ARTICLE IN PRESS: TRANSPORTATION RESEARCH RECORD

THE INFLUENCE OF DRIVING EXPERIENCE ON DISTRACTION ENGAGEMENT IN AUTOMATED VEHICLES

Dengbo He
Department of Mechanical and Industrial Engineering, University of Toronto
5 King's College Road, Toronto, ON, Canada M5S 3G8
Tel: 416-978-0881 Email: dengbo.he@mail.utoronto.ca

Birsen Donmez, Corresponding Author
University of Toronto, Department of Mechanical and Industrial Engineering
5 King's College Road, Toronto, ON M5S 3G8, Canada
Tel: +1 416-978-7399 Fax: +1 416-978-7753 Email: donmez@mie.utoronto.ca

Original Submission Date: August 1, 2018

1st Revision Date: October 25, 2018

Revision after Acceptance: March 6, 2019

ABSTRACT

State-of-the-art vehicle automation requires drivers to visually monitor the driving environment and the automation (through interfaces and vehicle's actions), and intervene when necessary. However, as evidenced by recent automated vehicle crashes and laboratory studies, drivers are not always able to step in when the automation fails. Research points to the increase in distraction or secondary task engagement in the presence of automation as a potential reason. However, previous research on secondary task engagement in automated vehicles mainly focused on experienced drivers. This issue may be amplified for novice drivers with less driving skill. In this paper, we compared secondary task engagement behaviors of novice and experienced drivers both in manual (non-automated) and automated driving settings in a driving simulator. A self-paced visual-manual secondary task presented on an in-vehicle display was utilized. Phase 1 of the study included 32 drivers (16 novice) who drove the simulator manually. In Phase 2, another set of 32 drivers (16 novice) drove with SAE-level 2 automation. In manual driving, there were no differences between novice and experienced drivers' rate of manual interactions with the secondary task (i.e., taps on the display). However, with automation, novice drivers had a higher manual interaction rate with the task than experienced drivers. Further, experienced drivers had shorter average glance durations toward the task than novice drivers in general, but the difference was larger with automation compared to manual driving. It appears that with automation, experienced drivers are more conservative in their secondary task engagement behaviors compared to novice drivers.

Keywords: Automated Driving, Driving Experience, Distraction, Workload, Driving Simulator

INTRODUCTION

The driving task, which remained relatively unchanged since the advent of motor vehicles, is now being transformed drastically. Within certain limits, today's vehicles are capable of detecting and reacting to hazards as well as maintaining lateral and longitudinal control; they can act in partial, high, or full automation, assuming some or all aspects of vehicle control (*1*). The rapid development of sensor, wireless communication, and computing technology has also given rise to a range of devices, such as smart phones, which are capable of entertaining and informing the driver. Although many of these devices raise concerns regarding driver distraction (*2*), they will continue to be a part of the vehicle cockpit (*3*).

Until a self-driving vehicle is capable of taking full responsibility (i.e., SAE-level 4 or level 5 as per SAE J3016_201401 (*4*)), the driver's task will increasingly transform into monitoring of, and coordination with, vehicle automation while being exposed to several information sources, both related and unrelated to the driving task. Current state-of-the-art systems, e.g., Tesla Autopilot, require drivers to visually monitor the driving environment and the automation (through the vehicle's actions and the autopilot interface), and intervene in a timely manner when necessary. However, as evidenced by recent automated vehicle crashes (*5, 6*), drivers are not always able to step in when the automation fails to perform its tasks. Laboratory studies also provide supporting evidence for potential issues with state-of-the-art automated vehicles.

Overall, laboratory studies suggest that drivers experience reduced workload when either the lateral, or longitudinal vehicle control, or both are assigned to an automated system. For example, in a driving simulator, Stanton and Young (*7*) found their participants to report lower levels of workload with an Adaptive Cruise Control (ACC) System. With a fully automated system, again in a driving simulator study, de Waard et al. (*8*) also found reduced workload levels indicated through self-reports and also through physiological measures (heart rate). Although a reduction in workload may have positive results on performance when workload levels are high, reducing workload when it is already low can impair performance, as suggested by the Yerkes-Dodson Law (*9, 10*). In fact, research in driving simulators has shown that driving with automated systems can lead to a loss of situational awareness (*7*) and an increase in fatigue (*11*). Further, a meta-analysis comparing manual driving, partial automation (in particular ACC), and highly automated driving found that the drivers of a highly automated car and to a lesser extent ACC drivers are likely to engage in non-driving tasks, which lead to distraction (*12*). On the other hand, if drivers are not engaged in non-driving tasks when they are driving with higher levels of automation, they would be more prone to experiencing fatigue compared to if they are not engaged in non-driving tasks when they are driving manually.

Although research points to potential issues regarding distraction or secondary task engagement in automated vehicles, previous research mainly focused on experienced drivers. Crash risk is known to decrease with driving experience (*13*), partly because of the improved skills to control the vehicle (*14*), but also because of the improved capability to distribute attention, even when being distracted (*15*). Thus, the negatives identified about automation so far for experienced drivers may be even more amplified for novice drivers. One vehicle automation study we could identify that focused on both novice and experienced drivers was a simulator study by Young and Stanton (*16*). Drivers were given a self-paced visual-manual secondary task as a measure of workload and attentional capacity, and how well they performed on this task was recorded across four levels of automation: manual control, ACC, active steering (AS), and ACC + AS. Data was collected from four experience groups: 'novices' who never drove, 'learners' who had a leaner's license, 'experts' who had a full license for more than a year, and 'advanced' who had passed an advanced driving training. It was found that the ‘novice’ group benefited more form the introduction of the ACC compared to other groups as indicated by a larger increase in their secondary task performance from manual driving to ACC. However, this study did not report how drivers engaged with the task and how they allocated their attention between the task and the road; the aim was not to simulate real-world distractions and the task was used as a tool to measure workload.

In this paper, we report a driving simulator study that compared novice drivers to experienced drivers in terms of their secondary task engagement behaviors. A self-paced visual-manual secondary task that mimics the operation of in-vehicle infotainment systems was utilized. We captured manual and visual interactions with this secondary task as well as the associated workload and perceived risk. Phase 1 of the study included 32 drivers who drove the simulator manually. In Phase 2, another set of 32 drivers drove the simulator with SAE-level 2 automation (SAE J3016_201401) that combined ACC and Lane Keeping Assist (LKA) systems.

METHODS

As mentioned earlier, the study consisted of two phases: Phase 1 focused on manual driving and Phase 2 focused on automated driving and was conducted after Phase 1 was completed (Table 1). Each phase used a 2×2 between subjects design, with driving experience (novice or experienced) and secondary task (with and without) as independent variables. Participants were randomly assigned to different secondary task conditions. Each participant completed four experimental drives in the simulator.

Participants

A total of 64 participants were recruited for the study with 32 participants completing each phase. Participants were mainly recruited through advertisements posted on the University of Toronto campus, in online forums, and in nearby residential areas. Within each phase, half of the participants were experienced drivers while the other half were novice drivers. The criteria for experienced and novice drivers were based on (*17*) and (*18*). Experienced drivers were required to have held a full driver’s license (G license in Ontario or equivalent ones in Canada or the U.S.) for at least 8 years with >20,000 km (12,427 miles) driven in the past year. Novice drivers were required to have held a driver’s license (at least G2 license in Ontario, Canada or equivalent ones in Canada or the U.S.) for less than 3 years with <10,000 km (6,213 miles) driven in the past year.

The mean ages of participants, as well as minimum, maximum, and standard deviation (SD) of the age within each study group are summarized in Table 1. Each study group (8 total with 8 participants each) was balanced for gender. As would be expected, novice drivers were younger than experienced drivers (mean difference: 12.7 years, $F(1, 56) = 63.2$, $p < .0001$). As expected based on experimental randomization, no difference of age was found across secondary task levels ($p = .15$) and also across phases ($p = .2$). However, experienced drivers in Phase 2 were estimated to be 6.1 years older on the average than experienced drivers in Phase 1 ($t(56) = 2.71$, $p = .009$). Further, most of the participants in Phase 2 reported to having never used ACC or LKA systems. One participant reported using the systems several times a week (an experienced driver randomly assigned to the no secondary task condition), and five participants reported using either an ACC or an LKA system less than several times a year (1 experienced driver in secondary task condition; 2 experienced drivers in no secondary task condition; 1 novice driver in secondary task condition; and 1 novice driver in no secondary task condition). ACC and LKA use information was not collected from participants of Phase 1, which is a limitation of the current study.

Participants were told they would be compensated at a rate of $14/hour. The participants who were assigned to the no secondary task condition were told that they could also earn a bonus of up to $8 based on their driving performance, to encourage participants to take the driving task seriously. The participants who were assigned to the secondary task condition were also told that they could earn a bonus of up to $8; however, their presented bonus scheme depended on both driving performance and secondary task performance. In addition to encouraging these participants to take the driving task seriously, we wanted to encourage them to care about the secondary task simulating real world scenarios (e.g., taking a work-related phone call while driving, or searching for a favorite song on the radio). All participants received the full bonus regardless of their performance.

Apparatus

The study was conducted on a NADS MiniSim Driving Simulator (Figure 1), which is a fixed-based simulator with three 42-inch screens, creating a 130º horizontal and 24º vertical field at a 48-inch viewing distance, with two speakers for stereo sound and a sub-woofer simulating vibration from the road surface. The centre screen displays the left and centre parts of the windshield; the right screen displays the rest of the windshield, the rear-view mirror, the right window, and the right-side mirror; the left screen displays the left window, and the left-side mirror. The simulator is able to simulate ACC and LKA systems simultaneously, creating a SAE level-2 automation (SAE J3016_201401). The driving data is recorded at 60 Hz.

A Surface Pro 2 laptop with a 10.6" touch screen was mounted to the right of the dashboard where half of the participants were presented with a secondary task. A Dikablis head-mounted eye tracking system by Ergoneers was used to record participants' visual attention when they were driving, with two cameras facing the eyes and one camera facing forward. Another camera was mounted under the dashboard to record feet movements.

Heart rate and Galvanic skin response were collected as workload measures (*19-21*). Electrocardiogram (ECG) and GSR sensors by Becker Meditec collected data at 240 Hz using the D-Lab software developed by Ergoneers. Solid gel foam electrodes were used for the ECG and GSR sensors. ECG was recorded with three electrodes placed on participant's chest. The GSR sensors were attached beneath the bare left foot with one sensor in the middle and the other under the heel.

Secondary Task

A visual-manual secondary task developed by Donmez, Boyle, and Lee (*22*) was utilized. The task mimicked the operation of in-vehicle information systems such as searching and selecting songs, and has been shown to degrade driving performance (*17, 22*). Participants were asked to select one out of 10 phrases to match either "Discover" with its first word, "Project" with its second word, or "Missions" with its third word. All phrases consisted of three words and there was only one correct answer in the list of 10 candidate phrases (e.g., "Project Discover Misguide" is not a match, whereas "Discover Missions Predict" is). Two phrases were displayed at one time and participants could tap up and down arrows to scroll through the 10 phrases. Participants tapped a submit button to enter their selection and received feedback on whether their entry was correct or not. A new set of 10 phrases then became available, regardless of the correctness of the submission. The task was available throughout the drive and participants decided when to start a new task and did so by hitting a start button.

Driving Task

Each participant completed four experimental drives, each around 5-minutes long. In total, the experimental drives took about 20 minutes, corresponding to approximately 29 km (18 miles) driven. The order of the four experimental drives was the same across participants. Drives 1 and 3 were on rural roads, whereas Drives 2 and 4 were on highways. The speed limits were 80.5 km/h (50 mph) for rural roads and 96.6 km/h (60 mph) for highways. Participants were instructed to maintain a comfortable distance from lead vehicles and drive around the speed limit.

The vehicle was controlled by the participants in Phase 1, whereas participants in Phase 2 were asked to use the vehicle automation (ACC and LKA combined) whenever possible. Participants in Phase 2 did use the vehicle automation throughout each drive except when they felt that they had to take over vehicle control in the presence of roadway events described below. All participants were found to use automation at least 80% of the total driving time. The vehicle automation was capable of handling all roadway events that occurred in the experimental drives. Both phases included a pre-experiment drive, similar to experimental drives, which provided additional training to the participants. Further, the pre-experiment drive in Phase 2 included an intense lead vehicle braking event that exceeded the capabilities of ACC and hence the participants were primed for potential automation failures.

There were a series of events within each drive representing normal driving conditions (e.g., lead vehicle braking events, vehicle behind taking over) to ensure that the participants attended to the driving task. Analysis of these roadway events are out of scope for the current paper. A subset of these events was specifically designed to investigate anticipatory driving behaviors, and preliminary findings are reported in He and Donmez (*17*).

Procedures

Upon arrival, participant eligibility was verified, and informed consent was obtained. The participants were then provided with experiment instructions and completed practice drives as detailed below. The practice drives were designed to be 5 or 10 minutes; however, if the participants indicated that they had not yet felt comfortable with the amount of practice they received, they were given additional practice time. The routes used in practice were similar (in terms of traffic density and road types) to the ones that were used in experimental drives.

For Phase 1, the participants were given a 5-minute practice drive to get familiar with the simulator. The participants who were assigned to the secondary task condition were then introduced to the secondary task and completed an additional 5-minute practice drive with the secondary task.

For Phase 2, the experimenter first introduced the vehicle automation to the participants. Participants were also informed about the limitations of both ACC (i.e., may not avoid a crash if intensive braking is required, does not respond to stationary objects) and LKA (may not work if lane markings are absent or not visible such as at an intersection). Participants then completed a 10-minute practice drive, the first half without automation and the second half with automation. The participants who were assigned to secondary task condition were introduced to the secondary task before the practice drive and practiced the task during their practice drives.

Following the practice, the participants in both phases were outfitted with the head-mounted eye tracking system and physiological sensors. Participants then completed one more practice drive for additional training (or pre-experiment drive) that lasted for about 6 minutes, but they were told that this was an experimental drive. As mentioned earlier, this drive was used in Phase 2 to introduce an ACC failure to prime participants for the possibility of automation failures.

All participants were told to prioritize driving safety. Participant preparation including practice drives took approximately one hour. Following the practice drives, participants completed the four experimental drives. Before each drive, the eye-tracker was calibrated. After each drive, participants completed questionnaires, including one on workload using the NASA Task Load Index (NASA-TLX) (*23*), and one on perceived risk (*24*), and were allowed a 5-minute rest. NASA-TLX captures workload through six constructs (i.e., mental demand, physical demand, temporal demand, performance, effort, and frustration) assessed on a scale ranging from "0: very low" to "100: very high". The perceived risk questionnaire (*24*) consists of a 10-point ordinal scale ranging from "1: as risky as driving on easy road with no traffic, pedestrians, or animals while perfectly alert" to "10: as risky as driving with my eyes closed; a crash is bound to occur every time I do this". Overall, each experiment took approximately 2.5 hours.

Dependent Variables and Statistical Model

Analysis were conducted on secondary task engagement metrics, as well as on measures of workload and perceived risk. Secondary task engagement was assessed through manual and visual engagement with the secondary task display. The specific metrics included manual interaction rate (number of taps per minute), average glance duration (ms), glance rate (number of glances per minute), long glance rate (number of glances >2 seconds per minute), and percent time looking at the secondary task display. Glance duration was defined from the gaze first intersecting with the secondary task display to it having moved away from the display. Glance durations shorter than 100 ms were not considered (*25*) and were excluded from analysis. Glances longer than 2 seconds were analyzed in particular as they have been found to increase crash risk significantly (*26*).

Workload was assessed through physiological measures of heart rate and GSR, as well as NASA-TLX. Heart rate and GSR were averaged across each drive excluding the periods of driving that corresponded to roadway events. Due to the noise in the signals, heart rate from 19 drives and GSR from one drive (out of 256 drives total) were excluded from analysis. NASA-TLX scores were calculated following the method proposed by Hart and Staveland (*23*). Perceived risk was a single value from a scale of 1-10, with larger values indicating greater perceived risk as discussed earlier.

All statistical models were built in SAS University Edition V9.4. Rate of manual interactions, glance rates, and long glance rates were analyzed through negative binomial regression given that over-dispersion was detected in the count data (*27*). Count data follow the Poisson distribution (if the mean is large, Poisson distribution is approximately normal, but for small means this approximation does not hold). The mean and the variance of the Poisson distribution are equal; if there is over-dispersion in the data (i.e., variance > mean), the negative binomial distribution is used. The number of manual interactions, glances, and long glances observed in our dataset were therefore analyzed with a negative binomial regression. Given that there was variability in how long each participant took to complete their drive, the length of each drive was used as the offset variable in our negative binomial models, transforming the model estimates from counts (number) to rates (number per minute). Further, repeated measures (four drives completed by each participant) were accounted through Generalized Estimating Equations (*28*). All other variables were analyzed through mixed linear models as the residuals met the model assumptions (e.g., normality), with participant as a random factor (to account for correlated observations) and the variance-covariance structure chosen based on the Bayesian Information Criterion.

Secondary task engagement models focused on the 32 participants who performed the secondary task in the experiment. Thus, these models included experience, experimental phase, and their interaction as predictor variables. Workload and perceived risk analysis included experience, secondary task, experiment phase, and their two-way interactions as predictor variables. Given that the data collection for manual driving (Phase 1) was completed before data collection for automated driving (Phase 2), results comparing manual to automated driving are confounded by time of data collection. Rather than performing a qualitative comparison between the findings of the two phases we included phase as a predictor variable to quantify potential differences. However, the readers should be cognizant of this potential confound in interpreting the statistical results generated from the comparisons of the two phases.

The model equations for secondary task engagement are presented below, where, Y is the response variable; t is the length of each drive or the offset variable in the negative binomial models; β_1, β_2, and β_3 are the coefficients for predictor variables: x_1 (=1 when Phase 2, 0 otherwise), x_2 (=1 when experienced, 0 otherwise), and $x_1 * x_2$ interaction.

Negative binomial models:

$$\log\left(\frac{\mathrm{E[Y]}}{t}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 * x_2$$

Mixed effects models for fixed factors:

$$\mathrm{E[Y]} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 * x_2$$

The model equation used for workload measures and perceived risk presented below, includes three additional coefficients, β_4, β_5, and β_6, corresponding to secondary task and its two-ways interactions with phase and experience: x_3 (=1 when with secondary task, 0 otherwise), and $x_1 * x_3$ and $x_2 * x_3$ interactions.

$$\mathrm{E[Y]} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 * x_2 + \beta_4 x_3 + \beta_5 x_1 * x_3 + \beta_6 x_2 * x_3$$

RESULTS

Table 2 summarizes the model results. In Figures 2 to 7, the boxplots present minimum, first quartile, median, third quartile, and maximum with the bottom whisker, lower edge of the box, bold horizontal line, upper edge of the box, and the top whisker, respectively. Raw data points are indicated with gray dots and the averages are indicated with hollow diamonds. “M” stands for mean and “SD” stands for standard deviation.

Secondary Task Engagement

An interaction effect was found for rate of manual interactions with the secondary task display ($\chi^2(1) = 4.31$, $p = .04$). As shown in Figure 2, in Phase 1, i.e., in manual driving, there were no differences between novice and experienced drivers ($p = .6$). However, in Phase 2, i.e., with automation, novice drivers had a 58% higher manual interaction rate with the display compared to the experienced drivers (95% CI: 7%, 133%, $\chi^2(1) = 5.24$, $p = .02$). When comparisons were made across phases, experienced drivers' manual interaction rate did not differ ($p = .09$), whereas novice drivers in Phase 2 who drove with automation had a 131% higher manual interaction rate with the display compared to the novice drivers in Phase 1 who drove manually (95% CI: 55%, 246%, $\chi^2(1) = 16.68$, $p < .0001$).

An interaction effect was also found for average glance duration toward the secondary task display ($F(1, 28) = 4.92$, $p = .03$). As can be seen in Figure 3, experienced drivers had shorter average glance durations than novice drivers in both phases (manual: $t(28) = 2.81$, $p = .009$; automated: $t(28) = 5.95$, $p < .0001$) but the difference between the experienced and novice drivers was bigger with automation. When comparisons were made across phases, experienced drivers' average glance durations did not differ ($p = .08$), whereas novice drivers in Phase 2 who drove with automation had longer average glance durations toward the display compared to the novice drivers in Phase 1 who drove manually ($t(28) = 4.93$, $p < .0001$).

Experienced drivers had a 34% higher rate of glances toward the secondary display than novice drivers (Figure 4, 95% CI: 13%, 60%, $\chi^2(1) = 10.91$, $p = .001$); however, their rate of long glances (> 2 seconds) was 62% lower than novice drivers (Figure 5, 95% CI: 27%, 80%, $\chi^2(1) = 8.41$, $p = .004$). When comparisons were made across phases, rate of long glances (> 2 seconds) were found to be 197% higher in Phase 2 - automation than Phase 1 - manual driving (95% CI: 54%, 473%, $\chi^2(1) = 10.59$, $p = .001$). Percent time looking at the display was also 14% higher in Phase 2 compared to Phase 1 (Figure 6, 95% CI: 6%, 22%, $F(1, 28) = 14.06$, $p = .0008$).

Workload Measures and Perceived Risk

No significant effects were found for GSR or heart rate ($p > .05$). For NASA-TLX, there was an interaction effect of secondary task and automation (Figure 7, $F(1, 57) = 4.15$, $p = .046$). In Phase 1, the presence of the secondary task increased self-reported workload ($t(57) = 3.48$, $p = .001$); whereas in Phase 2, it had no significant effect on self-reported workload ($p = .6$). When comparisons were made across phases, self-reported workload without secondary task did not differ ($p = .7$), whereas self-reported workload with secondary task decreased with automation ($t(57) = 2.54$, $p = .01$).

There was a main effect of secondary task on perceived risk ($F(1, 57) = 23.67$, $p < .0001$). Drivers in the secondary task conditions self-reported to perceive a higher level of risk compared to those in the no secondary task conditions (mean difference: 1.96, 95 CI% = 1.15, 2.77).

DISCUSSION

We conducted a driving simulator study that compared novice drivers to experienced drivers in terms of their secondary task engagement behaviors both in manual and automated driving settings. There were two phases for the study: in Phase 1, the experiment was conducted in manual driving mode, and in Phase 2, the experiment was conducted in automated driving mode. In line with previous research (*15*), experienced drivers showed what can be considered safer glance behaviors when provided with the secondary task compared to novice drivers. Compared to novices, experienced drivers had higher glance rates to the secondary task display but shorter glance durations and lower rate of long glances (> 2 seconds). Although experienced drivers looked at the display more frequently, their glances were shorter than novices, overall leading to similar percentage of driving time spent looking at the secondary task display. With automation, although both experienced and novice drivers appeared to engage with the task more as indicated by an increase in percent time spent looking at the display, experienced drivers' behavior appeared to be affected less than novice drivers' behavior: in Phase 1, i.e., manual driving, there were no differences between novice and experienced drivers in terms of their manual interactions with the task; however, with automation, novice drivers had a higher interaction rate than experienced drivers. Further, experienced drivers had shorter average glance durations toward the task than novice drivers in general, but the difference was larger with automation compared to that with manual driving.

It can be visually concluded from the figures that some observed behaviors were more variable in automated driving compared to manual driving, and the variability was particularly large for novice drivers. For example, as can be seen in Figures 2 and 3, novice driver data in Phase 2 showed the highest dispersion for manual interaction rate and average glance duration. These individual differences should be further explored to identify the type of novice drivers who might be more prone to being distracted in automated vehicles. Although our results suggest that experienced drivers' behavior was affected less by automation than novice drivers' behavior, nonetheless it was still affected. However, we did not have any driving performance assessment in this paper and cannot comment on whether these changes in secondary task engagement behaviors would translate to driving performance decrements. Similarly, real-world data also needs to be collected to assess how our findings would translate to crash risk in an automated vehicle. The criterion we used for defining long glances, i.e., 2 seconds, is based on crash risk assessments on manual driving (*26*) but a different criterion may be needed to tie glance measures recorded in a simulator to crash risk in automated vehicles in general, and different types and levels of automation in particular. This issue merits further research.

Our findings indicate that driving experience makes a difference for how drivers behave when interacting with an automated vehicle. It should be noted that our participants who used automation in the experiment were in general not frequent users of vehicle automation. Thus, their knowledge of how automated vehicle systems work may have been lacking. Further, although our participants were instructed about the limitations of automation and were presented with one automation failure so that they would be cognizant of the limitations of automation, they only completed four experimental drives. With longer interaction, drivers would form experience with such systems. In addition to manual driving experience, experience with automated systems would also affect drivers' behaviors. The change in behaviors as a result of automated driving experience can both be positive (e.g., formation of appropriate mental models (*29*)) and negative (e.g., complacency (*30*)).

Our participants in the no secondary task conditions appeared to perceive the same level of workload whether they were driving manually or with automation. One potential explanation is that our driving task may not have been very demanding to begin with. It is also possible that the automation failure in the pre-experiment drive may have resulted in the participants to monitor the automation with more effort than they would have if they had not experienced the failure. Previous research that found lower levels of self-reported workload with automation did not introduce such automation failures (*7, 8*). Thus, the participants in those studies may not have put as much effort into monitoring the automation as our participants. In Phase 1 (manual driving), the presence of the secondary task increased self-reported workload, whereas in Phase 2 (automated driving), it had no significant effect on self-reported workload. It appears that the drivers perceived to have more spare capacity to perform a secondary task when they were aided with automation.

Although our study provides interesting insights, it also has some limitations that should be pointed out. First, although both study phases used the exact same methodologies, the manual driving phase of the study was completed about four months before data collection began for the automated driving phase. Thus, time of data collection is a potential confound with participants not being randomly assigned to the two phases. Further, although we attribute our findings to driving experience, experience and age are inherently confounded in the driving population and thus our experienced participants were also older than our novice participants. Thus, the findings can be considered to be due to a combination of experience and age factors. Further, we focused on secondary task engagement behaviors, but other potential issues of automated driving should also be investigated which may or may not be exacerbated by lack of driving experience (e.g., fatigue, delayed reaction times).

CONCLUSION

We conducted a driving simulator study investigating the effects of driving experience on distraction (or secondary task) engagement in automated vehicles. The differences observed between the novice and the experienced drivers' manual and visual interactions with the secondary task display indicate that driving experience (in a non-automated or manual vehicle) leads to potentially safer secondary task engagement behaviors in the presence of vehicle control automation. However, it should be noted that most of our participants did not have experience with the type of vehicle automation investigated in our experiment (i.e., Adaptive Cruise Control combined with Lane Keeping Assist), therefore, further research is needed to understand whether these observed benefits of driving experience would sustain with longer term use of vehicle automation.

ACKNOWLEDGEMENT

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC). We gratefully acknowledge the members of the Human Factors and Applied Statistics Laboratory of the University of Toronto for their advice in experimental design and manuscript preparation, Kaiyang Chen for helping with experiment design and data collection, and Wenyi Peng, Songhui Xu, Meining Ji, David Dongsoon Min, Tanishq Singh and Amir Ganiev for their help in data cleaning.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm their contributions to the paper as follows: study conception and design: D. He, B. Donmez; study implementation and data collection: D. He; analysis and interpretation of results: D. He, B. Donmez; manuscript preparation: D. He, B. Donmez. Both authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Regan, M. A., A. Stevens, and T. Horberry. Driver Acceptance of New Technology: Overview. In Driver Acceptance of New Technology: Theory, Measurement and Optimisation (M. A. Regan, T. Horberry, and A. Stevens, eds.), Ashgate Publishing, Farnham, Surrey, England, 2014, pp. 3–8.

2. Regan, M. A., J. D. Lee, and K. Young. Driver Distraction: Theory, Effects, and Mitigation. CRC Press, Boca Raton, Fla., 2008.

3. Cornet, A., D. Mohr, F. Weig, B. Zerlin, and H. Arnt-Philipp. Mobility of the Future: Opportunities for Automotive OEMs. McKinsey & Company, Inc., 2012.

4. SAE On-Road Automated Vehicle Standards Committee. Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. In SAE Standard J, No. 3016, 2014, pp. 1–16.

5. Preliminary Report Highway HWY18FH011. National Transportation Safety Board, Washington, D.C., 2018.

6. Preliminary Report Highway HWY18MH010. National Transportation Safety Board, Washington, D.C., 2018.

7. Stanton, N. A. and M. S. Young. Driver Behaviour with Adaptive Cruise Control. Ergonomics, Vol. 48, No. 10, 2005, pp. 1294–1313.

8. de Waard, D., M. van der Hulst, M. Hoedemaeker, and K. A. Brookhuis. Driver Behavior in an Emergency Situation in the Automated Highway System. Transportation Human Factors, Vol. 1, No. 1, 1999, pp. 67–82.

9. Coughlin, J. F., B. Reimer, and B. Mehler. Monitoring, Managing, and Motivating Driver Safety and Well-Being. IEEE Pervasive Computing, Vol. 10, No. 3, 2011, pp. 14–21.

10. Yerkes, R. M. and J. D. Dodson. The Relation of Strength of Stimulus to Rapidity of Habit-Formation. Journal of Comparative Neurology and Psychology, Vol. 18, No. 5, 1908, pp. 459–482.

11. Jamson, A. H., N. Merat, O. M. J. Carsten, and F. C. H. Lai. Behavioural Changes in Drivers Experiencing Highly-Automated Vehicle Control in Varying Traffic Conditions. Transportation Research Part C: Emerging Technologies, Vol. 30, 2013, pp. 116–125.

12. de Winter, J. C. F., R. Happee, M. H. Martens, and N. A. Stanton. Effects of Adaptive Cruise Control and Highly Automated Driving on Workload and Situation Awareness: A Review of the Empirical Evidence. Transportation Research Part F: Traffic Psychology and Behaviour, Vol. 27, 2014, pp. 196–217.

13. Mayhew, D. R., H. M. Simpson, and A. Pak. Changes in Collision Rates among Novice Drivers During the First Months of Driving. Accident Analysis and Prevention, Vol. 35, 2003, pp. 683–691.

14. Bjørnskau, T. and F. Sagberg. What Do Novice Drivers Learn During the First Months of Driving? Improved Handling Skills or Improved Road User Interaction? Proc., International Conference of Traffic and Transport Psychology, Nottingham, pp. 129–140, 2005.

15. Wikman, A. S., T. Nieminen, and H. Summala. Driving Experience and Time-Sharing During in-Car Tasks on Roads of Different Width. Ergonomics, Vol. 41, 1998, pp. 358–372.

16. Young, M. S. and N. A. Stanton. What's Skill Got to Do with It? Vehicle Automation and Driver Mental Workload. Ergonomics, Vol. 50, No. 8, 2007, pp. 1324–1339.

17. He, D. and B. Donmez. The Effect of Distraction on Anticipatory Driving. Proc., Human Factors and Ergonomics Society 62nd Annual Meeting, Philadelphia, Pa., 2018.

18. Stahl, P., B. Donmez, and G. A. Jamieson. Supporting Anticipation in Driving through Attentional and Interpretational In-Vehicle Displays. Accident Analysis and Prevention, Vol. 91, 2016, pp. 103–113.

19. Carsten, O., F. C. Lai, Y. Barnard, A. H. Jamson, and N. Merat. Control Task Substitution in

Semiautomated Driving: Does It Matter What Aspects Are Automated? Human Factors: The Journal of the Human Factors and Ergonomics Society, Vol. 54, No. 5, 2012, pp. 747–761.

20. Cha, D. Driver Workload Comparisons among Road Sections of Automated Highway Systems. Presented at SAE 2003 World Congress & Exhibition, Detroit, Mich., 2003.

21. He, D., C. C. Liu, B. Donmez, and K. N. Plataniotis. Assessing High Cognitive Load in Drivers through Electroencephalography. Proc., Transportation Research Board 96th Annual Meeting (17-02615), Washington, D.C., 2017.

22. Donmez, B., L. N. Boyle, and J. D. Lee. Safety Implications of Providing Real-Time Feedback to Distracted Drivers. Accident Analysis and Prevention, Vol. 39, No. 3, 2007, pp. 581–590.

23. Hart, S. G. and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Journal of Chemical Information and Modeling, Vol. 53, 1988, pp. 1689–1699.

24. Tsimhoni, O., D. Smith, and P. Green. On-the-Road Assessment of Driving Workload and Risk to Support the Development of an Information Manager. Publication No. UMTRI-2003-08. University of Michigan Transportation Research Institute, 2003.

25. Crundall, D. and G. Underwood. Visual Attention While Driving: Measures of Eye Movements Used in Driving Research. In Handbook of Traffic Psychology (B. E. Porter, ed.), Academic Press, San Diego, 2011, pp. 137–148.

26. Dingus, T. A., S. G. Klauer, V. L. Ramsey, A. Neale, S. E. Petersen, J. Lee, M. A. Sudweeks, J. Perez, D. Hankey, and S. Gupta. The 100-Car Naturalistic Driving Study, Phase II-Results of the 100-Car Field Experiment. Publication No. DOT HS 810 593. National Highway Traffic Safety Administration, Washington, D.C., 2006.

27. Hilbe, J. M. Negative Binomial Regression. Cambridge University Press, 2011.

28. Liang, K. Y. and S. L. Zeger. Longitudinal Data Analysis Using Generalized Linear Models. Biometrika, Vol. 73, No. 1, 1986, pp. 13–22.

29. Bahner, J. E., A. D. Hüper, and D. Manzey. Misuse of Automated Decision Aids: Complacency, Automation Bias and the Impact of Training Experience. International Journal of Human Computer Studies, Vol. 66, 2008, pp. 688–699.

30. Hoff, K. A. and M. Bashir. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. Human Factors: The Journal of the Human Factors and Ergonomics Society, Vol. 57, 2015, pp. 407–434.

FIGURE 1 Experimental setup.

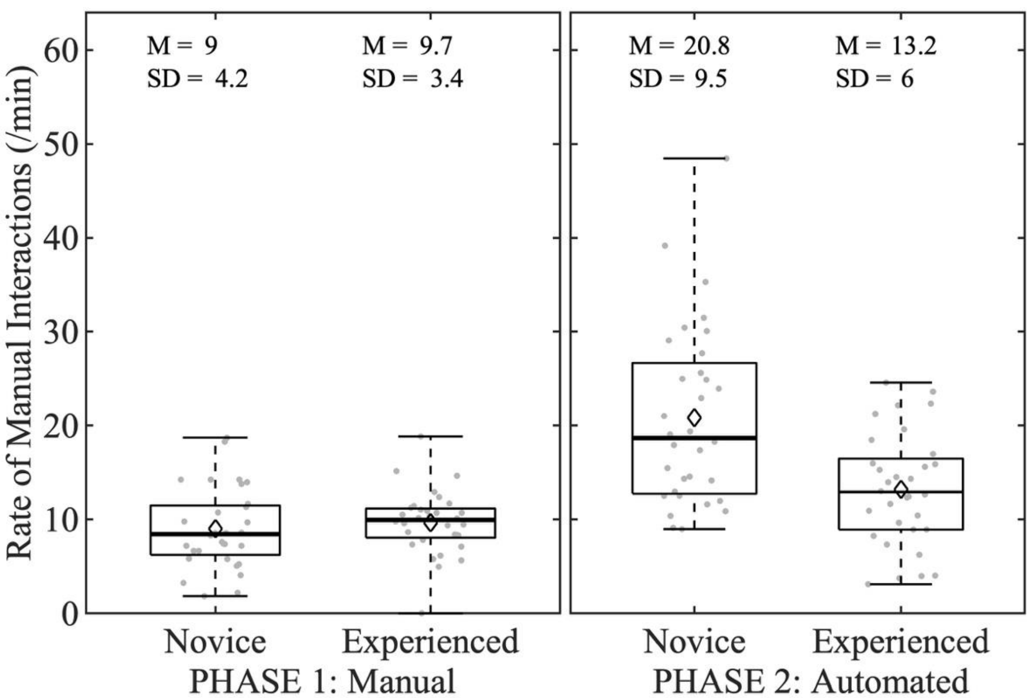

FIGURE 2 Rate of interaction (per minute) with the secondary task display.

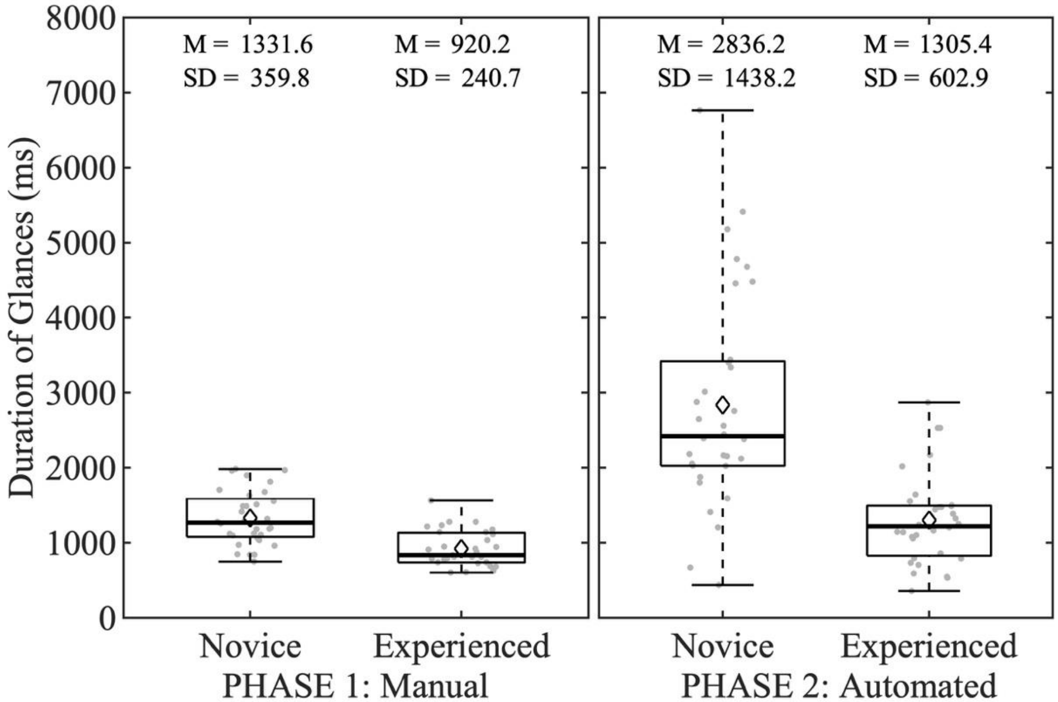

FIGURE 3 Average glance duration toward secondary task display.

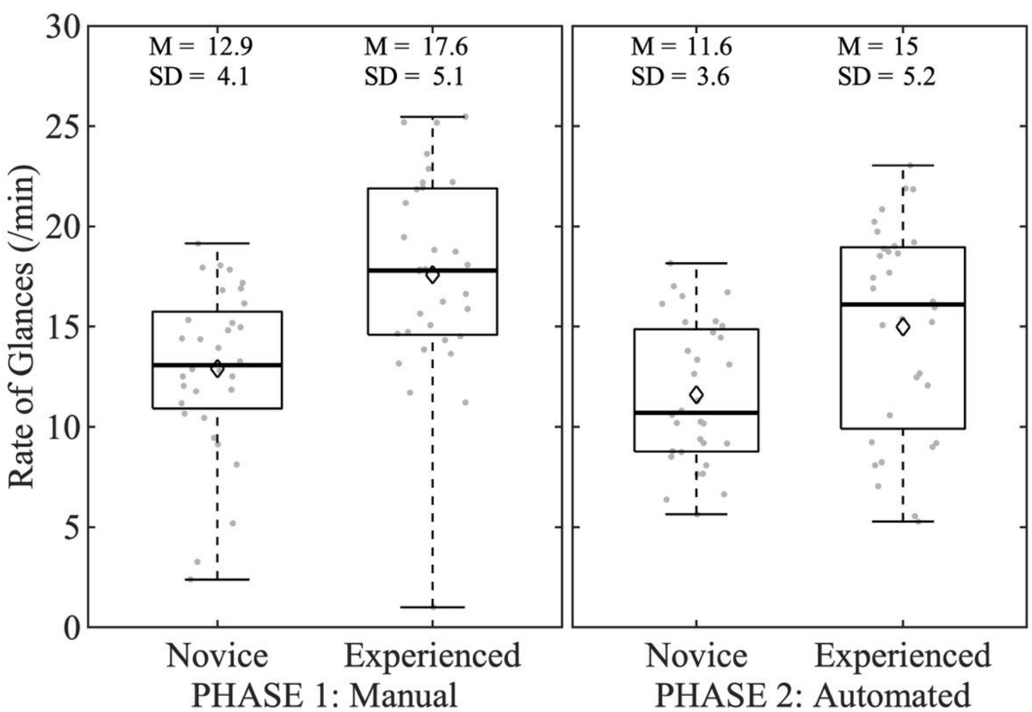

FIGURE 4 Rate of glances (per minute) toward secondary task display.

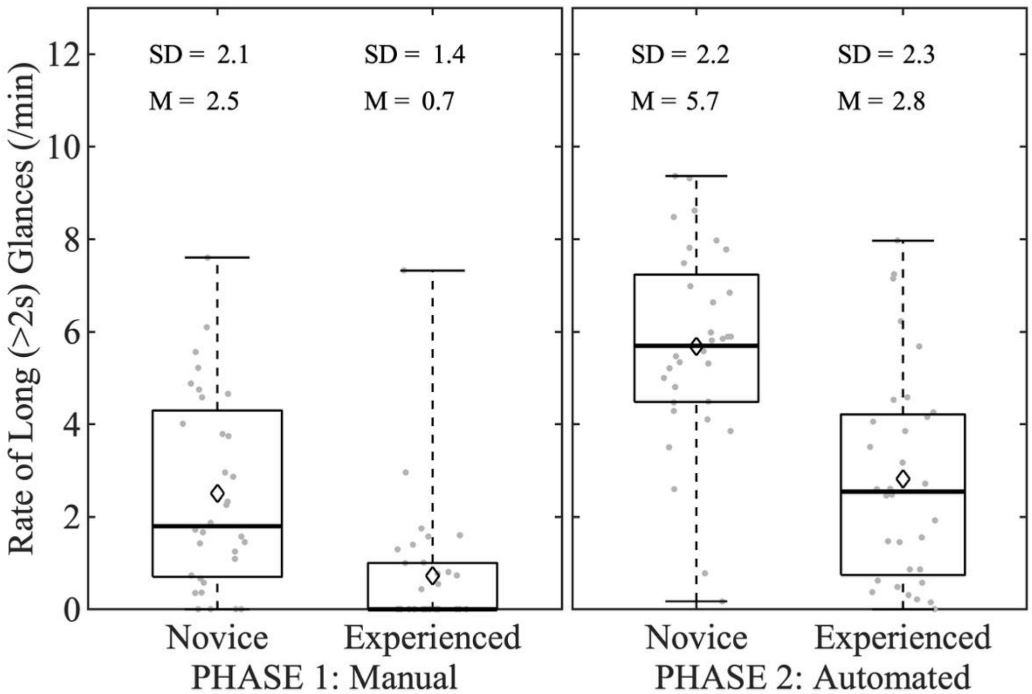

FIGURE 5 Rate of long (>2 seconds) glances (per minute) toward secondary task display.

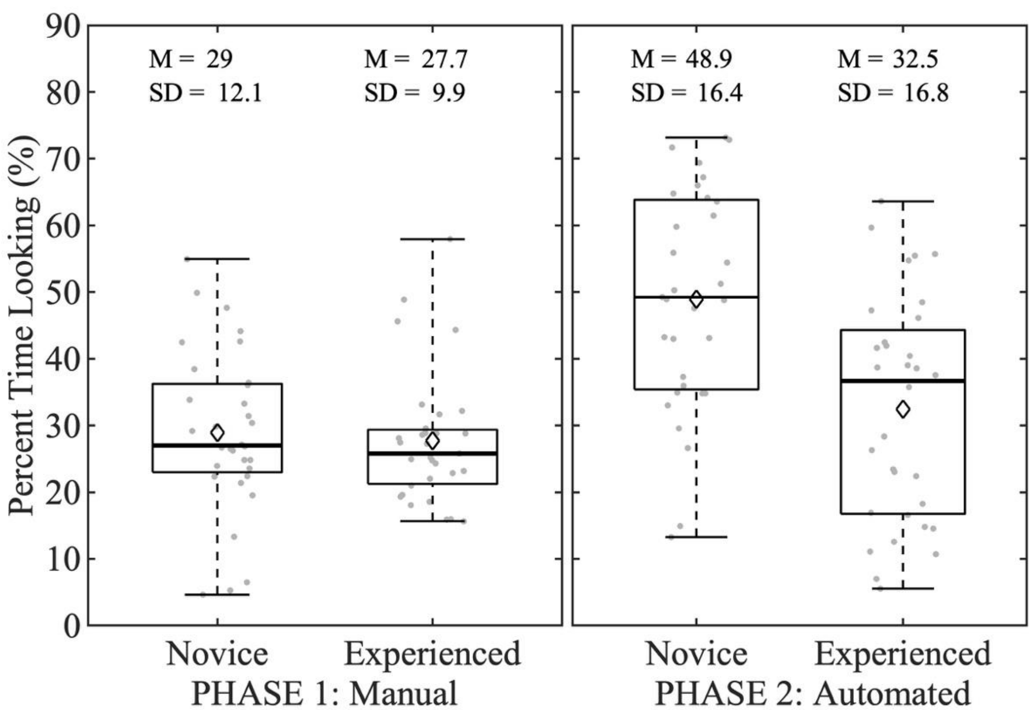

FIGURE 6 Percent time looking at secondary task display.

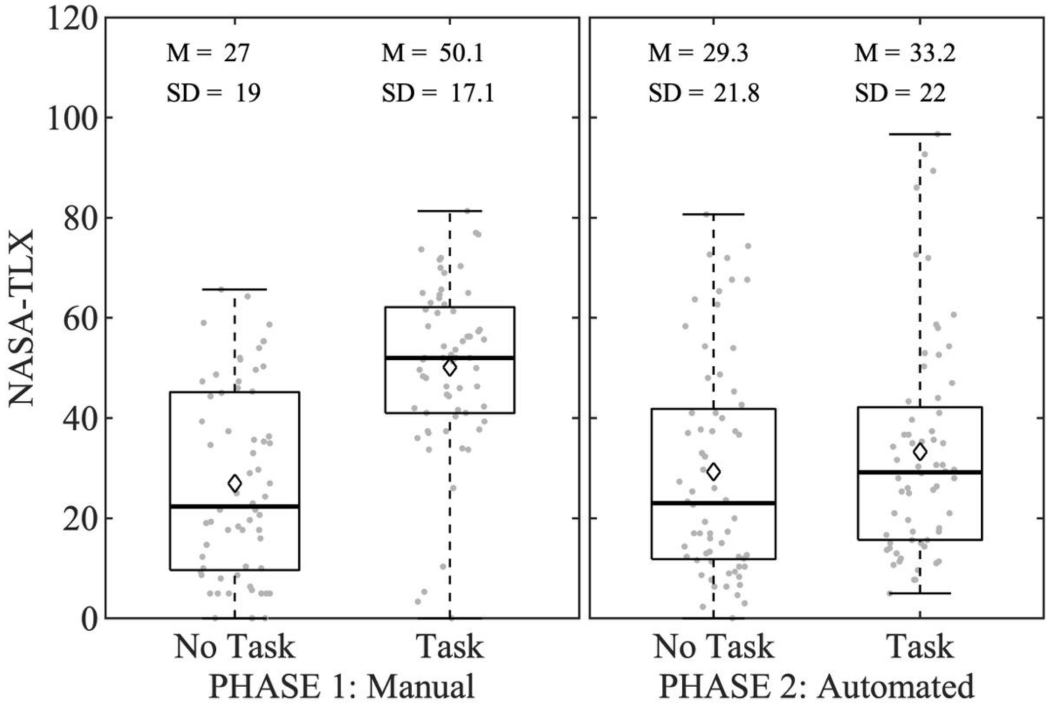

FIGURE 7 NASA-TLX ratings.

TABLE 1 Experimental Design and Participant Age

	Experience	Secondary Task	Mean Age (Minimum - Maximum, SD)
Phase 1: Manual Driving (n=32)	Experienced (n=16)	Yes (n=8)	30.3 (25 - 36, 3.9)
		No (n=8)	33.9 (26 - 47, 7.1)
	Novice (n=16)	Yes (n=8)	21.8 (19 - 27, 2.9)
		No (n=8)	25.3 (19 - 33, 5.2)
Phase 2: Automated Driving (n=32)	Experienced (n=16)	Yes (n=8)	37.4 (28 - 58, 9.4)
		No (n=8)	39.3 (28 - 52, 9.6)
	Novice (n=16)	Yes (n=8)	21.1 (18 - 27, 3.2)
		No (n=8)	21.6 (18 - 24, 1.9)

TABLE 2 Model results

Measure	Phase	Experience	Phase *Experience	Model Coefficients $\beta_0, \beta_1, \beta_2, \beta_3$				
Secondary Task Engagement								
Rate of manual interaction (/min)	$\chi^2(1)=17.82$ $p<.0001$	$\chi^2(1)=1.81$ $p=.2$	$\chi^2(1)=4.31$ $p=.04$	2.20, 0.83, 0.10, -0.55				
Duration of glances (ms)	$F(1,28)=22.55$ $p<.0001$	$F(1,28)=38.31$ $p<.0001$	$F(1,28)=4.92$ $p=.03$	7.85, -0.71, -0.86, 0.45				
Rate of glances (/min)	$\chi^2(1)=2.92$ $p=.09$	$\chi^2(1)=10.91$ $p=.001$	$\chi^2(1)=0.27$ $p=.6$	2.56, -0.11, 0.34, -0.09				
Percent time looking (%)	$F(1,28)=14.06$ $p=.0008$	$F(1,28)=1.97$ $p=.2$	$F(1,28)=1.41$ $p=.2$	0.44, -0.18, -0.10, 0.09				
Rate of long glances (/min)	$\chi^2(1)=10.59$ $p=.001$	$\chi^2(1)=8.41$ $p=.004$	$\chi^2(1)=0.66$ $p=.4$	0.92, 0.82 -1.24, 0.54				
Workload and Perceived Risk					**Secondary Task**	**Phase* Secondary Task**	**Experience* Secondary Task**	**Model Coefficients $\beta_4, \beta_5, \beta_6$**
NASA TLX	$F(1,57)=2.41$ $p=.13$	$F(1,57)=0.14$ $p=.71$	$F(1,57)=2.48$ $p=.12$	5.95, 4.85, 1.40, -2.96	$F(1,57)=8.29$ $p=.006$	$F(1,57)=4.15$ $p=.046$	$F(1,57)=0.21$ $p=.65$	-1.22, -3.83, 0.86
Perceived risk	$F(1,57)=3.83$ $p=.06$	$F(1,57)=0.14$ $p=0.71$	$F(1,57)=1.40$ $p=.24$	3.79, 1.84, 0.37, -0.95	$F(1,57)=23.8$ $p<.0001$	$F(1,57)=2.00$ $p=.16$	$F(1,57)=0.01$ $p=.92$	-1.35, -1.14, -0.08