DRIVERS' RELIANCE ON LANE KEEPING ASSISTANCE SYSTEMS: EFFECTS OF DIFFERENT LEVELS OF ASSISTANCE

Anke Popken (Chemnitz University of Technology, Germany),
Lena Nilsson (VTI, Sweden),
Josef F. Krems (Chemnitz University of Technology, Germany)

anke.popken@phil.tu-chemnitz.de

ABSTRACT: Overreliance on Advanced Driver Assistance Systems (ADAS) and a lack of involvement in the driving task reflect current safety-concerns in connection with the increasing automation of the driving task due to ADAS. It is hypothesised that those effects gain relevance the more ADAS intervene in the cognitive and regulatory processes underlying driving. In a driving simulator study drivers' preparedness to divert attention away from the driving task was investigated as a function of the level of assistance. Forty-five drivers drove 55 km on a simulated rural road with either one of three levels of lane keeping assistance: (1) a high level of assistance realised by a Heading Control system, (2) a low level of assistance realised by a Lane Departure Warning system, and (3) no assistance (control condition). Drivers' attention allocation to a visually demanding secondary task served as a measure for their reliance on the systems.

1 Introduction

In an attempt to increase road safety and to reduce the number of traffic deaths and injuries, active safety in-vehicle technologies gain more and more importance as a complement to passive safety technologies. The main aim is not only to reduce actual crash consequences, but also to minimise crash risk. Advanced Driver Assistance Systems (ADAS) are electronic in-vehicle systems that offer support to the driver while driving and aim at enhancing driver (and traffic) safety and comfort. ADAS intervene to different degrees in the driving task and in vehicle control. Although ADAS are supposed to positively influence driver behaviour and safety, the full range of possible adaptation effects in response to the introduction of these system is at present difficult to predict (e.g., [1]). Safety concerns with ADAS are primarily related to possible negative human performance consequences that arise in conjunction with the automation of task processes, as extensively studied in other domains like aviation ([2]). Because ADAS - according to their type and functionality - intervene in different levels of the control processes involved in driving, they change the way the driving task is performed and affect the information processing and action regulation processes involved in driving. The "relieving" function of ADAS may indeed positively influence driver performance and comfort in non-critical ("normal") driving situations, but may have negative impacts in situations that suddenly demand for drivers' intervention; for instance situations that lie beyond the ADAS' functional scope or disturbing events such as system failures or malfunctions.

In aviation research, evidence was found for problematic adaptation effects such as overreliance on systems and inappropriate trust in system capabilities ([3, 4, 5, 6, 7, 8, 9]), complacency ([10, 11, 12]), vigilance problems and high mental workload due to system monitoring demands ([13, 14, 15, 16, 17, 18]), underload ([19, 20, 21]), and reduced Situation Awareness ([22, 23, 24, 25, 26]).

It is suggested that the occurrence and size of possible problematic adaptation effects is related to the degree the operator is taken "out-of-the-loop" from the controlled processes (e.g. [25, 27, 28]). Thus, system characteristics like the design of the human-machine interaction and the systems' performance (e.g., reliability) play an important role in the evolution of adaptation effects. A number of studies demonstrated the strong relationship between system reliability and its effects on human trust and monitoring behaviour ([29, 30, 4, 6, 7, 9, 10]). People tend to rely on and to neglect monitoring of highly reliable systems, after they have gained a sufficient level of trust in their performance. Highly reliable systems reduce the necessity for human intervention in the automated processes. A reduced involvement in the task, coupled with changes in cognitive (information processing) and regulatory processes (e.g., vigilance) are most likely consequences. Similar effects are predicted for the impact of the level of automation, i.e. the level of control the operator has in the human-machine interaction (e.g. [22]). In the automotive area, overreliance on ADAS and a lack of involvement in the driving task (e.g., reduced Situation Awareness, attention shifts to driving-unrelated activities) reflect major safety-concerns in response to the introduction of ADAS. In this study drivers' preparedness to divert attention away from the driving task is investigated as a function of the level of (lane keeping) assistance. Thereby the focus laid on an initial interaction with a system under high demanding conditions that were designed to promote reliance on the system. Of special interest are potential changes in drivers' reliance after encountering critical situations where they have to override the system or to ignore the system's actions.

2 Methods

The VTI advanced moving base driving simulator III in Linköping, Sweden, with a 120 degrees horizontal field of view was used in this study (Figure 1). Participants drove six times on a rural road of 11 km length with one driving lane in each direction. The lane width was 3,5 m. The speed limit during the study was 70 km/h.

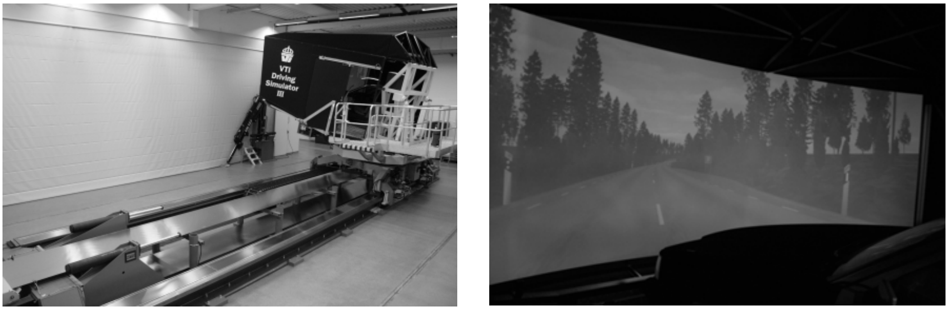

Fig. 1. The VTI advanced driving simulator III.

2.1 Participants

Forty-five participants (23 male, 22 female) aged between 25 and 45 years (*M* = 33,7; *SD* = 6,4) took part in this study. They had been driving at least 5000 km during the last year. Most of them had been driving in the VTI driving simulator before.

2.2 Levels of lane keeping assistance

The 45 participants were randomly assigned to one of three experimental conditions (15 participants in each condition) representing three different levels of assistance in lane keeping: (1) Driving with a Heading Control system (HC condition) corresponding to the highest level of lane keeping assistance, (2) driving with a Lane Departure Warning system (LDW condition) corresponding to a low level of lane keeping assistance, and (3) driving manually (MD condition) without any assistance in lane keeping (control condition). The LDW system warned the driver via a vibration in the steering wheel when he or she crossed the lane markings. The HC system guided the drivers within the lane by applying counter forces on the steering wheel as soon as they were approaching the left or right lane marking. Both systems were 100% reliable during the whole study.

2.3 Secondary Task

Concurrently to driving participants performed a visually demanding secondary task: The Arrows Task developed in the European HASTE project. The visual distraction caused by this task should provoke lateral path deviations and thus enable drivers to experience the actions of the lane keeping assistance systems and to develop trust in the system. Furthermore, drivers' attention allocation to the primary driving task and the secondary task was used as an indicator of their reliance on the assistance systems (compared to non-assisted driving).

The Arrows Task was presented on a 7 inches TFT touch screen mounted in the centre console of the simulator vehicle. The secondary task display was positioned on a horizontal eccentricity of about 22 to 37 degrees visual angle to the left of the driver's straight-ahead line of sight and on a vertical eccentricity of about 13 to 22 degrees visual angle below the driver's straight-ahead line of sight. The Arrows Task started on certain road marks which were the same for each driver in each condition. The duration of one Arrows Task was 30 s. One Arrows Task consisted of a series of displays presenting different configurations of 16 arrows pointing in different directions. For each display drivers had to decide if there was an arrow present pointing upward and to accordingly press the “yes” or “no” buttons in the upper or lower part of the screen, respectively. The time out for each display was 5 s. Response times above 5 s were recorded as missing responses. Each presentation of a new display was signalled by a sound. Additionally, drivers got an auditory and visual performance feedback.

2.4 Procedure

Participants were instructed to drive as they would usually drive on a real road with corresponding character concerning interaction with other traffic participants, speed choice, manoeuvring, and time gaps. Participants were instructed to use the turn indicator to signal their intention to depart from the driving lane, e.g. during lane change manoeuvres. They were informed that the activation of the indicator would suppress the lane departure warning and deactivate the HC system’s steering torque. Participants in the HC condition were additionally told that they could principally override the system if they applied a counter force to the steering wheel greater than the maximum steering force of the HC system.

Participants started with a training session in order to become accustomed to driving with the simulator and with the LDW and HC system. During the training session participants were instructed to perform some lane change manoeuvres in order to experience the systems’ functionality. After that participants practised the secondary task while the simulator vehicle was parked. Participants were instructed to perform the Arrows Task as well and as fast as possible while driving, but without jeopardising traffic safety.

Subsequently participants performed two experimental drives. During the first drive (about 15 min.) participants were supposed to build up a stable level of trust in the system and adapt their time-sharing strategies accordingly. During the second drive (about 25 min.) drivers encountered eight critical situations in which they had to react accordingly in order to not compromise traffic safety or violate the traffic rules. The LDW and the HC system were 100% reliable during the whole study (no system failures occurred). However, the critical situations lay partly outside the systems’ functional limits and thus, drivers had to override the systems. Of special interest are the potential changes in drivers’ trust, reliance, and workload as a function of their experience of these critical situations.

2.5 Dependent Measures

Performance-based measures of drivers’ time-sharing strategies (eye glance behaviour, secondary task performance) were used as indicators of their reliance on the lane keeping assistance systems. Additionally various subjective measures were collected after the two experimental sessions in order to assess amongst other things drivers’ alertness and mental workload, their perception of system performance, their trust in the systems, and their opinions about how such systems may influence their safety and driving behaviour.

3 Results

All measures were subject to a 3 x 3 factorial repeated measures ANOVA with the level of assistance (Heading Control, Lane Departure Warning, No Assistance) treated as between-subjects factor and the session/criticality of driving situation (session one – no critical driving situations, session two – no critical driving situations, session two – critical driving situations) treated as within-subjects factor. Some first results of the analysis are presented below.

3.1 Secondary Task Performance

Drivers performed 19 Arrows Tasks in the first experimental session and 25 Arrows Tasks in the second experimental session.

3.1.1 Response Accuracy

Response accuracy was generally very high for all experimental conditions. Unassisted drivers had a higher percentage of correct answers and a lower percentage of false and missed responses in the Arrows Task than LDW drivers and HC drivers in both experimental sessions. This difference was however not significant. The performance of the unassisted drivers improved from session one to session two (in terms of percentage of correct and false responses), whereas the performance of the LDW and HC drivers deteriorated from session one to session two. The interaction between level of assistance and session/criticality of driving situation was not significant. HC drivers generally missed more responses than LDW and unassisted drivers (n.s.). There was a highly significant effect ($p < 0.001$) for the occurrence of critical driving situations, in that response accuracy deteriorated considerably in critical situations compared to non-critical driving situations in both sessions.

3.1.2 Mean Number of responded displays

Each Arrows Task lasted for about 30 s. During one Arrows Task a number of displays representing different arrows configurations were shown. Dependent on how fast drivers responded to each single display, more or less displays were presented during the whole duration of one Arrows Task. Drivers in the HC condition on the average answered the smallest number of displays and LDW drivers answered the largest number of displays during one Arrows Task in non-critical driving situations. For HC and LDW drivers, performance in critical situations in session two was better than performance in non-critical situations in session one; whereas the performance of unassisted drivers was worse during critical situations in session two compared to non-critical situations in session one. There was neither a significant main effect for the level of assistance nor a significant interaction between the level of assistance and the number of session/criticality of driving situation. There was a significant performance improvement for drivers in all conditions in non-critical driving situations from session one to session two ($p < 0.001$).

3.1.3 Reaction time

Reaction times were calculated for all displays except those drivers missed to respond to within five seconds (recorded as missing responses). LDW drivers and drivers in the no assistance condition reacted faster than HC drivers in non-critical driving situations. This effect was not statistically significant. There was a significant main effect for the session/criticality of driving situation factor, in that reaction times were significantly longer during critical situations compared to non-critical situations in both sessions for all levels of assistance ($p < 0.001$). Reaction times decreased from session one to session two but this effect was not significant.

3.2 Eye glance behaviour

This analysis is based on a total number of more than 13000 display glances.

3.2.1 Mean single glance durations to the secondary task display

HC drivers had the longest mean single glance durations to the display in session one and the shortest mean single glance durations in session two. Mean single glance durations to the display decreased from session one to session two, but this difference did not yield statistical significance in post-hoc analyses. However, there was a significant effect for the criticality of the driving situation, in that mean single glance durations to the display were significantly shorter during critical situations compared to non-critical situations ($p < 0.001$). During critical situations, unassisted drivers had the longest display glance durations compared to HC and LDW drivers.

3.2.2 Mean display glance frequencies

LDW drivers had the highest frequency of glances to the display; and HC drivers had the lowest display glance frequency in session one. During session two, HC driver looked less often to the Arrows Task display than LDW and unassisted drivers. Unassisted drivers looked most often to the display in session two. There was no significant effect for the level of assistance and no significant interaction between the level of assistance and session number/ criticality of situation.

There was however a highly significant main effect ($p < 0.001$) for the session number/ criticality of situation factor: Drivers in all conditions looked significantly less often to the display in session two (compared to session one), and looked least often to the display during critical driving situations.

4 Discussion

The results show that drivers adapted their task related effort to driving task demands. In critical driving situations, drivers allocated fewer resources to the secondary task, confirmed by a significantly worse secondary task performance and significantly shorter and fewer glances to the secondary task display in critical driving situations.

There is also evidence for a general training effect for drivers' performance in the secondary task. Drivers answered more displays in session two than in session one, and reaction times slightly improved. Furthermore, drivers needed significantly fewer glances to extract the relevant information from the secondary task display in session two compared to session one.

The results for the different levels of assistance show a more complex picture. There was no significant main effect of the level of assistance for any of the measures.

For the Arrows Task performance, there was a surprising finding in that unassisted drivers performed better than HC and LDW drivers in terms of response accuracy. This difference was also obvious in critical driving situations. A further interesting finding was that HC and LDW drivers’ Arrows Task performance deteriorated from session one to session two (in terms of percentage of correct and false responses), whereas performance of unassisted drivers improved. HC drivers had the highest number of missing responses during both sessions and especially in critical driving situations. In critical driving situations there seems to be a tendency for LDW drivers of tempting to respond to displays rather than to omit them (even when the answer is false), whereas HC drivers rather drop the secondary task. One possible explanation for these findings could be that HC drivers had to invest some additional resources in system control, which was particularly necessary in session two when they had to recognise critical driving situations which required them partly to override the HC system.

There seems to be some further evidence for interpretation in the results of the eye glance measures. For HC drivers both duration and frequency of glances to the secondary task display decreased from session one to session two. HC had the longest display glance durations in session one, and the shortest display glance durations in session two. For LDW and unassisted drivers there was mainly a difference in glance frequencies from session one to session two (less frequent display glances in session two), whereas display glance durations were similar in both sessions. Unassisted drivers had the longest and also the most frequent glances to the display in critical driving situations.

Further analysis of other glance behaviour measures (e.g. the percentage of time spent looking to the display during one Arrows Task) and the relation of the behavioural measures with the subjective data (e.g. mental workload) will complete the picture of drivers resource allocation strategies when driving with different levels of lane keeping assistance.

5 Acknowledgements

This study was funded by a grant for infrastructure-sharing from the HUMANIST Network of excellence (HUMAN centred design for Information Society Technologies; www.noehumanist.org). We would like to thank the following persons at VTI for their overwhelming support in conducting this study: Anne Bolling (for a lot of fruitful discussions and the timely management of the study), Beatrice Söderström, Birgitta Thorslund and Jerker Sundström (for being experimental leaders), Håkan Jansson and Laban Källgren (for realising the simulator scenarios), Göran Palmkvist (for implementing the lane keeping assistance systems), Mikael Adlers (for implementing the secondary task), and

Jan Törnros (for translating the always newest version of the questionnaires and instructions into swedish).

6 References

[1] Saad, F., Hjälmdahl, M., Cañas, J., Alonso, M., Garayo, P., Macchi, L., Nathan, F., Ojeda, L., Papakostopoulos, V., Panou, M., Bekiaris, E.: 'AIDE Deliverable 1_2_1: Literature review of behavioural effects' (Subproject SP1, Workpackage 1.2), 2004, Available at: http://www.aide-eu.org/pdf/aide_d1-2-1final.pdf

[2] Parasuraman, R., and Mouloua, M.: 'Automation and human performance: Theory and applications' (Lawrence Erlbaum Associates, 1996)

[3] Muir, B. M.: 'Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems', Ergonomics, 1994, 37(11), pp. 1905-1922

[4] Muir, B. M. and Moray, N.: 'Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation', Ergonomics, 1996, 39(3), pp. 429-460

[5] Lee, J. D., and See, K. A.: 'Trust in automation: Designing for appropriate reliance', Human Factors, 2004, 46(1), pp. 50-80

[6] Lee, J. D., and Moray, N.: 'Trust, self-confidence, and operators' adaptation to automation', International Journal of Human-Computer Studies, 1994, 40, pp. 153-184

[7] Lee, J. D., & Moray, N.: 'Trust, control strategies and allocation of function in human-machine systems', Ergonomics, 1992, 35, pp. 1243-1270

[8] Parasuraman, R. and Riley, V.: 'Humans and automation: Use, misuse, disuse, abuse', Human Factors, 1997, 39(2), pp. 230-253

[9] Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., and Beck, H. P.: 'The role of trust in automation reliance', International Journal of Human-Computer Studies, 2003, 58, pp. 697-718

[10] Parasuraman, R., Molloy, R., and Singh, I. L.: 'Performance consequences of automation-included "complacency"', The International Journal of Aviation Psychology, 1993, 3(1), pp. 1-23

[11] Moray, N., and Inagaki, T.: 'Attention and complacency', Theoretical Issues in Ergonomics Science, 2000, 1(4), pp. 354-365

[12] Farrell, S., and Lewandowsky, S.: 'A connectionist model of complacency and adaptive recovery under automation', Journal of Experimental Psychology: Learning, Memory, and Cognition, 2000, 26(2), pp. 395-410

[13] Parasuraman, R., Molloy, R., Mouloua, M., and Hilburn, B.: 'Monitoring of automated systems', in Parasuraman, R. and Mouloua, M. (Eds.): 'Automation and human performance: Theory and applications' (Lawrence Erlbaum Associates, 1996), pp. 91-115

[14] Warm, J. S., Dember, W. N., and Hancock, P. A.: 'Vigilance and workload in automated systems', in Parasuraman, R. and Mouloua, M. (Eds.): 'Automation and human performance: Theory and applications' (Lawrence Erlbaum Associates, 1996), pp. 183-200

[15] Molloy, R., and Parasuraman, R.: 'Monitoring an automated system for a single failure: Vigilance and task complexity effects', Human Factors, 1996, 38(2), pp. 311-322

[16] Hancock, P. A., and Warm, J. S.: 'A dynamic model of stress and sustained attention', Human Factors, 1989, 31(5), pp. 519-537

[17] Desmond, P. A., and Hancock, P. A.: 'Active and passive fatigue states', in Hancock, P. A. and Desmond, P. A. (Eds.): 'Stress, workload, and fatigue' (Lawrence Erlbaum Associates, 2001), pp. 455-465

[18] Parasuraman, R.: 'Human-Computer monitoring', Human Factors, 1987, 29(6), pp. 695-706

[19] Matthews, G., and Desmond, P. A.: 'Task-induced fatigue states and simulated driving performance', The Quarterly Journal of Experimental Psychology, 2002, 55A (2), pp. 659–686

[20] Desmond, P. A., Hancock, P. A., and Monette, J. L.: 'Fatigue and automation-induced impairments in simulated driving performance', Transportation Research Record, 1998, 1628, pp. 8-14

[21] Carsten, O.M.J., and Nilsson, L.: 'Safety Assessment of Driver Assistance Systems', European Journal of Transport and Infrastructure Research, 2001, 1(3), pp. 225-243

[22] Endsley, M. R.: 'Automation and situation awareness', in Parasuraman, R. and Mouloua, M. (Eds.): 'Automation and human performance: Theory and applications' (Lawrence Erlbaum Associates, 1996), pp. 163-181

[23] Kaber, D. B., and Endsley, M. R., 'The effects of level of automation on human performance, situation awareness and workload in a dynamic control task', Theoretical Issues in Ergonomics Science, 2004, 5(2), pp. 113-153

[24] Endsley, M. R., and Kaber, D. B., 'Level of automation effects on performance, situation awareness and workload in a dynamic control task', Ergonomics, 1999, 42(3), pp. 462-492

[25] Endsley, M. E., and Kiris, E. O.: 'The out-of-the-loop performance problem and level of control in automation', Human Factors, 1995, 37(2), pp. 381-394

[26] Adams, M. J., Tenney, Y. J., and Pew, R. W.: 'Situation Awareness and the cognitive management of complex systems', Human Factors, 1995, 37(1), pp. 85-104

[27] Buld, S., and Krüger, H.-P.: 'Wirkungen von Assistenz und Automation auf Fahrerzustand und Fahrsicherheit. Projekt: EMPHASIS (Effort-Management und Performance-Handling in sicherheitsrelevanten Situationen, 19 S 9812 7, Bundesministerium für Bildung und Forschung), Würzburg: Interdisziplinäres Zentrum für Verkehrswissenschaften an der Universität Würzburg, 2002

[28] Parasuraman, R., Sheridan, T. B., and Wickens, C. D.: 'A model of types and levels of human interaction with automation', IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 2000, 30(3), pp. 286-297

[29] Wiegmann, D. A., Rich, A., & Zhang, H.: 'Automated diagnostic aids: The effects of aid reliability on users' trust and reliance', Theoretical Issues in Ergonomics Science, 2001, 2(4), pp. 352-367

[30] Moray, N., Inagaki, T., & Itoh, M.: 'Adaptive automation, trust, and self-confidence in fault management of time-critical tasks', Journal of Experimental Psychology: Applied, 2000, 6(1), pp. 44-59