Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations: The Role of Traffic Density

Christian Gold, **Moritz Körber**, **David Lechner**, and **Klaus Bengler**, Institute of Ergonomics, Technical University of Munich, Germany

Objective: The aim of this study was to quantify the impact of traffic density and verbal tasks on takeover performance in highly automated driving.

Background: In highly automated vehicles, the driver has to occasionally take over vehicle control when approaching system limits. To ensure safety, the ability of the driver to regain control of the driving task under various driving situations and different driver states needs to be quantified.

Methods: Seventy-two participants experienced takeover situations requiring an evasive maneuver on a three-lane highway with varying traffic density (zero, 10, and 20 vehicles per kilometer). In a between-subjects design, half of the participants were engaged in a verbal 20-Questions Task, representing speaking on the phone while driving in a highly automated vehicle.

Results: The presence of traffic in takeover situations led to longer takeover times and worse takeover quality in the form of shorter time to collision and more collisions. The 20-Questions Task did not influence takeover time but seemed to have minor effects on the takeover quality.

Conclusions: For the design and evaluation of human–machine interaction in takeover situations of highly automated vehicles, the traffic state seems to play a major role, compared to the driver state, manipulated by the 20-Questions Task.

Application: The present results can be used by developers of highly automated systems to appropriately design human–machine interfaces and to assess the driver's time budget for regaining control.

Keywords: vehicle automation, autonomous driving, driver behavior, human–automation interaction, mental workload, phoning while driving

Address correspondence to Christian Gold, Institute of Ergonomics, Technical University of Munich, Boltzmannstraße 15, D-85747 Garching, Germany; e-mail: gold@lfe.mw.tum.de.

HUMAN FACTORS
Vol. 58, No. 4, June 2016, pp. 642–652
DOI: 10.1177/0018720816634226

INTRODUCTION

Highly automated vehicles (Gasser, 2012), also known as Level 3—limited self-driving automated vehicles (National Highway Traffic Safety Administration [NHTSA], 2013)—or conditional automation (SAE International, 2014), are already being tested on public roads. By definition, the driver is "expected to be available for occasional control, but with sufficiently comfortable transition time" (NHTSA, 2013, p. 5) to resume control in situations that the automation is not able to handle. These TORs may occur also in those situations when complexity and traffic density are high. There are several indications that the density of the surrounding traffic affects driver performance. For manual driving, Heenan, Herdman, Brown, and Robert (2014) showed that higher traffic density results in an impaired ability to control vehicle speed and in lower situation awareness. Considering driving with adaptive cruise control (ACC), Stanton and Young (2005) found higher workload and demand in medium (approximately 15 vehicles per kilometer) and high traffic (approximately 23 vehicles per kilometer) compared to low traffic density situations (approximately seven vehicles per kilometer).

For highly automated driving, traffic density has already been identified to have distinct influence on takeover performance (Gold, Lorenz, & Bengler, 2014; Radlmayr, Gold, Lorenz, Farid, & Bengler, 2014). With a high traffic density of approximately 30 vehicles per kilometer in the neighboring lane, the authors found poorer takeover performance compared to a condition without traffic. Next to an extended takeover time and higher accelerations due to intensified braking and dynamic lane changes, the authors also found a larger number of crashes. The studies of Radlmayr et al. (2014) and Gold et al. (2014), however, lack a subdivision of traffic densities as they solely used an extremely high traffic density condition of 30 vehicles per kilometer. On the one hand, lower traffic densities could reveal nonlinearities, like the ceiling effect found by Stanton and Young (2005) for traffic densities above approximately 15 vehicles per kilometer. On the other hand, the investigated traffic density of 30 vehicles per kilometer is rather extreme, as the traffic flow gets unstable above 18 vehicles per kilometer at speeds of 120 km/h (cf. Schöpplein, 2013). Therefore, this study aims to close this gap and quantify the impact of different traffic densities on takeover performance.

By definition, highly automated driving allows drivers to engage in non-driving-related tasks, as the driver is not needed to permanently monitor automation. Studies have shown that with higher automation levels, the engagement in non-driving-related tasks increases (e.g., Jamson, Merat, Carsten, & Lai, 2013). Thus, takeover experiments should involve consideration of the role and impact of non-driving-related tasks. In previous studies, a wide range of non-driving-related tasks have been investigated, including the visual Surrogate Reference Task (e.g., Beller, Heesen, & Vollrath, 2013; Gold, Dambock, Lorenz, & Bengler, 2013; Kerschbaum, Lorenz, & Bengler, 2014), the cognitive *n*-Back Task (Radlmayr et al. 2014), the conversational 20-Questions Task (TQT; e.g., Merat, Jamson, Lai, & Carsten, 2012), and naturalistic tasks, like texting, Internet search (Zeeb, Buchner, & Schrauf, 2015), and talking on the phone (Neubauer, Matthews, & Saxby, 2012). It is remarkable that talking on the phone is not widely investigated in the context of automated driving, considering that drivers are engaged in speaking on the phone 6.7% of the driving time (Funkhouser & Sayer, 2012). This cognitive distraction is noteworthy as cognitive distraction impairs the takeover performance (Radlmayr et al., 2014), and the cognitive takeover process is considered more critical than the motoric reorientation to the driving task (Zeeb et al., 2015). Therefore, the TQT was introduced as a second factor to the experiment.

Merat et al. (2012) considered the TQT as a non-driving-related task to be similar to speaking on the phone. For manual driving, Heenan et al. (2014) used the TQT as a substitute for conversations on the phone and in combination with different traffic densities and stated that this task is "well suited for experimentation, as it is continuous and the pace of the task can be monitored by the experimenter" (p. 1090). Just like talking on the phone, the TQT affects working memory and situation awareness (Heenan et al., 2014). It is assumed that the TQT complicates the takeover by affecting drivers' workload, just like in manual driving (e.g., Strayer, Drews, & Johnston, 2003) and prolongs drivers' reaction times (Strayer & Drews, 2004; Strayer, Drews, & Crouch, 2006). The task further reduces drivers' attention (Strayer et al., 2003) and may, therefore, reduce situation awareness during highly automated driving. It is also connected to the factor of traffic density, as the effort of regaining situation awareness is dependent on the number of relevant vehicles to be considered in the takeover maneuver.

The driving simulator study presented in this publication assesses changes of timing and quality aspects of the takeover due to different traffic density conditions and the verbal TQT representing a phone conversation while driving automated.

METHOD

Participants

Seventy-two participants (14 women and 58 men), between 19 and 79 years of age (M = 45.0, SD = 22.2), recruited from two different age groups (36 young, M = 23.3, SD = 2.6; 36 elderly, M = 66.7, SD = 4.56) participated in this study. Participants of the elderly group were paid 30 euros compensation, as they had an extended effort to approach the laboratory, whereas the younger participants did not receive a compensation. All participants had a valid driver's license for at least 1 year and had normal or corrected-to-normal vision. The evaluation of age effects would exceed the desired paper length and are therefore presented in a separate paper (see Körber, Gold, Lechner, & Bengler, in press). In brief, there was no main effect for age, $F(1, 66) = 2.449$, $p = .055$, and no interaction between the main factors age group and task, $F(1, 66) = .489$, $p = .744$, or between age group and traffic density, $F(2, 132) = 1.680$, $p = .122$.

Figure 1. Wide-angle shot of the simulator from the driver's seat.

Apparatus

The experiment was conducted in a fixed-based, high-fidelity driving simulator, equipped with a full vehicle mockup and six projectors enabling a front viewing angle of more than 180° and the use of all driving mirrors (Figure 1). The implemented automation had a longitudinal capability similar to common ACC systems, followed the indicated speed limits, kept in the center of the current lane, and was able to independently pass slower vehicles. Data were recorded at a frequency of 100 Hz, including the vehicle's position, accelerations, steering wheel angle, and pedal positions. The head mounted eye-tracking system (Dikablis) was used to track driver's gaze behavior with a recording rate of 25 frames per second.

Traffic Density and TQT

With a 3 × 2 mixed design, the within-subject factor traffic density (zero, 10, 20 vehicles per kilometer) and the between-subject factor task (none, TQT) were assessed regarding their influence on the takeover performance. Task was distributed equally over the age groups. The traffic density changed naturally and fluidly over the experiment and was varied by manipulation of the distance between vehicles in the neighboring lanes. Just before the take-over request (TOR), the surrounding traffic, consisting of 10 vehicles in each lane, was given a speed equal to the automated vehicle and was arranged with constant and equal distance between the vehicles (100 or 50 m for density of 10 or 20 vehicles per kilometer, respectively), while the neighboring lanes were empty in the condition when there were zero vehicles per kilometer. The participant's vehicle was next to the middle of a gap between two neighboring vehicles to enable a lane change.

Half of the participants had no task to perform (baseline), whereas the other half had to engage in the TQT whereby participants had to guess animals by asking the experimenter questions that were answered with yes or no only. As soon as an animal was guessed correctly, the TQT continued with another animal. This cognitive task employed hands-free communication between participant and experimenter, and it did not require any additional (manual) driver input. It was continuously presented during the automated drive.

Each participant performed three takeovers with a traffic density of zero, 10, and 20 vehicles per kilometer in an automated drive on a three-lane highway, lasting approximately 20 min. The takeover situations occurred approximately every 6 min. The lane of travel was varied within the participants (left, center, right) to reduce effects of expectancy and to consider takeover behavior in different lanes. Traffic density and lane were counterbalanced across participants to avoid bias. The situations evolved within a normal traffic flow; for example, the participant's vehicle was passing slower vehicles and therefore changing to the lane where the takeover would occur, making it impossible to anticipate an approaching takeover situation. The cause of the TOR was a broken, stopped vehicle in the current lane. Together with the TOR, the vehicle suddenly popped up 233 m ahead; thus the system limit could not be detected before the TOR was announced. The TOR itself consisted of a double beep of 75 dB and about 2800 Hz, according to guidelines of NHTSA for crash warnings (Campbell, Richard, Brown, & McCallum, 2007). Subsequent to the TOR, the automation was deactivated, leading to slight deceleration. All TORs were applied on straight road sections that did not require immediate steering. In all situations, the automation speed was set to 120 km/h (33.3 m/s; 74.6 mph), and initially drivers were given 7 s to respond to the TOR. See Figure 2.

Procedure

Participants were welcomed at the institute, signed a consent form, and completed a demographic questionnaire that gathered information regarding age, gender, and driving experience.

Figure 2. Examples of the takeover situations from the ego vehicle's view at the moment of the takeover request. Inset zooms in on the stopped vehicle, representing the system limit.

Participants were briefed regarding the experiment and the automation with a written handout. They were informed that they would be able to take their hands off the steering wheel and feet off the pedals as the automation was capable of longitudinal and lateral control. Furthermore, they were told that the automation did not have to be monitored, although there were situations the automation could not handle. In these cases, the driver was warned (i.e., TOR) and had to take over control. Depending on the condition, participants were given instructions regarding the TQT. Subsequently, a second questionnaire was completed, recording subject's attitude toward automated driving. The experimenter explained the driving simulator, and the participants were given a test drive to familiarize themselves with the simulation and automation, including one TOR at the end, albeit without any representation of a system limit. As soon as the participants indicated that they were comfortable with the simulator, the experiment started, including the three takeover situations explained in the previous section. Following the experimental drives, a similar questionnaire assessing the attitude toward automated driving was completed, followed by participant debriefing.

Dependent Measures

Building upon our earlier work (e.g., Gold et al., 2013), we assessed timing and quality aspects of the takeover based on a number of dependent variables (see Table 1).

Regarding the timing aspects, hands-on time ($t_{\text{hands-on}}$) and takeover time (t_{takeover}) were evaluated. We defined $t_{\text{hands-on}}$ as the time between the TOR and moment when the hands grasped the steering wheel, indicative of how long it takes the driver to return to a driving position. We defined t_{takeover} as the time between the TOR and the first measurable response to the situation. Based on similar experimental data, Gold et al. (2013) defined thresholds of 2° for steering wheel angle and 10% for brake pedal position. As soon as driver's input exceeded one of these thresholds, it was considered to be an overt maneuver and t_{takeover} was measured. The variable t_{takeover} captures the cognitive processes of

TABLE 1: Dependent Variables

Dependent Variable	Unit	Category	Explanation
Hands-on time ($t_{\text{hands-on}}$)	Seconds	Timing aspects	Time between TOR and hands on steering wheel
Takeover time (t_{takeover})	Seconds	Timing aspects	Time between TOR and start of maneuver
Maximum longitudinal accelerations (a_{long})	m/s²	Quality aspects	Maximum braking acceleration during situation
Maximum lateral accelerations (a_{lat})	m/s²	Quality aspects	Maximum lateral acceleration during situation
Minimal time to collision (TTC)	Seconds	Quality aspects	Minimal TTC during situation
Horizontal gaze dispersion (HGD)	Pixels	Workload (due to task)	Deviation of horizontal gaze position

Note. TOR = takeover request.

regaining situation awareness and selecting a response to the TOR. It is therefore a reasonable measure for assessing the increased complexity of the situation due to higher traffic densities.

Quality aspects of the takeover were evaluated by the maximum longitudinal accelerations, including brake accelerations (a_{long}), and the maximum lateral accelerations (a_{lat}) of the vehicle. Another variable, time to collision (TTC), represented the time remaining until colliding with the obstacle, assuming a constant speed difference between vehicle and obstacle (see International Organization for Standardization, 2013). At the onset of the TOR, the TTC was 7 s and equivalent to the time budget. From that moment on, the TTC decreased until the participant braked or changed lanes. A TTC of zero represents a collision, either with the obstacle or vehicles in the neighboring lanes, whereas higher TTC values indicate a better handling of the situation, as the temporal distance to a collision is longer.

Accelerations were evaluated from the TOR up to 67 m after passing the obstacle, capturing the evasive maneuver and stabilization in the new lane. It was assumed that slower accelerations are equivalent to a less dynamic and therefore safer maneuver.

In order to measure the influence of task on participants' workload, the gaze behavior of drivers was assessed. Several studies suggest gaze behavior is a valid tool for measuring cognitive workload (e.g., Wang, Reimer, Dobres, & Mehler, 2014), for example, in aviation (Di Nocera, Camilli, & Terenzi, 2006) or road transportation (He, Becic, Lee, & McCarley, 2011). Visual scanning is also related to situation awareness (Ratwani, McCurry, & Trafton, 2010). Wang et al. (2014) concluded that "horizontal gaze dispersion is the most sensitive measure of cognitive demand" (p. 236). As cognitive workload should decrease horizontal gaze dispersion (HGD), this measure was chosen for assessing the differences between the task and baseline groups. We measured the HGD in 1-min intervals both at the beginning and at the end of the experiment and assessed the average HGD within those 2 min for each participant. As the traffic density was constantly changing during the experiment, the HGD could not measure differences resulting from traffic density.

Data Set and Statistical Analyses

For each participant, we tested three takeover situations, resulting in a total of 216 takeovers. A two-factorial MANOVA with repeated measures was performed, with traffic density as the repeated within-subject factor and task as the between-subject factor. As the lane was counterbalanced and was shown not to influence driver performance in similar experiments (cf. Radlmayr et al., 2014), lane was not included as a factor. In cases in which Mauchly's test for sphericity showed significance, values were corrected (Greenhouse-Geisser). All post hoc pairwise comparisons were Bonferroni corrected.

Although the TOR was designed according to common standards and was not expected to be missed, one participant (79 years) missed a TOR and was excluded from statistical analyses regarding the takeover performance. Lateral acceleration was not recorded properly in six trials but could be adequately derived from the vehicle's coordinates. In another four tests, participants ran into the median strip, leading to excessive peak lateral accelerations (>18 m/s²), which have been excluded from the calculation of a_{lat}. We measured $t_{hands\text{-}on}$ via the field-cam video from the eye-tracking system. In 13 situations—depending on participant's head position—hands were not inside the field of view of the eye-tracking system and $t_{hands\text{-}on}$ could not be assessed. Because of missing $t_{hands\text{-}on}$ values, a separate ANOVA was calculated for this measure. The other dependent variables, $t_{takeover}$, a_{long}, a_{lat}, and TTC, were analyzed using a 3×2 MANOVA for the main factors traffic density and task. The effect of task on HGD was analyzed using an unpaired t test. For all analyses, the significance level was set to .05.

RESULTS

We measured driver performance metrics within takeover situations in highly automated vehicles in varying traffic density and task conditions. Tables 2, 3, and 4 contain the results of the statistical analyses regarding the factors task and traffic density. Whereas the traffic density affected almost all performance metrics, the TQT did not show a significant effect on the takeover

TABLE 2: Results of the MANOVA for Task Considering $t_{takeover}$, a_{lat}, a_{long}, and TTC

Measure	*F*(1, 69)	*p*
Task	1.04	.393
$t_{takeover}$	0.732	.395
a_{lat}	0.187	.425
a_{long}	0.643	.667
TTC	4.305	.042

Note. $t_{takeover}$ = takeover time; a_{lat} = maximum lateral accelerations; a_{long} = maximum longitudinal accelerations; TTC = time to collision.

TABLE 3: Results of the MANOVA for Traffic Density and Interaction of Task and Traffic Density Considering $t_{takeover}$, a_{lat}, a_{long}, and TTC

Measure	*F*(2, 138)	*p*
Traffic density	8.27	<.001
$t_{takeover}$	15.21	<.001
a_{lat}	9.43	<.001
a_{long}	3.58	.034
TTC	22.62	<.001
Task × Traffic Density	0.41	.910
$t_{takeover}$	0.024	.976
a_{lat}	0.019	.982
a_{long}	1.854	.161
TTC	0.135	.874

Note. $t_{takeover}$ = takeover time; a_{lat} = maximum lateral accelerations; a_{long} = maximum longitudinal accelerations; TTC = time to collision.

TABLE 4: Results of the ANOVA Considering $t_{hands\text{-}on}$

Factor	*F*(1, 57)	*p*
Task	3.36	.072
Traffic density	0.33	.689
Task × Traffic Density	0.910	.394

Note. $t_{hands\text{-}on}$ = hands-on time.

performance. Neither $t_{hands\text{-}on}$, $F(1, 57) = 0.910$, $p = .394$, nor the other variables, $F(2, 138) = 0.41$, $p = .910$, showed an interaction between the factors traffic density and task.

Task did not show a significant main effect on takeover performance, $F(1, 69) = 1.04$, $p = .393$) but indicated an effect on TTC, $F(1, 69) = 4.31$, $p = .042$. As we performed a MANOVA, integrating the different dependent measures, significance of the single measures can be interpreted only when there is a significant main effect, which is not the case for the factor task. However, a multiple ANOVA, which is a common and less conservative method, also indicates a reduction of TTC with the TQT, $F(1, 69) = 5.53$, $p = .023$. There are also indications that the TQT may lead to longer $t_{hands\text{-}on}$, but the effect did not reach statistical significance, $F(1, 57) = 3.36$, $p = .072$. A strong effect (Cohen's $d = 1.18$; $r = .51$) was found in HGD between task and baseline group, $t(70) = 5.010$, $p < .001$ (cf. Figure 3), indicating a more-than-50% broader horizontal dispersion of gazes without the additional load of the TQT.

Traffic density showed a significant main effect on takeover performance, $F(2, 138) = 8.27$, $p < .001$, but not on $t_{hands\text{-}on}$, $F(1, 57) = .332$, $p = .689$ (cf. Figure 4a). Traffic density did influence the $t_{takeover}$, $F(2, 138) = 15.21$, $p < .001$; and post hoc pairwise comparisons showed that with a traffic density of zero, the maneuver started about 1 s earlier ($M = 2.57$ s, $SD = 1.02$) than with a traffic density of 10 ($M = 3.42$ s, $SD = 1.28$), $t(69) = -4.680$, $p < .001$, or 20 vehicles per kilometer ($M = 3.51$ s, $SD = 1.20$), $t(69) = -5.437$, $p < .001$ (cf. Figure 4b).

Figure 5 shows the dependent variables that measure takeover quality and the occurrence of crashes. Differences between the situation without traffic and the situations involving other vehicles are apparent. Considering the accelerations, the situations evolved more dynamically with traffic on the neighboring lanes. For the longitudinal accelerations, $F(2, 138) = 3.58$, $p = .034$ (cf. Figure 5a), post hoc tests showed higher accelerations with a traffic density of 10 ($M = 4.52$ m/s², $SD = 4.07$) vehicles per kilometer compared to the situation with no other vehicles ($M = 3.39$ m/s², $SD = 3.73$), $t(69) = -2.915$, $p = .014$. Similar effects were found with lateral accelerations, $F(2, 138) = 9.43$, $p < .001$ (Figure 5b), where the presence of 20 vehicles per kilometer yields higher lateral accelerations ($M = 4.31$ m/s², $SD = 2.21$) in comparison to the no-traffic condition ($M = 2.85$ m/s², $SD = 1.69$), $t(69) = -4.612$, $p < .001$.

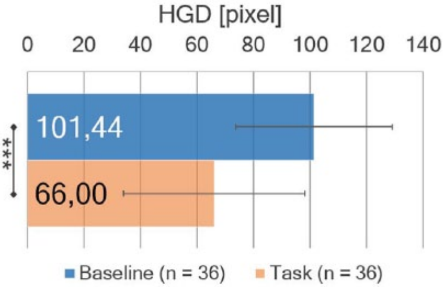

Figure 3. Means and standard deviations of horizontal gaze dispersion (HGD) as a function of task (***p < .001).

There was also an effect of traffic density on TTC, $F(2, 138) = 22.62$, $p < .001$ (Figure 5c), as the presence of traffic led to shorter TTC values and therefore more critical situations. We measured lower TTC values with a traffic density of 10 (M = 1.72 s, SD = 1.23), $t(69) = 5.560$, $p < .001$, and 20 (M = 1.59 s, SD = 1.23), $t(69) = 6.516$, $p < .001$, vehicles per kilometer compared to the no-traffic condition (M = 2.77, SD = 1.26). The minimum TTC values of 0 s in the conditions with 10 and 20 vehicles per kilometer indicate that crashes occurred. All crashes—including those with the broken-down, stopped vehicles representing the system limitation as well as with the surrounding vehicles in neighboring lanes—are shown in Figure 5d. Fisher's exact test, which is robust for small numbers of cases, showed that there was a larger number of crashes with 10 (p = .003) and 20 (p = .002) vehicles per kilometer compared to the no-traffic situation. Therefore, all quality measures of the takeover showed more critical characteristics under the presence of traffic. A supplementary chi-square test showed similar results, χ^2 (2, N = 71) = 9.33, p = .009, for crash occurrence. All post hoc performed pairwise comparisons are summarized in Table 5 along with effect sizes.

DISCUSSION

The results of the HGD supported the notion that the TQT was a cognitively loading task and appropriate for the experimental purpose to substitute phone conversations. With the conservative statistical methods we used, we did find indications for deteriorations in takeover performance due to the TQT but no significant effects. Neubauer et al. (2012) found a reduction in brake reaction times when participants were engaged in a phone call while engaged in automated driving, in comparison to a no-task condition. Although the experimental setup is very similar, such effects could not be found in the current study. A possible explanation is that in our experiment, the periods of automated driving between the TORs were only 5 to 6 min long, wherefore possible positive effects on driver arousal due to the TQT did not emerge although could become apparent when lengthening the automated driving time span.

It was hypothesized that the complexity of a takeover situation in highly automated vehicles impairs takeover performance at the perception, cognition, and execution levels. The time until participants placed their hands back on the steering wheel was found to be independent of traffic density and matched the range of other experiments (cf. Gold et al., 2013). This finding suggests that this motion pattern is either rule or skill based (cf. Rasmussen, Pejtersen, & Goodstein, 1994) and, as it is not influenced by the traffic density, can be executed without extensive situational assessment or understanding. On the other hand, traffic density does influence the takeover time, which is in line with the results from Radlmayr et al. (2014). A higher traffic density and therefore a greater number of objects led to delays in initiating the maneuver. This result may be caused by extended visual scanning and decision-making process. It can be argued that longer takeover times are actually indicative of a better reaction, as participants took the necessary time to regain situation awareness before starting a maneuver. Traffic density furthermore influenced the quality of the takeover in the form of higher accelerations, lower TTCs, and a higher crash probability in higher traffic densities, likely to be also caused by the delayed takeover. All of these variables indicate a more critical takeover, and the increase in the number of crashes matches findings and results from other research (cf. Radlmayr et al., 2014). Situational parameters that increase complexity of the situation must not be underestimated in order to ensure the safety of the driver-vehicle-automation system.

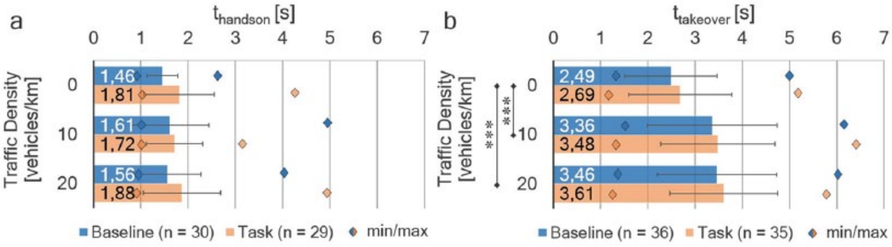

Figure 4. Timing measures hands-on time ($t_{\text{hands-on}}$) and takeover time (t_{takeover}) as a function of traffic density and task. Bars represent means, error bars represent standard deviations, and diamonds indicate maximum and minimum values (****p* < .001).

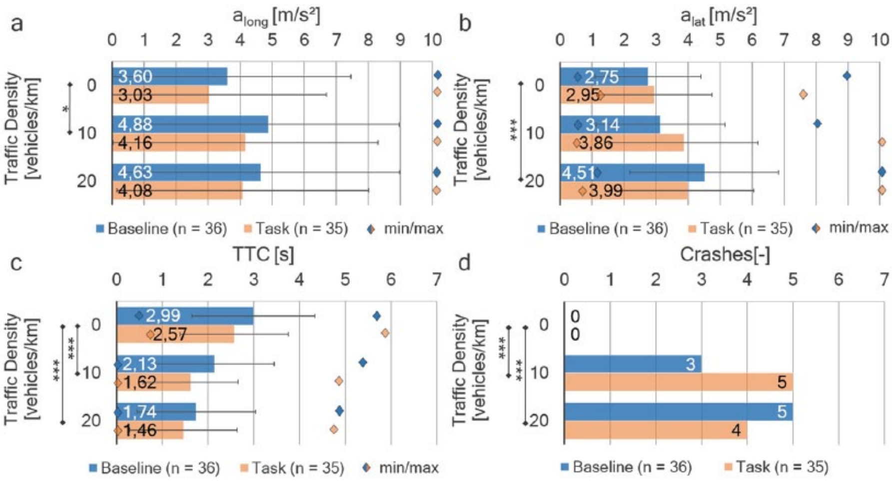

Figure 5. Quality measures longitudinal accelerations (a_{long}), lateral accelerations (a_{lat}), time to collision (TTC), and crashes occurred as a function of traffic density and task. Bars represent means, error bars represent standard deviations, and diamonds indicate maximum and minimum values (**p* < .05, ****p* < .001).

It is noteworthy that no variable showed significant differences between the medium and high traffic density conditions. The influence of traffic density on takeover performance seems to show ceiling effects, as found in the study of Stanton and Young (2005) with ACC systems. Further studies should include sampling points between zero and 10 vehicles per kilometer to further assess the correlation of traffic density and takeover performance. A higher object density of more than 20 vehicles per kilometer will primarily lead to stronger braking as changing lanes becomes harder or even impossible (cf. Gold et al., 2014; Radlmayr et al., 2014). It could be argued that effects of traffic density showing up in the dependent variables represent such changes from lane changing to braking and do not indicate a deterioration of takeover quality; however, the concurrent increase of lateral accelerations and crash probabilities as measured in this study suggest the contrary. Especially when looking at accelerations and crash probabilities, a distinct limitation of this study is that the transferability of driving simulator experiments to future automated vehicles is currently lacking evidence (Kemeny & Panerai Francesco, 2003) and therefore must be validated on the road in future research. Nevertheless, the measured accelerations and a crash risk of more than ~10% with a time budget of 7 s and common traffic densities of 10 to 20 vehicles per kilometer suggest that such takeover situations

TABLE 5: Results of Post Hoc Pairwise Comparisons of the MANOVA

Comparison Traffic Density	Difference of Means	*P* Value (Bonferroni)	*p* Value (LSD)	*T* Value	Cohen's *d*	Pearson *r*
$t_{takeover}$						
0 vs. 10	–0.833	0.000	.000	–4.684	–1.128	.491
0 vs. 20	–0.946	0.000	.000	–5.430	–1.307	.547
10 vs. 20	–0.113	1.000	.585	–0.545	–0.131	.065
a_{long}						
0 vs. 10	–1.201	0.014	.004	–2.914	–0.702	.331
0 vs. 20	–1.037	0.133	.043	–2.048	–0.493	.239
10 vs. 20	0.164	1.000	.756	0.308	0.074	.037
a_{lat}						
0 vs. 10	–0.651	0.129	.044	–2.060	–0.496	.241
0 vs. 20	–1.402	0.000	.000	–4.615	–1.111	.486
10 vs. 20	–0.751	0.103	.035	–2.159	–0.520	.252
TTC						
0 vs. 10	1.062	0.000	.000	5.562	1.339	.556
0 vs. 20	1.199	0.000	.000	6.509	1.567	.617
10 vs. 20	0.138	1.000	.115	0.657	0.158	.079

Note. LSD = least significant difference; $t_{takeover}$ = takeover time; a_{lat} = maximum lateral accelerations; a_{long} = maximum longitudinal accelerations; TTC = time to collision.

may lead to high risk states in future automated vehicles. It is assumed that such situations are challenging in the absence of automation as well and may also lead to high-risk maneuvers and crashes, as in the studies of Gold et al. (2014) and Radlmayr et al. (2014).

In addition, for measuring driver performance, we selected a generic takeover situation of a blocked lane, a situation that is likely to be handled by future automated vehicles without the necessity of requesting a takeover, similar to automated emergency braking available in current vehicles. It follows that the lack of realism may also have influenced the participants, but we feel that a takeover with evasive maneuver remains a valid tool for assessing takeover performance, as it includes the key aspects of rebuilding situation awareness, scanning the environment, and planning and executing a dynamic maneuver, including longitudinal and lateral control in interaction with other road users. One participant engaged in the TQT did not recognize the TOR and had to be warned by the experimenter, as even after 5 s no response was noticeable. At the same time, when the TOR was announced, the experimenter was answering a question of the TQT, which might have reduced the discriminability of the TOR. In the inquiry and a subsequent interview, no impaired hearing ability was reported. We can neither eliminate all doubt that the participant suffered from impaired hearing nor assume that 75 dB was insufficient for the warning. As a conclusion, a warning cascade instead of a short auditory stimulus may be beneficial in order to prevent misses of TORs.

The results of the current experiment suggest that, next to the factors driver state and human–machine interaction (HMI), the situational parameters are shown to be crucial for a safe and successful takeover. Although the cognitive TQT showed minor effects on the takeover performance, a strong influence of traffic density is evident. In consequence, a takeover might not be appropriate in every situation or has to be secured by guaranteeing good driver availability and/or providing a compensatory HMI, like suggested by Lorenz, Kerschbaum, and Schumann (2014). Reducing the vehicle's speed by automated braking could further reduce criticality (Gold et al., 2014) and enable a safer takeover.

KEY POINTS

- In a driving simulator experiment, 72 participants experienced takeover situations in a highly automated vehicle in three traffic density conditions, with or without the verbal 20-Questions Task.
- The experiment shows a distinct negative influence of traffic density on takeover performance.
- Situational complexity must not be underestimated when developing human–machine interaction for highly automated vehicles.
- The 20-Questions Task showed minor effects on the takeover performance and a distinct change of the horizontal gaze dispersion.

REFERENCES

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

Campbell, J. L., Richard, C. M., Brown, J. L., & McCallum, M. (2007). *Crash warning system interfaces: Human factors insights and lessons learned*. Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Di Nocera, F., Camilli, M., & Terenzi, M. (2006). Using the distribution of eye fixations to assess pilots' mental workload. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 63–65). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/154193120605000114

Funkhouser, D., & Sayer, J. (2012). Naturalistic census of cell phone use. *Transportation Research Record: Journal of the Transportation Research Board*, *2321*, 1–6. doi:10.3141/2321-01

Gasser, T. M. (2012). *Rechtsfolgen zunehmender Fahrzeugautomatisierung: Gemeinsamer Schlussbericht der Projektgruppe* [Legal consequences of an increase in vehicle automation]. Bergisch Gladbach, Germany: Bundesanstalt für Strassenwesen.

Gold, C., Dambock, D., Lorenz, L., & Bengler, K. (2013). Take over! How long does it take to get the driver back into the loop? In *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting* (pp. 1938–1942). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/1541931213571433

Gold, C., Lorenz, L., & Bengler, K. (2014, June). *Influence of automated brake application on take-over situations in highly automated driving scenarios*. Paper presented at the FISITA World Automotive Congress, Maastricht, Netherlands.

He, J., Becic, E., Lee, Y.-C., & McCarley, J. S. (2011). Mind wandering behind the wheel: Performance and oculomotor correlates. *Human Factors*, *53*, 13–21. doi:10.1177/0018720810391530

Heenan, A., Herdman, C. M., Brown, M. S., & Robert, N. (2014). Effects of conversation on situation awareness and working memory in simulated driving. *Human Factors*, *56*, 1077–1092. doi:10.1177/0018720813519265

International Organization for Standardization. (2013). *ISO 15623:2013. Intelligent transport systems—Forward vehicle collision warning systems—Performance requirements and test procedures*. Geneva, Switzerland: Author.

Jamson, A. H., Merat, N., Carsten, O. M., & Lai, F. C. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. *Transportation Research Part C: Emerging Technologies*, *30*, 116–125. doi:10.1016/j.trc.2013.02.008

Kemeny, A., & Panerai Francesco. (2003). Evaluating perception in driving simulation experiments. *TRENDS in Cognitive Sciences*, *7*, 31–37.

Kerschbaum, P., Lorenz, L., & Bengler, K. (2014). Highly automated driving with a decoupled steering wheel. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 1686–1690). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/1541931214581352

Körber, M., Gold, C., Lechner, D., & Bengler, K. (in press). The influence of age on the re-gain of vehicle control in highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*.

Lorenz, L., Kerschbaum, P., & Schumann, J. (2014). Designing take over scenarios for automated driving: How does augmented reality support the driver to get back into the loop? In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 1681–1685). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/1541931214581351

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. (2012). Highly automated driving, secondary task performance, and driver state. *Human Factors*, *54*, 762–771. doi:10.1177/0018720812442087

National Highway Traffic Safety Administration. (2013). *Preliminary statement of policy concerning automated vehicles*. Retrieved from http://www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated_Vehicles_Policy.pdf

Neubauer, C., Matthews, G., & Saxby, D. (2012). The effects of cell phone use and automation on driver performance and subjective state in simulated driving. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 1987–1991). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/1071181312561415

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 2063–2067). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/1541931214581434

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive systems engineering*. Wiley Series in System Engineering. New York, NY: Wiley.

Ratwani, R. M., McCurry, J. M., & Trafton, J. G. (2010). Single operator, multiple robots: An eye movement based theoretic model of operator situation awareness. In *5th ACM/IEEE International Conference on Human–Robot Interaction (HRI)* (pp. 235–242). Piscataway, NJ: IEEE.

SAE International. (2014). *SAE surface vehicle information report: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems* (No. J3016). Warrendale, PA: Author.

Schöpplein, S. E. (2013). *Integration fahrstreifenbezogener Kenngrößen und seitlicher Detektionsdaten in ein makroskopisches Verkehrsflussmodell für dreistreifige Richtungsfahrbahnen* [Integration of lane-based traffic parameters and data of sidewise detection into a macroscopic traffic flow model for three-lane carriageways] (Dissertation). Technische Universität München, Munich, Germany.

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, *48*, 1294–1313. doi:10.1080/00140130500252990

Strayer, D. L., & Drews, F. A. (2004). Profiles in driver distraction: Effects of cell phone conversations on younger and older drivers. *Human Factors*, *46*, 640–649. doi:10.1518/hfes.46.4.640.56806

Strayer, D. L., Drews, F. A., & Crouch, D. J. (2006). A comparison of the cell phone driver and the drunk driver. *Human Factors*, *48*, 381–391. doi:10.1518/001872006777724471

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, *9*, 23–32. doi:10.1037/1076-898X.9.1.23

Wang, Y., Reimer, B., Dobres, J., & Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26*, 227–237. doi:10.1016/j.trf.2014.08.003

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accident Analysis & Prevention*, *78*, 212–221. doi:10.1016/j.aap.2015.02.023

Christian Gold studied automotive and combustion engine technology with focus on automotive engineering and product development at the Technische Universität München. Since May 2012, he works at the Institute of Ergonomics as a graduate research associate in the field of automated driving. In driving simulator studies, he is predominantly researching controllability aspects of the driver-vehicle system in takeover situations of highly automated vehicles, modeling of takeover performance, and assessing trust in automation.

Moritz Körber is a graduate research assistant under Klaus Bengler at the Institute of Ergonomics at the Technische Universität München. In 2012, he earned his diploma (German equivalent to a master's degree) in psychology and business at the University of Regensburg. His thesis topic was ethical leadership and its influence on employees' challenging citizenship behavior. After working on several user experience projects, his primary research interests are vigilance, fatigue, and automation effects on driver attention.

David Lechner worked as a scientific assistant at the Institute of Ergonomics at the Technische Universität München in cooperation with Christian Gold from June 2014 until January 2015. His field of activity included the execution and evaluation of driving simulator studies and the research of takeover performance and trust in highly automated driving. At the moment he is finalizing his bachelor of science in mechanical engineering with focus on automotive engineering at the Technische Universität München.

Klaus Bengler graduated in psychology at the University of Regensburg in 1991 and received his PhD in 1995 in cooperation with BMW at the Institute of Psychology (Prof. Dr. Zimmer). After his PhD he was active on topics of software ergonomics and evaluation of human–machine interfaces (HMIs). He investigated the influence of additional tasks on driving performance in several studies within EMMIS EU project and in contract with BMW. Multifunctional steering wheels, touch screens, and adaptive cruise control functionality are examples for the topics of these investigations. In 1997 he joined BMW. From several projects he is experienced in experimental knowledge and experienced with different kinds of driving simulators and field trials. At BMW he was responsible for the HMI project of the MOTIV program, a national follow-up on the program of PROMETHEUS. Within BMW Research and Technology he was responsible for projects on HMI research and leader of the usability lab. Since May 2009 he has been leader of the Institute of Ergonomics at Technische Universität München, which is active in research areas like digital human modeling, human–robot cooperation, driver assistance HMI, and human reliability. He is leading the German Standardization Group (FAKRA) AK-10 "Mensch als Fahrzeugführer" and is an active member of ISO TC22 SC13 WG8 "Road vehicles—Ergonomic aspects of transport information and control systems," as well as a member of VDI working group "Menschliche Zuverlässigkeit."

Date received: May 5, 2015
Date accepted: January 25, 2016