate the effect of LKAS in different scenarios, the participant had to drive through all scenarios two times (laps) i.e. with and without LKAS. The order of introducing LKAS was counterbalanced. The routes are illustrated in Figure 1.

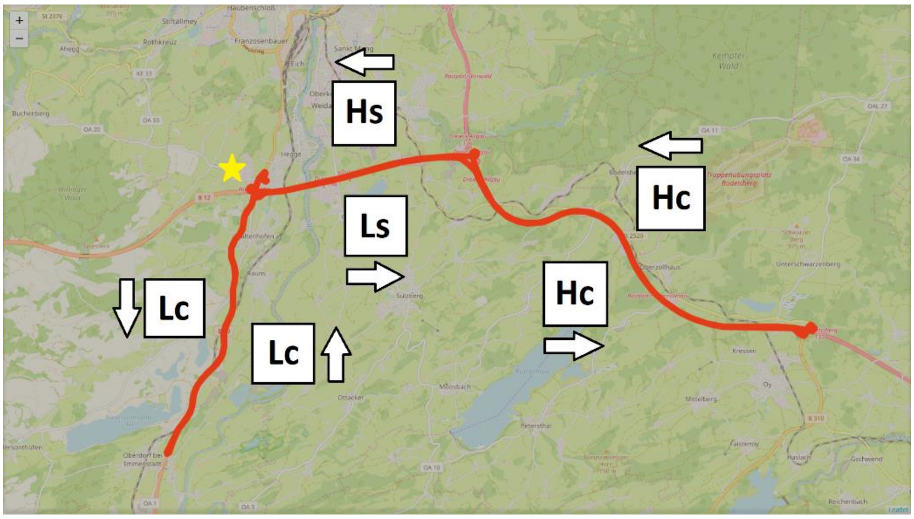

Figure 1. Four real-world scenarios (Lc: Low-speed-curvy, Ls: Low-speed-straight, Hc: High-speed-curvy, Hs: High-speed-straight). The yellow star indicates the starting and ending of one complete lap. (Figure adapted from openstreetmap.org)

Participants

In total, 21 volunteers between 19 and 65-year of age participated in the experiment (M: 32.6; S.D. 13.5). They had participated in a previous pilot study. All participants possessed a driving licence for at least three years, their self-reported average annual driving mileage was 16333 km (S.D. = 5365 km). The participants signed an informed consent form before taking part in the experiment. Due to adverse weather, traffic conditions and technical issues, the data of five participants were discarded.

Objective measures

To assess mental workload, the skin conductance response (SCR) was taken as an indicator of the activity of sweat glands. In the literature, both the SCR and its tonic counterpart (skin conductance level, SCL) have been used for measuring mental workload (Gris et al., 2012; Zangróniz et al., 2017). In this experiment, the count of SCR per kilometre was taken as a workload indicator. In addition, the heart rate (HR) and heart rate variability (HRV) were also taken as dependent variables. It has been shown that HR and HRV are sensitive to evaluate operators' effort (Aasman et al., 1987) and mental workload (De Waard & Brookhuis, 1991; Wilson & Eggemeier, 1991). For HR, these measures in the time-domain were included:

- Inter-beat-interval (IBI)
- Root-mean-square of successive R-R interval differences (RMSSD)
- Standard deviation of N-N intervals (SDNN)
- Percentage of successive R-R intervals that differ by more than 50 ms (pNN50)

To describe driver's performance and behaviour, the standard variation of lateral position (SDLP) and the steering reversal rate (SRR) were used. Before computing SDLP, as suggested by Östlund et al. (2005), the distance-to-line was filtered with second order Butterworth 0.1 Hz high-pass filter to ignore the variation within 10 seconds of observation window. In addition, the data that were 5 seconds before and after any lane-crossing events were excluded. The SRR was defined as the number of times per minute that the direction of steering movement was reversed through a small finite angle (3-degree). Finally, the steering effort was included to quantify the level of physical effort required to perform the steering task. It was calculated as the product of steering angle (degrees) and steering torque (Nm).

The LKAS tested in this study was equipped in a premium class vehicle. The vehicle parameters were assessed with a data acquisition system (DEWE2-A4, Dewetron) installed in the rear trunk. The physiological data were recorded with a wireless wearable system (BioNomadix, BIOPAC Systems Inc.).

Subjective measures

The NASA-TLX (Hart & Staveland, 1988) was used to measure subjective mental workload. A 21-point scale was used to map workload level from 0% to 100% for six subscales (mental demand, physical demand, temporal demand, effort, performance and frustration). The result (Raw-TLX, R-TLX) was obtained by averaging the ratings across subscales for four conditions (120/160 km/h x with/without LKAS).

Protocol

The experiment was conducted in late April until early May 2018 in the Allgäu region, Germany. Upon arrival, each participant was briefed about the routes and the goal of the study. Starting from the research centre, the driver used the first 8 km to become familiar with the test vehicle before the starting point (the yellow star in Figure 1). One research staff member sat on the passenger seat to operate the measurement devices. After finishing the first lap, the participant parked the car in a parking lot nearby and filled the questionnaire (NASA-TLX) before starting the second lap. The LKAS was then switched either on or off here. For safety reasons, the driver had their hands on the steering wheel all the time. In case of an unexpected system failure occurred, the driver should perform counter steering or any other measures to correct the vehicle's trajectory. It took about one hour for each participant to finish one complete test run (two laps). A summary of the each scenario is listed in Table 1.

Table 1. Experiment design for one complete lap (Lc-Lc-Ls-Hc-Hc-Hs). Each participant had to drive two laps: with and without LKAS.

Scenarios	LKAS	Velocity	Length (km)	Route
Lc	with / without	120 km/h	7	Curvy
Lc			7	Curvy
Ls			5	Straight
Hc		160 km/h	10	Curvy
Hc			10	Curvy
Hs			5	Straight

Results

For data analysis, the objective data were submitted to 2 (LKAS: ON, OFF) x 4 (scenarios: Lc, Ls, Hc, Hs) analysis of variance (ANOVA) with repeated measures. Greenhouse-Geisser corrections were applied in case where the data failed to pass Mauchly-Test. Post-hoc test with pairwise comparisons were corrected by Bonferroni corrections. The p-value for significance test was 0.05.

Physiological measures

In terms of HR, neither a main nor an interaction effect of LKAS was found. In contrast, a main effect of scenarios was found significant, $F(3, 45) = 7.931$, $p < .005$. The post-hoc pairwise comparison showed that the HR in scenario Ls (75.10 bpm) was significantly lower than the curvy scenarios (Lc = 76.67 bpm, Hc = 76.61 bpm), both $p < .01$. In other words, HR was in general higher in the winding route than on the straight motorway. For other dependent variables, only a significant main effect of scenarios on IBI was found, $F(3, 45) = 7.65$, $p < .005$.

Apart from the effect of LKAS and scenarios, the learning effect between the HR with groups (between-subject effect) and number of trials (within-subject effect) was investigated. The data were submitted to a two-way mixed ANOVA. With respect to HR, a main effect of trial numbers was observed, $F(2.1, 29.4) = 5.35$, $p=.01$. In contrast, the difference between groups was not significant. This result suggests that the group, which experienced LKAS in the first lap, showed a lower HR in the second lap. In contrast, the HR of another group (without LKAS in the first lap) remained at a similar level in the second lap where LKAS was switched on (Figure 2).

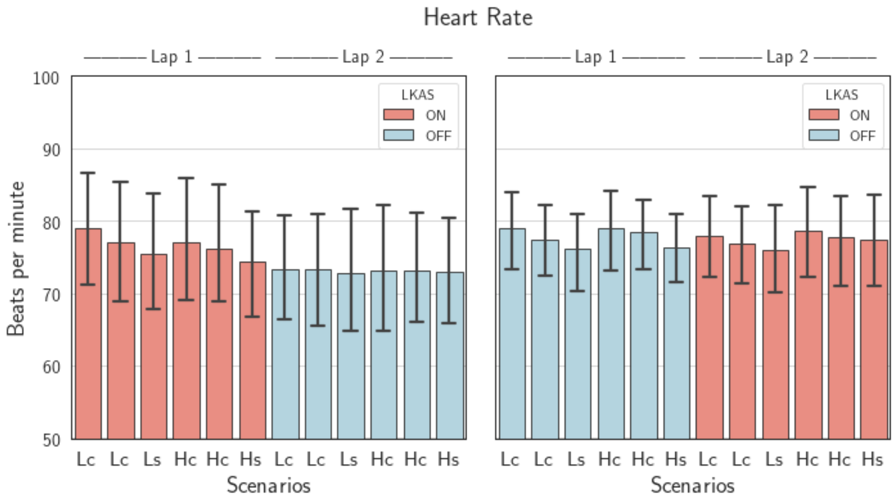

Figure 2. The heart rate over scenarios in a chronological order (left to right). The group on the left started with LKAS, while the group on the right started without LKAS. (Lc: Low-speed-curvy, Ls: Low-speed-straight, Hc: High-speed-curvy, Hs: High-speed-straight)

For SCR, a main effect of LKAS on the average count per kilometre was found, $F(1, 15) = 4.62$, $p = .048$. However, no difference was found with scenarios as well as their interactions. The result of average SCR / km over the scenarios is shown in Figure 3.

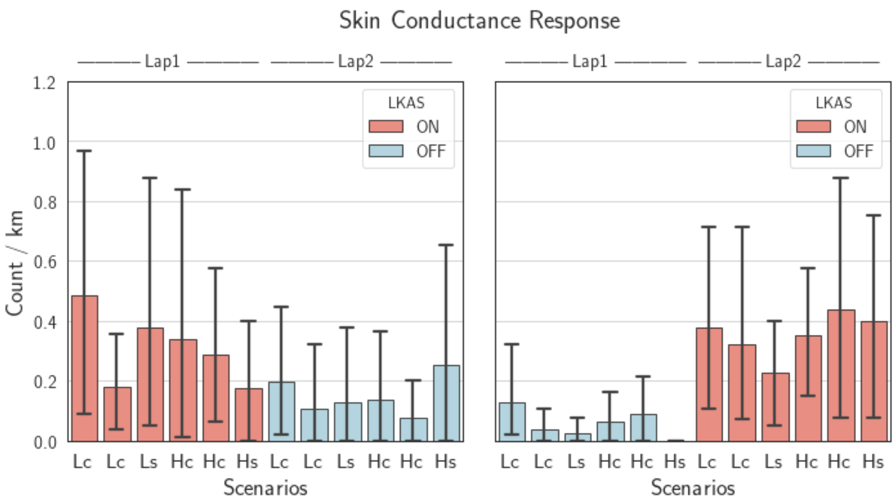

Figure 3. Averaged SCR / km over all scenarios in a chorological order. (Lc: Low-speed-curvy, Ls: Low-speed-straight, Hc: High-speed-curvy, Hs: High-speed-straight)

workload evaluation of effects of a lane keeping assistance system

Performance measures

The analysis of SRR revealed a significant main effect of LKAS ($F(1, 15) = 55.24$, $p < .005$) as well as of four scenarios ($F(3, 45) = 679.1$, $p < .005$). An interaction effect was also found between LKAS and scenarios ($F(1.7, 26.1) = 9.07$, $p = .002$). A pairwise t-test showed that the SRR was always higher when driving with LKAS in the curvy scenarios (in Lc, $t = 4.767$, $p<.005$; and Hc, $t = 5.475$, $p < .005$), whereas the difference was not significant in straight scenarios (Ls and Hs). On the other hand, the SRR in curvy scenarios (Lc vs. Hc) was significantly different from each other irrespective of LKAS (all $p < .005$), while no difference between straight scenarios (Ls vs. Hs) was found. This result is illustrated in Figure 4a.

In terms of steering effort, ANOVA showed that significant main effects of LKAS ($F(1, 15) = 242.3$; $P < .005$) and scenarios ($F(1.74, 26.1) = 635.6$, $p < .005$) were found. In addition, the interaction effect between two factors ($F(1.45, 21.8) = 165.0$, $p < .005$) was also significant. In contrast to SRR, the steering effort in every scenario was greater when driving without LKAS ($p < .05$), except of scenario Ls. However, when driving with LKAS, no significant difference was found between curvy scenarios (Lc vs. Hc), as well as between straight scenarios (Ls vs. Hs). Overall, the steering effort in the curvy scenarios (Lc and Hc) was significantly greater than straight scenarios (Ls and Hs). This observation is illustrated in Figure 4b.

For SDLP, no difference was found between scenarios, and between LKAS. The post-hoc paired-sample t-test indicated that the SDLP was only significantly higher in scenario Hc ($M = .120$ m) than scenario Lc ($M = .107$ m) when driving without LKAS. The rest comparisons were all not different from each other.

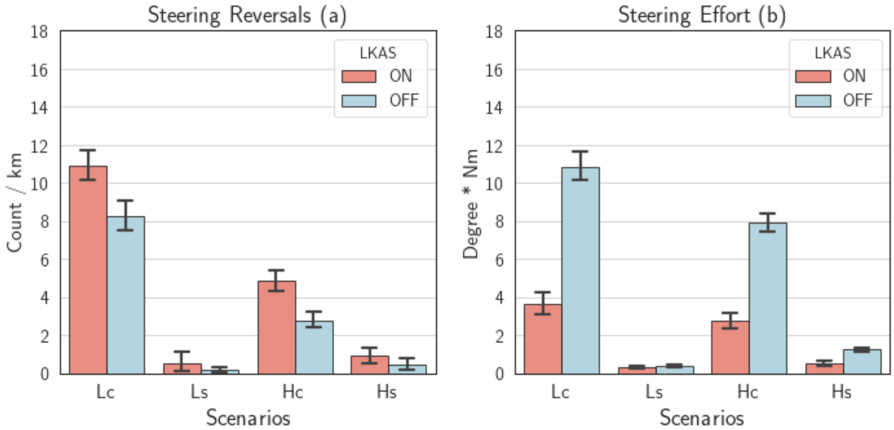

Figure 4. Driving performance measures. a) Steering Reversals b) Steering Effort (Lc: Low-speed-curvy, Ls: Low-speed-straight, Hc: High-speed-curvy, Hs: High-speed-straight)

Subjective measures

A two-way (LKAS x velocity) ANOVA was performed on the results of R-TLX. It was observed that the LKAS had a main effect on the subjective rating of mental workload ($F(1, 60) = 6.17$, $p = .016$). In contrast, no difference in mental workload was found between two velocity settings. There was also no interaction effect. The result of NASA-TLX is presented in Figure 5.

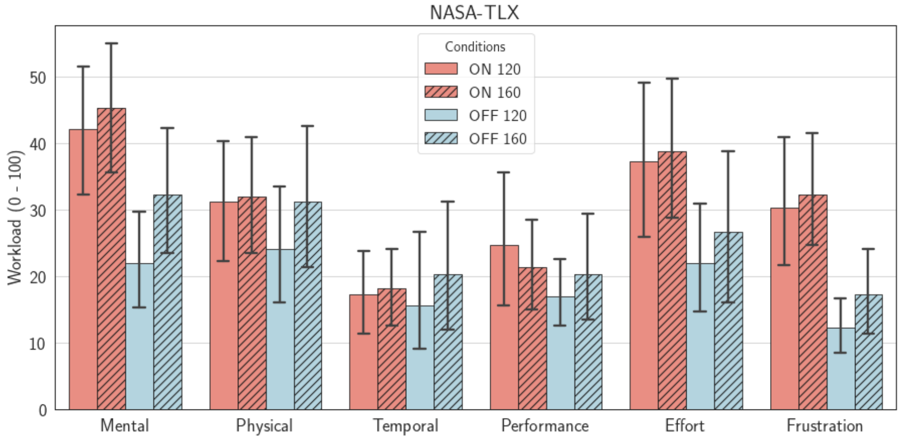

Figure 5. These subjective ratings were collected after the each lap. Depends the order of introducing LKAS, the participant would answer in each lap for either LKAS ON or OFF, in both velocity conditions (120 and 160 km/h). For analysis, the R-TLX was obtained by averaging the score of subjective workload over six subscales.

Discussion

The result of R-TLX shows that the participants rated their mental workload overall higher when the LKAS was switched on. This result supports H1 that the LKAS increases drivers' perceived workload, which is also in line with our pilot study (Seidler & Schick, 2018). However, the perceived mental workload was not different in low and high-speed scenario. This different result might be due to a small sample size, or driver's experience with automation (Stapel et al., 2017). Out of six subscales from the NASA-TLX, it can be seen in Figure 5 that the difference between two LKAS conditions (ON vs. OFF) was particularly huge in the subscale mental demand, effort and frustration. The reasons may be explained by the objective data.

The analysis of SRR (Figure 4a) reveals that, in the curvy scenarios, drivers performed more counter-steering to correct the trajectory while driving with LKAS. This suggests that constantly correcting LKAS's output might be annoying and disturbing, which results in frustration and mental effort, regardless of the steering effort under the same LKAS setting was actually lower (Figure 4b). This is an interesting result, as reducing physical workload is one of the goals of ADAS (Tanaka et al., 2000). However, this contradictory observation shows that the drivers might prefer applying more steering torque (manual driving) than having an assistance system that reduces physical workload but requires more mental effort.

In contrast to performance measures, the physiological data only partially supports the hypothesis H1 that driving with LKAS induces mental workload. On one hand, the SCR in Figure 3 demonstrats that the LKAS introduced a significant effect on the average SCR/km in different laps. On the other hand, the HR did not show a statistical difference between LKAS conditions (ON vs. OFF). Instead, it is only found that the HR was higher in the curvy scenarios (Lc, Hc) than the straight scenario (Ls). This result is however expected, because the task of keeping a vehicle between lanes depends highly on a psychomotor eye-hand coordination of the driver (De Waard, 1996). Although the observation in Figure 2 could be another evidence that LKAS induces mental workload (as decreasing HR over time was not found in both groups), it is known that higher HR does not necessarily correlate with increasing mental workload, since HR is also sensitive to physical workload e.g. as a result of steering reversals (Jahn et al., 2005).

In terms of driving performance (H2), the objective data reveals that LKAS did not improve the lane keeping performance. Even though the SDLP was significantly higher in the curvy scenario (Hc) than other three scenarios when driving without LKAS, it is difficult to conclude that the driver experienced more mental workload here, since the result of HR did not support this observation. Moreover, it is still an open question whether the measure SDLP could truly reflect mental workload in a field study, despite the fact that data during overtaking and lane changing events were excluded. As Östlund et al. (2005) points out, the width of the route and observation window may heavily influence the reliability of this measure.

Finally, it is realized that certain driving performance measures e.g. SRR and SDLP, though may be helpful interpreting the driving behaviour, are not ideal for examining mental workload associated with LKAS. The argument is that LKAS's performance (whether it applies enough torque or counter steers at the right time) is often associated with curves, in which the lateral position/control is heavily influenced by the system itself. This means that even if the performance measure can truly mirror the variation of mental workload, the interpretation would also not be easy. In this case, subjective measure (NASA-TLX) is a relatively robust way to assess mental workload.

Conclusion

Overall, the steering effort has shown that the LKAS has reduced physical workload significantly, particularly in the curvy scenarios. However, the reduced physical effort did not result in a better lane keeping performance as no difference of SDLP was observed. Moreover, the result of NASA-TLX shows that drivers experienced higher mental demand and frustration while interacting with LKAS. This could be explained by the frequent steering reversals required to correct the driving trajectory while driving with LKAS, as shown in the SRR. This increasing physical activity possibly led to an increase in HR, which results in difficulties in assessing changes in mental workload from this psychophysiological measure. However, the difference in count of SCR suggests that the driver could be annoyed or surprised by the LKAS behaviour. Therefore, the results from this study suggests that mental workload is higher when driving with this tested LKAS.

References

Aasman, J., Mulder, G., & Mulder, L.J.M. (1987). Operator Effort and the Measurement of Heart-Rate Variability. *Human Factors, 29*, 161–170.

Banks, V.A., & Stanton, N.A. (2016) Keep the driver in control: Automating automobiles of the future. *Applied Ergonomics, 53*, 389-395

Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C., & Winner, H. (2014). Three Decades of Driver Assistance Systems: Review and Future Perspectives. *IEEE Intelligent Transportation Systems Magazine, 6*, 6–22.

Brookhuis, K.A., De Waard, D., & Janssen, W.H. (2001). Behavioural impacts of Advanced Driver Assistance Systems - an overview. *European Journal of Transport and Infrastructure Research. 1*. 245–253

Carsten, O.M.J., Lai, F.C.H., Barnard, Y., Jamson, A.H., & Merat, N. (2012) Control Task Substitution in semiautomated driving: Does it matter what aspects are automated? *Human Factors, 54*, 747–761

De Waard, D., & Brookhuis, K.A. (1991). Assessing driver status: A demonstration experiment on the road. *Accident Analysis & Prevention, 23*, 297–307.

De Waard, D. (1996). *The measurement of drivers' mental workload*. PhD thesis, University of Groningen. Groningen, the Netherlands: The Traffic Research Centre.

De Winter, J.C.F., Happee, R., Martens, M.H., & Stanton, N.A. (2014) 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, 27*, 196–217

Hancock, P.A., & Parasuraman, R. (1992). Human factors and safety in the design of intelligent vehicle-highway systems (IVHS). *Journal of Safety Research, 23*, 181–198.

Hart, S.G., & Staveland, L.E. (1988) Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, *Advances in Psychology, 52*, 139-183.

Ishida, S., & Gayko, J.E. (2004) Development, evaluation and introduction of a lane keeping assistance system. *IEEE Intelligent Vehicles Symposium.* IEEE

Jahn, G., Oehme, A., Krems, J.F., & Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*, 255–275

Östlund, J., Peters, B., Thorslund, B., Engström, J., Markkula, G., Keinath, A., Horst, D., Juch, S., Mattes, S., & Foehl, U. (2005). Driving performance assessment methods and metrics. Deliverable 2.2.5, AIDE Integrated Project, IST-1-507674-IP.

Sarter, N.B., Woods, D.D., & Billings, C.E. (1997) Automation surprises. In G. Salvendy (Eds.), *Handbook of Human Factors & Ergonomics, second edition* (pp. 1926-1943). Wiley, New York

Schick B., Seidler C., Aydogdu S., & Kuo Y.J. (2019) Driving experience vs. mental stress with automated lateral guidance from the customer's point of view. In P. Pfeffer (Eds.) *9th International Munich Chassis Symposium 2018*. (pp. 27-44). Wiesbaden: Springer Verlag

workload evaluation of effects of a lane keeping assistance system

Seidler, C., & Schick, B. (2018). Stress and workload when using the lane keeping assistant: Driving experience with advanced driver assistance systems. *27th Aachen Colloquium - Automobile and Engine Technology*. Aachen. Germany

Stanton, N.A., & Marsden P. (1996) From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science, 24*, 35-49

Stapel, J., Mullakkal Babu, F.A., & Happee, R. (2017). Driver Behavior and Workload in an On-road Automated Vehicle. In *Proceedings Road Safety & Simulation International Conference 2017*.

Tanaka, J., Ishida, S., Kawagoe H., & Kondo, S. (2002). Workload of using a driver assistance system. *ITSC2000. 2000 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.00TH8493).* IEEE

Wilson, G.F., & Eggemeier, F.T. (1991). Physiological measures of workload in multi-task environments. In D. Damos (Eds). *Multiple Task Performance*. (pp. 329-359) London: Taylor & Francis.

Young, M.S., & Stanton, N.A. (2002) Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors, 44*, 365-375

Zangróniz, R., Martínez-Rodrigo, A., Pastor, J.M., López, M.T., & Fernández-Caballero, A. (2017). Electrodermal Activity Sensor for Classification of Calm/Distress Condition. *Sensors (Basel, Switzerland), 17*, 2324.