Accident Analysis and Prevention 128 (2019) 132–138

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Evaluate driver response to active warning system in level-2 automated vehicles

Jonathan R. Atwood[a,b], Feng Guo[a,b,*], Myra Blanco[c]

[a] Virginia Tech Transportation Institute, 3500 Transportation Research Plaza (0536), Blacksburg, 24061, VA, USA
[b] Department of Statistics, Virginia Tech Virginia Tech Transportation Institute, 406-A Hutcheson Hall (MC 0439), Blacksburg, 24061, VA, USA
[c] Center for Public Policy, Partnerships, & Outreach Virginia Tech Transportation Institute, 3500 Transportation Research Plaza (0536), Blacksburg, 24061, VA, USA

A R T I C L E I N F O

Keywords:
Level-2 automated vehicle
Driver response
Active safety warning
Driver adaptation
Traffic safety

A B S T R A C T

As vehicles with automated functions become more prevalent on U.S. roadways, maintaining driver attention while the vehicle is engaged in automation will be an important consideration for safe operation of these vehicles. The objective of this paper is to evaluate how drivers respond and adapt to active safety warning signals in a Level 2 automatic vehicle. Specifically, statistical analysis was conducted to evaluate whether the amount of inattention prompts that drivers received changed over time, possibly indicating a change in the amount of inattention that drivers exhibited. The driving performance data was collected from sixteen participants who drove a Level 2 vehicle in an experimental setting, as part of the study Human Factors Evaluation of Level 2 and Level 3 Driving Concepts. A proprietary driver inattention warning system was installed on the experiment vehicles. The system would send a warning signal if the driver's attention was not on the primary driving task for a pre-specified duration. This study focuses on driver's response when experiencing prompts after two seconds of inattention while operating a Level 2 vehicle in automated mode. The results show that on average, the frequency of prompts the participants received decreased over the course of the experiment from 29.9 in the first ten minutes to 18.1/10 min after 110 min. The decrease levelled off after about two hours. The fact that participants received fewer prompts over time suggests that they had fewer instances of inattention lasting at least two seconds as the experiment progressed. This suggests that drivers would adapt to the alert and adjust their behavior to avoid triggering the inattention alert. The results of this study provide evidence for a potential benefit of incorporating a prompting system in vehicles with automated functions.

1. Introduction

The introduction of automated vehicle technology into public roadway system creates a need to understand how drivers will interact with and adapt to vehicle systems that allow ceding some level of control to the vehicle, yet still require driver involvement. According to the National Highway Traffic Safety Administration (NHTSA), Level 2 (L2) automated vehicles, many of which are available now, are designed to ease the driver's burden while on the road. However, drivers must still monitor the road and keep alert in order to retake control of the vehicle at a moment's notice (Marinik et al., 2014). This is reflected in NHTSA's (2013) description of L2 vehicle capabilities, which states, "The system can relinquish control with no advanced warning and the driver must be ready to control the vehicle safely" (p. 5).

Recent automated technological advances that have facilitated the relinquishing of some of the driving tasks from the driver to the car may include such features as, for example, adaptive cruise control (ACC), ACC lane-centering, automatic braking, collision warnings, and lane-keeping assistance (Trimble et al., 2014). With the possibility of different features being incorporated into vehicle automation in mind, NHTSA has defined five levels of automation, from Level 0, where the driver has complete control, to Level 4, where the driver's role will be limited to providing destination/navigation information (Trimble et al., 2014). One of the main goals of L2 vehicle designs is to relieve drivers' physical stress by allowing them to simultaneously disengage from the pedal and steering wheel.

While potentially capable of increasing drivers' comfort, vehicles with L2 automation systems still require constant attention from drivers (Trimble et al., 2014). As such, drivers need to maintain situational awareness even when L2 vehicle features are active. Thus, one of the key questions regarding increasingly automated vehicles is how, if at all, these vehicles affect a driver's situational awareness, and what

* Corresponding author.
E-mail addresses: jatwood@vtti.vt.edu (J.R. Atwood), feng.guo@vt.edu (F. Guo), mblanco@vtti.vt.edu (M. Blanco).

https://doi.org/10.1016/j.aap.2019.03.010
Received 13 July 2018; Received in revised form 26 February 2019; Accepted 19 March 2019
Available online 18 April 2019

effect that change in situational awareness has on driver performance. Endsley (1988) originally defined situational awareness as “the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” (p. 97). This definition contains three subconcepts that are hierarchical in nature. The first is “perception” of the events, in which individuals are able to accurately recognize the basic structures of the situation. The second is “comprehension,” in which individuals use the knowledge gained from “perception” to form an understanding of the overall current significance of the events and their relevance to the individuals’ goals. Finally, the third is “projection,” in which individuals use knowledge gained from “perception” and “comprehension” to determine how the current events will shape the near future (Endsley, 1995). Though this conception of situational awareness was originally applied to aircraft, Walker et al. (2008) applied it to driving by defining it as “knowing about the vehicle’s current position in relation to its destination, the relative positions and behavior of other vehicles and hazards, and also knowing how these critical variables are likely to change in the near future” (p. 283).

Loss of situational awareness on the road can lead to insufficient driver performance, which in turn can lead to a greater potential for collisions. With regard to L2 vehicles’ effect on situational awareness, Shen and Neyens (2014) found in an experimental setting that response times to automation failures were significantly longer in vehicles with lane-keeping assistance and ACC than in vehicles with only lane-keeping assistance. In a study also conducted in an experimental setting, Merat and Jamson (2009) found that drivers of vehicles that were in automated mode had decreased anticipation of lead vehicle deceleration compared to those driving vehicles that were in manual mode. Additionally, Tsai et al. (2007) found that participants performing an auditory task while driving reduced their scanning range, rearview mirror, and odometer glances.

Driver distraction may also cause loss of situational awareness. With the increasing use of L2 automated vehicles, effort should be made to ensure that increased instances of distraction are minimized, especially if drivers are performing a non-driving task while in automated mode. Driver distraction has been defined in a few different ways across the literature (Regan et al., 2011). Ranney et al. (2000) defined distraction as “any activity that diverts driver’s attention away from the task of driving” (p. 1). NHTSA further identifies and defines three different subtypes of driver distraction (2010). Firstly, there is visual distraction, or "tasks that require the driver to look away from the roadway to visually obtain information.” Secondly, there is manual distraction, or “tasks that require the driver to take a hand off the steering wheel and manipulate a device.” Finally, there is cognitive distraction, or “tasks that require the driver to avert their mental attention away from the driving task.” Note that these three distraction types are not mutually exclusive. For example, typing a text message requires the driver to look away from the road, use one or two hands to type the message, and focus their mental energy on the contents of the message instead of the road.

One direct effect of driver distraction, and in particular visual distraction, is an increase in the amount of time that drivers spend looking away from the forward roadway. For example, Olson et al. (2009) found that, on average, texting was associated with an average of 4.6 s of eyes off the roadway in a 6-second interval. Other tasks that involved extended time of eyes off the forward roadway were calculator use (4.4 s), writing (4.2 s), reading (4.3 seconds), dialing a cell phone (3.8 s), and reaching for an object (2.9 s). The increased eye glance time off the forward roadway may in turn increase the risk of crashes, particularly when the time of a glance off the road is greater than 2 s in short intervals, as previous research contends. Olson et al. (2009) found that driving scenarios with instances of eyes off the forward roadway for at least 2 s in a 6-second interval were 2.93 times more likely to result in a safety critical event than those with glances less than 2 s off the forward roadway. Guo et al. (2017) showed that the prevalence and impact of distraction varies substantially by age. Klauer et al. (2006) found that driving instances with total eye glance time off the roadway of at least 2 s in a 6-second interval resulted in an odds ratio of 2.19 of experiencing a crash or near-crash, compared to baseline instances of driving.

Previous research has also uncovered deleterious effects of various types of driver distraction on the risk of crashes while driving. NHTSA (2016) reported that around 16% of police-reported crashes, 10% of fatal crashes, and 18% of injury crashes were reported as related to driver distraction in 2014. Dingus et al. (2016), using the Second Strategic Highway Research Program Naturalistic Driving Study (SHRP 2 NDS), which tracked over 3500 participants for over 34 million miles of driving in naturalistic settings, estimated that 68.3% of injury or property damage crashes involved some form of distraction. Guo (2019) provides a detailed review on the methodologies for analyzing driver behavior and driving risk using NDS data.

Because of the potential loss of situational awareness associated with driver distraction and the performance of non-driving tasks (and the subsequent potential risks), it is important to understand whether operating vehicles in L2 mode will potentially increase the frequency with which drivers perform non-driving tasks. Recent research suggests that when allowed to operate vehicles with L2 levels of automation (i.e. ACC and lane centering simultaneously), drivers are more likely to engage in non-driving tasks (especially risky ones), and make longer off-road glances. In an experimental setting, Llaneras et al. (2013) found that more participants performed 13 out of 15 non-driving tasks during a driving session with ACC and lane-centering active then they did during a driving session with only ACC active. Further, many of the tasks that showed the greatest increase in percentage of participants performing them during ACC and lane-centering active have been identified as likely to increase crash risk (Dingus et al., 2016; Klauer et al., 2014; Olson et al., 2009), including texting and e-mailing (42% participant increase), reading (25% participant increase), other phone interactions (34% participant increase), and cell phone answering (25% participant increase). Finally, when participants were operating with ACC and lane centering active, they made 27% more off-road glances of greater than 2 s than they did when under ACC driving conditions alone. Note that during this experiment, participants were not provided with alerts or prompting, and were free to perform non-driving tasks as they chose.

Although there are concerns with the amount of attention drivers may pay to the roadway with advanced automated features in use, there also is evidence that direction from the L2 vehicle, and the manner in which that direction is presented, may influence the ability of drivers to recover from a state in which they are not directly engaged in the driving task. For example, Blanco et al. (2015) found in an experimental setting that, when engaged in secondary tasks while the vehicle was in automated mode (with ACC and lane-centering active), 98% of participants regained control of the vehicle after an experimentally-imposed lane drift before the vehicle completely exited the lane if the lane drift was accompanied by a visual and haptic alert, whereas only 53% did so if no alert was present. Additionally, in a separate experimental setting with different participants within Blanco et al. (2015), participants were found to react to an alert while engaged in non-driving tasks when L2 automation was active about 4.6 times faster with a visual/haptic alert (mean = 0.66 s) than they did with only a visual alert (mean = 3.04 s). Participants also regained control significantly faster after a visual/haptic alert than they did after a visual alert only; the difference was also affected by whether the alert was imminent (a flashing red light with “take steering” instruction), cautionary (a yellow light with “take steering” instruction), or staged (a cautionary alert that shifted to an imminent alert after 10 s). These results suggest that manufacturers of L2 vehicles can potentially affect and improve driver performance in L2 vehicles if the automation is turned on and provides suitable direction when participants are engaged in secondary tasks.

This study focuses on the effect of an active warning system on driver behavior in an L2 automated vehicle in an experimental setting. Specifically, we investigated at least 2 s of inattention that drivers experienced while performing secondary tasks during 10-minute intervals of operating an L2 vehicle, and whether or not the number of instances of inattention changed over time.

2. Methods

A single, long exposure experiment was conducted in which participants operated an L2 vehicle on a test track with simulated highway driving conditions for three 60-minute sessions. The vehicles were operated with adaptive cruise control (ACC) and lane centering active. Participants were divided into three groups: those who received prompts from a warning system after 2 s of inattention (the 2-second group), those who received prompts after 7 s of inattention (the 7-second group), and those who never received prompts (the no-prompt group). For the purposes of this study, inattention refers to attention off-road. The prompting system alerted participants when their attention was off road for 2 s, 7 s, or did not alert at all, depending on the group.

2.1. Experimental setup

This study uses data from an experiment performed as part of the NHTSA sponsored, “Human Factors Evaluation of Level 2 and Level 3 Driving Concepts” (Blanco et al., 2015), conducted by the Virginia Tech Transportation Institute (VTTI). The purpose of that study was to investigate the ways in which drivers interact with and respond to L2 and L3 automated vehicles. This paper used the data from one of the experiment (Experiment 2), in which the objective was to measure how drivers in an L2 vehicle in mixed-traffic conditions responded to attention-state prompts, and how this response changed over time (if at all). Participants received the prompts after a certain amount of time of inattention. This paper focuses on the frequency of prompts as a surrogate for the frequency of instances of inattention.

The experiment was conducted at General Motor’s Milford Proving Ground circle track in Milford, Michigan. The experiment was conducted in Lane 2, which allowed speeds between 48 and 80 kph (30 and 50 mph). The circle track has been used in previous studies on automated driving to simulate freeway conditions.

A 2010 model year Cadillac SRX equipped with a prototype L2 automated driving system was used as the experimental vehicle. As part of the automated driving system, several human machine interface (HMI) components were installed. These included an instrument panel binnacle-mounted screen providing information on the automated driving system, and two steering wheel buttons to control the automation: one ACC button and one button for the lane-centering system.

Because non-driving tasks in particular are likely to lead to extreme instances of inattention, participants were given tasks of this nature to perform during the experiment. These tasks included navigation, email, and web browsing. A total of up to 90 different non-driving tasks, 30 in each category, were completed. These tasks were similar in terms of the visual/manual demand required and they were presented in a random order. All participants received all different types of tasks at one point or another so that one type of task would not cluster together in time. All tasks were performed using a tablet computer, which was equipped with a standard QWERTY touchscreen keyboard. Each task was typed on a notecard. The pace of these tasks was not forced but at the participants’ leisure. Once a task was completed, the experimenter would provide a new task 30–60 s later. Each task was repeated the same number of times and the tasks were presented in a random order. Details of each general type of task are given below.

- Navigation. The navigation tasks required the participant to open an application on the device, choose the new destination option, and enter the address provided on the notecard in a printed number-street-city-state format.
- Email. The email tasks required the participant to compose an email using the tablet.
- Web browsing. The web-browsing tasks required the participant to open the web browser on the tablet. Participants were asked to determine the answer to a specific question or task that required searching or interacting with the internet.

Prior to proceeding to the test vehicle, participants were provided with the tablet. They were given a brief introduction to the types of tasks they would be performing, and were allowed time for any questions and to practice one of each of the different types of tasks, if needed. Participants were then provided with a static orientation to the experimental vehicle, which included basic controls and the L2 automation features. Following this, participants received an on-track orientation consisting of four laps on the test track. During this training, participants were instructed on the tasks they needed to perform but were not given an explanation of the levels of automation.

The purpose of this on-track orientation was to allow participants to acclimate to the vehicle and the test-track environment. Participants were asked to drive to the circle track entrance, enter the second travel lane, and maintain a speed of 50 mph (80 kph). The first lap was completed under manual driving. During the second lap, ACC was activated, which allowed participants to release longitudinal control to the automated system while they maintained lateral control of the vehicle. The third and fourth laps were completed using the L2 automation (ACC plus lane centering), which allowed participants to release both longitudinal and lateral control of the vehicle. Using a proprietary algorithm detection method, the system tracked and alerted participants when they had reached a certain amount of time of inattention (after 2 s or 7 s, depending on the group participants were assigned to, as described below). For this paper, only valid prompts (as defined by the researchers) were included in the analysis. After completing the on-track orientation, participants were provided with additional instructions pertaining to the experimental session and were given the opportunity to ask any additional questions. Experimenters followed a set script and protocol in order to ensure consistency between experimenters.

2.2. Participants

Participants in Experiment 2 were divided into three groups: those who received prompts after 2 s of inattention, those who received prompts after 7 s of inattention, and those who did not receive prompts at all. There were 16 participants in each of the three groups. However, because this study used prompts to indicate instances of inattention longer than 2 s, only those experiment participants who received prompts were used. Additionally, since Klauer et al. (2006, 2010) identified 2 s as a potential point at which inattention can lead to safety critical event risk, it was of interest in this study to identify how often 2 s instances of inattention occur. Therefore, the sample used in this study was further reduced to only the 2 s prompt group. In this group, the prompts began after 2 s of inattention, when a flashing yellow light would occur. If the participant did not react within 5 s, the light would change to red and be accompanied by both a haptic and an audio alert for another 5 s. Of the 16 participants included in this study, there were nine females (mean age = 37.8 years, S.D. = 14.8 years) and seven males (mean age = 40.6 years, S.D. = 17.5 years).

2.3. Experimental time intervals

The experiment was conducted in three 60-minute driving sessions with a 15-minute rest period between sessions, for a total of 180 min driving time. Participants activated L2 automation (lane-centering and ACC), and completed laps around the test track.

After every 10 min of the driving, participants (while continuing to

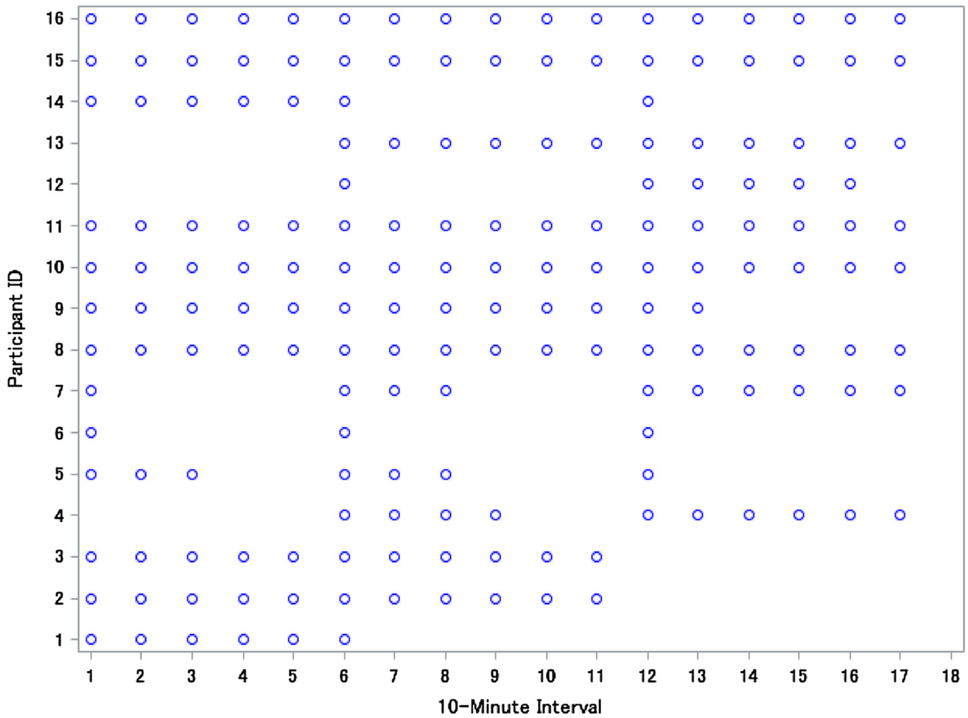

Fig. 1. Plot of Participants' Data Availability throughout the Experiment.

drive) rated their level of trust in the system according to a 7-point Likert-type scale (strongly disagree to strongly agree that they trusted the system). Because of this intervention, the total number of prompts was counted within these 10-minute intervals.

There were a total of 18 10-minute intervals in the experiment for each participant. In some time intervals, the data are not valid because of data recording issues. In these instances, data was excluded from the subsequent analysis. Fig. 1 displays the intervals with used data. A blue circle indicates that data are available for the participant at the specific time interval.

Because the prompts were designed to go off after 2 s of inattention, the total number of prompts can be used as a pseudo variable for the total number of instances of inattention that lasted at least 2 s. Naturally, it is desirable that the total amount of these instances be as low as possible. To ask whether or not drivers adapt to this system in a beneficial way, therefore, is to ask whether or not the number of these instances decreases as participants have more experience with the system. This question, in principle, can be addressed within the current study by modeling the total count of prompts as a function of time. If there is a significant decreasing trend in total prompts over time, for example, then this may provide evidence that drivers will adapt favorably to the prompting system by decreasing the number of instances of extreme inattention. This study applied this approach to the data described above to determine if there is evidence of such a trend.

In the event of a significant decrease in prompts over time, the question arises as to whether the decrease might be due to an overall decrease in inattention, or an increase in the length of time that prompts last due to participants failing to respond to them. To address this question, the analysis of prompt frequency was followed up with an analysis on the amount of time (in seconds) that participants took to react to the prompt. This analysis was also intended to detect any significant time trend. Because determining prompt reaction time required data reduction, one prompt per 10-minute interval per participant was sampled for this analysis. "A stratified random sampling method was used to sample one prompt randomly within each 10-minute interval. Reductionists examined video of the sampled prompts to determine how quickly the participants reacted after the first onset of the prompt. Reaction time was defined as the total time (in seconds) from the onset of the prompt to the first action that could be considered a reaction (i.e. looking forward, reaching for the steering wheel, etc.).

3. Models

3.1. Prompt frequency

A generalized linear mixed model for count data was used to model the change in the total number of instances of inattention per 10-minute interval in the study. A negative-binomial regression model with random driver terms was used due to evidence of overdispersion, a phenomenon where the observed variability in the data is greater than that which is predicted by the standard Poisson model for count data, which assumes that the mean is equal to the variance. As described in Stroup (2012), the negative binomial regression model assumes that the marginal distribution of the dependent variable (in this case, the total prompt count) follows a negative binomial, with an expected value

$$E(Y) = \mu$$

The negative binomial regression model, then, models the mean, μ, as a function of the covariates. The model used involves terms that allow both for a possible increase/decrease in the amount of prompts, and curvature to evaluate the temporal profile of increase/decrease in prompts. Specifically, the linear predictor for this model is

$$\log(\mu_{ij}) = \beta_0 + \beta_1 * T_i + \beta_2 * T_i^2 + \beta_3 * Gender_j + \beta_4 * Age_j + \alpha_j \tag{1}$$

where $\log(\mu_{ij})$ is the natural log of the expected total count of prompts (and therefore, instances of inattention lasting 2 s) at the i^{th} time interval for the j^{th} participant; β_0 is an intercept term, representing the first time interval; β_1 is regression coefficient for time trend; T_i is a variable for the i^{th} time interval (0:17); β_2 evaluates if there is the time trend is quadratic; β_3 and β_4 are coefficients for gender and age variables, which were included as covariates to adjust for possible effects of these variables; and α_j is a random effect term that incorporates the correlations among observations from the same participant, and we assume

$\alpha_j \sim N(0, \sigma^2)$.

Of particular interest from this model are the coefficients β_1 and β_2, which jointly provide the time-interval component of the model.

Note that this is conditional on the participant-specific intercept, or the amount of prompts participants received in the first time interval, as well as the individual specific covariate levels (age and gender).

3.2. Reaction time to prompts

The model for the reaction time to the prompt is a linear mixed effect regression model with the reaction times were log-transformed. The model is expressed as

$$\log(RT_{ij}) = \beta_0 + \beta_1 * T_i + \beta_2 * T_i^2 + \beta_3 * Gender_j + \beta_4 * Age_j + \alpha_j + \varepsilon_{ij}$$

where RT_{ij} is the reaction time to the prompt for the ith participant in the jth time interval,

$\varepsilon_{ij} \sim N(0, \sigma_1^2)$

and everything else is as before.

4. Results

4.1. Prompt frequency over Time

The mean amount of prompts during the first 10-minute time interval was 29.9 prompts (S.D. = 21.2,). The mean amount of prompts got down to 18.1 prompts during the 11th time interval (S.D. = 11.4)). Boxplots of prompt frequency in each time interval are displayed in Fig. 2. The box plot provides the mean, median, 25th and 75th percentile. The whiskers lines represent the maximum/minimum or 1.5 time the Interquartile range in the presence of outliers. We added a line linking the means of each time interval. As can be seen from the plot, there is a general trend of decreasing number of prompts over time. The frequency reached the lowest point about time interval #10 and rebounded. The frequency never go back to the initial period though.

Model estimates are displayed in Table 1. Note that these are estimates of the parameters of the linear predictor on the log count scale and should be interpreted accordingly. The output indicates a statistically significant change over time. A decreasing linear trend ("Linear Time Trend" = -0.0713) and positive quadratic trend ("Quatritic Time Trend" = 0.003). This indicates that, in general, the amount of prompts decreased with overtime and the rate of decrease is not linear. The most distinct pattern is the quick decrease at the initial phase. The rate of decrease definitely slow down after time interval 5 and level off after time interval 10. The difference after time interval 10 is relatively small. This suggests that participants would adjust their behavior to adapt to the system. This adaptive effect is more pronounced at the initial phase of exposure to the system."

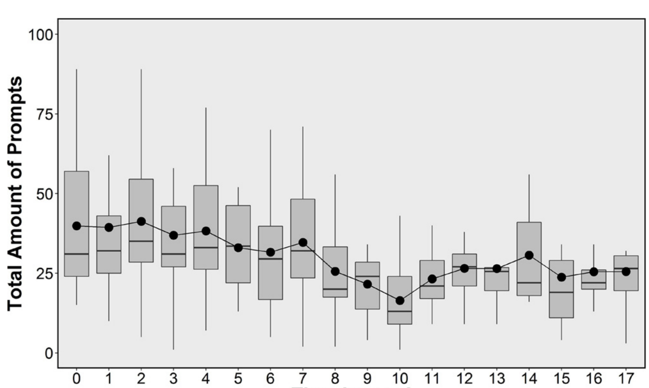

Fig. 2. Boxplots of prompt frequency by time interval.

In the above model, the respective linear and quadratic time trend coefficients of -0.07 and 0.003 taken together indicate evidence of a decreasing trend in the number of prompts that levels out.

This model can be back-transformed to represent a mean change in prompts over time in the experiment. This equation can be represented in the following formula:

$$e^{3.99 - .07 * T_i + .003 * T_i^2 + \Omega_j}$$

where $\Omega_j = (\hat{\beta}_3 * Gender_j + \hat{\beta}_4 * Age_j + \tilde{\alpha}_j)$ is the subject-specific term, including age and gender effects, and predicted random effect $\tilde{\alpha}_j$. It is important to understand that the above equation is a conditional mean of the change over time, given a participant's individual-specific starting point, which is represented by the random intercept term, and individual-specific age and gender.

In general, the results derived from generalized linear models are subject to strong influence from extreme values, in particular when the number of subjects is small. As such, a sensitivity analysis was performed where the same model was run with each participant removed, in turn. The purpose of this analysis was to investigate whether the overall results are sensitive to the removal of any one participant. If the overall trend changes dramatically with the removal of one participant, it may suggest that the trend found in the results is driven by that one participant. Fig. 3 plots each mean equation of mean prompt over time with one participant removed (gray lines), along with the equation with all participants (the bold dash line). While results show certain degree of variation, the overall pattern is consistent across all fitted curves i.e. rapid decrease in the initial phase and level off at later stage.

4.2. Significant change of prompt frequency over time

One key question of interest is which time periods contain a statistically significant change in the number of prompts. Such change will provide information on whether and when drivers would adapt to the prompt system. For this purpose, we use the estimated derivative of the function for the linear predictor with respect to the time interval in Eq. (1). The estimated derivative with respect to time at the i^{th} time interval is expressed as

$$\hat{D}_i = \frac{\partial \log(\mu_{ij})}{\partial (T_i)} = \hat{\beta}_1 + 2 * \hat{\beta}_2 * T_i.$$

The $\hat{D}_i$ is the estimated derivative of the logarithm of the expected number of prompts with respect to time. The number of prompts is not change around time T_i if $\hat{D}_i$ equals zero. Therefore, $\hat{D}_i$ can be used to test whether there is a significant change in the frequency of prompt at T_i. Note that terms in Eq. (1) that are unrelated to the time interval equal to zero when taken derivative with respect to T_i. Therefore, the estimated change in the linear predictor as a function of the time interval is the same for all participants within all age groups and genders.

Inference on the derivative requires computing the 95th percent confidence interval for the estimated derivative. This requires calculating the standard error for the derivative at each time interval, and the standard error in turn requires the estimated variance. The estimated variance of the derivative at any time interval T_{ij} is calculated as

$$var(\hat{D}_i) = \hat{\sigma}^2_{\hat{\beta}_1} + 2 * T_i * \hat{\sigma}_{\hat{\beta}_1, \hat{\beta}_2} + 4 * T_i^2 * \hat{\boldsymbol{\sigma}}^2_{\hat{\boldsymbol{\beta}}_2}{}^2$$

where $\hat{\sigma}^2_{\hat{\beta}_1}$ and $\hat{\sigma}^2_{\hat{\beta}_2}$ are the estimated variance of $\hat{\beta}_1$ and $\hat{\beta}_2$ respectively, and $\hat{\sigma}_{\hat{\beta}_1, \hat{\beta}_2}$ is the estimated covariance of $\hat{\beta}_1$ and $\hat{\beta}_2$.

The standard error SE_i for the derivative is calculated by taking the square root of the variance $var(\hat{D}_i)$. The confidence interval is calculated by implementing the following formula

$$CI = (\hat{D}_i - 1.96 * SE_i,\ \hat{D}_i + 1.96 * SE_i)$$

If the confidence interval for the derivative at a time interval does not contain 0, then the number of prompts is considered to be changing

Table 1
Parameter Estimates for Statistical Model for Prompt Frequency.

Effect	Estimate	S.E.	T-Value	P-Value	Lower Confidence Limit	Upper Confidence Limit
Intercept	3.987	0.190	20.95	< .0001	3.568	4.406
Linear Time Trend	**−0.0713**	**0.024**	**−3.03**	**0.0028**	**−0.118**	**−0.025**
Quadratic Time Trend	**0.003**	**0.001**	**2.03**	**0.0442**	**0.0001**	**0.0055**
Age: 18-24	0.131	0.221	0.59	0.5548	−0.305	0.566
Age: 25-39	−0.605	0.224	−2.70	0.0076	−1.048	−0.163
Age: 40-54	0.014	0.223	0.06	0.9508	−0.426	0.454
Age: 55+	0	.	.	.	.	.
Gender: Female	−0.337	0.172	−1.96	0.0515	−0.676	0.002
Gender: Male	0	.	.	.	.	.

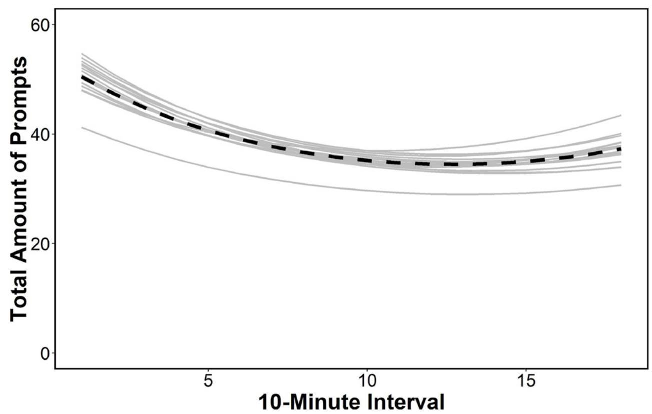

Fig. 3. Predicted conditional mean of prompt change over time, black bold dash line includes all participants, each grey line represents the results after removing one participant.

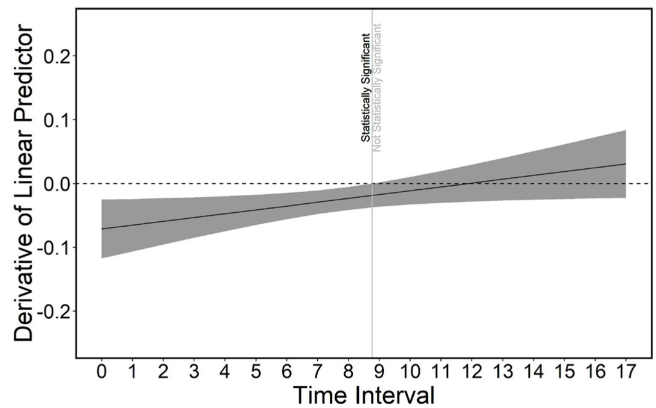

Fig. 4. Estimated Derivative of Linear Predictor with Respect to Time Interval, with Pointwise 95 Percent Confidence Interval.

at a statistically significant level at that time interval.

Using the estimated coefficients given in Table 1, the estimated derivative of the linear predictor with respect to time interval is

$$\hat{D}_i = -0.0713 + 2 * T_i * 0.003$$

The variances/covariance of $\hat{\beta}_1$ and $\hat{\beta}_2$ are provided in Table 2.

This leads to the variance estimate at a particular time interval as

$$var(\hat{D}_i) = 5.53*10^{-4} + 4*T_i*(-3.0*10^{-5}) + 4*T_{ij}{}^{2}*1.92*10^{-6}$$

Fig. 4 shows the estimated derivative and 95 percent confidence interval. Prior to the 9th time interval, the confidence interval is completely below 0, indicating a significant decrease. After that, the confidence interval contains 0, indicating no statistically significant change. Solving numerically, the time point at which the derivative loses its statistical significance is around 8.8. The interpretation is that before 8.8, the number of prompts is decreasing at a statistically significant level, and after 8.8, the number of prompts is not changing at a statistically significant level.

4.3. Reaction time to prompts

The mean time to react was 2.7 s (S.D. = 2.78 s). The model results were not statistically significant for change in the log of time to react over the course of the study. Model estimates are displayed in Table 3.

Table 2
Variance Estimates for Time Interval Coefficients.

Variance Parameter	Estimate
$\hat{\sigma}^2_{\hat{\beta}_1}$	$5.53*10^{-4}$
$\hat{\sigma}^2_{\hat{\beta}_2}$	$1.92*10^{-6}$
$\hat{\sigma}_{\hat{\beta}_1,\hat{\beta}_2}$	$-3.0*10^{-5}$

5. Summary and conclusion

The study performed sought to determine if there was a change in the number of instances of driver inattention/distraction over time when drivers were operating an L2 automated vehicle in an experimental setting. In this study, the number of prompts that participants received in a previous NHTSA study was used as a surrogate for inattention, as the system used for the experiment from which the study drew data was designed to provide a prompt after 2 s of inattention. The results found that, when the total number of prompts was examined after 10-minute intervals, the number of prompts the participants received followed a temporarily decreasing trend over time and level off at some point. Additionally, the study found no significant increase in the time to react to the prompts over the course of the experiment, suggesting that the decrease in prompt frequency may not be due to a general decease in the participants' ability to respond to the prompts. The small sample size of participants (n = 16) raises the question of whether or not the lack of a significant trend in reaction time is due to a lack of data to detect a significant trend. Future studies may need larger sample sizes to rule out this possibility.

The results of this study may suggest that drivers in an L2 vehicle who are prompted to return focus to the road after instances of at least 2 s of inattention may experience an overall change in driving behavior as their experience with such a system increases. Specifically, the results suggest that drivers will adapt to the warning system of L2 vehicles and alter their behavior to avoid the warning trigger. This suggests a potential benefit in a prompting system for L2 vehicles in helping to reduce the extended roadway inattention that could potentially accompany the increase in usage of L2 vehicles. This research cannot conclusively rule out other explanations of decreased inattention, such as the possibility drivers may have decreased inattention because of increased familiarity to the non-driving tasks over time (thus plausibly

Table 3
Parameter Estimates for Statistical Model for Time to React to Prompt.

Effect	Estimate	Standard Error	T-Value	P-Value	Lower Confidence Limit	Upper Confidence Limit
Intercept	0.1881	0.4640	0.41	0.6899	−0.7863	1.1625
Linear Time Trend	0.008221	0.05048	0.16	0.8708	−0.09123	0.1077
Quadratic Time Trend	0.000386	0.002646	0.15	0.8841	−0.00483	0.005598
Age: 18-24	−0.2312	0.5314	−0.43	0.6711	−1.3859	0.9236
Age: 25-39	0.1617	0.5104	0.32	0.7566	−0.9464	1.2699
Age: 40-54	−0.2581	0.5303	−0.49	0.6351	−1.4113	0.8951
Age: 55+	0	.	.	.	.	.
Gender: Female	0.2616	0.4030	0.65	0.5285	−0.6168	1.1400
Gender: Male	0	.	.	.	.	

lowering time to complete the tasks). However, the results do present evidence that is consistent with behavioral adaptation, warranting further investigation into systems like these for behavioral modification.

There were several limitations to this study. Because the participants were given secondary tasks to perform, they were not in a position to operate the L2 vehicle as they might naturally, making it difficult to generalize the results. The results should be interpreted within the context of drivers potentially performing non-driving tasks while operating increasingly-automated vehicles. A second limitation is that this study does not account for instances of inattention, if any, that the prompt system did not pick up. It is at this point unknown how many such instances, if any, occurred, nor if there was any time trend in these instances that skewed the results. Finally, comparison to baseline (participants without prompts) is theoretically impractical. The response variable in this study was instances of inattention lasting at least 2 s. Without a prompting system to interrupt these instances, drivers may have fewer instances of inattention that nevertheless last longer, thereby resulting in more total inattentive time.

This study provides evidence that over time, a prompting system within an L2 automated vehicle can help in reducing the number of instances of driver distraction/inattention that last at least 2 s in cases where drivers are engaged in non-driving tasks. Although there may be more instances of these non-driving tasks when drivers are operating vehicles with increased automation of driving tasks, prompting systems may help restore attention to the driving environment before there is a significant increase in the risk of a serious event.

Acknowledgements

This analysis was performed by the Virginia Tech Transportation Institute using data obtained from a research study funded by the National Highway Traffic Safety Administration (NHTSA). Details about the original study, data collection process, and other analyses related to the same dataset could be found in Blanco et al. (2015). This document represents the opinion of the authors and not of NHTSA or other authors in the original study report.

References

Blanco, M., Atwood, J., Vasquez, H.M., Trimble, T.E., Fitchett, V.L., Radlbeck, J., Morgan, J.F., 2015. Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts. (Report No. DOT HS 812 182). August. National Highway Traffic Safety Administration, Washington, DC.
Dingus, T.A., Guo, F., Lee, S., Antin, J.F., Perez, M., Buchanan-King, M., Hankey, J., 2016. Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proceedings of the National Academy of Sciences of the United States of America 113 (10), 2636–2641. https://doi.org/10.1073/pnas.1513271113.
Endsley, M.R., 1988. Design and evaluation for situation awareness enhancement. Proceedings of the Human Factors Society 32 Annual Meeting 97–101. https://doi.org/10.1177/154193128803200221.
Endsley, M.R., 1995. Toward a theory of situation awareness in dynamic systems. Human Factors 37 (1), 32–64. https://doi.org/10.1518/001872095779049543.
Guo, F., 2019. Statistical methods for naturalistic driving studies. Annual Review of Statistics and Its Application. In press.
Guo, F., Klauer, S.G., Fang, Y., Hankey, J.M., Antin, J.F., Perez, M.A., Lee, S.E., Dingus, T.A., 2017. The effects of age on crash risk associated with driver distraction. Int. J. Epidemiol. no. 46 (1), 258–265.
Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., 2006. "The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data" Report Number DOT HS 810 594. Retrieved from. National Highway Traffic Safety Administration. http://www.nhtsa.gov/Research/Human+Factors/Naturalistic+driving+studies:.
Klauer, S.G., Guo, F., Sudweeks, J., Dingus, T.A., 2010. An Analysis of Driver Inattention Using a Case-Crossover Approach on 100-Car Data. Retrieved from. http://www.nhtsa.gov/Research/Human+Factors/Naturalistic+driving+studies:.
Klauer, S.G., Guo, F., Simons-Morton, B.G., Ouimet, M.C., Lee, S.E., Dingus, T.A., 2014. Distracted driving and risk of road crashes among novice and experienced drivers. N. Engl. J. Med. 370, 54–59. https://doi.org/10.1056/NEJMsa1204142.
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. Proceedings of the International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design 7, 92–98.
Marinik, A., Bishop, R., Fitchett, V., Morgan, J.F., Trimble, T.E., Blanco, M., 2014. Human Factors Evaluation of Level 2 And Level 3 Automated Driving Concepts of Operation. National Highway Traffic Safety Administration (NHTSA). Department of Transportation (DOT).
Merat, N., Jamson, A.H., 2009. How do drivers behave in a highly automated car. Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design. pp. 514–521.
National Highway Traffic Safety Administration, 2010. Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. Washington, DC: National Highway Traffic Safety Administration (NHTSA). Retrieved from. Department of Transportation (DOT). http://www.distraction.gov/downloads/pdfs/visual-manual-nhtsa-driver-distraction-guidelines-for-in-vehicle-electronic-devices.pdf.
National Highway Traffic Safety Administration, 2013. Preliminary Statement of Policy Concerning Automated Vehicles. Retrieved from. National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT), Washington, DC. http://www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated_Vehicles_Policy.pdf.
National Highway Traffic Safety Administration, 2016. Traffic Safety Facts: Research Note. Distracted Driving 2014. Report Number DOT HS 812 260. Retrieved from, 2019 Retrieved from. National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT), Washington, DC. http://www-nrd.nhtsa.dot.gov/Pubs/812260.pdf.
Olson, R.L., Hanowski, R.J., Hickman, J.S., Bocanegra, J., 2009. Driver Distraction in Commercial Vehicle Operations (Document No. FMCSA-RRT-09-042). Federal Motor Carrier Safety Administration, USDOT, Washington, DC.
Ranney, T.A., Mazzae, M., Garrott, R., Goodman, M.J., 2000. NHTSA driver distraction research: past, present, and future. Driver Distraction Internet Forum (Vol. 2000). Retrieved from. http://www-nrd.nhtsa.dot.gov/departments/Human%20Factors/driver-distraction/PDF/233.pdf.
Regan, M.A., Hallett, C., Gordon, C.P., 2011. Driver distraction and driver inattention: definition, relationship and taxonomy. Accid. Anal. Prev. 43 (5), 1771–1781. https://doi.org/10.1016/j.aap.2011.04.008.
Shen, S., Neyens, D.M., 2014. Assessing drivers' performance when automated driver support systems fail with different levels of automation. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 58 (1), 2068–2072. https://doi.org/10.1177/1541931214581435. SAGE Publications.
Stroup, W.W., 2012. Generalized Linear Mixed Models: Modern Concepts, Methods and Applications. CRC Press.
Trimble, T.E., Bishop, R., Morgan, J.F., Blanco, M., 2014.). Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts: Past Research, State of Automation Technology, and Emerging System Concepts. (Report No. DOT HS 812 043). National Highway Traffic Safety Administration, Washington, DC.
Tsai, Y., Viirre, E., Strychacz, C., Chase, B., Jung, T.-P., 2007. Task performance and eye activity: predicting behavior relating to cognitive workload. Aviat. Space Environ. Med. 78 (Suppl. 1), B176–B185.
Walker, G.H., Stanton, N.A., Young, M.S., 2008. Feedback and driver situation awareness (SA): a comparison of SA measures and context. Transp. Res. Part F: Traffic Psychol. Behav. 11 (4), 282–299. https://doi.org/10.1016/j.trf.2008.01.003.