Accident Analysis and Prevention 123 (2019) 150–158

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

The effects of takeover request modalities on highly automated car control transitions

Sol Hee Yoon, Young Woo Kim, Yong Gu Ji*

Department of Industrial Engineering, Yonsei University, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Takeover request
Highly automated driving
Alert modality
Non-driving related task
Control transition

ABSTRACT

This study investigated the influences of takeover request (TOR) modalities on a drivers' takeover performance after they engaged in non-driving related (NDR) tasks in highly automated driving (HAD). Visual, vibrotactile, and auditory modalities were varied in the design of the experiment under four conditions: no-task, phone conversation, smartphone interaction, and video watching tasks. Driving simulator experiments were conducted to analyze the drivers' take-over performance by collecting data during the transition time of re-engaging control of the vehicle, the time taken to be on the loop, and time taken to be physically ready to drive. Data were gathered on the perceived usefulness, safety, satisfaction, and effectiveness for each TOR based on a self-reported questionnaire. Takeover and hands-on times varied considerably, as shown by high standard deviation values between modalities, especially for phone conversations and smartphone interaction tasks. Moreover, it was found that participants failed to take over control of the vehicle when they were given visual TORs for phone conversation and smartphone interaction tasks. The perceived safety and satisfaction varied for the NDR task. Results from the statistical analysis showed that the NDR task significantly influenced the takeover time, but there was no significant interaction effect between the TOR modalities and the NDR task. The results could potentially be applied to the design of safe and efficient transitions of highly controlled, automated driving, where drivers are enabled to engage in NDR tasks.

1. Introduction

Research efforts and discussions are currently in progress regarding approaches to providing safe transitions of control between drivers and vehicles in highly automated vehicles (Naujoks et al., 2014; Gold et al., 2013; Zeeb et al., 2015). Experts on automated driving systems argue that while most vehicles are not completely autonomous, safety issues will arise in highly automated vehicles (Merat et al., 2014; Gold et al., 2016). For example, relevant research topics may include the transition of control between the vehicles and driver, the drivers' roles in the vehicles, degradation of driving skills, engagement in non-driving related (NDR) tasks, and "out of the loop" situations with decreased driver awareness (Banks and Stanton, 2016; Kun et al., 2016; Meschtscherjakov et al., 2015; Clark and Feng, 2016; Miller et al., 2014). In the interim, research studies and car manufacturers are expected to focus on vehicles with partial automation. Automation is classified in five levels (levels 0–4) (NHTSA, 2013) based on the definition by the United States National Highway Traffic Safety Administration (NHTSA). Highly automated driving (HAD) refers to NHTSA level three (limited self-driving automation) where drivers are not expected to be constantly fully aware, and to support drivers with lateral and longitudinal control. For level three, it is only necessary for the driver to regain control when the vehicle requires it (NHTSA, 2013). Methods for transmitting information to the driver regarding actions to be taken are highly relevant in HAD because drivers are not expected to be constantly paying attention and monitoring the vehicle. Studies need to be aimed at enhancing the development of safety systems for highly automated vehicles with emphasis on the drivers' behavior for HAD. It should be noted that several factors are related to the takeover task which can be categorized in accordance with the driver, their environment, and condition of the vehicle (Radlmayr et al., 2014; Gold et al., 2016).

Therefore, the goal of this study was to investigate the effect of the NDR task and takeover request (TOR) modality on the drivers' transition to manual driving in HAD. Consequently, we evaluated the takeover performance of the driver based on takeover time, the time taken to fix drivers' eyes on the loop, and the time taken to reach the steering wheel. This was done in seven different alert modalities under four different conditions, which included three NDR tasks and a no-task condition. We then investigated if the perceived satisfaction,

* Corresponding author at: Dept. of Industrial Engineering, Yonsei University, 50 Shinchon-ro, Seodaemun-gu, Seoul, 03722, Republic of Korea.
E-mail address: yongguji@yonsei.ac.kr (Y.G. Ji).

https://doi.org/10.1016/j.aap.2018.11.018
Received 13 September 2017; Received in revised form 20 November 2018; Accepted 22 November 2018
Available online 29 November 2018

effectiveness, safety, and usefulness of the TOR modalities varied with the NDR task that was engaged by the drivers in HAD. Thus, we examined the relationship between the TOR modalities and NDR tasks in HAD.

1.1. Alert modalities for highly automated vehicles (HAD)

Consistent with earlier findings, the type of modality provided to drivers for TOR in HAD greatly influenced the performance and quality of reengagement of control (Naujoks et al., 2014; Petermeijer et al., 2016; Petermeijer et al., 2017; Bazilinskyy et al., 2018). Several research studies have been conducted in this field by analyzing the means of alerting, the type of information that needed to be provided for successful takeover, and the modality that needed to be selected for each scenario. Most research studies focused on a combination of modalities to alert drivers to take over and classified them as unimodal or multimodal alert types. Previous studies reported that multimodal alerts were more efficient than unimodal alerts (Naujoks et al., 2014; Petermeijer et al., 2016, 2017; Politis et al., 2015). Naujoks et al. (2014) investigated the effects of visual and visual-auditory TOR in complex scenarios of takeover (straight and curved roads) by analyzing the reaction times of drivers. Results showed that the reaction times of drivers varied significantly in regard to the type of modality presented, eliciting faster reactions for visual-auditory requests. Conversely, differences in complex takeover scenarios did not yield significant effects on takeover performance (Naujoks et al., 2014). Additionally, Petermeijer et al. (2017) analyzed driver takeover reaction time, usefulness, and satisfaction of unimodal versus bimodal TOR, while performing a surrogate reference task (SuRT) in HAD. By doing so, they suggested that bimodal TOR presented faster takeover times than unimodal. Bazilinskyy et al. (2018) conducted a survey to investigate drivers' preferences for TOR in HAD, depending on the driving situation. The highest preference was recorded for modalities presenting all three alerts (auditory, vibrotactile, and visual) irrespective of the TOR scenario.

Research studies conducted on the efficiency of a single modality were also of interest. Variables such as static, dynamic, continuous, discrete, and the type and length of information, were also investigated. Upon testing efficiency within the unimodal alert type, vibrotactile research studies were considered relevant. Vibrotactile modes are efficient ways of alerting drivers (Petermeijer et al., 2016; Telpaz et al., 2015), and could constitute a solution that allows fast extraction of relevant information in each situation (Petermeijer et al., 2017). However, it was suggested that multimodal feedback could assist in the takeover process (Petermeijer et al., 2016). Given that alert preferences varied in the published results of different research studies, it was important to examine and investigate the effect of alert type modalities by comparing unimodal as well as multimodal alerts. In addition, other factors, such as the state of the driver when the alert occurs, might affect these findings.

1.2. Non driving-related (NDR) task in the transition of control for HAD

The effective transition of control between vehicle and driver depends on several factors, such as the state of the driver, road type, weather conditions, request lead time, and vehicle request type. Several research studies argued that depending on the level of the awareness of the driver for the situation, the performance of the takeover varied (De Winter et al., 2014; Gold et al., 2016). Since then, several studies focused on the state of drivers before the transition of control. Depending on the type of task the drivers were involved in during automated driving, the drivers' abilities to reengage in the driving task efficiently varied. Visual, auditory, physical, and cognitive demands of the secondary task influenced the preference, efficiency, and safety of the transition of control of the vehicle (Zeeb et al., 2016; Marberger et al., 2017). Therefore, previous research investigated the relationship and influence of the NDR task on drivers' performances for the reengagement of control after a TOR (Walch et al., 2015; Eriksson and Stanton, 2017), where NDR tasks were classified in accordance with the type of demand they required from the driver (Winter et al., 2016; Funkhouser and Drews, 2016; Petermeijer et al., 2017).

Despite the amount of literature on NDR tasks in HAD, it was difficult to find consistency in results of previously published studies. Some studies indicated that a 5–8 s interval was a sufficient takeover interval for drivers (Gold et al., 2013; Mok et al., 2015; Eriksson and Stanton, 2017). Zeeb et al. (2016) studied the influences of writing emails, reading the news, and watching videos on takeover time and quality of takeover. Additionally, other studies focused on non-naturalistic tasks, such as SuRT (Surrogate Reference Task) and n-back, to investigate their influence on takeover (Radlmayr et al., 2014; Gold et al., 2013). Conversely, some research revealed that the takeover performance, that is, the reaction time of drivers on the transition of control, did not vary based on whether drivers were involved in an NDR task or not (Eriksson and Stanton, 2017). Thus, in order to analyze the effect of NDR tasks on TOR efficiency, in this study we investigated the drivers' behaviors in HAD when drivers engaged in natural NDR tasks.

2. Method

2.1. Participants

Twenty drivers (5 females and 15 males) participated in our experiment. They all had a valid driver's license and their ages varied from 24 to 42 years old (M = 30.1, SD = 4.77). We collected basic demographic data on participants with respect to their driving experience and driving frequency (Table 1). Participation in this study was voluntary.

2.2. Apparatus

Fig. 1 shows the laboratory setup and driving simulator used for this experiment. The simulator consisted of three 27″ screens placed within a horizontal field-of-view of approximately 110°, and two sets of Logitech® MOMO® racing force feedback wheels and pedals. We used the Wizard of Oz to implement the autonomous driving environment. The Wizard of Oz is a technique by which the experimenter manipulates the system to make it look realistic to the participant (Bernsen et al., 1994). Because the driving simulation software used in this experiment did not actually have an auto-driving mode, one of the experimenters drove manually during autonomous driving, making participants believe that they were in autopilot mode. Figuratively, the wizard drove the car with the second steering wheel located in a separate room. The transition between the two steering wheels was done through a switching button that the experimenter pressed, which indicated the reengagement of control of drivers. We simulated virtual driving using City Car Driving software (Version 1.5). Participants operated the vehicle with an automatic gear shift and both side and rear-view mirrors through the three 5760 × 1080 pixel resolution monitors. An Apple iPad Pro with a 9.7″ display located on the right side of the driving seat was used to present visual alerts and video tasks. A 2.0 channel sound bar speaker was placed below the display for auditory alerts. The haptic seat used for vibrotactile alerts consisted of four Mabuchi RE-280 DC motors at a

Table 1
Demographic information on driving experience and frequency.

Driving Experience	n	Driving Distance per Year	n	Driving Frequency	n
1–5 years	11	~ 5,000 km	8	1– 2 times per week	14
6–10 years	4	5,000 – 10,000 km	5	3–4 times per week	1
More than 10 years	5	More than 10,000 km	4	More than 5 times per week	6

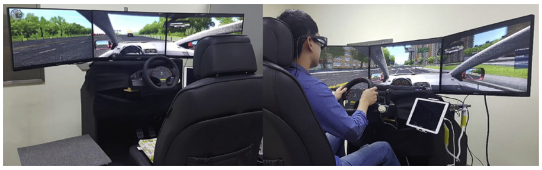

Fig. 1. Experimental setup and driving simulator.

rated voltage of 12 V, which were placed on the car seat in a 2 × 2 arrangement. In addition, we used a SensoMotoric Instruments (SMI) mobile eye tracker with a frame rate of 30 frames per seconds (fps) to collect the participants' gaze movements, and an LG G2 smartphone for phone conversation and smartphone interaction tasks.

2.3. Experimental design

A two-factor repeated measure design (4 × 7) was used to evaluate the effect of NDR tasks (4) and TOR modalities (7). Each participant attended four sessions that were set depending on the task condition: 1) no-task, 2) phone conversation, 3) smartphone interaction, and 4) video watching. The tasks selected for the experiment were based on previous studies related to NDR tasks on HAD, and were classified by the existence of physical, visual, and auditory demands (Llaneras et al., 2013). All participants completed the no-task condition first, and followed up by conducting the three tasks in a counterbalanced order. For each session, seven different TORs were assigned in random order. Specifically,

1) No-task session (baseline): Participants were asked to observe the driving situation during autonomous driving and to press a button placed on the right side of the handle when the alert sounded to trigger the transfer from autonomous to manual driving. During this session, no specific task was given.
2) Phone conversation (physical-auditory demand): Twenty task questions were used for the phone conversation task. Participants were asked to call one of the experimenters and hold the phone device constantly during their conversations.
3) Smartphone interaction (physical-visual demand): A practical application involving typing was selected for the task. Participants were asked to type sentences or words on the touch screen. The device was set to the silent mode to avoid auditory demands.
4) Video watching (visual-auditory demand): A video was selected from an entertainment show, and was presented on the tablet located on the right side of the handle. The participants were instructed to watch the video while they were in an automated driving status.

2.4. Driving scenario and design of TOR

For the experiments, participants drove on a two-way highway, which mostly consisted of straight roads, for approximately 18–20 min for each session. The highway consisted of four lanes with no traffic and participants were asked to drive in the second lane using manual driving at a speed that ranged between 60 km/h to 80 km/h. Participants traveled a driving course of approximately 24 km and seven TOR. After the reengagement of control, drivers drove the car for about 500 m and then shifted to automated mode. Each trial consisted of the same driving course with seven takeover points at the same distance during the course. The takeover scenario consisted of a non-emergency situation, but took place before increased road curvature, or before the road sections where changes in the number of lanes occurred (i.e. between 3 and 4 or 4 and 3 lanes). The lead time for the TOR was 7 s. The design of TORs consisted of seven types of alerts, three types of unimodal (i.e., either visual, vibrotactile, or auditory), and four multimodal. Visual alert was provided on the tablet PC through the signage of a steering wheel with drivers' hands indicating that participants had to take control of the vehicle. The image appeared when the TOR was issued, and was sustained until the participant pressed the takeover button, or until the lead time of 7 s ended. In the video watching session, the image was overlaid on the video. The vibrotactile alerts were implemented by a static pattern that lasted 2 s with vibration motors placed on the car seat. The auditory alerts consisted of a beep sound that lasted 2 s. The multimodal alerts were generated through combinations of unimodal alerts, namely, visual-vibrotactile, visual-auditory, vibrotactile-auditory, and visual-vibrotactile-auditory. Each alert included in the multimodal alerts was presented simultaneously when the TOR occurred.

2.5. Dependent variables

2.5.1. Objective measures

We gathered three objective datasets that were used in previous studies to measure the drivers' takeover performance: takeover time, hands-on time, and time to fixation. Takeover time refers to the time taken by drivers to reengage control of the vehicle (Borojeni et al., 2016). In this study, we measured the time it took for participants to press the disengagement button next to the steering wheel from the time the TOR was issued. In accordance with previous studies on takeover time, we selected a lead time of 7 s before the takeover. Thus, for participants who missed the alert, we treated this data as failures. Hands-on time measured the time between the TOR and the time taken for drivers to reach the steering wheel (Petermeijer et al., 2017). The hands-on time is the physical readiness of the driver to reengage control. In this study, we measured the time it took participants to first grab the steering wheel after the TOR. Finally, the time to first fixation indicates the time from the TOR to the first fixation on the road. This measure is important as it indicates how fast the drivers can assimilate information on the road and are ready to drive. Data gathered on fixation time was only analyzed for smartphone interaction and video watching tasks. In the case of no-tasks and phone conversation tasks, the fixation time was not measured because participants with visual fixation were mostly aware.

2.5.2. Self-reported measures

Self-reported data were collected to assess participants' subjective attitude towards the TOR. The questionnaire consisted of four variables: 1) satisfaction: *"I felt satisfied with the alert type provided as a takeover request"*, 2) usefulness: *"I think that this alert type can be considered useful as a takeover request"*, 3) effectiveness: *"The type of alert provided as a takeover request was perceived effective"*, and 4) safety: *"I think that the alert type provided as a takeover request is safe"*. Each variable was assessed in accordance with a single question based on a 7-point Likert scale ranging from -3 (strongly disagree) to +3 (strongly agree). The questionnaire was conducted after each TOR. Perceived satisfaction and usefulness were obtained to assess the acceptance of TOR (Van der Laan et al., 1997). The perceived effectiveness was gathered to the extent possible, and based on the participants' perception of each alert for each task condition. Finally, the perceived safety was selected to assess the drivers' reactions towards the alert in the effort to securely regain control of the vehicle.

2.6. Experimental procedures

Participants were first asked to fill out a questionnaire listing demographic information on the participant and their driving experience. Then, a detailed explanation was provided on the driving simulator, HAD situation, and TOR in HAD. After that, participants were given approximately 5 min or more (if needed) to become familiar with the driving simulator. A practice session followed where participants could experience HAD, manual driving, the transition between manual control to autonomous control, and the takeover task. The practice session consisted of three TOR with three different unimodal alerts (visual, auditory, vibrotactile) for each takeover. Additionally, it was explained to the participants that during the experiments, multimodal alerts were also planned to be provided. The experiment began with the no-task session for all participants. The other three tasks, smartphone interaction, video watching, and phone conversation tasks, were assigned in a counterbalanced order. At the beginning of each session, the eye tracker was calibrated and a brief description with instructions of the task was given. Participants were asked to drive manually first, and to immediately transfer to autonomous driving. During autonomous driving, participants were asked to perform the predefined NDR task until the TOR was issued. After the takeover, participants drove in a manual mode, and after 30 s, they were asked to move to HAD. After each takeover, a short questionnaire on the TOR alert was collected. After the questionnaire, participants were asked to reengage in NDR tasks until the next TOR. Participants experienced all seven alerts for each session. Thus, 28 total takeover datasets were collected for each participant.

2.7. Data analysis

Data obtained for each of the dependent variables were tested to see if they were normally distributed using the Kolmogorov–Smirnov test. The results of the test revealed that the distribution of the data was not normally distributed. Thus, we selected a nonparametric test for the descriptive analysis of data. For the nonparametric analysis, an aligned-rank transform method was selected to statistically analyze the main effect of the task, TOR modality, and interaction effect (Wobbrock et al., 2011). For the eye-tracking data analysis, we only utilized data from participants with a tracking ratio above 90%. This justified excluding data from two participants in the results of the video watching task.

3. Results

3.1. Takeover time

Based on the total duration taken by drivers to take over control of the vehicle, we examined the effect of NDR tasks and TOR modalities. Successful reengagement of control was observed for all task conditions and through the six TOR alerts, with the exception of visual modalities. When only visual modalities were presented, the success rate of takeover for phone conversations was 55% (n = 11), for smartphone interaction tasks, it was 85% (n = 17), and for no-task conditions, it was 95% (n = 19). In the case when all three TOR modalities were presented together, the reaction time from takeover to another task ranged between 0.77 s to 5.05 s. The mean reaction time and SD for each of the cases are reported in Table 2.

Because there were some failure trials with visual modalities, statistical analysis was conducted by excluding the failure trials only. The examination of the statistical analysis on the effect of the task and TOR modalities indicated the main effects on the task (F(3546) = 7.510, $p < 0.001$, $\eta^2 = 0.042$). Post-hoc analysis with Bonferroni correction indicated that there was a significant difference between phone conversation and smartphone interaction ($p = 0.023$), and smartphone and video watching tasks ($p < 0.001$), showing a longer takeover reaction time when interacting with a smartphone. Results also showed a significant main effect on alert type (F(6, 546) = 4.193, $p < 0.001$, $\eta^2 = 0.046$). Post-hoc analysis on the main effect on alert type indicated a statistically significant difference between visual alert with auditory ($p = 0.008$), visual-tactile ($p = 0.034$), vibrotactile-auditory ($p = 0.010$), visual-auditory ($p < 0.001$), and visual-vibrotatiale-auditory ($p < 0.001$), except for the unimodal vibrotactile alert ($p = 0.117$). Among the others TOR modalities, there were no statistically significant difference. Furthermore, there was no significant interaction effect (F(18,546) = 0.416, $p = 0.985$, $\eta^2 = 0.014$).

3.2. Fixation time

Analysis of the fixation time was only conducted for smartphone interactions and video watching tasks. For no-tasks and phone conversation tasks, the time taken by drivers to fix their eyes on the road was not analyzed because those tasks were not visually demanding, and because most of the participants had their eyes on the loop (Fig. 2).

The analysis of the effect of TOR modalities showed significant effects on TOR modalities (F(6,245) = 3.603, $p = 0.002$, $\eta^2 = 0.081$). Based on Table 3, the longest fixation time was for the visual modality. Bonferroni post-hoc analysis on the effect of TOR modalities indicated that there was a significant difference between visual alert conditions with unimodal vibrotactile ($p = 0.017$) and auditory ($p = 0.002$); and multimodal vibrotactile-auditory ($p = 0.012$), visual-auditory ($p = 0.005$), and visual-vibrotactile-auditory ($p = 0.011$). There was no significant difference between the visual-tactile and visual alert ($p = 0.067$). Among the other six conditions, there was no significant difference in alert modality (Fig. 3). Statistical analysis of the fixation time showed significant effects on tasks (F(1,245) = 10.286, $p = 0.002$, $\eta^2 = 0.040$), showing shorter fixation time for the video watching task than for smartphone interaction. Conversely, there was no significant interaction effect between the two variables (F(6,245) = 1.056, $p = 0.390$, $\eta^2 = 0.025$).

3.3. Hands-on time

We investigated the effect of NDR tasks and alert types on hands-on time (Table 4). The range of time taken for participants to grab the steering wheel was between 0.32 s and 5.40 s for the no-task condition, 0.47 s and 4.78 s for the phone conversation task, 0.82 s and 5.56 s for smartphone interaction task, and 0.40 s and 5.09 s for the video watching task.

The task types had a significant effect on hands-on time (F(3, 459) = 48.576, $p < 0.001$, $\eta^2 = 0.241$). Bonferroni post-hoc analysis showed a significant difference in mean time between the no-task condition and phone conversation ($p < 0.001$) and no-task condition and smartphone interaction ($p < 0.001$). Also, there was a significant difference between the phone conversation and video watching tasks ($p < 0.001$) and the phone conversation and smartphone interaction tasks ($p < 0.001$). Statistical analysis showed that the TOR modality had a significant main effect (F(6, 459) = 4.012, $p < 0.001$, $\eta^2 = 0.050$). According to the Bonferroni post-hoc analysis, the unimodal visual TOR was longer than the other modalities: vibrotactile ($p = 0.020$), auditory ($p = 0.010$), visual-vibrotactile ($p = 0.006$), vibrotactile-auditory ($p = 0.001$), visual-auditory ($p = 0.001$), and visual-vibrotactile-auditory ($p < 0.001$) (Fig. 4). The interaction effect between NDR tasks and TOR modalities was not significant (F(18, 459) = 0.540, $p = 0.939$, $\eta^2 = 0.021$).

3.4. Self-reported questionnaires

Fig. 5shows the results of the self-reported questionnaire on the perceived usefulness, effectiveness, satisfaction, and safety of each TOR modality for each of the task conditions. Results showed that the only modality that had a negative score was the visual modality for all the tasks expected for the video watching tasks.

Statistical analysis demonstrated that there was a significant effect of TOR modalities (F(6,518) = 19.033, $p < 0.001$, $\eta^2 = 0.181$) on perceived safety. The results of the Bonferroni post-hoc test showed a low score for visual modality compared with vibrotactile ($p < 0.001$), auditory ($p < 0.001$), visual-vibrotactile ($p < 0.001$), vibrotactile-auditory ($p < 0.001$), visual-auditory ($p < 0.001$), and visual-vibrotactile-auditory ($p < 0.001$) (Fig. 5). However, there was no significant main effect on task (F(3,518) = 2.194, $p = 0.088$, $\eta^2 = 0.013$) and no significant interaction effect (F(18,518) = 1.515, $p = 0.079$, $\eta^2 = 0.050$). Thus, unimodal visual TOR yielded the lowest score for perceived safety

Table 2
Descriptive statistical analysis of the takeover time in accordance with NDR task conditions for each alert type.

	No task		Phone Conversation		Smartphone		Video Watching		M. Avg.
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	
V	2.86[a] (1.006)	[1.56, x > 7]	2.56[b] (1.098)	[1.56, x > 7]	3.45[c] (1.414)	[1.68, x > 7]	2.42 (1.089)	[1.13, 4.83]	2.82
T	2.37 (0.845)	[1.01, 3.81]	2.49 (1.347)	[1.18, 4.93]	2.54 (1.002)	[1.26, 4.82]	2.18 (1.328)	[0.74, 5.94]	2.40
A	2.10 (1.119)	[0.68, 5.62]	2.32 (1.291)	[1.18, 5.69]	2.58 (1.209)	[0.84, 5.30]	2.23 (1.266)	[0.84, 4.63]	2.31
V&T	2.50 (1.193)	[1.12, 6.04]	2.15 (0.910)	[0.98, 4.33]	2.47 (1.002)	[0.93, 4.26]	1.88 (0.804)	[0.88, 3.68]	2.25
T &A	2.33 (1.118)	[0.74, 4.97]	2.18 (0.986)	[1.16, 4.49]	2.48 (1.320)	[0.91, 5.87]	2.05 (0.985)	[1.07, 4.44]	2.26
V&A	2.04 (1.014)	[0.88, 5.45]	2.11 (1.079)	[0.87, 4.75]	2.41 (1.131)	[1.05, 4.92]	1.97 (0.948)	[0.83, 3.76]	2.13
V, T&A	1.98 (1.032)	[0.77, 4.31]	2.22 (1.070)	[1.14, 4.60]	2.50 (1.168)	[0.96, 5.05]	1.95 (1.065)	[0.96, 4.46]	2.16
M. Avg.	2.31		2.29		2.63		2.10		

V = visual, T = Vibrotactile, A = Auditory. Mean and SD (s), Range = [min (s), max (s)], M. Avg. = marginal average.

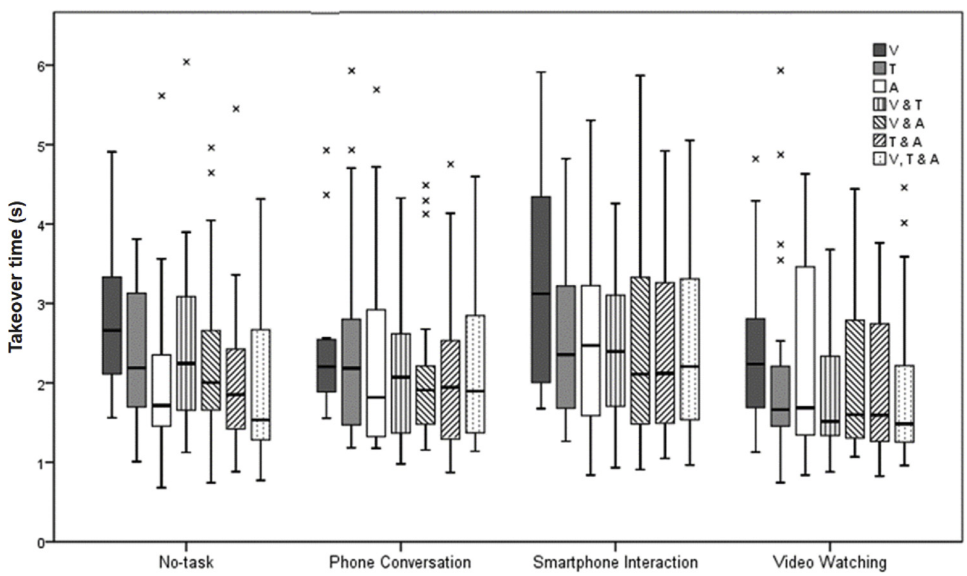

Fig. 2. Box plot of TOR modality reaction time for each of the NDR tasks (V = Visual, T = Vibrotactile, A = Auditory; excluding failure TOR on visual modality).

Table 3
Descriptive statistical analysis of the time taken by drivers to fix their eyes on the loop display in smartphone interactions and video watching tasks.

Modality	Smartphone task		Video watching task		M. Avg.
	Mean (SD)	Range	Mean (SD)	Range	
V	2.71 (1.042)[c]	[1.00, 4.17][c]	1.83 (0.946)	[0.73, 4.30]	2.27
T	1.67 (0.724)	[0.69, 2.79]	1.33 (0.928)	[0.26, 4.12]	1.50
A	1.55 (0.701)	[0.40, 3.12]	1.27 (0.850)	[0.43, 3.54]	1.41
V&T	1.64 (0.857)	[0.51, 3.15]	1.50 (0.758)	[0.47, 3.00]	1.57
T &A	1.58 (0.768)	[0.41, 3.07]	1.38 (0.794)	[0.40, 2.97]	1.48
V&A	1.52 (0.807)	[0.68, 3.36]	1.45 (0.864)	[0.50, 3.05]	1.49
V, T&A	1.54 (0.723)	[0.59, 3.55]	1.41 (0.762)	[0.47, 3.40]	1.48
M. Avg.	1.74		1.45		

V = visual, T = Vibrotactile, A = Auditory. Mean and SD (s), Range = [min (s), max (s)], M. Avg. = marginal average.

[c] = mean takeover time for successful trials for smartphone interaction tasks (n = 17).

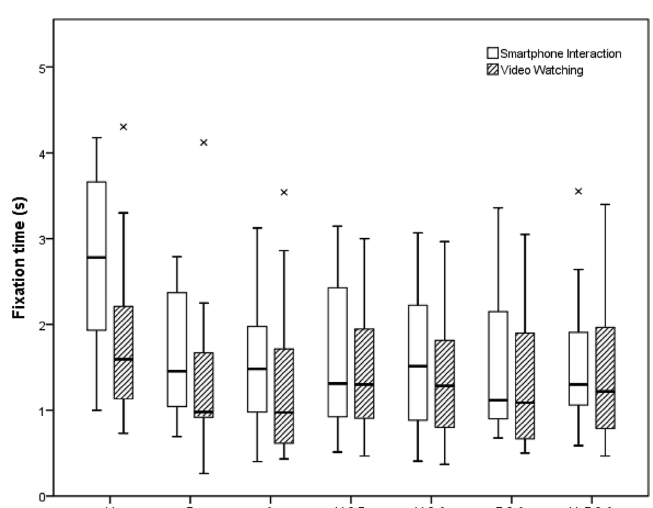

Fig. 3. Box plot of fixation time for each TOR modality comparing smartphone and video watching tasks.

Table 4
Descriptive statistical analysis on the hands-on time in each of the task conditions for each alert modality.

	No Task		Phone Conversation		Smartphone		Video Watching		M. Avg.
	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	
V	2.16[a] (0.996)	[0.92, 5.40][a]	2.70[b] (1.098)	[1.16, 4.78][b]	3.27[c] (1.158)	[1.53, 5.56][c]	1.84 (0.790)	[1.07, 3.67]	2.49
T	1.47 (0.712)	[0.32, 3.64]	2.30 (0.917)	[0.57, 4.58]	2.30 (0.778)	[0.82, 3.48]	1.64 (1.067)	[0.67, 5.09]	1.93
A	1.49 (0.866)	[0.67, 4.08]	2.19 (0.886)	[0.77, 4.30]	2.41 (0.867)	[1.17,5.00]	1.61 (0.940)	[0.67, 4.03]	1.93
V&T	1.54 (0.735)	[0.65, 3.76]	1.85 (0.875)	[0.47, 3.43]	2.45 (0.802)	[1.33, 4.33]	1.54 (0.688)	[0.73, 2.94]	1.85
T &A	1.56 (0.941)	[0.57, 4.25]	2.25 (0.793)	[0.90, 3.97]	2.24 (0.802)	[1.17, 3.93]	1.30 (0.518)	[0.77, 3.09]	1.84
V&A	1.36 (0.708)	[0.63, 3.57]	2.09 (0.803)	[0.68, 3.18]	2.29 (0.928)	[1.13, 4.30]	1.50 (0.842)	[0.40, 3.07]	1.81
V, T&A	1.24 (0.687)	[0.57, 3.63]	2.08 (0.932)	[0.57, 4.21]	2.43 (1.001)	[1.18, 5.18]	1.26 (0.587)	[0.63, 2.43]	1.75
M. Avg.	1.55		2.21		2.48		1.53		

V = visual, T = Vibrotactile, A = Auditory. Mean and SD (s), Range = [min (s), max (s)], M. Avg. = marginal average.
[a] = mean takeover time for successful trials for no-task condition (n = 19), b = mean takeover time for successful trials for phone conversations (n = 11), and c = mean takeover time for successful trials for smartphone interaction tasks (n = 17).

no matter the task being engaged in during automated driving.

The analysis of the perceived satisfaction showed a significant main effect of TOR modalities ($F(6,518) = 14.924$, $p < 0.001$, $\eta^2 = 0.147$). The post-hoc analysis with Bonferroni correction indicated a significant difference between visual with the rest of TOR modalities ($p < 0.001$). There was no significant main effect on task for satisfaction ($F(3,518) = 1.112$, $p = 0.344$, $\eta^2 = 0.006$), neither was found a significant interaction effect ($F(18,518) = 1.534$, $p = 0.073$, $\eta^2 = 0.051$). Statistical analysis on effectiveness demonstrated that there was a significant effect on NDR tasks ($F(3,518) = 3.997$, $p = 0.008$, $\eta^2 = 0.023$) and main TOR modalities ($F(6,518) = 35.537$, $p < 0.001$, $\eta^2 = 0.292$), and significant interaction effect between both of these ($F(18,518) = 3.223$, $p < 0.00$, $\eta^2 = 0.101$) (Fig. 5). Thus, participants rated the effectiveness of the alert type depending on the NDR task being done. Finally, statistical analysis of the perceived usefulness showed that there was a significant effect on TOR modality ($F(6,518) = 20.267$, $p < 0.001$, $\eta^2 = 0.190$). Post-hoc analysis on the main effect of TOR modalities with Bonferroni correction showed significant difference between visual alert with the other modalities ($p < 0.001$) and vibrotactile with visual-vibratactile-auditory ($p = 0.043$). There were no significant main effects on NDR tasks ($F(3,518) = 1.023$, $p = 0.382$, $\eta^2 = 0.006$) or interaction effects ($F(18,518) = 1.412$, $p = 0.120$, $\eta^2 = 0.047$).

4. Discussion

Conventional alert modalities include visual alerts, mostly presented in an instrument cluster and used to provide information on the status of the car, auditory alerts for emergency occasions, such as collision alerts upon parking, and vibrotactile alerts for alerting driving behaviors, such as lane deviation. Research led to interesting findings regarding the differences in performances due to the perception of the alert. For instance, for visual TOR, we expected the elicited outcomes to be failures in reference to smartphone interaction tasks compared to phone call interactions. The visual and attentional behaviors of the participants were different upon interactions with the smartphone from those in the case of the phone. Although visual attention was focused most of the time on the phone, the participants continuously shifted their attention to monitor the vehicle as they were expecting a TOR. On the other hand, participants in the phone call tasks had their visual attention on the loop that let them decide if they needed to reengage control or not, irrespective of whether an alert was provided. The unstable takeover performance of drivers with a visual alert was also demonstrated on the results of the questionnaire for all the tasks except for video watching. The video task required drivers' visual attention on the same screen as the visual alert and this yielded a positive score for perceived safety, satisfaction, usefulness, and effectiveness for only the video watching task. Also, most of the participants that failed to take

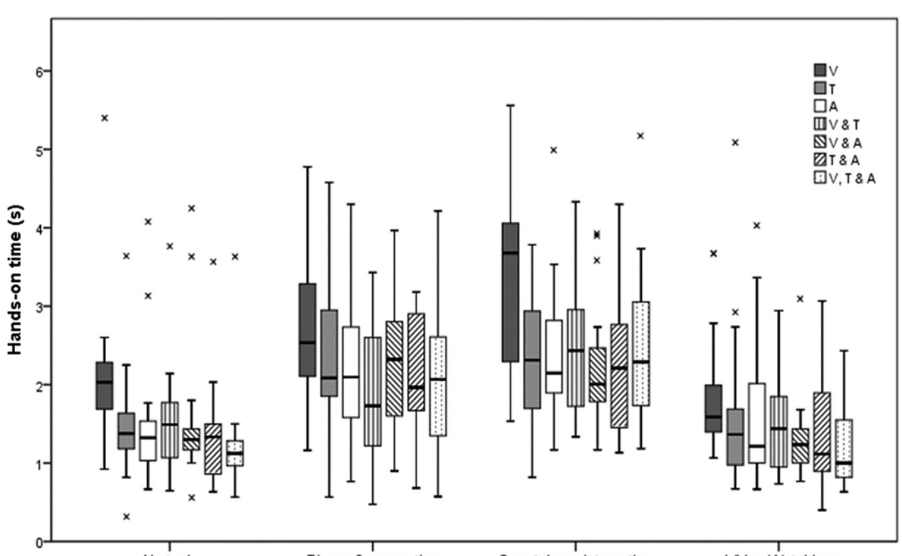

Fig. 4. Box plot of hands-on time for each of the TOR modalities comparing the four task conditions.

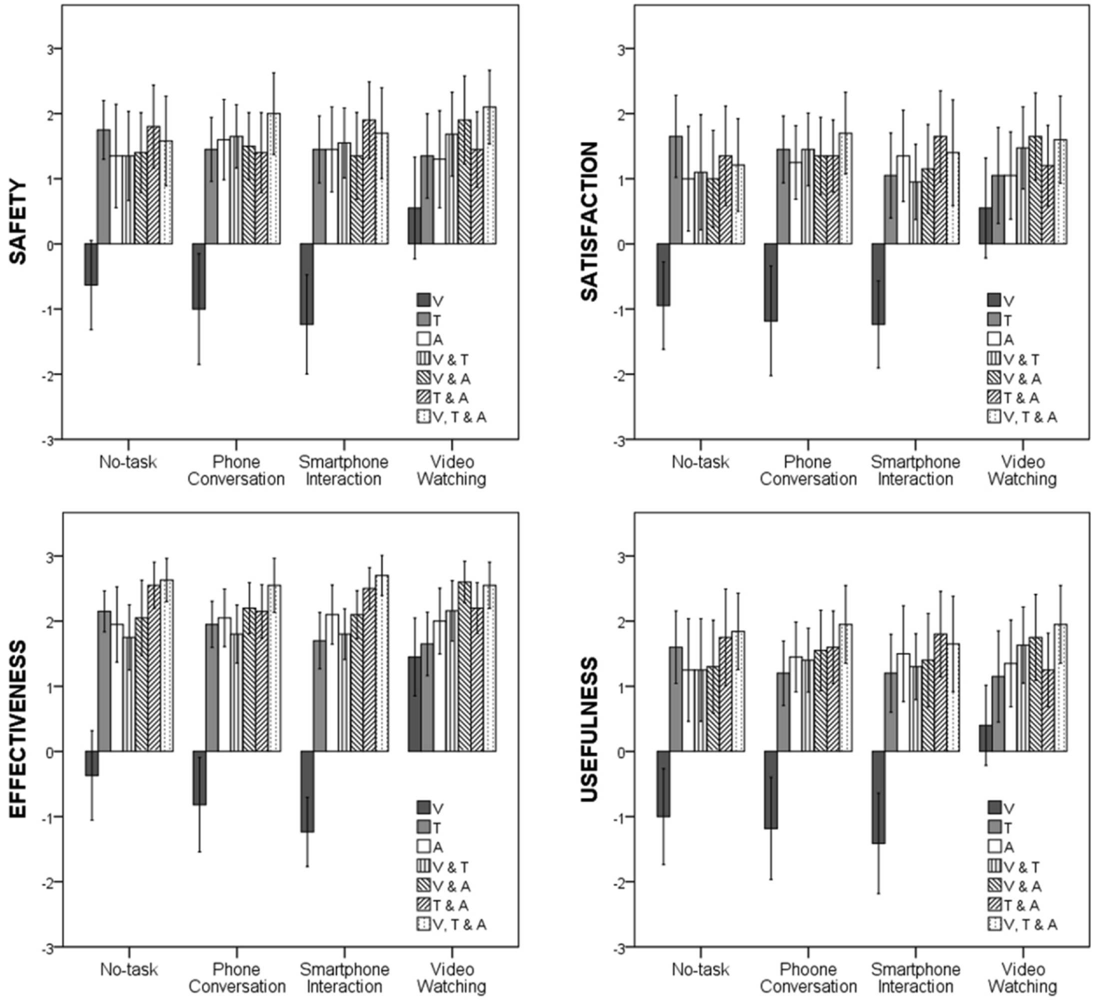

Fig. 5. Mean scores of self-reported questionnaires on perceived safety, satisfaction, effectiveness, and usefulness for each TOR modality and each task condition.

over control of the vehicle with visual alert, they did not even know that the alert was issued. That is, the limitation of the visual alert is that we could not expect drivers to be aware of the alert as fast as other modalities when they were issued. Therefore, it is interesting to note that it is important to consider the visual field of drivers while performing the NDR task during automated driving, the location where the visual alert is presented to the drivers, and how drivers monitor the system to be aware of what is happening. Thus, providing insight for further research on how to design visual alerts for automated driving. As part of the limitation of this study, participants expecting the TOR led to participants monitoring the system while being engaged in the smartphone interaction task. That is, we might think that during the no-task condition and the phone conversation task, drivers were enabled to see how the auto-pilot mode was driving the vehicle, however, during the smartphone interaction task, drivers have to actively decide to monitor the system while performing the NDR task to be aware of the situation. This provides insight on the importance of the analyzing the NDR task characteristic being conducted during the automated mode and the design of alert for TOR.

Visual TOR had longer hands-on time than vibrotactile or auditory alerts not only for smartphone interaction and phone conversation tasks, but also for video watching tasks and the no-task condition. According to previous research on alerts in vehicles, this finding can be easily explained because participants tended to perceive and assign lower priority to urgencies from visual alerts compared to other modalities (Petermeijer et al., 2017). That is, the design of TOR based on only visual modality might not be sufficient as an indication of driver's need to reengage control of the vehicle as fast as possible, which was also shown in the results of the questionnaire on perceived safety, usefulness, effectiveness, and satisfaction. On the other hand, providing multiple modalities simultaneously led to faster hands-on times (Petermeijer et al., 2017). As expected, hands-on times for smartphone interactions and phone conversation tasks were longer than no-task and video watching tasks, which indicates that the requirement of physical demands influenced the time needed for the participants to be physically ready for the takeover.

It is interesting to see that the highest score for usefulness, safety, effectiveness, and satisfaction, were all multimodal with combinations of two or three modalities, but all of them included auditory alerts. Thus, enabling understanding the importance of drivers to tell that the alert type was adequate for the TOR. That is, drivers in automated driving are given the opportunity to be engaged in tasks not related to driving, which might decrease the possibility of successful takeover control of the vehicle when necessary. We might suggest that the inclusion of the three modalities as an alert for TOR can make drivers perceived as useful as the perception of the alert is secure no matter in which NDR task is involved. Although most of the research studies focus on the bimodal (auditory-visual) or unimodal (vibrotactile) alert separately, we suggest that there is a need to concentrate on a combination of multimodal alert design. A limitation of the experiment design was that it was design so as participants were familiarized with the unimodal alert in the practice section while multimodal alerts were not. Participant scoring of subjective rating of perceived safety, usefulness, effectiveness, and satisfaction might have been influence, which might suggest further research with participants having experience of all types of alerts previous to the experimental section.

Our results further suggest that TOR, including the visual modalities, take a longer time to achieve the first fixation, and thus lead to longer overall takeover times. Eye tracking data showed that participants shifted their eyes to fix on the takeover alert icons, even though these stimuli were presented in combination with other modalities, such as auditory and vibrotactile. We can deduce that the time taken to fix their eyes on the icon led to longer times for fixing their eyes on the road, thereby introducing a delay in the process of re-engaging control of the vehicle. This result also implies that the perceived effectiveness of the alert modality for drivers varied depending on the related NDR task being engaged during automated driving. In accordance with the results of the survey results published by Bazilinskyy et al. (2018), drivers preferred alerts presented as a combination of the three alert types (auditory, vibrotactile, and visual), results from this study on perceived effectiveness also supported this finding for all NDR task conditions except for the smartphone interaction task.

Nevertheless, even if there was a significant main effect on alert modality and on NDR tasks for performance measures of takeover, it was difficult to find an interaction effect between both variables. We observed that the range of values for the driving simulator data varied between values of less than a second to more than 7 s. Although our research included only one trial per section, it is important to take into consideration all of the individual data because of the dependence on the lead time provided to the drivers proving to be an extremely dangerous situation and lead to critical accidents. In other words, there are documented previous research studies arguing that a lead time of 5 s is sufficient for the reengagement of control. However, we believe that there is need of further work focusing on what variables influence on the reaction time of takeover depending on the modalities and NDR tasks.

5. Conclusions

This study investigated the effects of NDR tasks and TOR alert modalities in HAD. Because previous research mostly focused on the comparison between unimodal versus multimodal modalities, in this study, we investigated all the possible combinations for visual, auditory, and vibrotactile alerts, presented as unimodal and multimodal. Results suggested that NDR tasks influenced the takeover performance of drivers. Moreover, different alert types were found to be efficient for each task scenario in HAD. Furthermore, for this study, there were some limitations. When the visual TOR was provided in the central part of the entertainment system of the car, it influenced the time needed to discover that the alert was activated since the visual cue might have been located in a subordinate part of the field-of-view of the driver. Thus, only for visual alerts, the effect of the alert type might have significantly influenced the drivers, compared to other alert modalities performing the same selected tasks. Further research on the location of visual alerts might be of importance. Additionally, we noted that the deviation of response times is an important issue to be taken into consideration for the implementation and design of alerts, such as the reengagement of control of the vehicle, as it can lead to serious, dangerous, and unsafe situations. This study had highlighted the effect of NDR tasks on drivers' performance on takeover. Not only did it show that the performance of engagement of control is different depending on what the driver is doing in HAD, but it also presented findings on the differences of TOR alert modality efficiencies. Recommendations for basic standards for the design of TOR have also been proposed that aimed to provide safe and adequate alerts to the drivers. In addition, although we observed only the aspects of takeover performance according to TOR modalities, the drivers' cognitive workload and driving performance may vary depending on their individual differences such as their age and cultural characteristics (Son et al., 2010). For further research, it would be interesting to take into consideration the effect of drivers' characteristics on the engagement of NDR task and thus, the takeover performance during HAD.

Acknowledgement

This research was partially supported by the Graduate School of YONSEI University Research Scholarship Grants in 2017.

References

Banks, V.A., Stanton, N.A., 2016. Keep the driver in control: automating automobiles of the future. Appl. Ergon. 53, 389–395.
Bazilinskyy, P., Petermeijer, S.M., Petrovych, V., Dodou, D., de Winter, J.C., 2018. Takeover requests in highly automated driving: a crowdsourcing survey on auditory, vibrotactile, and visual displays. Transp. Res. Part F Traffic Psychol. Behav. 56, 82–98.
Bernsen, N.O., Dybkjær, H., Dybkjær, L., 1994. Wizard of oz prototyping: how and when. Proc. CCI Working Papers Cognit. Sci./HCI.
Borojeni, S.S., Chuang, L., Heuten, W., Boll, S., 2016. Assisting drivers with ambient take-Over requests in highly automated driving. October. Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, pp. 237–244.
Clark, H., Feng, J., 2016. Age differences in the takeover of vehicle control and engagement in non-driving-related activities in simulated driving with conditional automation. Accid. Anal. Prev.
De Winter, J.C., 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. Transp. Res. Part F Traffic Psychol. Behav. 27, 196–217.
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.
Funkhouser, K., & Drews, F., (2016, September). Reaction times when switching from autonomous to manual driving control: A pilot investigation. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 60, 1, 1854–1858). Sage CA: Los Angeles, CA: SAGE Publications.
Gold, C., Damböck, D., Lorenz, L., & Bengler, K., (2013, September). Take over!?. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 57, 1, pp. 1938–1942). Sage CA: Los Angeles, CA: SAGE Publications.
Gold, C., Körber, M., Lechner, D., Bengler, K., 2016. Taking over control from highly automated vehicles in complex traffic situations: the role of traffic density. Hum. Factors 58 (4), 642–652.
Kun, A.L., Boll, S., Schmidt, A., 2016. Shifting gears: user interfaces in the age of autonomous driving. IEEE Pervasive Comput. 15 (1), 32–38.
Llaneras, R.E., Salinger, J., Green, C.A., 2013. Human factors issues associated with limited ability autonomous driving systems: drivers' allocation of visual attention to the forward roadway. June. Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design. Public Policy Center, University of Iowa, Iowa City, pp. 92–98.
Marberger, C., Mielenz, H., Naujoks, F., Radlmayr, J., Bengler, K., Wandtner, B., 2017. Understanding and applying the concept of "Driver availability" in automated driving. July. International Conference on Applied Human Factors and Ergonomics. Springer, pp. 595–605 Cham.
Merat, N., Jamson, A.H., Lai, F.C., Daly, M., Carsten, O.M., 2014. Transition to manual: driver behaviour when resuming control from a highly automated vehicle. Transp. Res. Part F Traffic Psychol. Behav. 27, 274–282.
Meschtscherjakov, A., McCall, R., Louveton, N., Engel, T., Tscheligi, M., Koenig, V., 2015. MaDSAV: maintaining driving skills in semi-autonomous vehicles. September. Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, pp. 136–139.
Miller, D., Sun, A., Ju, W., 2014. Situation awareness with different levels of automation. 2014 IEEE International Conference on Systems, Man and Cybernetics (SMC) 688–693.
Mok, B.K.J., Johns, M., Lee, K.J., Ive, H.P., Miller, D., Ju, W., 2015. Timing of unstructured transitions of control in automated driving. June. Intelligent Vehicles Symposium (IV), 2015 1167–1172.
National Highway Traffic Safety Administration, 2013. Preliminary Statement of Policy Concerning Automated Vehicles. Washington, DC. pp. 1–14.
Naujoks, F., Mai, C., Neukum, A., 2014. The effect of urgency of take-over requests during highly automated driving under distraction conditions. Adv. Human Asp. Transp. 7 (Part I), 431.
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, 204–2150.
Petermeijer, S.M., De Winter, J.C., Bengler, K.J., 2016. Vibrotactile displays: a survey with a view on highly automated driving. IEEE Trans. Intell. Transp. Syst. 17 (4), 897–907.
Politis, I., Brewster, S., Pollick, F., 2015. Language-based multimodal displays for the handover of control in autonomous cars. September. Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications 3–10.
Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K., (2014, September). 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 Annual Meeting (Vol. 58, 1, pp. 2063-2067). Sage CA: Los Angeles, CA: Sage Publications.
Son, J., Reimer, B., Mehler, B., Pohlmeyer, A.E., Godfrey, K.M., Orszulak, J., Long, J.,

Kim, M.H., Lee, Y.T., Coughlin, J.F., 2010. Age and cross-cultural comparison of drivers' cognitive workload and performance in simulated urban driving. Int. J. Automot. Technol. 11 (4), 533–539.

Telpaz, A., Rhindress, B., Zelman, I., Tsimhoni, O., 2015. Haptic seat for automated driving: preparing the driver to take control effectively. September. Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, pp. 23–30.

Van Der Laan, J.D., Heino, A., De Waard, D., 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. Transp. Res. Part C Emerg. Technol. 5 (1), 1–10.

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

Winter, J.D., Stanton, N.A., Price, J.S., Mistry, H., 2016. The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. Int. J. Veh. Des. 70 (4), 297–324.

Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J., 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. May. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 143–146.

Zeeb, K., Buchner, A., Schrauf, M., 2015. What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accid. Anal. Prev. 78, 212–221.

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.