Compliance with Monitoring Requests, Biomechanical Readiness, and Take-Over Performance
Video Analysis from a Simulator Study

Zhang, Bo; Lu, Zhenji; Happee, Riender; de Winter, Joost; Martens, Marieke

Publication date
2019

Document Version
Final published version

Published in
13th ITS European Congress

Paper number ITS-1977

Compliance with Monitoring Requests, Biomechanical Readiness, and Take-Over Performance: Video Analysis from a Simulator Study

Bo Zhang[1,*], Zhenji Lu[2], Riender Happee[2], Joost de Winter[2], Marieke Martens[1,3]

1. Centre for Transport Studies, University of Twente, Enschede, the Netherlands

2. Department of Cognitive Robotics, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, the Netherlands

3. TNO Traffic & Transport, The Hague, the Netherlands

*Corresponding author: b.zhang@utwente.nl

Abstract

In the context of automated driving, a monitoring request (MR) is a means to prepare drivers for a take-over event. However, driver compliance may be an issue because not all MRs require a take-over. In this study, we investigated how drivers' compliance with MRs was associated with previously experienced scenarios. The compliance level was measured based on drivers' eye, hand, and foot preparatory behaviours retrieved from manual video observation. Although drivers showed good overall compliance by looking up to the road in response to MRs in all cases, hand and foot preparatory behaviour appeared to deteriorate after experiencing an MR without a critical event, and increased after a take-over event. Results further showed a positive association between preparatory behaviour and take-over performance.

Keywords: Automated driving, monitoring request, compliance

Introduction

Unless automated driving is fully reliable, drivers of an automated vehicle will have to take over control in case the automated driving system reaches its operational limits. Many human factors play a role in the driver's response to take over-events (Lu, Happee, Cabrall, Kyriakidis, & De Winter, 2016), including vigilance (Körber, Schneider, & Zimmermann, 2015; Carsten & Martens, 2019), workload (De Winter, Happee, Martens, & Stanton, 2014), and distraction (Louw, Kountouriotis, Carsten, & Merat, 2015; Zeeb, Härtel, Buchner, & Schrauf, 2017). A sufficient time budget is essential for a safe and comfortable control transition (Zhang, De Winter, Varotto, Happee, & Martens, 2019).

An automated driving system cannot always know in advance when the driver has to take over. The automation relies on radar or cameras to detect critical situations, which implies that the achievable time budget may often be short. One possible countermeasure to inadequate driver responses is to inform drivers about automation uncertainty (e.g., Beller, Heesen, & Vollrath, 2013; Helldin, Falkman, Riveiro, & Davidsson, 2013; Dziennus, Kelsch, & Schieben, 2016). With such systems, drivers are requested to supervise the driving environment when the automation is no longer certain about its capabilities. If an intervention is required, drivers would then benefit from increased situation awareness and respond properly within a shorter time. In a recent study, Lu, Zhang, Feldhütter, Happee, Martens, and De Winter (2019) examined the effects of a system that provides a visual-auditory monitoring request (MR) when approaching a location where driver take-over is likely to be requested. The MR asked the driver to pause the non-driving task, monitor the traffic environment, and be prepared for a potential take-over. If a critical event was detected, the system provided a TOR as well. Lu et al. (2019) found that, compared to a conventional system that only issued a TOR, the MR+TOR system improved participants' take-over performance in terms of shorter take-over response time and longer minimum time collision (TTC), and yielded more positive subjective ratings regarding workload, trust, and acceptance.

Despite the positive findings of Lu et al. (2019), questions related to driver compliance need to be raised, because only a small portion out of all MRs require an actual driver take-over. For a warning system to be effective, it is essential to "provide the opportunity for protective behaviour to occur before a threat materializes" (Breznitz, 1984). Compliance plays a central role because it reflects the operator's willingness to perform protective/preparatory behaviour in response to a warning signal (Meyer, 2004; Lees & Lee, 2007). When a warning is issued for something that does not actually occur, a decrease in compliance in the following warning phases might take place, a phenomenon also known as the *false alarm effect* or *cry-wolf effect* (Breznitz, 1984; Sorkin, 1988). In Breznitz' experiments that extensively explored the cry-wolf effect, an electric shock threat was announced by a three-minute warning, then cancelled out by the experimenter (i.e., the experimenter announced that the shock would not occur). During the warning phase, the participants could reduce the intensity of the impending electric shock by pressing a pedal at some monetary cost, which represented protective behaviour against the pain. The session was repeated three times. Using three indices to quantify protective behaviour, Breznitz found reductions in the probability and amplitude of protective behaviour (i.e., whether the participants pressed the pedal; if so, how many times the pedal was pressed), as well as an increase in latency between the alarm onset and the initiation of the protective action. Besides being one of the first demonstrations that false alarms reduce compliance, Breznitz further pointed to dynamic patterns of the compliance level: compliance may decrease with consecutive false alarms, and increase significantly after a true alarm.

In the automotive domain, studies have mainly investigated drivers' compliance with imperfect collision warning systems (e.g., Bliss & Acton, 2000, 2003; Cotté, Meyer, & Coughlin 2001; Lees & Lee, 2007; Naujoks, Kiesel, & Neukum, 2016). In line with Breznitz (1984), results suggest that false alarms substantially decrease driver compliance (e.g., lower response frequency, slower braking response, and smaller reductions of speed). A few researchers have distinguished between true false alarms and unnecessary alarms based on the context in which the alarm occurs. While true false alarms are caused by detection errors, unnecessary alarms are issued as intended (e.g., associated with the driving context) but the threat resolves before intervention is needed (e.g., a pedestrian stood on the roadside but later decided not to cross the road). According to Breznitz (1984), both alarm types will cause cry-wolf effects because either the warning or the threat loses credibility. The magnitude of the cry-wolf effect, however, may differ between the two alarm types. In a comparison of false alarms and unnecessary alarms, Lees and Lee (2007) reported more frequent brake responses and larger speed reductions with unnecessary collision alarms and concluded that unnecessary alarms do not diminish compliance because they are comprehensible to the driver. Naujoks et al. (2016) documented a similar differential influence of the two alarm types.

The MR implemented in Lu et al. (2019) notified drivers while entering locations where a take-over was likely to be requested; therefore the MRs that are not followed by TORs are likely to be perceived as unnecessary alarms rather than false alarms. MRs are also different from the aforementioned collision warnings in that MRs request attention and preparation, rather than driver intervention. In this study, we revisit the results of Lu et al. (2019) to examine drivers' compliance with MRs. More specifically, we first looked into how drivers prepared themselves for a potential take-over event in response to the MR. Second, we examined whether drivers' preparatory behaviour decreased after unnecessary MRs (i.e., no take-over required) and increased after experiencing a take-over event. Third, we examined whether drivers' preparatory behaviour was associated with take-over performance.

Methods

Participants

Forty-one persons (of which six females) participated in this study. Their mean age was 29.6 years (*SD* = 7.0, min = 20, max = 57). Participants had a valid driving license for 11.2 years on average (*SD* = 7.2). All participants provided written informed consent, and each was compensated with 10 Euros for their participation. The research was approved by the Human Research Ethics Committee (HREC) of the Delft University of Technology.

Apparatus

The study was conducted in a static driving simulator consisting of a BMW 6-Series full vehicle mock-up with a 180-degree field of view, located at the Technical University of Munich. The rear views were presented using three projectors. SILAB from WIVW GmbH was used to simulate the driving scenarios, and to record vehicle data at a frequency of 120 Hz. The implemented automated driving system carried out longitudinal and lateral control, and could be activated and deactivated by pressing a button on the steering wheel. The sound effects of the engine, passing vehicles in the opposite lane, as well as warnings were provided via the speakers of the vehicle cabin. A 9.5 by 7.31-inch handheld tablet (iPad 2) was provided to the participants for performing a non-driving task. Three cameras were mounted to record the participants' body movement, foot movement, and hand movement, respectively.

Test scenario and automation system

Three test sessions were implemented in the original study of Lu et al. (2019), using a within-subject design: (1) an MR + TOR session, in which both MRs and TORs were provided, (2) a TOR-only session, in which only TORs were provided, and (3) an MR-only session, in which an MR was provided but the system failed to issue a TOR upon the critical event. The order of the first two sessions was counterbalanced, whereas the MR-only session was always presented last. The present study only concerns the MR + TOR session; the other two sessions will not be described. Before the formal test, a training session was conducted for the participants to familiarize themselves with the simulator, the automation system, and the presentation of MRs and TORs.

The track consisted of urban and rural segments with one lane in each direction and moderate traffic in the opposite lane, with no traffic in the ego lane. The participant started on a short rural segment with manual driving. Upon approaching the urban area, the automation system issued a notification indicating the availability of automation. The participant pressed a button on the steering wheel to switch on the automation, and then take the hands off the wheel and feet off the pedals. The automation drove at a constant speed of 50 km/h in the urban segments and 80 km/h in the rural segments. During automated driving, the participants were instructed to play Angry Birds or Candy Crush (visual-motor tasks without sound) on a handheld tablet PC provided by the experimenter. These games are self-paced and interruptible so that the participants could pause the game whenever they felt necessary to look up to the road.

The areas that triggered an MR contained a zebra crossing without traffic lights. In the following text, such areas are described as “MR blocks”. The critical events that triggered the TOR and required driver intervention were pedestrians walking onto the zebra crossing. In total, five MR blocks were implemented in the session; the MR was issued 12 s before reaching the zebra crossing. The MR consisted of a verbal notification “Please monitor” following a gong sound and a yellow icon, while the automation remained active. The participants had been instructed to pause the non-driving task and monitor the driving environment after receiving an MR.

At three MR blocks no pedestrians were present (Figure 1a), and the MR was dismissed after passing the zebra crossing. The participants did not need to intervene. In the other two MR blocks, two pedestrians started to cross the road 7 s after the onset of the MR (Figure 1b). At the same time, a TOR was provided (i.e., time budget = 5 s) and the automation was deactivated. The auditory TOR warning was a sharp double beep (75 dB, 2800 Hz) followed by a verbal request “Please take-over” accompanied with an orange icon. The participant had to take over by braking to avoid colliding with the pedestrians. After taking over control, the participant had to drive manually until the automation became available again; the participant could then reactivate the automation. In front of each zebra crossing, either a truck parked on the side of the road or a bus stop was placed to prevent the participants from seeing whether there were pedestrians or not before the TOR onset. The order of the five MR blocks was randomized between participants. The duration of the session was approximately 14 min.

Figure 1 - MR blocks with (a) and without (b) two pedestrians crossing the road from the observer's view.

Grouping of Participants

To explore variability in driver compliance, the first three of five MR blocks were analysed. The participants were divided into four groups based on four possible combinations in their first two trials:

a) 1. MR-only, 2. MR-only

b) 1. MR-only, 2. MR+TOR

c) 1. MR+TOR, 2. MR-only

d) 1. MR+TOR, 2. MR+TOR

Accordingly, their prior experience when receiving an MR in the first three trials was as follows:

a) 1. First MR, 2. After one MR-only, 3. After two MR-only (see Figure 2 for illustration)

b) 1. First MR, 2. After one MR-only, 3. After TO

c) 1. First MR, 2. After TO, 3. After one-MR-only

d) 1. First MR, 2. After TO, 3. After TO.

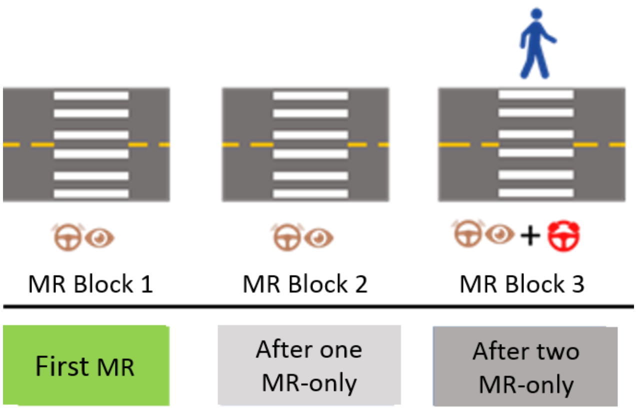

Figure 2 - Example of the coding of blocks.

Dependent measures

The compliance with the MR is the main dependent variable in this study. Because the MR aims to prepare the driver for a critical event, we measured the compliance with MR based on drivers' preparatory behaviours, which involved eye, hand, and foot movements. Breznitz's three indices of protective behaviour (see Introduction) were adopted to quantify drivers' preparatory behaviour, namely probability, latency, and amplitude. Detailed descriptions are listed in Table 1. The data were retrieved from manual video annotations. All observations started from the onset of the MR, with a duration of 7 s. Two independent raters (the first and second authors) rated the level of preparatory behaviour. The raters came to a consensus on all disagreements.

Table 1 - Measures of preparatory behaviour

Measures	Description		
	Eyes	Hands	Feet
Probability	The percentage of participants who looked up (any glance counts, regardless of the duration).	The percentage of participants whose hands were in a more convenient position for taking the steering wheel compared to when no preparation was made at all (i.e., Level of preparation > 0, as described below).	The percentage of participants whose feet were in a more convenient position for taking the steering wheel compared to when no preparation was made at all (i.e., Level of preparation > 0, as described below).
Latency	How long after the MR onset the participant started looking up.	How long after the MR onset the participant started moving their hands to be in a more convenient position for taking the steering wheel.	How long after the MR onset the participant started moving their foot to be in a more convenient position for pressing the braking pedal.
Amplitude	-	Level 0: Both hands were holding/interacting with the iPad/phone at the MR onset, with no obvious movement to put the iPad/phone away. Level 1: The iPad was already placed on the lap at the MR onset, or the participant lowered the iPad on the lap after receiving the MR. Both hands were still holding/touching the edge of the	Level 0: The feet were placed far from the pedals at the MR onset, with no obvious indication to move one foot closer to the pedals. Level 1: One foot was already placed close to the pedals at the MR onset, or the participant moved one foot closer to the pedals after receiving the MR, but not hovering above the brake pedal.

	iPad/phone, with no obvious movement to free the hand(s) to be prepared to take the wheel.	
	Level 2: There was an obvious attempt to free at least one hand to be prepared to grab the steering wheel (e.g., putting the iPad on the passenger seat, or having one hand in the air).	Level 2: The foot was hovering above the brake pedal at the MR onset, or the participant put one foot on the braking pedal after receiving the MR.
	Level 3: At least one hand was touching the steering wheel at the MR onset, or the participant put at least one hand on the steering wheel after receiving the MR.	

This study attempted to determine associations between the level of preparatory behaviour and subsequent take-over performance. Three performance measures were compared between preparation levels: 1) take-over time, measured from the onset of the TOR until the moment the driver started to press the brake pedal, same as in Lu et al. (2019); 2) minimum time to collision, calculated after the moment the driver pressed the brake pedal; 3) maximum deceleration calculated during the braking process. A lower maximum deceleration can be indicative of earlier braking.

Results

Data availability

Two participants experienced strong simulator sickness, one participant did not understand how the automation system worked, and video data for one participant were unavailable. The data from these four participants were excluded from the analysis. Among the remaining 37 participants, several video recordings were unavailable due to technical problems or poor visibility. The number of participants with complete probability/amplitude data for the first three blocks was 26 for hand preparation and 27 for foot preparation.

Probability of preparatory behaviour

All participants responded to the MR by looking to the road. Hence, the ‘eyes’ probability was 1 for each MR block. Figure 3 shows the percentages of participants per block in which a hand (left) or foot (right) preparatory behaviour occurred. Results indicate an increase in hand and foot preparation in *After TO* blocks as compared to the previous block. If the participant had experienced no TOR at all, the probability of preparation was overall low and/or showed a decreasing trend (red circular markers).

Regarding hand preparations in the third block: From the three participants who had experienced two take-overs before, all three were prepared (black diamond marker), and from the six participants who had experienced no take-overs before, only three were prepared (red circular marker). Similarly, regarding foot preparations in the third block: From the four participants who had experienced two take-overs before, all four were prepared (black diamond marker), and from the four participants who had experienced no take-overs before, none were prepared (red circular marker).

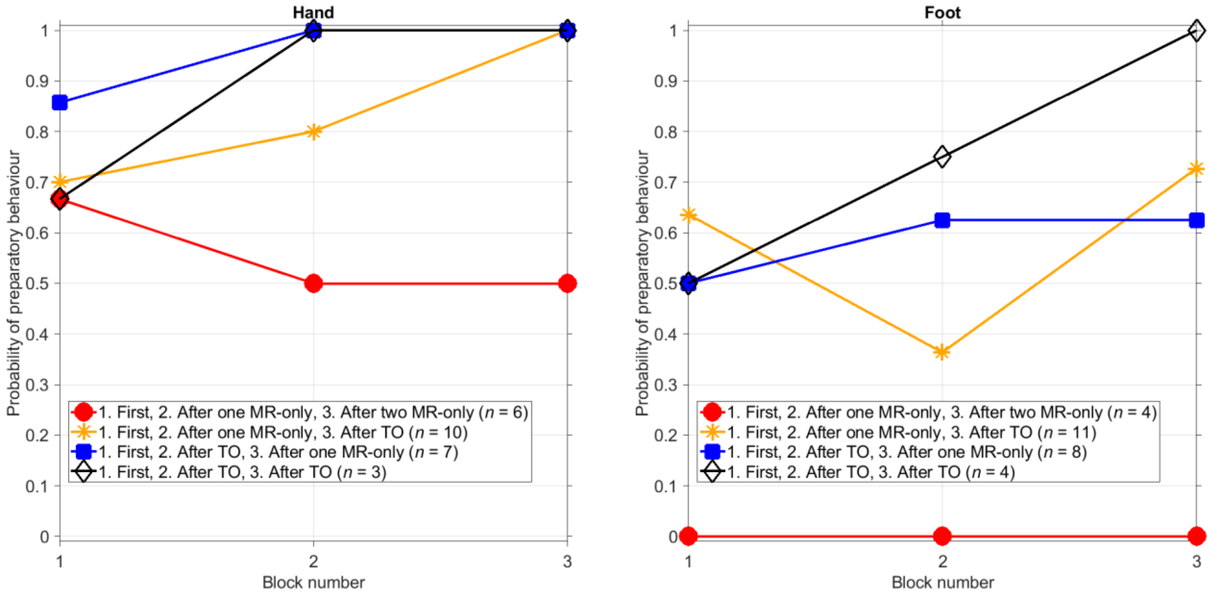

Figure 3 - Probability of hand (left) and foot (right) preparatory behaviour for the first three blocks.

Latency of Preparatory Behaviour

Averaged over all five MR blocks, participants started looking to the road on average 0.70 s (*SD* = 0.38) after the MR. The start of hand and feet movement (if available) occurred on average 2.13 s (*SD* = 1.09) and 3.17 s (*SD* = 1.55) after the MR, respectively. These effects are illustrated in Figure 4.

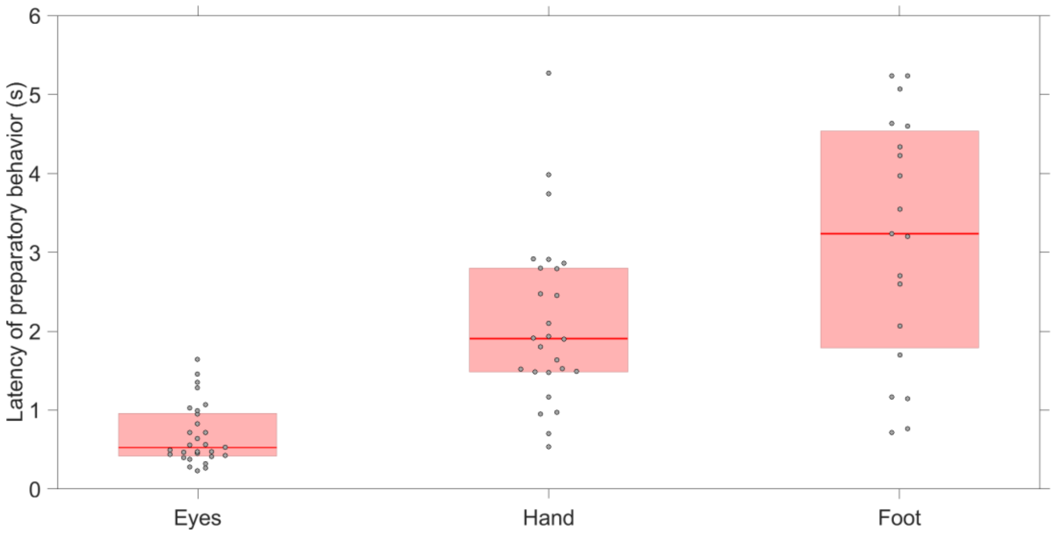

Figure 4 - Latency of participants' preparatins. Eyes (29 participants based on 115 observations), Hand (26 participants based on 87 observations), Foot (19 participants based on 46 observations)

Amplitude of Preparatory Behaviour

Figure 5 shows the means of hand and foot preparation levels. Results support the above observations that the mean level of preparation showed an increasing trend in the case of *After TO* blocks.

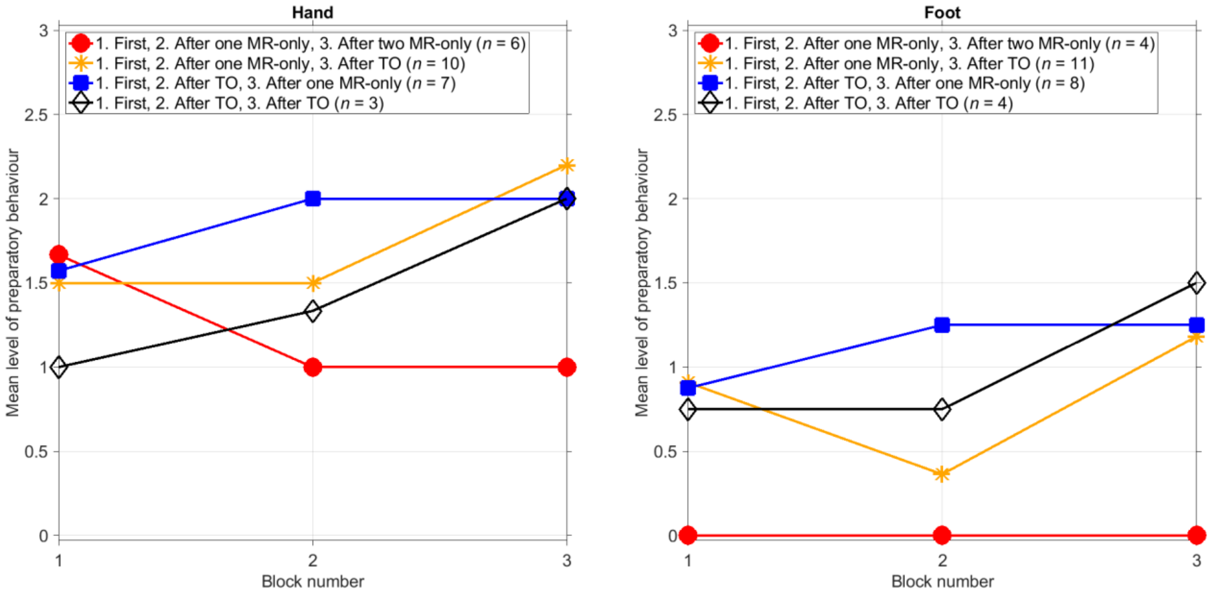

Figure 5 - Mean level of hand (left) and foot (right) preparation for the four MR block categories.

Level of Preparation and Subsequent Take-Over Performance

Table 3 shows the means and standard deviations of the take-over performance measures per level of hand and foot preparation. A tendency can be observed that take-over performance improved with higher levels of preparation, where better take-over performance is characterized by a shorter take-over time, a longer minimum TTC, and a lower maximal deceleration. Also, learning effects were observed from the first take-over event to the second, in line with Lu et al. (2019). It is worth noting that a larger number of participants showed the highest level of preparation before the second take-over event as compared to the first event. The trends of preparation level and take-over attempt are illustrated in Figure 6.

Table 3 - Means with standard deviations in parentheses of take-over response time, minimum TTC, and maximum deceleration for each level of hand and foot preparation.

Hand preparation	**Level 0**	**Level 1**	**Level 2**	**Level 3**
First take-over	***n* = 6**	***n* = 7**	***n* = 7**	***n* = 4**
Take-over time (s)	2.49 (0.78)	1.94 (0.79)	2.10 (0.64)	1.89 (0.30)
minTTC (s)	2.45 (0.75)	2.53 (0.52)	2.64 (0.60)	2.76 (0.51)
maxDEC (m/s^2)	9.49 (0.57)	9.36 (0.27)	8.47 (1.05)	8.79 (0.86)
Second take-over	***n* = 4**	***n* = 7**	***n* = 8**	***n* = 7**
Take-over time (s)	1.90 (0.60)	1.97 (0.80)	2.07 (0.68)	1.53 (0.41)
minTTC (s)	2.98 (0.57)	2.81 (0.70)	2.55 (0.70)	2.93 (0.45)
maxDEC (m/s^2)	7.93 (1.33)	8.64 (1.00)	8.38 (1.66)	7.42 (1.81)

Foot preparation	Level 0	Level 1	Level 2
First take-over	***n* = 16**	***n* = 6**	***n* = 4**
Take-over time (s)	2.03 (0.82)	2.01 (0.45)	1.64 (0.39)
minTTC (s)	2.73 (0.79)	2.72 (0.37)	2.94 (0.39)
maxDEC (m/s^2)	9.14 (0.51)	8.94 (0.56)	9.26 (0.01)
Second take-over	***n* = 12**	***n* = 7**	***n* = 9**
Take-over time (s)	2.00 (0.59)	1.66 (0.42)	1.20 (0.37)
minTTC (s)	2.83 (0.60)	2.95 (0.47)	3.16 (0.51)
maxDEC (m/s^2)	8.27 (0.96)	7.55 (1.88)	7.56 (1.62)

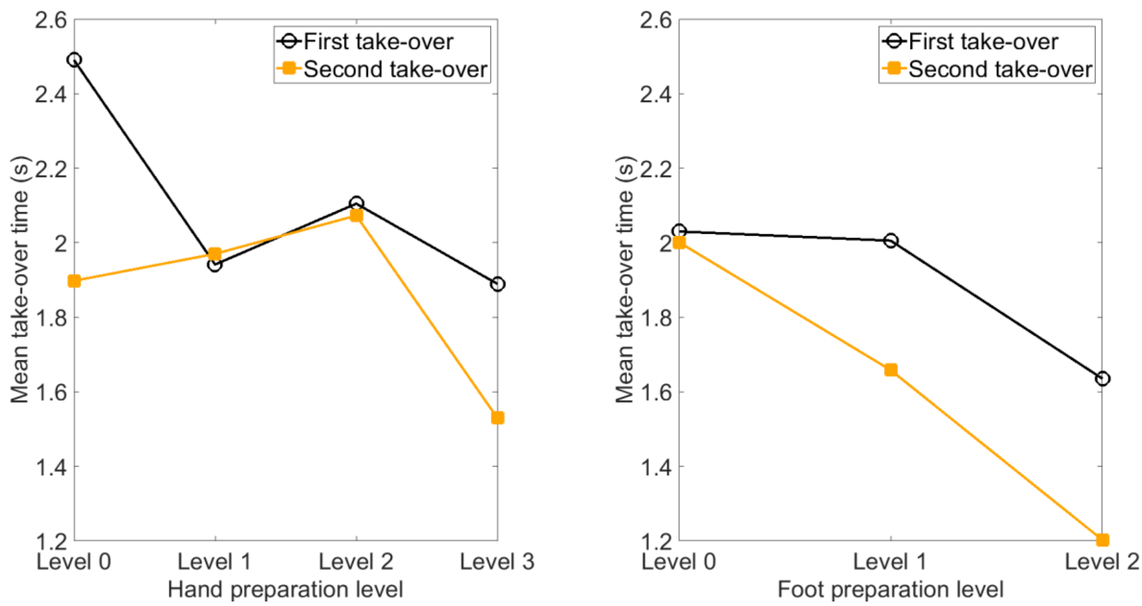

Figure 6 - Association between hand and foot preparation levels and take-over time. Standard deviations are shown in Table 3.

Discussion and Conclusion

In this study, we investigated how drivers complied with MRs, how the level of compliance changed with experience within one test session, and how the level of compliance associated with subsequent take-over performance. Compliance was measured by the three indices adapted from Breznitz (1984): probability, latency, and amplitude of eye, hand, and foot preparatory behaviour, retrieved from manual video observation.

The results indicated high overall compliance with MR. Participants looked up onto the road within a short time in all cases, and moved their hands to be better prepared for a possible take-over (e.g., putting down the iPad) in the majority of the cases. In several cases, participants also moved their feet closer to the pedal, but the probability of feet movements was smaller, and the associated latency was higher, than hand preparatory behaviour.

We found a higher preparation probability in response to MRs after a take-over event as compared to MRs after two consecutive MRs without take-over event, in line with the statements of Breznitz (1984) that threat cancellation reduces protective behaviour, whereas hit alarms increase compliance. We observed that after experiencing one or two MRs without critical event, a few participants did not look up as soon as possible, but continued with the non-driving task for a short time. Some participants monitored the road only shortly, and continued with the non-driving task even before the MR was dismissed. Such potentially risky behaviour may increase as more consecutive MRs without event are experienced in a prolonged drive. Also, drivers may develop alarm fatigue (Cvach, 2012) by excessive 'unnecessary' MRs, and turn off the notification system.

Furthermore, our findings suggest a substantial influence of preparation behaviour on take-over times: when monitoring with a hand on the wheel or with a foot hovering above the pedal, drivers responded faster to the TOR as compared to without any hand or foot preparation action. This is possibly due to a reduction of hand/foot travelling distance (Zhang, Wilschut, Willemsen, & Martens, 2019).

There are several limitations to be mentioned. First, drivers' behaviour was only observed within one short experimental session. The maximum number of successive cancellations was two. It is likely that a more severe reduction in preparatory behaviour would occur if more MR blocks were implemented. Moreover, as pointed out in a review by Green (2000), expectancy is an important determinant of brake response times. In our experiment, there were no true surprises. That is, the participants were probably expecting at least one take-over event in the entire session. If no event has occurred yet, each transition from one MR to another implies that the take-over event is approaching: Successive MRs without TOR implied greater proximity of danger, which may counteract the decrease in compliance (cf. Chapter 5 in Breznitz, 1984). Also, in general, participants in an experimental setting tend to 'behave well'; preparatory behaviour may be less in real life settings. In future studies, various reliabilities of TOR and MR should be investigated to reach conclusions that are more credible.

In summary, our study showed high driver compliance with the MRs in terms of the probability of eye and hand response. Furthermore, a situation-dependent change in the compliance level was observed. Finally, we showed that the level of preparation was associated with driving performance during a subsequent take-over event. Our findings suggest that the cry-wolf effects may be a concern in the use of MR. Measures such as training and education could be applied to counteract the cry-wolf effect (Breznitz, 1984; Zabyshny & Ragland, 2003).

Acknowledgements

The authors were involved in the Marie Curie Initial Training Network (ITN) HFauto – Human Factors of Automated Driving (PITN-GA-2013-605817), and would like to thank the TU Munich for hosting the experiment.

References

1. Breznitz, S. (1984). *Cry wolf: The psychology of false alarms.* Hillsdale, NJ: Lawrence Erlbaum Associates.
2. Beller, J., Heesen, M., & Vollrath, M. (2013). Improving the driver–automation interaction: An approach using automation uncertainty. *Human Factors, 55*, 1130–1141. doi:10.1177/0018720813482327
3. Bliss, J. P., & Acton, S. A. (2000). An evaluation of the safety, utility, and reliability of three-dimensional alarm systems for automotive use. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 45*, 1689–1693. doi:10.1177/154193120104502327
4. Bliss, J. P., & Acton, S. A. (2003). Alarm mistrust in automobiles: how collision alarm reliability affects driving. *Applied Ergonomics, 34*, 499–509. doi:10.1016/j.apergo.2003.07.003
5. Carsten, O., & Martens, M. H. (2019). How can humans understand their automated cars? HMI principles, problems and solutions. *Cognition, Technology & Work, 21*, 3–20. doi:10.1007/s10111-018-0484-0
6. Cotté, N., Meyer, J., & Coughlin, J. F. (2001). Older and younger drivers' reliance on collision warning systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 45, 277–280. doi: 10.1177/154193120104500402
7. 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. doi:10.1016/j.trf.2014.06.016
8. Dziennus, M., Kelsch, J., & Schieben, A. (2016). Ambient light based interaction concept for an integrative driver assistance system – a driving simulator study. *Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2015 Annual Conference*, 171–182.
9. Green, M. (2000). “How long does it take to stop?” Methodological analysis of driver perception-brake times. *Transportation Human Factors, 2*, 195–216. doi:10.1207/STHF0203_1
10. Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013). Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 210–217. doi: 10.1145/2516540.2516554
11. Körber, M., Schneider, W., & Zimmermann, M. (2015). Vigilance, boredom proneness and detection time of a malfunction in partially automated driving. *Proceedings of the 2015 International Conference on Collaboration Technologies and Systems*, 70–76. doi:10.1109/CTS.2015.7210402
12. Lees, M. N., & Lee, J. D. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems, *Ergonomics, 50*, 1264–1286, doi:10.1080/00140130701318749
13. Louw, T., Kountouriotis, G., Carsten, O., & Merat, N. (2015). Driver inattention during vehicle automation: How does driver engagement affect resumption of control? *4th International Conference on Driver Distraction and Inattention*, Sydney, Australia.
14. Lu, Z., Happee, R., Cabrall, C. D. D., Kyriakidis, M., & De Winter, J. C. F. (2016). Human factors of transitions in automated driving: A general framework and literature survey. *Transportation Research Part F: Traffic Psychology and Behaviour, 43*, 183–198. doi:10.1016/j.trf.2016.10.007

15. Lu, Z., Zhang, B., Feldhütter, A., Happee, R., Martens, H. M., & De Winter, J. C. F. (2019). Beyond mere take-over requests: The effects of monitoring requests on driver attention, take-over performance, and acceptance. *Transportation Research Part F: Traffic Psychology and Behaviour, 63*, 22–37. doi:10.1016/j.trf.2019.03.018
16. Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors, 46*, 196–204.
17. Naujoks, F., Kiesel, A., & Neukum, A. (2016). Cooperative warning systems: The impact of false and unnecessary alarms on drivers' compliance. *Accident Analysis & Prevention, 97*, 162–175. doi:10.1016/j.aap.2016.09.009
18. Cvach, M. (2012). Monitor alarm fatigue: an integrative review. *Biomedical Instrumentation & Technology, 46*, 268–277. doi:10.2345/0899-8205-46.4.268
19. Sorkin, R. D. (1988). Why are people turning off our alarms? *The Journal of the Acoustical Society of America, 84*, 1107–1108. doi:10.1121/1.397232
20. Zabyshny, A. A., & Ragland, D. R. (2003). False alarms and human-machine warning systems. UC Berkeley: *Safe Transportation Research & Education Center*. Retrieved from http://escholarship.org/uc/item/88w8b8g3
21. Zeeb, K., Härtel, M., Buchner, A., & Schrauf, M. (2017). Why is steering not the same as braking? The impact of non-driving related tasks on lateral and longitudinal driver interventions during conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 50*, 65–79. doi:10.1016/j.trf.2017.07.008
22. Zhang, B., De Winter, J. C. F., Varotto, S., Happee, R., & Martens, M. H. (2019). Determinants of take-over time from automated driving: a meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour.*
23. Zhang, B., Wilschut, E. S., Willemsen, D. M. C., & Martens, M. H. (2019). Transitions to manual control from highly automated driving in non-critical truck platooning scenarios. *Transportation Research Part F: Traffic Psychology and Behaviour.* doi:10.1016/j.trf.2019.04.006