Designing the Transition from Highly Automated Driving to Manual Driving: A Study of Drivers' Preferences

Stefan Brandenburg, Technische Universität Berlin, Germany, *stefan.brandenburg@tu-berlin.de*,

Sandra Epple, Technische Universität Berlin, Germany

ABSTRACT

One important aspect of highly automated driving is the transition of control between automated vehicle and driver. Previous research showed that the design of take-over requests (TORs) influences the success of this shift of control. In an online study with 53 participants (26 females), we examined drivers' preferences regarding TOR modality (auditory and visual) and TOR procedure (one or two steps). In one-step TOR procedures a TOR is presented at a single point in time. In contrast, in two-step TOR procedures drivers are informed about the transition of control at two points in time: an initial warning followed by an alarm. The study's findings show that a two-step TOR procedure is preferred to a one-step TOR procedure. Two-step TORs are rated as more intuitive, useful, attractive and more appropriate in displaying information than one-step procedures. Moreover, participants preferred verbal auditory TORs (speech) to non-verbal TORs (tone). Implications on the design of TOR interfaces for highly automated driving are discussed.

Keywords: Highly automated driving, HMI design, take-over request, subjective preferences, TOR modality, two-step TOR procedure.

1 INTRODUCTION

Highly automated cars are estimated to be on European roads between 2020 and 2025 (ERTRAC, 2015). They were anticipated for a long time and became more likely through the advances in computerisation, the development of on-board sensors (Walker, Stanton, & Young, 2001) and changes in legal restrictions (ECE/TRANS/WP.1/145, 2014). Amendments made to the 1968 Vienna Convention on road traffic, allow for the transition of vehicle control to an automated system, as long as drivers can resume control whenever needed (ECE/TRANS/WP.1/145, 2014). The transition of control between driver and automated vehicle is an important and safety-critical aspect of highly automated driving (e.g. Merat, Jamson, Lai, Daly, & Carsten, 2014). In order to make this transition of control as safe and comfortable as possible, drivers have to be provided with an understanding of the driving situation and of the steps necessary to resume control of the vehicle in a safe and comfortable manner. The design of take-over requests (TOR) can accelerate drivers' understanding of the situation and of necessary actions. The human-machine-interface (HMI) issues take-over requests and mediates the interaction between vehicle and driver. Therefore, the design of the HMI has a significant impact on safety outcomes of automated systems (Casner, Hutchins, & Norman, 2016). For instance, HMI design influences drivers' reaction times when resuming control of the vehicle (Forster, Naujoks, Neukum, & Huestegge, 2017). Ideally, drivers are able to intuitively understand the interface that issues a take-over request because intuitive interaction is fast, unconscious and automatic (Macaranas, Antle, & Riecke, 2015). But how should take-over requests be designed? Is there a combination of HMI design aspects that results in take-over requests with high usability, usefulness, and attractiveness? In the past, various different HMI designs have been used to issue take-over requests to drivers. Design aspects that have been examined in this context include TOR modality (e.g. auditory and visual TORs) and TOR procedure (e.g. one-step and two-step procedures). The term *TOR procedure* refers to the number of take-over requests that are presented within one take-over situation. In one-step TOR procedures a TOR is presented at a single point in time. In contrast, in two-step TOR procedures drivers are informed about the transition of control at two points in time: an initial warning followed by an alarm. Two-step TOR procedures have the potential to create a more gradual take-over process, by providing a time frame between warning and alarm, within which the driver can take-over control. According to Walch, Lange, Baumann, and Weber (2015), two-step TOR procedures are preferred to one-step TOR procedures. However, the authors did not investigate the impact of TOR modality on two-step TOR procedures. Research on TOR modalities has focused on auditory, visual, and tactile modalities in one-step TOR procedures (e.g. Forster et al., 2017; Politis, Brewster, & Pollick, 2015). The results of these studies suggest that verbal auditory TORs (speech) can be valuable for the design of TOR interfaces. According to Forster et al. (2017), adding verbal auditory information (speech) to non-verbal visual-auditory TORs lead to shorter reaction times and more positive subjective ratings. These findings are in contrast with many industrial prototypes that mostly rely on written text or pictograms appearing on the dashboard in combination with a single tone. Moreover, past research on TORs has focused on one-step TOR procedures only. Yet, no studies tested whether these findings are valid for two-step TOR procedures. The present study aims to close this gap by systematically varying TOR procedure and TOR modality. The first objective of this study is to examine the effect of TOR procedure (one-step or two-step) on drivers' preferences. The second objective is to investigate the impact of TOR modalities (auditory and visual) on drivers' preferences.

2 MATERIAL AND METHODS

2.1 Participants

A total of 53 participants (26 females) took part in the study. Their mean age was 32 years (*SD* = 9.86 years) and ranged from 20 to 64 years. Forty-eight participants (91%) possessed a valid drivers' license. The study was subject to evaluation of the local ethics committee.

2.2 Design of the human-machine interfaces for take-over requests

Participants experienced 8 different human-machine interfaces issuing take-over requests. These TOR interfaces differed regarding three dimensions: (1) TOR procedure (one-step vs. two-step), (2) the presentation of auditory information (via tone vs. via speech), and (3) the presentation of visual information (via text vs. via text and pictogram). In the condition with a one-step TOR procedure, an alarm was issued at a single point in time ("Alert - Take-over vehicle control now"). A two-step TOR procedure, on the other hand, consisted of an initial warning ("Warning - Roadworks ahead - Take-over vehicle control soon") followed by the alarm ("Alert - Take-over vehicle control now"). All TOR interfaces were multimodal as they contained auditory as well as visual information. Auditory information was either presented by a single tone, or by a mechanical voice reading the warning and the alarm (speech). Visual information was presented by a written text, or by text and additional pictograms appearing on the dashboard. The pictograms displayed a standard road works sign for the warning and the vehicles' pedals and steering wheel for the alarm if applicable. TOR interfaces were dynamic in the sense that they were presented as short film clips. For the one-step TOR the film clip was 8 seconds long and for the two-step TOR 16 seconds long. Each film clip followed the same plot. First, the dashboard was displayed for 3 seconds. Then the TOR appeared and lasted for 3 seconds. Finally, the TOR faded and the dashboard screen was displayed again for 2 seconds. In the two-step TOR procedure, two clips were displayed: an initial warning followed by the alarm. Examples of the TOR interfaces used in this study can be found in table 1.

Table 1 – Screenshots of 2 examples of the take-over request interfaces

<table>
<tr><th>Interface 1
one-step procedure with text and tone</th><th colspan="2">Interface 8
two-step procedure with text, pictogram, and speech</th></tr>
<tr><td>alarm</td><td>warning</td><td>alarm</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
</table>

2.3 Procedure

Participants were invited to this online-study via social media platforms like facebook™ and via university mailing lists. When they clicked on the invitation link they were forwarded to SociSurvey, the online tool that was used for data assessment. The online study consisted of four parts: (1) an introduction, (2) a general evaluation of each TOR interface, (3) a ranking of TOR interfaces, and (4) a detailed evaluation of the personal best and worst TOR interfaces. Part (4) will not be discussed any further in this paper due to limited space. First, participants read information about the general procedure and gave their informed consent and data assessment agreement. Then, a short introductory text instructed participants to imagine sitting in a highly automated car and reading text messages on their smart phone. After that, they experienced each of the 8 TOR interfaces separately and in randomized order. Participants were asked to evaluate each TOR interface regarding the following four items: (a) This TOR interface is intuitive, (b) I find this TOR interface useful, (c) I find this TOR interface attractive, and (d) I find the amount of information appropriate. Finally, participants sorted the interfaces in descending order with the best interface on rank 1 and the worst interface on rank 8.

3 RESULTS

3.1 General Evaluation of TOR interfaces

A within-subjects MANOVA revealed significant main effects for each of the independent variables: TOR procedure, visual modality, and auditory modality. Participants rated two-step TOR procedures as more intuitive ($F(1, 50) = 29.22$, $p < .001$, $\eta^2_{part} = .36$), more useful ($F(1, 50) = 24.87$, $p < .001$, $\eta^2_{part} = .33$), more attractive ($F(1, 50) = 21.35$, $p < .001$, $\eta^2_{part} = .29$), and more appropriate for displaying the information ($F(1, 50) = 25.80$, p

< .001, η^2_{part} = .34) than one-step TOR procedures. Moreover, participants indicated that speech was more intuitive (F(1, 50) = 5.14, p = .02, η^2_{part} = .09), more useful (F(1, 50) = 7.04, p = .01, η^2_{part} = .12) and provided information more appropriately than a single tone (F(1, 50) = 5.54, p = .02, η^2_{part} = .10). Lastly, text and pictogram received higher ratings with respect to intuitiveness (F(1, 50) = 6.38, p = .01, η^2_{part} = .11), and attractiveness of the TOR interface (F(1, 50) = 7.86, p = .01, η^2_{part} = .13) than mere text. Moreover, the MANOVA revealed significant interactions between auditory and visual modality on usefulness (F(1, 50) = 4.62, p = .03, η^2_{part} = .08) and appropriateness of information (F(1, 50) = 4.56, p = .03, η^2_{part} = .08). Text and pictogram were rated as more useful and as more appropriate in displaying information than pure text, when a tone was used as auditory information. However, when speech was used as auditory information, there was no difference between the two visual modality conditions with respect to usefulness and appropriateness of information. All other main effects and interactions were not significant.

3.2 TOR interface ranking

Analysing the ranking of the 8 different TOR interfaces, results showed that each interface appeared on all possible ranks at least once. However, clear preferences became evident when analysing the median rank of each TOR interface. Friedman's ANOVA for non-parametric data showed a significant difference indicating that the rank distributions of the interfaces are not similar, p < .001. Table 2 shows each interface, its median rank and the corresponding interquartile range as measure of spread.

Table 2 – Median ranks of take-over request (TOR) interfaces

			Design aspect		
Rank	**Interface**	**Median rank (interquartile range)**	**One-step or two-step**	**Text or text + pictogram**	**Tone or speech**
1 (best)	No. 7	2 (1-2.5)	Two-step	Text	Speech
2	No. 8	3 (1-5)	Two-step	Text + pictogram	Speech
3	No. 5	4 (2-5)	Two-step	Text	Tone
4	No. 6	4 (3-5)	Two-step	Text + pictogram	Tone
5	No. 3	5 (3.5-6)	One-step	Text	Speech
6	No. 4	6 (5-6)	One-step	Text + pictogram	Speech
7	No. 1	6 (4-7)	One-step	Text	Tone
8 (worst)	No. 2	7 (6-8)	One-step	Text + pictogram	Tone

Note. The interquartile range is a measure of spread that captures 50% of the values around the median.

Table 2 shows that TOR interface No. 7 received the lowest median rank and therefore was rated the best interface. It combined text with speech in a two-step procedure. Interface No. 8, which contained an additional pictogram, received the second lowest median rank. The TOR interface with the highest rank (worst interface) was interface No.2. It consisted of a single tone and a combination of text and pictogram in a single step. Moreover, table 2 shows that all two-step TOR conditions outranked all one-step TOR conditions. Within the two-step and one-step TOR procedure conditions, speech outranked pure tone.

4 DISCUSSION

The present study had two objectives. The first objective was to examine the effect of TOR procedure (one-step or two-step) on drivers' preferences. The second objective was to investigate the impact of TOR modalities (auditory and visual) on drivers' preferences.

Regarding the first objective, our findings show that the two-step TOR procedure is preferred to the one-step TOR procedure. The two-step procedure was rated as more intuitive, useful, attractive and more appropriate in displaying the relevant information than the one-step procedure. Moreover, median ranks indicate that all two-step TOR conditions outrank all one-step TOR conditions. In line with the study of Walch et al. (2015), our findings suggest that two-step TOR procedures are in fact superior to one-step TOR procedures with regard to preference ratings. Two-step TOR procedures have the potential to create a more gradual take-over process, by providing a time frame between warning and alarm, within which the driver can take-over control. Therefore, they are suited for uncritical take-over situations with enough time available to switch control (e.g. at the end of a phase of highly automated driving on an autobahn). However, two-step TOR procedures might not be applicable in highly dynamic situations where driver reactions have to follow the TOR promptly. Future studies should examine the applicability of two-step TOR procedures in varying driving situations.

Concerning the second objective, our results indicate that a visual TOR consisting of text and pictograms was rated as more intuitive and attractive than a TOR consisting of text only. Moreover, a visual TOR consisting of text and pictograms was rated as more useful and appropriate in displaying the information in the auditory condition with a single tone. This effect could not be found in the auditory condition with speech. These findings suggest that adding pictograms to a text can be more intuitive and aesthetic but it does not increase the usefulness of a TOR containing speech any further. During periods of highly automated driving, the drivers' visual attention is likely to be involved in secondary tasks or in monitoring the driving situation. His or her capacity to process additional visual information of the TOR might be very limited. Hence, dashboard displays should not contain excessive verbal information. Instead, verbal information could be conveyed via the auditory channel. Our findings show that an auditory TOR consisting of speech was rated as more intuitive and more useful than a single tone. Moreover, speech outranked pure tone within two-step and one-step procedure conditions. These results are in line with Forster et al. (2017), who found auditory TORs containing verbal information to be superior to non-verbal TORs for one-step TOR procedures. Future studies should further investigate the potential of non-visual car-to-driver interaction (Gellatly, Hansen, Highstrom, & Weiss, 2010). Pictograms, text, and beeps might no longer be the primary way of communicating information from vehicle to driver. Instead, the interaction of drivers and vehicles should be based on meaningful auditory information.

The current study examined subjective preferences of TOR interfaces. It demonstrated that a TOR design containing a two-step procedure and verbal auditory information (speech) is preferred by drivers. However, shortcomings of the study should also be noted here. Firstly, the study was conducted as an online study. Thus, there was no controlled testing environment (e.g. screen size), no real driving experience, and no assessment of driving performance. Future studies should apply the TOR interfaces in real driving scenarios. Moreover, the manipulation of TOR modalities was not exhaustive. Future studies could investigate the impact of non-verbal visual (pictogram only) und hybrid auditory (speech and tone) TORs on drivers' preferences. Despite the shortcomings of the study, the current investigation addresses an important aspect in the development of highly automated vehicles. The success of automated vehicles partially depends on their human-machine interfaces (Casner et al., 2016). The open question on how to design HMI for take-over requests in highly automated driving is therefore one of the very important research questions in the developing research field. The present study added some information on the design of human-machine interfaces. It highlighted the value of two-step TOR procedures and verbal auditory information when issuing take-over requests. Future studies could examine their characteristics and how they influence driver behaviour.

REFERENCES

Casner, S. M., Hutchins, E. L., & Norman, D. (2016). The challenges of partially automated driving. *Communications of the ACM*, *59*(5), 70–77. https://doi.org/10.1145/2830565

ECE/TRANS/WP.1/145. (2014). *Report of the sixty-eighth session of the Working Party on Road Traffic Safety*. Geneva.

ERTRAC. (2015, July 21). Automated Driving Roadmap Technical Report of the European Road Transport Research Advisory Council. European Union. Retrieved from http://www.ertrac.org/uploads/documentsearch/id38/ERTRAC_Automated-Driving-2015.pdf

Forster, Y., Naujoks, F., Neukum, A., & Huestegge, L. (2017). Driver compliance to take-over requests with different auditory outputs in conditional automation. *Accident Analysis & Prevention*, *109*, 18–28.

Gellatly, A. W., Hansen, C., Highstrom, M., & Weiss, J. P. (2010). Journey: General Motors' move to incorporate contextual design into its next generation of automotive HMI designs. In *Proceedings of the 2nd international conference on automotive user interfaces and interactive vehicular applications* (pp. 156–161). ACM.

Macaranas, A., Antle, A. N., & Riecke, B. E. (2015). What is Intuitive Interaction? Balancing Users' Performance and Satisfaction with Natural User Interfaces. *Interacting with Computers*, *27*(3), 357–370.

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 274–282. https://doi.org/10.1016/j.trf.2014.09.005

Politis, I., Brewster, S., & Pollick, F. (2015). Language-based multimodal displays for the handover of control in autonomous cars. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 3–10). Nottingham: ACM Press.

Walch, M., Lange, K., Baumann, M., & Weber, M. (2015). Autonomous driving: investigating the feasibility of car-driver handover assistance. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 11–18). Nottingham: ACM Press.

Walker, G. H., Stanton, N. A., & Young, M. S. (2001). Where Is Computing Driving Cars? *International Journal of Human-Computer Interaction*, *13*(2), 203–229. https://doi.org/10.1207/S15327590IJHC1302_7