Accident Analysis and Prevention 124 (2019) 202–209

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Effects of scheduled manual driving on drowsiness and response to take over request: A simulator study towards understanding drivers in automated driving

Yanbin Wu*, Ken Kihara, Yuji Takeda, Toshihisa Sato, Motoyuki Akamatsu, Satoshi Kitazaki

Automotive Human Factors Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Japan

ARTICLE INFO

Keywords:
Automated driving
Drowsiness
Eye-blink duration
Takeover
Task switching
Age difference

ABSTRACT

Because current automated vehicles have operational limitations, it is important to ensure that the fallback-ready driver is able to perform appropriately when required to take over control of the vehicle. However, time-related increase in driver drowsiness is well-known, and drowsy driving can affect response to take-over request (TOR). It was previously reported that a scheduled period of manual driving during automated driving was beneficial in maintaining driver arousal level. The present driving simulator study investigates the effects of scheduled manual driving on driver drowsiness and performance, as well as age differences therein. A total of 115 participants, whose gender was balanced and age was distributed uniformly from 20 to 70 years, drove an automated vehicle for 31 min, and a TOR was prompted before a collision event. A between-subjects design comprised two conditions: with versus without a scheduled 10-min interval of manual driving that ended 10 min before TOR. The Karolinska Sleepiness Scale and eyeblink durations estimated from electrooculograms (EOG) were used to subjectively and objectively measure participant's drowsiness. Reaction time, standard deviation of steering wheel angle, and minimum Time-to-Collison (TTC) were extracted to measure driver performance in response to TOR. The alleviating effect on drowsiness of 10-min scheduled manual driving became non-significant after another 10-min period of automated driving. Although the scheduled manual driving had no significant effect for younger drivers, older drivers reacted significantly more slowly in both steering and braking at the critical event. These findings provide essential insights for human-vehicle interactions: Scheduled manual driving cannot maintain drivers' arousal level for 10 min afterwards, and for older drivers, it would be better to avoid unnecessary task-switching between manual and automated driving.

1. Introduction

Of the many social expectations related to the launch of automated vehicles, the most compelling is the significant role that automation can play in reducing traffic accidents. Although it is agreed that the number of traffic fatalities would significantly decrease if the human driver's tasks were completely automated, it remains unknown when autonomous vehicles can be delivered that perfectly operate in all traffic situations without any human intervention. Major manufacturers are still at the stage of improving their products to deliver so-called automated driving at system SAE level 3 or higher. According to SAE definitions (SAE International, 2016), the level 3 system will take over the dynamic driving tasks and allow the driver to be less attentive to the dynamic driving environment, with hands and feet free from the steering wheel and pedals (Eriksson and Stanton, 2017). The driver's role changes from an active operator to a fallback-ready driver with the ability to engage temporarily in non-driving related tasks. However, there are certain situations (e.g., an unexpected construction zone on a highway) which the highly automated system would fail to handle, prompting a Take-Over Request (TOR). Fallback-ready drivers must be able to take over control and handle the critical situation within a reasonable transition time.

Drowsiness has been found to be correlated with the deterioration of driving performance, such as lane-keeping abilities (Stephan et al., 2006). This is one of the precursors of traffic accidents and accounts for 10%–30% of crash incidents (Philip and Åkerstedt, 2006; Vicente et al., 2016). Driver drowsiness may intensify during automated driving, because the driver's role changes from an active operator to a passive observer, a task that is very monotonous (Aria et al., 2016). Schömig et al. (2015) reported a more rapid increase of driver drowsiness when

* Corresponding author.
E-mail address: wu.yanbin@aist.go.jp (Y. Wu).

https://doi.org/10.1016/j.aap.2019.01.013
Received 28 June 2018; Received in revised form 19 December 2018; Accepted 10 January 2019
Available online 18 January 2019
0001-4575/

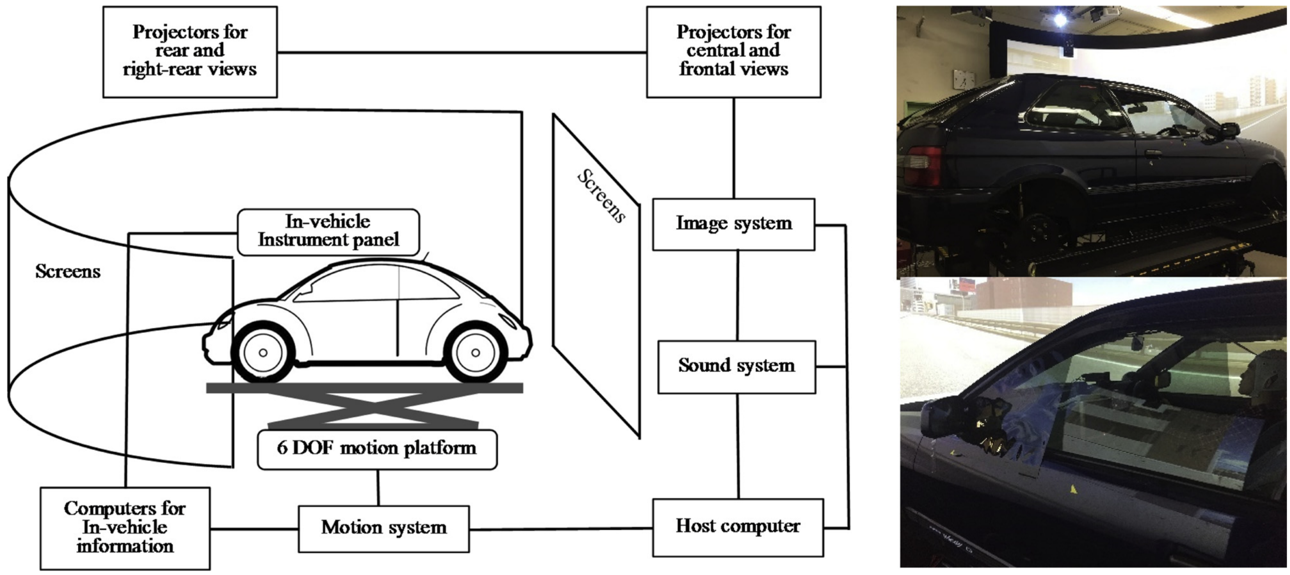

Fig. 1. Experimental apparatus: Motion based driving simulator.

highly automated driving systems were activated, but they also found that drivers' drowsiness levels stayed at a relatively low level when they performed a quiz task. Drowsiness is considered to be one of the most important factors that define driver's readiness for TOR (Schmidt et al., 2016). Previous research has reported that eye activity indices, such as percentage of eyelid closure (Wierwille et al., 1994) and eye blink duration (Ingre et al., 2006), are effective in measuring drivers' drowsiness levels. These objective measures have also demonstrated how automated driving intensifies effects of drowsiness. For instance, Jamson et al. (2013) reported that percentage of eyelid closure was greater during simulated automated driving compared to manual driving, and a heavy traffic environment did not moderate this effect. Takeda et al. (2016) designed an experiment in which they mimicked the state of the driver of an automated vehicle by assigning the monitoring task to a passenger in an actual manual driving task. Participants had a longer eyeblink duration when acting as a passenger (i.e., mimicking automated driving) than when driving, which indicated that their level of drowsiness was higher during automated than manual driving.

To improve a driver's readiness for TOR, one of the core considerations for human factors research is to develop methods that can alleviate the accumulation of drowsiness and thus increase the driver's arousal level. In an experimental study, Ralph et al. (2017) found that both rest breaks and demanding intervening tasks can alleviate the arousal decrement in prolonged monotonous tasks. Considering the characteristics of highly automated driving, one straightforward idea is to actively switch the driving mode to manual and engage the driver in the driving task in a scheduled manner to maintain the driver's arousal level. Blommer et al. (2015) designed a repeated engagement strategy for distracted drivers and examined the effects of this strategy on response performance to TOR. In their simulator study, a collision event was arranged and a TOR was prompted 38 s after a 1-min engagement in manual driving. Blommer et al. found that the intermittent manual driving slightly shortened the participants' reaction times to take-over. Although the authors speculated that their strategy increased participant arousal and mitigated fatigue, the actual mechanism of this performance improvement remains unknown. It is also unclear whether or not the positive effect of scheduled manual driving can persist for a longer duration (i.e., longer than 38 s).

The present study had two main purposes. First, we measured electrooculograms (EOG) to compare the continuous time-related changes in driver drowsiness during automated driving under two conditions: with versus without scheduled manual driving. Second, we examined the effects of scheduled manual driving on driver performance in response to TOR. We also examined the effect of age, because age-related decline in physical, sensory, and cognitive capabilities may place older drivers at a higher risk of accidents even when they ride in an automated vehicle (Young et al., 2017). In this study, the manipulation of scheduled manual driving involved task switching between an automated driving task and manual driving task. For instance, a driver in a manual vehicle serves as an active operator and continuously performs longitudinal and lateral control, while a driver in an automated vehicle serves as a fallback-ready user and is required to monitor the system status. Since it is well known that task-switching performance decreases with age (Kray and Lindenberger, 2000), it is possible that driving mode switching could interact with driver age and affect response to TOR.

2. Methods

2.1. Participants and apparatus

A total of 115 participants (55 females) who were active licensed drivers and in normal health were recruited from the local community. To study the linear correlation between age and TOR performance, we recruited participants ranging from 18 to 75 years old (mean = 44.6, SD = 15.6), with age distributed as uniformly as possible. We divided the participants into five age groups: 20 or younger, 30, 40, 50, and 60 and older. For each age group, 11 female and 12 male participants were recruited (11*5 + 12*5 = 115). All participants had normal vision without correction and were paid for their participation. The National Institute of Advanced Industrial Science and Technology (AIST) Safety and Ethics committee approved the study, and each participant gave written informed consent before the experiment.

All experiments were conducted in the AIST high-fidelity driving simulator (Mitsubishi Precision Co.), which was mounted on an electric motion platform with six degrees of freedom (DOF) (see Fig. 1). The driving simulator provided real time and high-fidelity tactile feedback to participants. The front and side views were projected on a cylindrical screen surrounding the motion platform with a field of view greater than 180°, and the rear view, which could be seen via the rearview mirror and side mirrors, was projected on a flat screen located behind the motion platform. The reactive force against the steering action was modulated by a steering control loading system (Moog Inc.).

2.2. Experiment

Three experimental conditions were designed: automated driving for 3 min (denoted as Auto-3); automated driving for approximately 31 min (Auto-31); and automated driving for approximately 10 min, followed by approximately 10 min of manual driving, then followed by automated driving for another approximately 10 min (A-M-A). The Auto-3 condition was considered as a practice session for driving an automated vehicle. The traffic scenario was built on a bidirectional loop motorway with two lanes in each direction, with 50 simulated vehicles flowing at a speed of 30 km/h, 60 km/h, or 90 km/h in both directions. Participants were instructed to run in the left lane. The cruising speed in the automated mode and the speed limit of the manual mode were both set to 60 km/h, which is the upper limit of Japanese legal speed in rural traffic. The constraint of a constant speed eliminated the effect of speed on driving performance measures and allowed us to focus on performance variations caused by experimental conditions. At the end of each experimental condition, a stopped van unfolded in front of the driver's vehicle and TOR was prompted at the same time. The stopped van had a headway of 100 m, which together with the speed of 60 km/h resulted in a time budget of 6 s. The choice of 6 s was made according to previous studies (Eriksson and Stanton, 2017) and our pilot study. After TOR, the drivers had to manually drive the car, and by braking and/or steering, had to change to the right lane to avoid a collision with the stopped van. The experiment ended 20 s after the TOR was prompted.

Before the experiment, the experimenters first introduced the participants to the experimental scenario. The participants were then instructed to complete a questionnaire about their driving habits. They were required not to put their hands on the steering wheel and not to put their feet on the pedal during automated driving mode. In the experiment, participants were not allowed to activate or deactivate the automated system by themselves. The activation and deactivation of the automated system were programmed based on the timing of the scenario, so that we could accurately control the duration of the automated driving and manual driving. The two types of transitions from automated to manual were explained, and participants were instructed to take over control when the scheduled manual driving came online or a TOR was prompted (i.e., before the critical collision event). At the beginning of each condition, a verbal message ('Automated driving starts' in Japanese) was played. For the scheduled manual driving, an auditory notice (verbal message of 'The system will transition to manual driving soon; please prepare to take over control' in Japanese) was played 20 s before control transition. At the end of the scheduled manual driving, an auditory notice (verbal message of 'Automated driving starts' in Japanese) was played, and the drivers were instructed to remove their hands and feet from the wheel and pedal, and control was transitioned back to the vehicle. TOR was prompted in both auditory (verbal message of 'Take over control' in Japanese) and visual (a symbol changed on the vehicle central control panel) information modalities. No prenotice information was presented before a TOR, and no further detailed information about the TOR situation was given to the participants.

All 115 participants first practiced driving the simulator manually for 8 min. Two participants reported simulator sickness and one participant reported fear of the motion-based driving simulator and quit the experiment. After practicing, the remaining participants first performed the Auto-3 test, followed by either the Auto-31 or the A-M-A test. Of these 112 participants, 84 participants performed the Auto-31 test, and the remaining 28 participants completed the A-M-A test. This unbalanced design was a consequence of the availability of the experimental apparatus. Finally, the participants completed the Karolinska Sleepiness Scale (KSS) immediately after each driving condition to report their subjective estimation of drowsiness.

2.3. Measures

2.3.1. Driving performance measures

An ideal performance in response to TOR would be characterized by a fast reaction, smooth maneuvers, and maintenance of an adequate safety margin. In this study, the following measures were used.

(1) Time-Steer-1: time consumed after TOR until the steering wheel is turned right by 1°. This reflects the reaction time of effective steering. Steering right more than 1° represents an active reaction that begins a change in the running lane.
(2) Time-Brake-0.1: time consumed until the brake pedal is pressed to 10 percent of full braking. Similar to the measure Time-Steer-1, this represents the reaction time of the braking operation and is defined as "the time between the take-over request and the point when the standardized brake pedal travel was larger than 10%" (Zeeb et al., 2016).
(3) Reaction time: the smaller value of the above two measures. In the present experimental scenario, steering and braking are the two alternative take-over actions to prevent a collision. The minimum value between Time-Steer-1 and Time-Brake-0.1 represents the reaction time of taking over control (Petermeijer et al., 2017).
(4) sdSteer: the standard deviation of the steering wheel angle after the lane change. sdSteer is the measure of maneuvering smoothness. In the TOR of the present experiment, the goal of the driving operation is to regain control of the vehicle and bypass an obstacle smoothly. This measure is especially sensitive to the driver's performance variation in handling such abrupt events and explicitly reflects the stability of steering operations. For example, Brookhuis et al. (1991) found that when the driver actively responded to a ringing mobile phone, the standard deviation of the steering wheel position could be 10 times larger than that on a quiet motorway.
(5) TTC: time to collision during the moment of the lane change. This is the quotient of headway divided by vehicle velocity. TTC at the moment of lane change reflects the safety margin in response to TOR.

2.3.2. Eyeblink duration

Eye activity indices such as eye blink duration (Caffier et al., 2003; Ingre et al., 2006) have been shown to be effective in measuring drowsiness level. Caffier et al. concluded that the prolongation of eyeblink duration mainly resulted from lengthening of eyelid reopening time with drowsiness. EOG allows the detection of eyeball movements and eye blink activities with a high temporal resolution and accuracy. In this study, a digital amplifier (Brain Products, BrainAmp Standard system) and silver-silver chloride electrodes with a sampling rate of 1000 Hz were used to record participant EOG. Electrodes were placed on the upper and lower eyelids to record the vertical EOG. The method of estimating eyeblink durations from the vertical EOG was similar to that used in Takeda et al. (2016). Vertical EOG signals were band-pass filtered offline at 1–30 Hz using a second-order Butterworth filter. A custom MATLAB program was used to detect eyeblinks based on the velocity of the vertical EOG. The time interval between initiation and termination at 50% of the blink amplitude was considered as the duration of each eyeblink. Time series of eyeblink duration detected from the original EOG record were cropped into 2-min epochs, and epoch average was calculated to estimate participant's drowsiness during each 2-min epoch.

2.3.3. Karolinska Sleepiness Scale

The Karolinska Sleepiness Scale (KSS) is a 9-point verbally anchored self-report scale ranging from *extremely alert* (scale 1), *alert* (scale 3), *neither alert nor sleepy* (scale 5), *sleepy but no difficulty remaining awake* (scale 7), to *extremely sleepy and fighting sleep* (scale 9) (Åkerstedt and Gillberg, 1990). KSS is widely used as a subjective measure of drowsiness; it has been found to be correlated with changes in EOG and is reflective of significant effects of sleepiness on task performance (Åkerstedt and Gillberg, 1990). A Japanese version of KSS (Kaida et al., 2006) was used to measure participants' drowsiness after each experimental condition.

2.4. Statistical analysis

First, to estimate the drowsiness level after each experiment condition, the means and standard deviations of KSS and eyeblink duration (for the final 2-min epoch before TOR) were calculated. Cross-subject *Pearson's* correlation coefficients were calculated between KSS and eyeblink duration within each condition. Unpaired t-tests were also conducted to compare Auto-31 and A-M-A conditions in each time epoch. To control the false discovery rate (i.e. type I error), significant level of the multiple comparisons was corrected based on Benjamini-Hochberg method (Benjamini and Hochberg, 1995). To investigate the interaction effects of driver age and driving condition, a one-way analysis of covariance (ANCOVA) was conducted on performance measures using the MATLAB aoctool function. ANCOVA partials out the effects of a covariate and examines the effect of treatment within the framework of multiple regression (Howell, 2010); here, the participant's age was the covariate and test condition was considered the treatment factor. The significance level p was set to 0.05.

3. Results

3.1. KSS and eyeblink duration

KSS was used as a subjective drowsiness measure, and average eyeblink duration extracted from the last 2-min epoch was used as an objective measure. Significant linear correlations were found between KSS and average eyeblink duration after Auto-31 ($r = 0.53$, $p < 0.001$), and a tendency was found after A-M-A ($r = 0.36$, $p = 0.07$), indicating that these measures are closely related to each other in terms of drowsiness. These results are in accordance with previous studies that found that eyeblink duration is closely associated with subjective KSS ratings in driving (e.g., Åkerstedt et al., 2005), and therefore, we used eyeblink duration as a continuous drowsiness measure. Fig. 2 depicts average KSS and average eyeblink durations for Auto-31 and A-M-A conditions. The one-way ANCOVA of KSS (Fig. 3a) revealed no effect of driving condition (treatment), $F(1,108) = 0.26$, $p = 0.61$, and a significant effect of participant age (covariate), $F(1,108) = 7.87$, $p = 0.006$, while their interaction was not significant ($p = 0.83$). A very similar pattern of ANCOVA was found for the objective measure of eyeblink duration (Fig. 3b). Only participant age, $F(1,104) = 12.7$, $p < 0.001$ had a significant effect on eyeblink duration, and neither the effect of driving condition nor their interaction effect was significant. These results for subjective reports and eyeblink index indicate that driver drowsiness levels were quite high after 30 min of automated driving. A scheduled 10-min manual driving could not moderate driver drowsiness at the end of the drive.

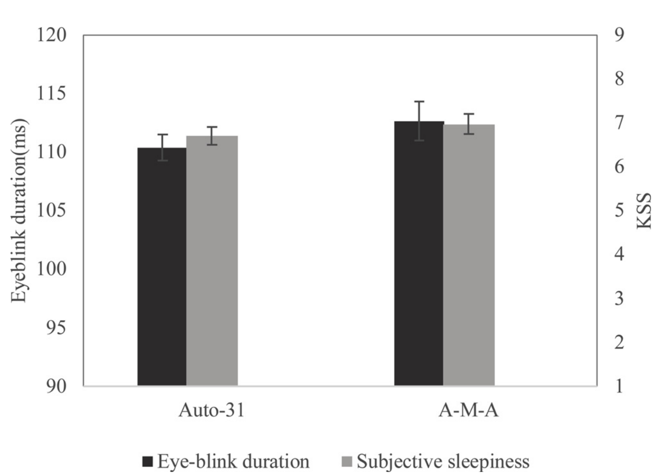

Fig. 2. Average KSS scores after each experimental condition and Average eyeblink duration of the last two-minute epoch before TOR. Error bars indicate standard errors.

We also examined the temporal variations of eyeblink duration extracted from each 2-min epoch in Auto-31 and A-M-A conditions. To eliminate individual differences in eyeblink durations, each participant's first epoch was used as baseline and was subtracted from the data. The results are shown in Fig. 4, with the grey shadows indicating the 95% confidence intervals estimated from all participants' data. An unpaired *t*-test was used to statistically test the difference between Auto-31 and A-M-A conditions for each epoch. After Benjamini-Hochberg correction, it was found that for the 6th, 7th, 8th, 10th, and 11th epoch, eyeblink duration was significantly shorter for the A-M-A condition than for the Auto-31 condition. In the Auto-31 condition, participants' eyeblink duration gradually increased to ceiling level after about 15 min. In the A-M-A condition, eyeblink durations decreased when the scheduled manual driving began (11th min), and eyeblink durations were smaller during manual driving epochs than during automated driving epochs. After control was relinquished to the automated vehicle (21st min), eyeblink durations increased back to ceiling level within 2–3 epochs (4–6 min). Note that eyeblink duration in the final epoch decreased in both conditions because this epoch included driving after TOR, which should have increased drivers' arousal level.

3.2. Interaction effects of driving condition and age on performance

The effects of driving condition and participant age on performance measures were examined using ANCOVA.

3.2.1. Reaction time

Average Time-Steer-1 after Auto-31 was 2.20 ± 0.71 s and after A-M-A was 2.54 ± 0.71 s. The ANCOVA of Time-Steer-1 (Fig. 5a) found a significant effect of driving condition, $F(1,108) = 5.37$, $p = 0.02$. Participant age had no significant effect, $F(1,108) = 2.08$, $p = 0.15$. More importantly, there was a significant interaction between driving condition and age, $F(1,108) = 6.5$, $p = 0.01$. The interaction effect is reflected in the difference in slope of the regression lines shown in Fig. 5a. The slope of Auto-31 (solid line) condition is almost 0, whereas the slope of A-M-A (dashed line) is positive.

The pattern of results for measure Time-Brake-0.1 was very similar to that for Time-Steer-1, although only about one third of all participants adopted a braking maneuver of more than 10% to avoid crashing into the stopped van. Average Time-Brake-0.1 after Auto-31 and A-M-A was 2.16 ± 0.69 and 2.27 ± 1.0 s, respectively. ANCOVA of Time-Brake-0.1 (Fig. 5b) revealed a significant effect of age, $F(1,36) = 12.3$, $p = 0.001$, and a significant interaction effect of condition and age, $F(1, 36) = 9.5$,$p = 0.004$.

Average reaction time to TOR in Auto-31 and A-M-A conditions was 2.05 ± 0.61 and 2.42 ± 0.77 s, respectively. ANCOVA of reaction time (Fig. 6) revealed significant effects of condition, $F(1,108) = 7.34$, $p = 0.008$, a non-significant effect of participant age, $F(1,108) = 2.95$, $p = 0.09$, and a significant interaction, $F(1,108) = 8.46$, $p = 0.004$. The interaction effect is reflected in the difference in slope of the regression lines shown in Fig. 6, which indicates that older participants tended to react significantly more slowly in the A-M-A condition, while this effect was not observed in Auto-31.[1]

[1] To account for the unequal sample size between the Auto-31 and A-M-A conditions, a bootstrap method was further used to check this interaction effect statistically. By randomly sampling 28 participants from the Auto-31 condition (with the number of participants belonging to the five age groups controlled to correspond to that of A-M-A) without preplacement 1000 times, about 75% of

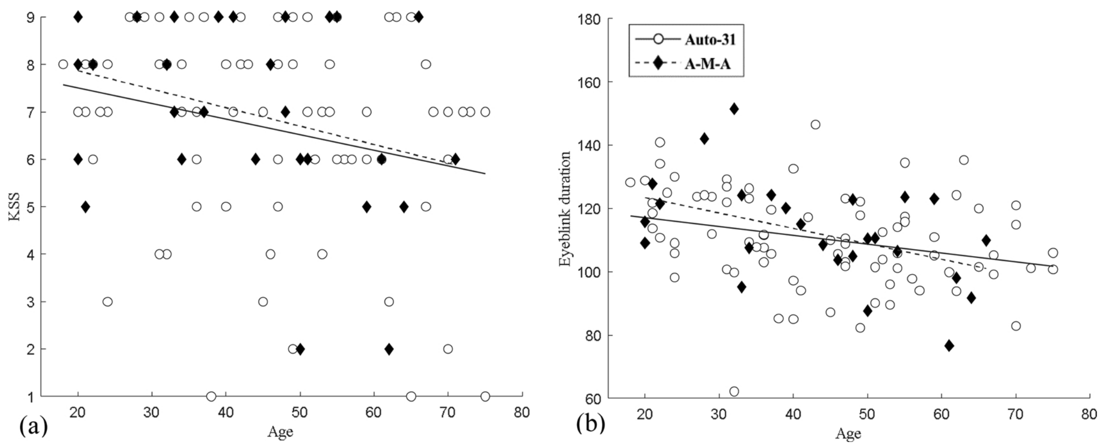

Fig. 3. ANCOVA: effects of participant's age (covariate) and driving condition (treatment) on (a) KSS and (b) eyeblink duration. Each point represents the data of one participant, and a line is linearly fitted across participants under each of the three driving conditions. The difference between the slope of those fitted lines reflects the interaction effect of covariate and treatment. The solid and dashed lines indicate the regression lines for Auto-31 and A-M-A respectively.

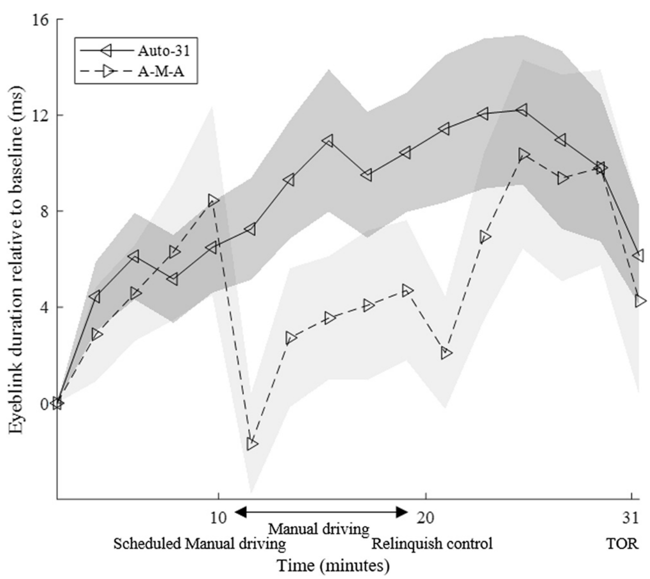

Fig. 4. Time series of average eyeblink durations (baseline subtracted) in Auto-31 and A-M-A conditions. The first 2-min epoch was used as baseline and therefore set to zero. In the Auto-31 condition, participants drove the automated vehicle for approximately 31 min. In the A-M-A condition, scheduled manual driving began from approximately the 11th min, the driver manually drove the car from approximately the 11th min to the 20th min, and at about the 21st min the driver relinquished control to the vehicle. For both conditions, a TOR was prompted at approximately the 31st min. The grey shadows show the 95% confidence intervals. Unpaired t-tests (Benjamini-Hochberg adjusted) indicate that eyeblink duration was significantly shorter for the A-M-A condition than for the Auto-31 condition from the 6th to 11th epoch with only the 9th epoch as an exception. Epochs with significant difference were marked by *.

3.2.2. Standard deviation of steering wheel angle

Average sdSteer in Auto-31 and A-M-A condition was 5.20 ± 2.06 and 5.02 ± 1.51 degrees. The results of ANCOVA (Fig. 7) indicated no effect of driving condition on sdSteer, $F(1,108) = 0.09$, $p = 0.76$. Participant age had a significant effect on sdSteer, $F(1,108) = 14$, $p < 0.001$, but there was no significant interaction between driving time and age, $F(1,108) = 0.19$, $p = 0.67$. These results suggest that older drivers performed worse than younger drivers in smoothly regaining control of an automated vehicle, and this effect was not affected by intervention of manual driving.[2]

3.2.3. Time to collision

The average minimum TTC in Auto-31 and A-M-A conditions was 1.92 ± 0.60 and 1.70 ± 0.60 s. ANCOVA revealed no significant effects of age or driving condition, or their interaction, on this performance measure. As shown in Fig. 8, a downward slope of A-M-A condition suggested a declining tendency of TTC with an increase in participant age, although this was not statistically significant.

While significant effects of participant age and driving condition on reaction time were found (Fig. 6), ANCOVA revealed only a non-significant tendency for TTC. The significantly slower reaction may have been compensated by harder braking and slower velocity adopted by older drivers. A similar result was reported by Michaels et al. (2017), who found that older drivers in manual driving preferred to drive more slowly than younger drivers to compensate for their slower reactions.

4. Discussion

The purposes of this paper were to compare the continuous time-related changes of driver drowsiness in prolonged automated driving with versus without scheduled manual driving, and to examine the effects of scheduled manual driving on performance in response to TOR.

4.1. Effects on drowsiness

Eyeblink duration and KSS were consistent in measuring the increased drowsiness caused by longer driving durations. The average KSS score after Auto-31 and A-M-A conditions was around 7, representing a state of drowsiness in which drivers experience involuntary dozing behavior and frequent lapse episodes (Azizan et al., 2017). A significant decrease in neither KSS nor eyeblink duration could be found for the A-M-A condition compared to Auto-31. The significant correlations between KSS and eyeblink duration supported previous indications that eyeblink duration is a very informative eyeblink parameter in inferring drowsiness (Caffier et al., 2003).

As shown in Fig. 4, eyeblink durations significantly decreased in the A-M-A condition when the scheduled manual driving began and were

(footnote continued)
the samples reached a significant level ($p < 0.05$) and about 90% reached a marginally significant level ($p < 0.1$).

[2] The same collision event was arranged at the end of the practice session for manual driving, and these data were recorded and analyzed. No significant correlations were found between driver age and driving performance measures. This suggests that the age-associated decline of steering smoothness was specific to the event related to the transition from automated to manual driving.

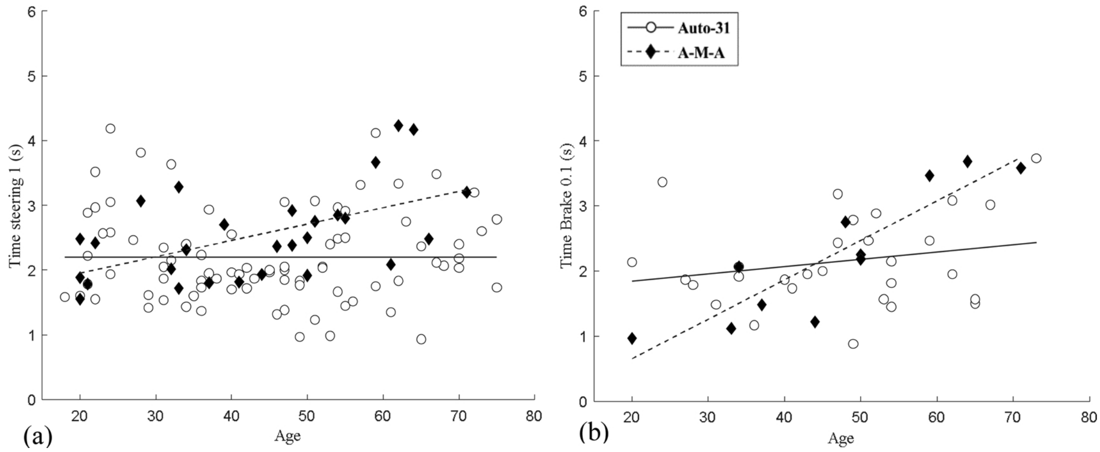

Fig. 5. ANCOVA: effects of participant's age (covariate) and driving condition on (a) Time-Steer-1 and (b) Time-Brake-0.1.

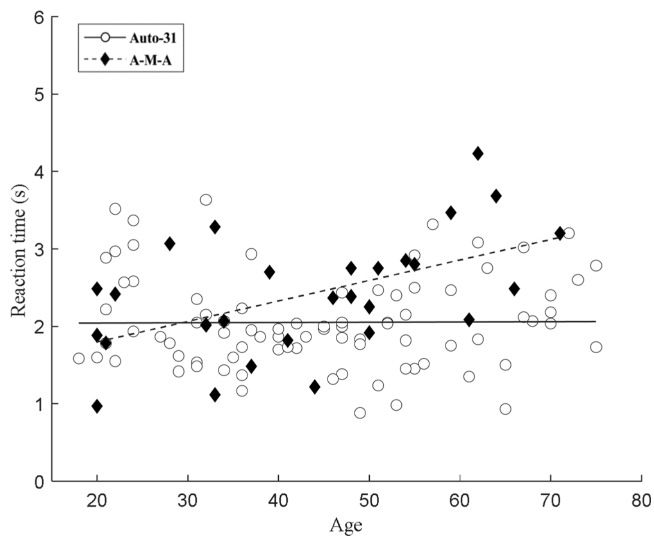

Fig. 6. ANCOVA: effects of participant age (covariate) and driving condition on reaction time.

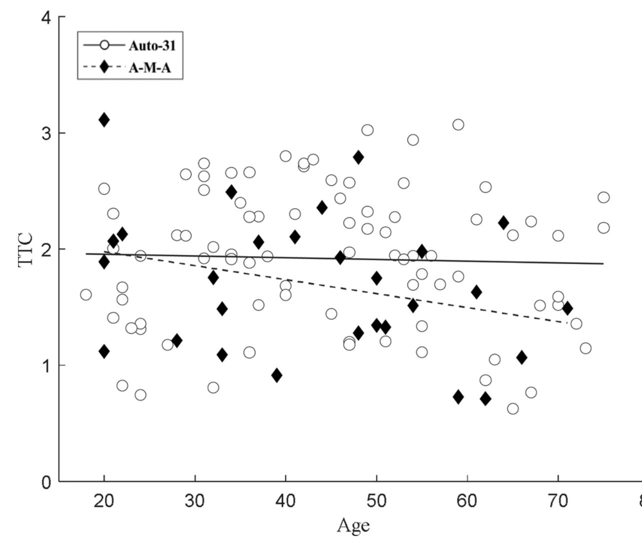

Fig. 8. ANCOVA: effects of participant's age (covariate) and driving condition on Time-To-Collision at lane change.

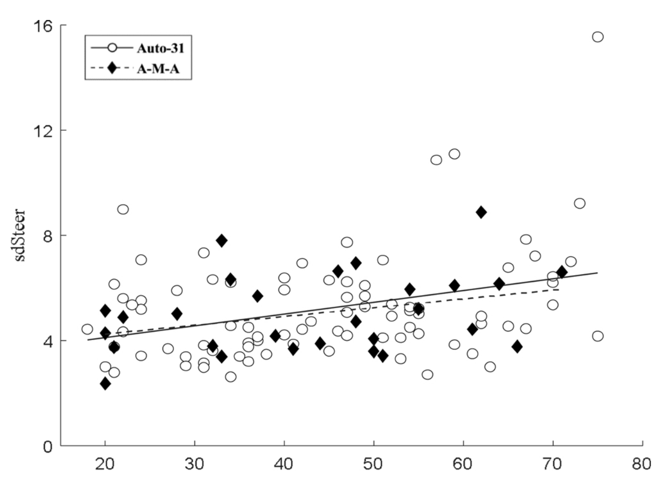

Fig. 7. ANCOVA: effects of participant age (covariate) and driving condition on standard deviation of steering wheel angle.

significantly smaller during the following manual driving epochs than during the same epochs in the Auto-31 condition. This result is in accordance with previous studies (Jamson et al., 2013; Takeda et al., 2016) in which drivers showed lower arousal levels in automated driving than in manual driving. In the A-M-A condition, driver drowsiness increased to a similarly higher level (at ceiling) within about two to three 2-min epochs after relinquishing control back to the vehicle, compared to the Auto-31 condition. In other words, the positive effects of 10-min manual driving on alleviating drowsiness faded away within about 4–6 min. Therefore, we conclude that the positive effects of the scheduled 10-min manual driving on drowsiness do not persist for as long as 10 min. It may be argued that driver arousal can be maintained at a higher level if the driver were required to manually drive the car for an interval of less than 10 min. However, drivers' experience and satisfaction with the automated vehicle may be negatively affected if they are required to engage in manual driving tasks more frequently.

Driver age had significant effects on both KSS and eyeblink durations, indicating a higher level of drowsiness in younger drivers. This finding is in line with an earlier survey study in which younger drivers appeared to be associated with increased frequency of drowsy driving (McCartt et al., 1996). This effect may result from a less regular sleep pattern for younger participants or may reflect a “generation gap” that requires further study.

4.2. Effects on response performance to TOR

The present experimental study failed to observe any significantly positive effects of scheduled manual driving on performance in response to TOR. By contrast, Blommer et al. (2015) observed a weak but significant performance improvement in response to a critical TOR prompted 38 s after a 1-min manual driving epoch; they speculated that the intermittent manual driving increased participant arousal levels during the event. This inconsistency may be explained by the longer interval (approximately 10 min) between relinquishing control back to vehicle and the TOR used in present study. Be that as it may, the present finding of time-related changes in drowsiness (Fig. 4) can to some degree support Blommer et al.'s speculation, because participant arousal levels were indeed higher during the first 2-min epoch after the scheduled manual driving.

The most interesting finding of this paper is the significant negative effect of scheduled manual driving on older drivers' reaction time to TOR. Specifically, older drivers reacted more slowly (as shown in Fig. 5) than younger drivers in the A-M-A condition. This negative effect was not significant for Auto-31, indicating that the effect cannot simply be attributed to the driving duration of 30 min. This negative effect also cannot be attributed to drowsiness, because older drivers did not show a higher level of drowsiness in the A-M-A condition than in the Auto-31 condition.

It may be that older drivers are more vulnerable to the mental fatigue caused by the task switching that requires a functionality shift, and thus they react more slowly. Lorist et al. (2000, 2009) found that task switching causes an increase in mental fatigue, leading to a slower reaction to stimulus; further, older adults are more vulnerable to mental fatigue induced by a cognitively demanding task (Arnau et al., 2017). Scheduled manual driving requires the driver to switch between manual and automated driving two times, and older drivers may be more vulnerable to the mental fatigue resulting from this task switching. Previous studies have shown that performance degradation of task-switching varies across the life span, with significant age-associated increments in costs, suggesting that older adults are less able to efficiently maintain and coordinate two alternating task sets (Kray and Lindenberger, 2000; Cepeda et al., 2001). In addition, unnecessary task switching may increase the problem of mode confusion (Lee and Ahn, 2015), further negatively affecting older drivers' performance. No matter the mechanism, the present result provides insight on one implication for the design of TOR logic: to prevent performance degradation in response to TOR, it may be better to avoid triggering unnecessary task switching between manual driving and automated driving, especially for older drivers.

5. Conclusions

In summary, the present study provides the following insights on the effects of scheduled manual driving on drowsiness and performance in taking over an automated vehicle.

a) Driver drowsiness significantly decreases when scheduled manual driving begins, but the 10-min period of scheduled manual driving can only alleviate drowsiness for a very short time (4–6 min) after control is relinquished to the vehicle.
b) Although scheduled manual driving has no significant effect on younger drivers' reaction time to TOR, with scheduled manual driving, older drivers react significantly more slowly in both steering and braking. This slower reaction may be attributed to the negative effects of task switching.

Finally, it is important to acknowledge the limitations of the simple design of the 10-min-interval task-switching scenario. It is possible that in a real automated vehicle, compared to a simulated one, drivers' drowsiness level could evolve differently. In addition, because the present paper intended to study driver drowsiness across driving time, the participants were not allowed to engage in a secondary task, which may have alleviated drowsiness. It is likely that in an actual situation, drivers would tend to engage in secondary tasks when they are in automated driving modes. Furthermore, in the present study, to reduce the surprise effects of the occurrence of TOR, each participant performed the Auto-3 condition before starting the A-M-A or Auto-31 condition. Although this manipulation may have been useful in reducing the surprise effect, it may have induced a learning effect that facilitated the response to the TOR. In real driving scenes, it is expected that drivers will not experience TOR so frequently. Future experiments are needed to clarify these issues.

Acknowledgements

This work was supported by Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), "Large-scale Field Operational Test for Automated Driving Systems" (funding agency: NEDO).

References

Åkerstedt, T., Gillberg, M., 1990. Subjective and objective sleepiness in the active individual. Int. J. Neurosci. 52 (1–2), 29–37.
Åkerstedt, T., Peters, B., Anund, A., Kecklund, G., 2005. Impaired alertness and performance driving home from the night shift: a driving simulator study. J. Sleep Res. 14 (1), 17–20.
Aria, E., Olstam, J., Schwietering, C., 2016. Investigation of automated vehicle effects on driver's behavior and traffic performance. Transp. Res. Procedia 15, 761–770.
Arnau, S., Möckel, T., Rinkenauer, G., Wascher, E., 2017. The interconnection of mental fatigue and aging: an EEG study. Int. J. Psychophysiol. 117, 17–25.
Azizan, A., Fard, M., Azari, M.F., Jazar, R., 2017. Effects of vibration on occupant driving performance under simulated driving conditions. Appl. Ergon. 60, 348–355.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57 (1), 289–300.
Blommer, M., Curry, R., Kochhar, D., Swaminathan, R., Talamonti, W., Tijerina, L., 2015. The effects of a scheduled driver engagement strategy in automated driving. September. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 59(1) 1681–1685.
Brookhuis, K.A., de Vries, G., De Waard, D., 1991. The effects of mobile telephoning on driving performance. Accid. Anal. Prev. 23 (4), 309–316.
Caffier, P.P., Erdmann, U., Ullsperger, P., 2003. Experimental evaluation of eye-blink parameters as a drowsiness measure. Eur. J. Appl. Physiol. 89 (3-4), 319–325.
Cepeda, N.J., Kramer, A.F., Gonzalez de Sather, J., 2001. Changes in executive control across the life span: examination of task-switching performance. Dev. Psychol. 37 (5), 715.
Eriksson, A., Stanton, N.A., 2017. Takeover time in highly automated vehicles: noncritical transitions to and from manual control. Hum. Factors 59 (4), 689–705.
Howell, D.C., 2010. Statistical Methods for Psychology, international edition. Wadsworth, Cengage Learning, Belmont, CA.
Ingre, M., Akerstedt, T., Peters, B., Anund, A., Kecklund, G., 2006. Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. J. Sleep Res. 15 (1), 47–53.
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. Transp. Res. Part C Emerg. Technol. 30, 116–125.
Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., Fukasawa, K., 2006. Validation of the Karolinska Sleepiness Scale against performance and EEG variables. Clin. Neurophysiol. 117 (7), 1574–1581.
Kray, J., Lindenberger, U., 2000. Adult age differences in task switching. Psychol. Aging 15 (1), 126.
Lee, S.H., Ahn, D.R., 2015. Design and verification of driver interfaces for adaptive cruise control systems. J. Mech. Sci. Technol. 29 (6), 2451–2460.
Lorist, M.M., Klein, M., Nieuwenhuis, S., De Jong, R., Mulder, G., Meijman, T.F., 2000. Mental fatigue and task control: planning and preparation. Psychophysiology 37, 614–625.
Lorist, M.M., Bezdan, E., ten Caat, M., Span, M.M., Roerdink, J.B.T.M., Maurits, N.M., 2009. The influence of mental fatigue and motivation on neural network dynamics: an EEG coherence study. Brain Res. 1270, 95–106.
McCartt, A.T., Ribner, S.A., Pack, A.I., Hammer, M.C., 1996. The scope and nature of the drowsy driving problem in New York State. Accid. Anal. Prev. 28 (4), 511–517.
Michaels, J., Chaumillon, R., Nguyen-Tri, D., Watanabe, D., Hirsch, P., Bellavance, F., et al., 2017. Driving simulator scenarios and measures to faithfully evaluate risky driving behavior: a comparative study of different driver age groups. PLoS One 12 (10), e0185909.
Petermeijer, S., Bazilinskyy, P., Bengler, K., de Winter, J., 2017. Take-over again: investigating multimodal and directional TORs to get the driver back into the loop. Appl. Ergon. 62, 204–215.
Philip, P., Åkerstedt, T., 2006. Transport and industrial safety, how are they affected by sleepiness and sleep restriction? Sleep Med. Rev. 10 (5), 347–356.
Ralph, B.C., Onderwater, K., Thomson, D.R., Smilek, D., 2017. Disrupting monotony while increasing demand: benefits of rest and intervening tasks on vigilance. Psychol. Res. 81 (2), 432–444.
SAE International, 2016. Surface Vehicle Recommended Practice, Taxonomy and

Definitions for Terms Related to Driving Automation Systems for On-road Motor Vehicles, J3016. Revised September 2016.

Schmidt, J., Braunagel, C., Stolzmann, W., Karrer-Gauß, K., 2016. Driver drowsiness and behavior detection in prolonged conditionally automated drives. June. 2016 IEEE Intelligent Vehicles Symposium (IV). pp. 400–405.

Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., Othersen, I., 2015. The interaction between highly automated driving and the development of drowsiness. Procedia Manuf. 3, 6652–6659.

Stephan, K., Hosking, S., Regan, M., Verdoorn, A., Young, K., Haworth, N., 2006. Monash University Accident Research Centre Report Series, Report 252. The relationship between driving performance and the Johns Drowsiness Scale as measured by the Optalert System.

Takeda, Y., Sato, T., Kimura, K., Komine, H., Akamatsu, M., Sato, J., 2016. Electrophysiological evaluation of attention in drivers and passengers: toward an understanding of drivers' attentional state in autonomous vehicles. Transp. Res. Part F Traffic Psychol. Behav. 42, 140–150.

Vicente, J., Laguna, P., Bartra, A., Bailón, R., 2016. Drowsiness detection using heart rate variability. Med. Biol. Eng. Comput. 54 (6), 927–937.

Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., Kirn, C.L., 1994. Research on Vehicle-Based Driver status/performance Monitoring; Development, Validation and Refinement of Algorithms for Detection of Driver Drowsiness. Final Report (DOT HS 808 247). National Highway Traffic Safety Administration., Washington, DC.

Young, K.L., Koppel, S., Charlton, J.L., 2017. Toward best practice in human machine interface design for older drivers: a review of current design guidelines. Accid. Anal. Prev. 106, 460–467.

Zeeb, K., Buchner, A., Schrauf, M., 2016. Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. Accid. Anal. Prev. 92, 230–239.