Evaluating Effects of Cognitive Load, Takeover Request Lead Time, and Traffic Density on Drivers' Takeover Performance in Conditionally Automated Driving

NA DU, University of Michigan
JINYONG KIM, University of Michigan
FENG ZHOU, University of Michigan, Dearborn
ELIZABETH PULVER, State Farm Mutual Automobile Insurance Company
DAWN M. TILBURY, University of Michigan
LIONEL P. ROBERT, University of Michigan
ANUJ K. PRADHAN, University of Massachusetts, Amhest
X. JESSIE YANG, University of Michigan

In conditionally automated driving, drivers engaged in non-driving related tasks (NDRTs) have difficulty taking over control of the vehicle when requested. This study aimed to examine the relationships between takeover performance and drivers' cognitive load, takeover request (TOR) lead time, and traffic density. We conducted a driving simulation experiment with 80 participants, where they experienced 8 takeover events. For each takeover event, drivers' subjective ratings of takeover readiness, objective measures of takeover timing and quality, and NDRT performance were collected. Results showed that drivers had lower takeover readiness and worse performance when they were in high cognitive load, short TOR lead time, and heavy oncoming traffic density conditions. Interestingly, if drivers had low cognitive load, they paid more attention to driving environments and responded more quickly to takeover requests in high oncoming traffic conditions. The results have implications for the design of in-vehicle alert systems to help improve takeover performance.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: conditionally automated driving, takeover transition, cognitive load, traffic density, TOR lead time

ACM Reference Format:
Na Du, Jinyong Kim, Feng Zhou, Elizabeth Pulver, Dawn M. Tilbury, Lionel P. Robert, Anuj K. Pradhan, and X. Jessie Yang. 2020. Evaluating Effects of Cognitive Load, Takeover Request Lead Time, and Traffic Density on Drivers' Takeover Performance in Conditionally Automated Driving. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '20), September 21–22, 2020, Virtual Event, DC, USA*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3409120.3410666

1 INTRODUCTION

The Society of Automotive Engineers (SAE) defines driving automation from L0 (no automation) to L5 (full automation) [28]. Today, automated vehicles (AVs) of L2 partial automation are already on the road and AVs of L3 conditional automation, such as the Audi Traffic Jam Chauffeur, could be purchased in Germany from 2019 [5]. Conditional automation, compared to partial automation, is not an improvement of automated driving features, but rather a revolution. The distinction between them lies in who handles the object and event detection and response (OEDR) task of driving. In L2, the driver completes the OEDR task and intervenes when necessary; In L3, OEDR is handled by the AV and the driver resumes responsibility when receiving a takeover request (TOR). In conditionally automated driving, drivers are allowed to engage in other tasks and therefore become increasingly out of the control loop. Drivers decoupled from the operational control of the vehicle have difficulty taking over control of the vehicle when requested [1, 24, 35]. In response to the difficulty, research has been conducted to examine factors that influence drivers' performance in takeover transitions.

One of the most important factors in takeover transitions is the non-driving-related tasks (NDRTs) themselves. A wide range of NDRTs have been utilized in experimental studies as reported in the literature, including both naturalistic tasks (e.g., text messaging) and artificial tasks (e.g., n-back memory task) [2, 26, 31, 32]. Prior research compared drivers' takeover performance when performing versus not performing an NDRT, and revealed that NDRTs deteriorated takeover quality, resulting in more crashes in high-traffic situations [26], shorter minimum time to collision (TTC) [12, 18], larger lateral acceleration [19], and larger standard deviation of lane position [33]. In addition, several studies examined the effects of performing different types of NDRTs. For example, Wandtner et al. [32] compared NDRTs in visual (i.e. Surrogate Reference Task, [15]) and auditory modality (i.e. N-back task, [26]). They found that operation of the visual task with handheld devices degraded takeover performance and led to a higher collision rate, while the auditory task led to comparable performance to a baseline without any task. Moreover, Zeeb et al. [33] and Wan et al. [31] examined drivers' takeover performance while they were typing, reading, watching a video clip, playing a game or taking a nap. Results showed that different NDRTs led to few differences on takeover reaction time, while watching a video and taking a nap resulted in a worse takeover quality as it occupied more sensory modalities or induced a very low arousal level.

The above mentioned studies shed some light on the influence of performing NDRTs on takeover performance. However, most of the studies did not directly manipulate the workload of NDRTs. One exception is the study of Zeeb et al. [34], wherein the drivers' cognitive task load was manipulated via the difficulty of NDRTs (i.e. reading vs. proofreading a text) and manual task load via the tablet location (i.e. handheld vs. mounted). Results revealed that high manual task load increased reaction times and deteriorated takeover quality, while the effect of cognitive task load on takeover ability was dependent on the type of driver intervention. High cognitive load lengthened the reaction times and deteriorated takeover quality in steering maneuver but not braking maneuver. In contrast, Bueno et al. [6] manipulated workload by asking drivers to identify and type out the association between three pictures and found a non-significant effect of workload on takeover time and quality. Further research is need to elucidate the mixed results and to examine the impact of NDRT workload.

Researchers also investigated how traffic density impacted drivers' takeover performance [8, 11, 12, 18, 26]. Traffic density refers to the average number of vehicles that occupy one mile or one kilometer of road space. High traffic density resulted in longer takeover time [12, 18], higher crash rate [12, 18], and higher accelerations [12]. In the above-mentioned studies, high traffic density was coupled with fewer escape paths. With a high traffic density, there were more vehicles in the same or neighboring lanes of the automated ego vehicle, hence restricting the drivers' action opportunities. There is a need to decouple traffic density and the availability of escape paths.

Another factor that influences takeover performance is the takeover request (TOR) lead time. Lead time refers to the time to collision at the time of the TOR [20]. Research demonstrated that shorter TOR lead time degraded takeover quality in the form of shorter minimum TTC, higher crash rates, greater maximum accelerations and greater standard deviation of steering wheel angle [22, 29–31]. It would be interesting to investigate whether there are any interaction effects between TOR lead time and traffic density/cognitive load.

In the present study, we aimed to examine the relationships between drivers' takeover performance and their cognitive load, TOR lead time, and traffic density in conditional automated driving. Using a high-fidelity driving simulator, we conducted a human-subject experiment with 80 participants. In the experiment, we manipulated drivers' cognitive load using the n-back memory task [16]. In addition, in order to decouple traffic density and the availability of escape paths, other vehicles only appeared in the oncoming lanes.

2 METHOD

This research complied with the American Psychological Association code of ethics and was approved by the Institutional Review Board of the University of Michigan.

2.1 Participants

Eighty university students with normal or corrected-to-normal vision participated in the study (mean age = 22.8 years, SD = 3.4 years; 33 females and 47 males). All the participants had a valid driver's license (mean year = 4.8 years, SD = 3.1 years). We used a 5-point Likert scale to measure participants' experience with driver assistance system (1 indicates "never" and 5 indicates "always"). Participants' average experience values were listed as follows: cruise control – 3.1, adaptive cruise control – 1.5, lane-departure warning – 1.7, lane-keeping assistance – 1.5, collision warning – 1.8, emergency braking – 1.4. Table 1 showed participants' average annual mileage and weekly mileage. Participants received $30 in compensation for an hour of participation.

Table 1. Participants' annual mileage and weekly mileage

Annual mileage	# participants	Weekly mileage	# participants
Less than 5,000 miles	24	Less than 50 miles	41
5,000 - 10,000 miles	25	50 - 100 miles	22
10,000 - 15,000 miles	23	100 - 150 miles	8
15,000 - 20,000 miles	2	150 - 200 miles	6
20,000 - 25,000 miles	4	200 - 250 miles	1
More than 25,000 miles	2	More than 250 miles	2

2.2 Experimental apparatus and stimuli

The study was conducted in a fixed-based driving simulator from Realtime Technologies Inc. (RTI, MI, USA) located in a dedicated lab space. The virtual world was projected on three front screens (16 feet away), one rear screen (12 feet away), and two side mirror displays (See Figure 1a). The simulated vehicle was controlled by a steering wheel and pedal system embeded in a Nisan Versa car model. The vehicle was programmed to simulate an SAE Level 3 automation, which handled the longitudinal and lateral control, navigation, and responded to traffic elements. Participants could press the button on the steering wheel to activate the automated mode, which was indicated by a green highlight on the dashboard. Once the AV reached its performance limit, an auditory TOR ("Takeover") would be issued, and the automated mode would be deactivated simultaneously for the driver to take over control of the vehicle.

The NDRT was a visual N-back memory task, adapted from the study of [16]. The stimulus consisted of nine (3×3) squares with two out of nine squares containing the image of a person. Each stimulus was presented for 500 ms in

(a) Driving simulator (b) In-vehicle setting (c) Smarteye system

Fig. 1. Driving simulator and Smarteye system

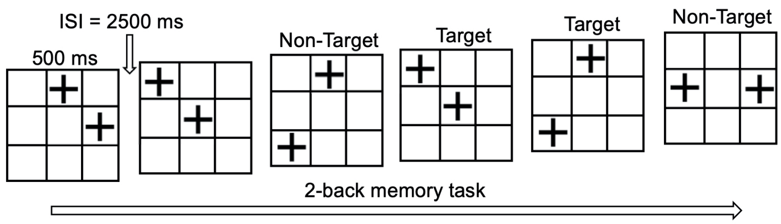

Fig. 2. N-back task

sequence with a 2500-ms interval. Participants were required to press the "Hit" button when the current stimulus was the same as the one presented N steps back in the sequence, and press the "Reject" button otherwise (Figure 2). The task was running on a 11.6 inch touch screen tablet mounted in the vehicle (Figure 1b).

This simulator was equipped with the Smart Eye (Smart Eye, Sweden) four-camera eye-tracking system that provided live head-pose, eye-blink, and gaze data (Figure 1c). The sampling rate of the eye-tracking system is 120 Hz.

2.3 Experimental design

The study employed a within-subjects design with the driver's cognitive load, TOR lead time, and traffic density as independent variables. The cognitive load was manipulated via the difficulty of the NDRTs (low: 1-back memory task; high: 2-back memory task). The heavy- and no- traffic conditions had 15 and 0 oncoming vehicles per kilometer respectively. The TOR lead time was 4 or 7 seconds. Based on prior literature [17, 21, 23, 27], eight takeover events were designed in urban and rural drives with typical roadway features (Table 2). The difficulty of the scenarios was designed to be approximately the same. The order of cognitive load, traffic density and TOR lead time was counterbalanced via an 8×8 balanced Latin Square across participants. Based on standard programming practices for the simulator, the order of scenario presentations was counterbalanced by having half of the participants drive from Event 1 to 8, and the other half from Event 8 to 1. During the entire session, there were no other vehicles in the driver's direction and participants could avoid the objects in their lane by changing to the adjacent lane.

2.4 Dependent variables

We collected drivers' subjective, vehicle-related and NDRT-related measures (Table 3). After each takeover event, participants reported their takeover readiness for each takeover event using a 0-100 scale, with 0 indicating not ready

Table 2. Descriptions of takeover events

Event	World	Scenario type	Event descriptions
Event 1	Urban	Lane changing	Bicyclist in the lane ahead
Event 2	Urban	Lane keeping	Construction zone on the left
Event 3	Urban	Lane changing	Construction zone ahead
Event 4	Urban	Lane keeping	Sensor error on the curve
Event 5	Rural	Lane changing	Swerving vehicle ahead
Event 6	Rural	Lane keeping	No lane markings on the curve
Event 7	Rural	Lane keeping	Sensor error on the curve
Event 8	Rural	Lane changing	Police vehicle on shoulder

Table 3. Dependent Variables

Dependent Variables	Unit	Category	Explanation
Takeover readiness	Scale from 0-100	Subjective rating	How ready participants were to takeover control of the vehicle
Takeover reaction time	Seconds	Takeover timing	Time between TOR and start of maneuver
Eyes-on-road reaction time	Seconds	Takeover timing	Time between TOR and eyes on road
Maximum resulting acceleration/jerk	m/s^2, m/s^3	Takeover quality	Maximum resulting acceleration/jerk during takeover situation
Minimum time to collision (TTC)	Seconds	Takeover quality	Minimum time to collision during takeover situation
Road deviation standard error	Centimeters	Takeover quality	Standard error of center road deviation during takeover situation
Reaction time in N-back task	Seconds	NDRT	The reaction time for the N-back memory task
Accuracy in N-back task	Percentage	NDRT	The accuracy for the N-back memory task

at all and 100 indicating absolutely ready. Takeover readiness was defined in the present study as the extent to which participants were ready to takeover control of the vehicle when an TOR was issued [20].

Takeover performance was measured in two aspects: timeliness and quality. With regard to takeover timeliness, takeover reaction time and eyes-on-road reaction time were measured. Takeover reaction time was defined as the time between the TOR and the start of maneuver. According to Gold et al. [12], the threshold of maneuver was set as a 2 degree change of the steering wheel angle or a 1% change of the pedals, whichever is quicker. Eyes-on-road reaction time was the time between the TOR and the driver's first gaze vector on the road [8].

Takeover quality was assessed by three measures: maximum resulting acceleration, maximum resulting jerk, and minimum time to collision (TTC_{min}) within the time window between the TOR and the end of the takeover process. The end of the takeover process was defined as either the endpoint defined for each takeover scenario, or when participants re-engaged the vehicle, whichever was earlier. Consistent with prior research [12], the endpoint of Event 1, 3, 5, and 8 was when the vehicle's center of gravity reached the boundary of the neighboring lane; the endpoint of Event 4, 6, and 7 was when the driver passed the exit point of the curve; and the endpoint of Event 2 was when the vehicle passed the construction zone location. However, participants were instructed to re-engage the vehicle as soon as they thought the vehicle was able to drive on its own.

Following prior research [7], maximum resulting acceleration is calculated as $max\ acceleration_{resulting} = \max_t \sqrt{acceleration^2_{longitudinal} + acceleration^2_{lateral}}$. A smaller acceleration represents a smoother and safer reaction to TORs. In addition, we calculated the maximum resulting jerk as $max\ jerk_{resulting} = \max_t \sqrt{jerk^2_{longitudinal} + jerk^2_{lateral}}$. Jerk is the derivative of acceleration and has been utilized to evaluate shift quality, ride comfort [14] and driving aggressiveness [3, 4, 9]. Similarly, a smaller jerk represents higher takeover quality. For the four takeover event that require drivers to change lanes, TTC was measured. TTC is a time-based event criticality indicator for detecting rear-end collision risk and is defined as the time taken for two objects to collide if maintaining their present speeds and trajectories [13]. A larger minimum TTC means lower risk of collision and better takeover quality. If the vehicle hit the objects ahead or ran off the road, we defined the scenario as a collision and set the minimum TTC as "NA".

In the present study, participants were informed that the NDRT would not stop automatically when an TOR was issued, and were instructed to end the NDRT by themselves. This setting was similar to the "no lockout" condition in the study of [32], where participants had to make a trade-off decision between terminating the NDRT and taking over control of the vehicle immediately [32]. Drivers' reaction time and accuracy in NDRT were used as a check for cognitive load manipulation.

2.5 Experimental procedure

After participants signed an informed consent form and completed an online demographics questionnaire, they were asked to track six targets on the front screen for eye-tracking calibration. They were then introduced the N-back task and automated driving features of the simulator. Participants were told that there was no need to actively monitor the driving environments or takeover control of the vehicle while no TOR was issued as the vehicle was able to handle the situations by itself.

Participants had a 2-minute practice for the N-back memory task, followed by a 5-minute practice drive using the AV. In the practice drive, they were asked to change lanes, engage the AV, perform the N-back task, and take over control of the vehicle. The takeover event was a scenario where traffic lights did not work and required drivers to observe the surroundings and take over control of the vehicle. During driving, participants were instructed to maintain the speed limit of 35 mph in urban and rural environments and 65 mph in highway environments and obey all the traffic rules.

Each participant drove two experimental drives (15-20 minutes each), each containing four takeover events. At the beginning of the drive, participants were asked to activate the AV mode and then start the N-back task when the audio command "Please start the NDRT" was issued. Participants were told that an extra $20 bonus could be earned if their NDRT performance was ranked among the top 10. After about 90-second NDRT, a TOR was issued unexpectedly and participants were required to terminate the NDRT manually and take over the control immediately. When participants thought they had negotiated the takeover event, they were free to activate AV mode. The experimenter would remind participants to re-engage the vehicle if they did not turn on the automation after the takeover event. The survey on takeover readiness was administered after each takeover event with AV mode activated. The operation of NDRT, takeover, AV mode activation and question process were repeated for each takeover event.

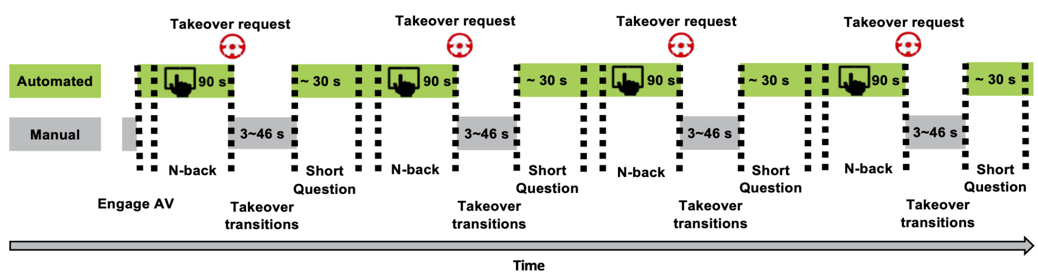

Fig. 3. Sequence of events in one drive

3 RESULTS

Each participants experienced 8 takeover events, so a total of 640 takeovers were available for data analysis. Due to technical malfunction, 21 takeover events were excluded for eye-tracking data analysis. We used a linear mixed model to analyze how cognitive load, TOR lead time, and traffic density affect takeover readiness and takeover performance (timeliness and quality). The cognitive load, TOR lead time, traffic density, and their two-way interactions were set as fixed effects. Random intercept (participants had their own intercepts) rather than random slope (participants had their own slopes) was used in the model development. Pearson correlation coefficients were examined to explore the relationships between drivers' subjective takeover readiness and objective takeover performance. α was set at .05 for results to be reported as significant. Table 4 and 5 show the mean and standard error of dependent variables.

Table 4. Mean and standard error of dependent variables under 4s TOR lead time ($Mean \pm SE$)

	4s TOR lead time			
	No traffic		Heavy oncoming traffic density	
	Low cognitive load	High cognitive load	Low cognitive load	High cognitive load
Takeover readiness	74.6 ± 2.9	65.8 ± 2.8	72.6 ± 2.8	68.1 ± 3.0
Takeover RT (s)	2.32 ± 0.14	2.26 ± 0.14	2.2 ± 0.13	2.23 ± 0.11
Eyes-on-road RT (s)	1.05 ± 0.11	1.17 ± 0.09	0.84 ± 0.1	1.33 ± 0.09
Max resulting acceleration (m/s^2)	3.29 ± 0.31	3.21 ± 0.26	3.22 ± 0.29	3.71 ± 0.3
Max resulting jerk (m/s^3)	44.1 ± 10.1	36.2 ± 8.0	43.5 ± 8.6	45.2 ± 8.0
Minimum TTC (s)	1.39 ± 0.21	1.46 ± 0.16	1.3 ± 0.16	1.22 ± 0.18
Road deviation SE (cm)	1.00 ± 0.15	0.84 ± 0.07	0.81 ± 0.08	0.85 ± 0.07
Accuracy in N-back task (%)	94.5 ± 0.6	88.8 ± 1.1	95.3 ± 0.4	89.6 ± 0.8
Reaction time in N-back task (s)	1.20 ± 0.02	1.24 ± 0.02	1.20 ± 0.02	1.23 ± 0.02

Table 5. Mean and standard error of dependent variables under 7s TOR lead time ($Mean \pm SE$)

	7s TOR lead time			
	No traffic		Heavy oncoming traffic density	
	Low cognitive load	High cognitive load	Low cognitive load	High cognitive load
Takeover readiness	79.9 ± 2.1	69.0 ± 2.8	78.9 ± 2.2	74.5 ± 2.3
Takeover RT (s)	2.33 ± 0.15	2.43 ± 0.14	2.11 ± 0.15	2.28 ± 0.14
Eyes-on-road RT (s)	1.16 ± 0.10	1.23 ± 0.10	1.00 ± 0.10	1.30 ± 0.09
Max resulting acceleration (m/s^2)	2.44 ± 0.25	2.46 ± 0.25	2.76 ± 0.28	2.39 ± 0.24
Max resulting jerk (m/s^3)	28.8 ± 7.1	28.0 ± 7.0	36.5 ± 7.7	26.4 ± 5.7
Minimum TTC (s)	3.34 ± 0.24	3.2 ± 0.25	2.68 ± 0.25	2.72 ± 0.23
Road deviation SE (cm)	0.78 ± 0.06	0.72 ± 0.06	0.83 ± 0.09	0.79 ± 0.05
Accuracy in N-back task (%)	95.1 ± 0.6	89.0 ± 1.0	95.2 ± 0.4	89.4 ± 1.0
Reaction time in N-back task (s)	1.20 ± 0.02	1.26 ± 0.02	1.18 ± 0.02	1.24 ± 0.02

3.1 Manipulation check

There were significant main effects of workload on drivers' reaction time and accuracy in the NDRTs, $F(1, 552) = 37.5$, $p < .001$; $F(1, 553) = 158$, $p < .001$. Drivers had longer reaction time and lower accuracy in 2-back task than 1-back task, which indicated the success of our manipulation.

3.2 Takeover readiness

There were significant main effects of cognitive load, $F(1, 554) = 23.50$, $p < .001$, and TOR lead time, $F(1, 554) = 12.89, p < .001$, on subjective ratings of takeover readiness. Drivers said they were more ready to take over control of the vehicle when they had lower cognitive load and when they had a larger time budget for takeover transitions. No other significant effects were found.

3.3 Takeover timeliness

There were no significant effect of cognitive load, $F(1, 554) = .44$, $p = .51$; traffic density, $F(1, 554) = 2.28, p = .13$; and TOR lead time, $F(1, 555) = .30, p = .59$, on takeover time. None of their interaction effects were significant.

For eyes-on-road reaction time, there was a significant effect of cognitive load, $F(1, 534) = 21.00$, $p < .001$. As shown in Figure 4, lower cognitive load led to shorter eyes-on-road reaction time. Moreover, there was a significant interaction effect between cognitive load and traffic density, $F(1, 534) = 7.11, p = .01$. Simple effect analysis showed that lower cognitive load led to a shorter eyes-on-road time in heavy oncoming traffic condition, $p < .001$, while cognitive load did not influence eyes-on-road time when there was no oncoming traffic, $p = .18$.

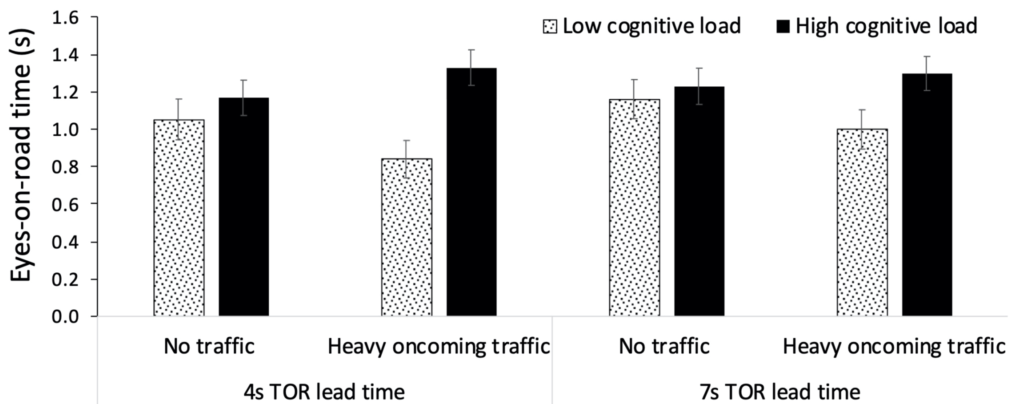

Fig. 4. Eyes-on-road reaction time (s). Error bar indicates one standard error

3.4 Takeover quality

The 4s TOR lead time resulted in a larger maximum resulting acceleration, $F(1, 555) = 23.23, p < .001$, and a larger maximum resulting jerk, $F(1, 555) = 5.64, p = .02$ (Figure 5). No other significant effects were found on maximum resulting acceleration, maximum resulting jerk, or standard error of road deviation.

In the four takeover events that required drivers to change lanes, drivers had shorter minimum TTC when oncoming traffic density was heavy, $F(1, 305) = 5.77, p = .02$, and when the TOR was 4s, $F(1, 305) = 115.32, p < .001$ (Figure 6).

3.5 Correlations between drivers' subjective takeover readiness and objective takeover performance

As shown in Table 6, there were significant correlations between drivers' subjective takeover readiness and objective takeover performance including eyes-on-road reaction time, maximum resulting acceleration/jerk, minimum TTC, and standard error of road deviation. To be specific, the more ready drivers said they were for takeover requests, the larger minimum TTC drivers had. The more takeover readiness drivers reported, the lower eyes-on-road reaction time, maximum resulting acceleration/jerk, and standard error of road deviation drivers had.

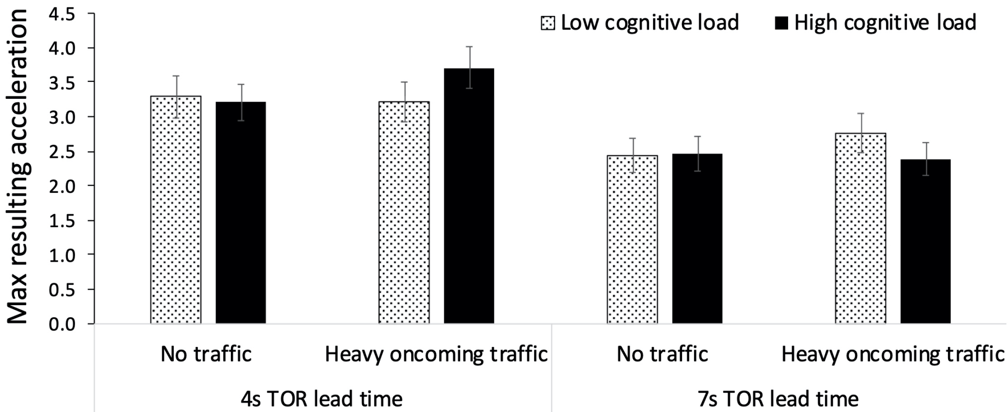

Fig. 5. Maximum resulting acceleration (m/s^2). Error bar indicates one standard error

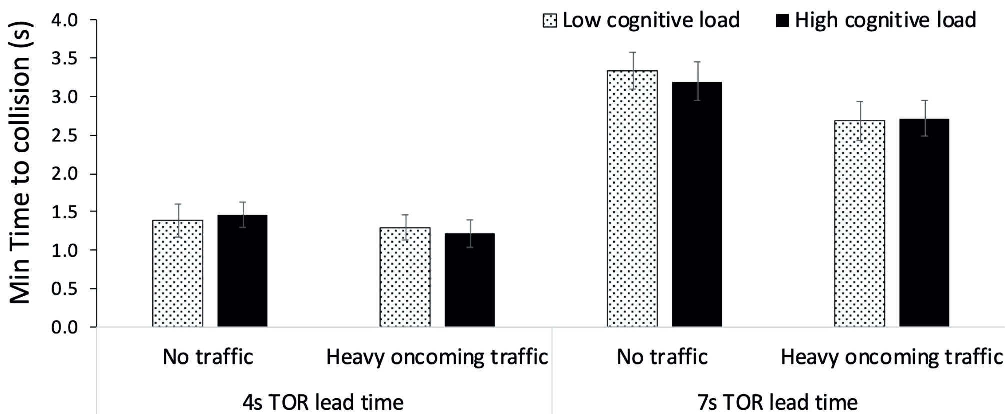

Fig. 6. Minimum time to collision (TTC) (s). Error bar indicates one standard error

Table 6. Correlation Matrix between drivers' subjective takeover readiness and objective takeover performance; * Difference is significant at the 0.05 level; ** Difference is significant at the 0.01 level

	TOR reaction time	Eyes-on-road reaction time	Max resulting acceleration	Max resulting jerk	Minimum TTC	Road deviation SE
Takeover readiness	-.074	-.137**	-.126**	-.086*	.192**	-.080*

4 DISCUSSION

In the present study, we systematically investigated the relationships between a driver's takeover performance and cognitive load, TOR lead time, and traffic density. We discuss how the three factors influenced different aspects of takeover transitions.

4.1 Effects of cognitive load, TOR lead time, and traffic density

We manipulated drivers' cognitive load using a N-back memory task and found that with a lower cognitive load, drivers had a higher level of perceived takeover readiness, and shorter eyes-on-road reaction time. We speculate that cognitive load influenced takeover actions by changing the monitoring behavior at the NDRT stage. When the cognitive load was low, drivers had the bandwidth to monitor the driving environment, stay prepared and therefore responded quickly once the TOR was issued. More gaze behaviors could be analyzed in the future to validate the explanations.

Prior research on traffic density indicated that heavy traffic reduced takeover performance [8, 11, 12, 18, 26]. However, in these studies, traffic density was confounded with the availability of escape paths. In the present study, all the other vehicles were in the oncoming traffic. This design allowed us to decouple traffic density and the availability of escape paths. We found that oncoming traffic density did not influence either takeover reaction time or maximum resulting acceleration/jerk. This result implies that heavy traffic density per se did not lead to increased use of braking or steering but the available escape paths did. Nevertheless, as represented by minimum TTC, heavy oncoming traffic density did increase drivers' risk of collision through longer decision making time.

Interestingly, we found an interaction effect between oncoming traffic density and cognitive load, that lower cognitive load shortened eyes-on-road reaction time only during heavy oncoming traffic. This suggests that participants adjusted their monitoring behavior according to the complexity of driving environments. When oncoming traffic density was heavy, low cognitive load allowed drivers to allocate their attention on the road in order to increase situational awareness and be prepared for potential events.

There were no effects of TOR lead time on takeover reaction time and eyes-on-road time. Generally, longer takeover lead time lead to longer takeover reaction times, but takeover reaction time is also influenced by other variables such as scenario emergency [20]. A possible reason for no significance here is that all the scenarios looked urgent so drivers reacted reflectively and equally quickly regardless of takeover request lead time.

When TOR lead time was 7s, participants had higher takeover readiness, smaller maximum resulting acceleration and jerk, and larger minimum time to collision. Maximum resulting acceleration and jerk have been utilized to identify shift safety, ride comfort [14] and driving aggressiveness [3, 9]. The results indicate that 7s TOR lead time led to safer and and more comfortable takeover behavior and lower collision risk, which supports previous studies [10, 12, 18, 31].

This study uncovered significant correlations between drivers' perceptions of takeover readiness and objective takeover performance. When drivers perceived themselves more ready to take over control the vehicle, it took them less time to move their eyes back on the roads upon takeover requests. During takeover transitions, the more takeover readiness drivers reported, the less maximum resulting acceleration/jerk drivers had, indicating more smooth and comfortable driving. Larger takeover readiness was also related with larger minmim TTC and smaller standard error of road deviations, suggesting lower collision risk and better lane keeping behaviors.

4.2 Design implications

Our findings have implications for the design of in-vehicle alert systems. While different levels of cognitive load, traffic density, and TOR lead time influence drivers' takeover performance differently, in-vehicle alert systems can be designed adaptively to match drivers' takeover readiness. For example, when oncoming traffic density is heavy, a gaze guidance system can be designed to highlight the most important features and support drivers noticing the potential hazards. When TOR lead time is short, a multimodal display combining auditory, visual and tactile alerts can be issued to alarm drivers intensively and increase their alertness [25]. If driver monitoring systems detect that drivers are in high cognitive load induced by NDRTs, a regulatory warning can be issued to remind drivers to stop current NDRTs, keep relaxed and prepare for potential takeover requests.

4.3 Limitations and future study

First, considering the experiment goals and time limitation, eight takeover scenarios were designed within about 40-minutes drive course. Future research can extend experiment duration to see the potential findings caused by drivers' drowsiness or boredom in automated driving mode. Second, only driving behaviors were reported in this study. Future studies can collect and report drivers' physiological measures such as heart rate and galvanic skin conductance to indicate their cognitive and emotional states and situational awareness non-intrusively before a takeover event. Third, drivers' individual demographic information and personality may affect takeover readiness and performance. Future studies can include more driver groups with different characteristics. Fourth, signal filtering can be applied to extract jerk (third derivatives of acceleration) to reduce numeric noise in future studies.

5 CONCLUSION

Existing studies on the effects of drivers' cognitive load on takeover performance have mixed findings and didn't comprehensively explore the interaction effects of cognitive load with other road situations and TOR lead time.

This study systematically investigated the effects of a driver's cognitive load, TOR lead time, and traffic density on takeover readiness and performance. Generally, the results showed that drivers had worse takeover readiness and performance when they had a high cognitive load, short takeover request lead time, and heavy oncoming traffic density. Interestingly, if drivers had low cognitive load, they paid more attention to driving environments and responded more quickly to takeover requests in high oncoming traffic density condition.

The results have important implications for the design of in-vehicle alert systems to monitor driver behaviors, improve takeover readiness and optimize takeover performance. It is highly recommended that appropriate self-regulatory behavior is required on NDRT selection, especially in a complex driving environment. The findings will enhance the interaction between drivers and conditionally automated vehicles.

ACKNOWLEDGMENTS

This work was supported by the University of Michigan Mcity and in part by the National Science Foundation. The views expressed are those of the authors and do not reflect the official policy or position of State Farm®.

REFERENCES

[1] Jackie Ayoub, Feng Zhou, Shan Bao, and X Jessie Yang. 2019. From Manual Driving to Automated Driving: A Review of 10 Years of AutoUI. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 70–90.

[2] Su Jin Baek, Hanna Yun, and Ji Hyun Yang. 2019. How do humans respond when automated vehicles request an immediate vehicle control take-over?. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*. 341–345.

[3] Omar Bagdadi and András Várhelyi. 2011. Jerky driving—an indicator of accident proneness? *Accident Analysis & Prevention* 43, 4 (2011), 1359–1363.

[4] Omar Bagdadi and András Várhelyi. 2013. Development of a method for detecting jerks in safety critical events. *Accident Analysis & Prevention* 50 (2013), 83–91.

[5] Richard Bishop. 2019. Is 2020 The Year For Eyes-Off Automated Driving? *Forbes* (2019). https://www.forbes.com/sites/richardbishop1/2019/10/01/is-2020-the-year-for-eyes-off-automated-driving

[6] M Bueno, E Dogan, F Hadj Selem, E Monacelli, S Boverie, and A Guillaume. 2016. How different mental workload levels affect the take-over control after automated driving. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2040–2045.

[7] Na Du, Feng Zhou, Elizabeth Pulver, Dawn M Tilbury, Lionel P Robert, Anuj K Pradhan, and X Jessie Yang. 2020. Examining the Effects of Emotional Valence and Arousal on Takeover Performance in Conditionally Automated Driving. *Transportation research part C: emerging technologies* 112 (2020), 78–87.

[8] Alexander Eriksson, Sebastiaan M Petermeijer, Markus Zimmermann, Joost CF De Winter, Klaus J Bengler, and Neville A Stanton. 2018. Rolling out the red (and green) carpet: supporting driver decision making in automation-to-manual transitions. *IEEE Transactions on Human-Machine Systems* 49, 1 (2018), 20–31.

[9] Fred Feng, Shan Bao, James R Sayer, Carol Flannagan, Michael Manser, and Robert Wunderlich. 2017. Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data. *Accident Analysis & Prevention* 104 (2017), 125–136.

[10] Christian Gold, Daniel Damböck, Lutz Lorenz, and Klaus Bengler. 2013. "Take over!" How long does it take to get the driver back into the loop?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 57. SAGE Publications Sage CA: Los Angeles, CA, 1938–1942.

[11] Christian Gold, Riender Happee, and Klaus Bengler. 2018. Modeling take-over performance in level 3 conditionally automated vehicles. *Accident Analysis and Prevention* 116 (nov 2018), 3–13. https://doi.org/10.1016/j.aap.2017.11.009

[12] Christian Gold, Moritz Körber, David Lechner, and Klaus Bengler. 2016. Taking over control from highly automated vehicles in complex traffic situations: the role of traffic density. *Human factors* 58, 4 (2016), 642–652.

[13] John C Hayward. 1972. Near miss determination through use of a scale of danger. (1972).

[14] Quanan Huang and Huiyi Wang. 2004. *Fundamental study of jerk: evaluation of shift quality and ride comfort.* Technical Report. SAE Technical Paper.

[15] IORS. 2012. Road vehicles-Ergonomic aspects of transport information and control systems-Calibration tasks for methods which assess driver demand due to the use of in-vehicle systems. ISO/TS 14198:2012. (2012).

[16] Susanne M Jaeggi, Martin Buschkuehl, John Jonides, and Walter J Perrig. 2008. Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences* 105, 19 (2008), 6829–6833.

[17] Jeamin Koo, Dongjun Shin, Martin Steinert, and Larry Leifer. 2016. Understanding driver responses to voice alerts of autonomous car operations. *International journal of vehicle design* 70, 4 (2016), 377–392.

[18] Moritz Körber, Christian Gold, David Lechner, and Klaus Bengler. 2016. The influence of age on the take-over of vehicle control in highly automated driving. *Transportation research part F: traffic psychology and behaviour* 39 (2016), 19–32.

[19] T Louw, G Kountouriotis, O Carsten, and N Merat. 2015. Driver Inattention During Vehicle Automation: How Does Driver Engagement Affect Resumption Of Control? (2015).

[20] Anthony D McDonald, Hananeh Alambeigi, Johan Engström, Gustav Markkula, Tobias Vogelpohl, Jarrett Dunne, and Norbert Yuma. 2019. Toward computational simulations of behavior during automated driving takeovers: A review of the empirical and modeling literatures. *Human factors* 61, 4 (2019), 642–688.

[21] David Miller, Mishel Johns, Brian Mok, Nikhil Gowda, David Sirkin, Key Lee, and Wendy Ju. 2016. Behavioral measurement of trust in automation: the trust fall. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 60. SAGE Publications Sage CA: Los Angeles, CA, 1849–1853.

[22] Brian Mok, Mishel Johns, Key Jung Lee, David Miller, David Sirkin, Page Ive, and Wendy Ju. 2015. Emergency, automation off: Unstructured transition timing for distracted drivers of automated vehicles. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2458–2464.

[23] Lisa J Molnar, Lindsay H Ryan, Anuj K Pradhan, David W Eby, Renée M St Louis, and Jennifer S Zakrajsek. 2018. Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving. *Transportation research part F: traffic psychology and behaviour* 58 (2018), 319–328.

[24] Luke Petersen, Lionel Robert, Jessie Yang, and Dawn Tilbury. 2019. Situational awareness, driver's trust in automated driving systems and secondary task performance. *SAE International Journal of Connected and Autonomous Vehicles, 2(2), DOI:10.4271/12-02-02-0009* (2019).

[25] Ioannis Politis, Stephen Brewster, and Frank Pollick. 2013. Evaluating multimodal driver displays of varying urgency. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 92–99.

[26] Jonas Radlmayr, Christian Gold, Lutz Lorenz, Mehdi Farid, and Klaus Bengler. 2014. How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 58. Sage Publications Sage CA: Los Angeles, CA, 2063–2067.

[27] Tara Rezvani, Katherine Driggs-Campbell, Dorsa Sadigh, S Shankar Sastry, Sanjit A Seshia, and Ruzena Bajcsy. 2016. Towards trustworthy automation: User interfaces that convey internal and external awareness. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 682–688.

[28] Society of Automotive Engineers. 2018. Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems.

[29] Arie P van den Beukel and Mascha C van der Voort. 2013. The influence of time-criticality on situation awareness when retrieving human control after automated driving. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. IEEE, 2000–2005.

[30] Marcel Walch, Kristin Lange, Martin Baumann, and Michael Weber. 2015. Autonomous driving: investigating the feasibility of car-driver handover assistance. In *Proceedings of the 7th international conference on automotive user interfaces and interactive vehicular applications*. 11–18.

[31] Jingyan Wan and Changxu Wu. 2018. The Effects of Lead Time of Take-Over Request and Nondriving Tasks on Taking-Over Control of Automated Vehicles. *IEEE Transactions on Human-Machine Systems* 99 (2018), 1–10.

[32] Bernhard Wandtner, Nadja Schömig, and Gerald Schmidt. 2018. Effects of non-driving related task modalities on takeover performance in highly automated driving. *Human factors* 60, 6 (2018), 870–881.

[33] Kathrin Zeeb, Axel Buchner, and Michael Schrauf. 2016. Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident analysis & prevention* 92 (2016), 230–239.

[34] Kathrin Zeeb, Manuela Härtel, Axel Buchner, and Michael Schrauf. 2017. Why is steering not the same as braking? The impact of non-driving related tasks on lateral and longitudinal driver interventions during conditionally automated driving. *Transportation research part F: traffic psychology and behaviour* 50 (2017), 65–79.

[35] Feng Zhou, X Jessie Yang, and Xin Zhang. 2019. Takeover Transition in Autonomous Vehicles: A YouTube Study. *International Journal of Human–Computer Interaction* 0, 0 (2019), 1–12. https://doi.org/10.1080/10447318.2019.1634317