EMERGING TECHNOLOGIES
TECHNICAL REPORT

FOUNDATION FOR TRAFFIC SAFETY

A Multi-Method Approach to Understanding Drivers' Experiences and Behavior Under Partial Vehicle Automation

SEPT 2023

607 14th Street, NW, Suite 701
Washington, DC 20005
202-638-5944
AAAFoundation.org

Title

A Multi-Method Approach to Understanding Drivers' Experiences and Behavior Under Partial Vehicle Automation

(September 2023)

Authors

David L. Strayer, Joel M. Cooper, David M. Sanbonmatsu, and Amy S. McDonnell

University of Utah

Foreword

Advanced driver assistance systems are becoming more commonplace in vehicles, leading to changes in the role and responsibilities of drivers using these technologies. Vehicle automation that controls vehicle speed, headway and lane position, call for drivers to continuously monitor the road and traffic environment—a scenario that carries implications for driver workload and attention.

The AAA Foundation for Traffic Safety has invested a notable amount of resources during the recent years to better understand impacts of vehicle technologies and automation on users. The work presented in this technical document, in collaboration with the University of Utah, is a good example of such effort. This report summarizes a novel research approach to study drivers' workload, arousal and attentiveness when driving vehicles equipped with Level 2 automation. Drivers were tracked for several weeks to observe how increased familiarity and use of the systems impacted their behaviors and perceptions. Results summarized in this report should help researchers, automobile industry and government entities better understand driver performance, behavior and interactions in vehicles with advanced technologies.

C. Y. David Yang, Ph.D.

President & Executive Director
AAA Foundation for Traffic Safety

About the Sponsor

AAA Foundation for Traffic Safety
607 14th Street, NW, Suite 701
Washington, D.C. 20005
202-638-5944
www.aaafoundation.org

Founded in 1947, the AAA Foundation for Traffic Safety in Washington, D.C., is a nonprofit, publicly supported charitable research and educational organization dedicated to saving lives by preventing traffic crashes and reducing injuries when crashes occur. Funding for this report was provided by voluntary contributions from AAA/CAA and their affiliated motor clubs, individual members, AAA-affiliated insurance companies, and other organizations or sources.

Table of Contents

EXECUTIVE SUMMARY

A multi-method approach was used to better understand driver behavior when using Level 2 partial vehicle automation. Three methods, experimental, naturalistic, and survey, were implemented as part of a single longitudinal study. The complete details for each method and the corresponding results are fully contained in the different parts of this report. In short, participants were trained on a research vehicle that supported Level 2 automation, then, in a controlled experimental trial, they drove under manual and automation modes on two sections of highway with varying driving demands. A researcher was present in the vehicle to facilitate the collection of behavioral and electrophysiological measures of driver cognition. After the experimental session, participants took the vehicle home with them to use on their regular commute to work.

During the 6- to 8-week naturalistic portion of the study, video was recorded to observe driver's behavior and system use. Periodically, participants completed a series of surveys to assess changes in their perceptions, attitudes, and beliefs regarding vehicle automation. Following the naturalistic driving phase, participants completed a second experimental session using the same protocol as in their first encounter.

In Part 1 (Experimental Study), behavioral results suggest that drivers paid more attention to the driving environment under partial vehicle automation than when they were manually driving the vehicle. However, after a 6- to 8-week familiarization period, there was a significant decrease in attention paid to the driving task under partial vehicle automation in the simpler highway driving environment. The spectral electroencephalogram (EEG) measures did not show evidence of decreased attention or workload when participants were driving under partial vehicle automation. This highlights the importance of including multiple dependent measures and a variety of roadway conditions in the testing protocol, as the relationship between driver cognitive state and vehicle automation may depend on the specific demands of a given driving environment.

Part 2 (Naturalistic Study) used a unique approach that combined naturalistic driving research and different control benchmarks, allowing for a more robust comparison of manual and automated driving. The study addressed four research areas: (i) automation use, (ii) system warnings and driving demand, (iii) fatigue and fidgeting, and (iv) secondary task engagement. The results showed that drivers used partial vehicle automation more than 70% of the time, and that there was an increase in system warnings as drivers became more experienced with the system, suggesting that drivers exhibited a tendency toward a more relaxed automation monitoring strategy over time. The study also found that drivers were less likely to use partial vehicle automation when driving demands were higher. Compared to manual driving, partial vehicle automation did not affect fidgeting or fatigue. As drivers grow more familiar with the system, they

were more likely to engage in secondary tasks over time; however, the lack of interaction with driving condition suggests that these behaviors may not be a direct consequence of over-reliance on the automation system. The study underscored the importance of leveraging different benchmarks in understanding the role and impact of vehicle automation on driver behaviors.

Part 3 (Survey Study) used periodic surveys taken throughout the study to assess changes in driver's beliefs and subjective experience. Participants reported that partial vehicle automation improved the experience of driving, reduced the stress of driving and made traveling more enjoyable. These positive effects of vehicle automation on the driving experience increased over time and were strongly correlated with intentions to use and purchase automated vehicles in the future. Participants reported that they engaged in more activities unrelated to driving under partial vehicle automation and become increasingly comfortable allowing their minds to wander as familiarity with the system increased. However, participants reported that they were cognizant of the risks of automated driving and were selective in their usage of the system. Surprisingly, trust in automated systems did not influence evaluations of the partial vehicle automation or the automated driving experiences.

The current multi-method approach combining experimental, naturalistic, and survey data provides important and comprehensive insight into driving behavior under partial vehicle automation; more than what would have been obtained with any single methodology. The data provided a positive perspective on the driver's behavior and experiences with automation. Across all three threads of the study, participants tended to stay engaged with the driving task when partial vehicle automation was activated. Participants reported that they trusted partial vehicle automation, but that they continued monitoring the system in case they needed to take over control. This continued monitoring of the system was largely reflected in their survey data, naturalistic data, and behavioral and neural data collected during on-road driving. Drivers tended to adjust their usage of automation as driving demands changed.

INTRODUCTION

Automated vehicles have the potential to transform society through improvements to safety, mobility, sustainability, and general quality of life for billions of drivers around the world. The promise of automation is that it can improve safety by reducing human error, which remains a significant factor in motor vehicle crashes (NHTSA, 2017). Automated vehicles may also provide a mobility solution to people that cannot drive due age or disability (Alessandrini et al., 2015; Fisher et al., 2016), and they may significantly reduce highway and city congestion (Makridis et al., 2018; Sener & Zmud, 2019).

However, the development of vehicle automation is a difficult problem with near infinite degrees of freedom (Musk, 2021). Given these challenges, many automakers have opted to take small and incremental steps. Indeed, it is now common for vehicles to include a suite of automation functions that provide safety and convenience features to drivers without removing the driver from the role of operator. To characterize the role of the driver in relation to the automation technology, the Society of Automotive Engineers (SAE) has distinguished six levels of automation that gradually transition from full manual control (Level 0) to full vehicle autonomy (Level 5; SAE, 2021). Adaptive Cruise Control (ACC) is characterized by the vehicle's ability to maintain a pre-set speed and following distance, and dynamically adjust speed based on the flow of traffic. Lane Centering Assist (LCA) is characterized by the vehicle's ability to maintain its position within the lane of travel. ACC, which has been available since 1995 (Mitsubishi Diamante), is now commonplace (Level 1). When paired with lane centering, the two technologies form what SAE defines as Level 2 driving automation. Technology to automate lateral and longitudinal vehicle control is now widely available.

With Level 2 automation, drivers are required to supervise the automation and be prepared to steer, brake, or accelerate as needed to maintain safety. Thus, the driver's role shifts from an active controller to that of a passive monitor of the automated system. In Level 2, the driver must pay enough attention to detect "edge cases," which the automated technology is not equipped to handle and human input is needed. In other words, the driver must remain sufficiently engaged to be able to take control of the vehicle if the automation were to fail at any given moment.

The consequences of these technologies on driver behavior are not yet fully understood. Decades of research on various aspects of vehicle automation strikes a cautionary note but real-world testing outcomes are not as clear. While a substantial body of research exists that has used driving simulation to evaluate various aspects of automation, only a handful of studies have been conducted on the road, and these rarely provide definitive answers. The promises of automation are compelling, but several safety concerns have been raised regarding the effects of automation on driver fatigue and secondary task engagement. Additional concerns relate to unintended consequences and changes in system use over time.

To address critical research gaps in our understanding of vehicle automation, the current study recruited participants who had no prior experience with advanced assistance systems. The study employed a hybrid research design that combined both naturalistic and experimental elements. Three methods, experimental, naturalistic, and survey, were implemented as part of a single longitudinal study as shown in Figure 1. This innovative approach allowed for a more comprehensive investigation of how drivers interact with and adapt to vehicle automation systems in real-world scenarios. By examining factors such as changes in system use over time, driver fatigue, and secondary task engagement, this study aimed to provide valuable insights into the safety concerns associated with automation use.

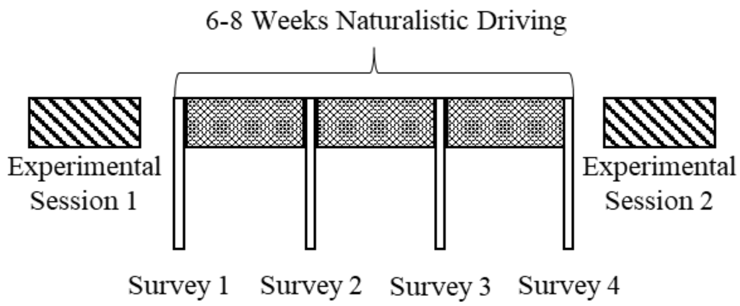

Figure 1. Research design overview.

In the controlled experimental session, participants were trained on a specific research vehicle that supported Level 2 automation, then they drove the vehicle under manual and partial vehicle automation modes in different highway environments. Behavioral and physiological measures of driver cognition and performance were gathered during on-road driving. After the experimental session, participants took the vehicle home with them to use on their regular commute to work. Driver behaviors and systems usage were continuously monitored through video recordings over a 6- to 8-week period. They were also surveyed periodically during their extensive on-road usage of the automation regarding their perceptions, beliefs, and trust in the technology. Following the naturalistic driving period, participants returned to the lab and the protocol used in the first experimental session was repeated in a second experimental session.

The parts of this report align with the major components of the study: the experimental study, describing the two experimental sessions (Part 1); the naturalistic study, describing outcomes observed during the 6 to 8 weeks of naturalistic driving (Part 2); and the survey study, describing the subjective outcomes measured at different points during the study (Part 3). More specific background information and unique details are provided in each of the sections below, along with a comprehensive treatment and discussion of the results.

PART 1 (Experimental Study): The Influence of Practice on Driver Workload and Engagement During Partially Automated Driving[1]

Humans are generally bad at monitoring for rare events (Wolfe et al., 2005) and perform poorly on driving tasks that require sustained attention (Greenlee et al., 2018). Thus, there is concern that the driver's new role as a passive monitor while using vehicle automation may lead to under-arousal and an increased likelihood to disengage from the driving environment (either by zoning out or engaging in secondary tasks), leading to unintended consequences associated with partially automated driving (Casner et al., 2016; Fisher et al., 2016).

This concern has been validated by some empirical evidence that demonstrate a decrease in driver situation awareness (Endsley & Kiris, 1995) and increase in driver drowsiness (Dufour, 2014) while utilizing automated technology. There is a related concern that as driver workload decreases with automation (de Winter et al., 2014), drivers may be more likely to engage in non-driving related tasks (NDRTs) such as talking on a cellphone, which are known to divert attention from the roadway (e.g., Strayer & Johnston, 2001). Research has found that drivers are more likely to engage in NDRTs when the cognitive demand of the primary driving task is low and proneness to boredom is high (Sanbonmatsu et al., 2013; Schroeter et al., 2015). Data from the human–automation interaction literature shows that engagement in NDRTs increases incrementally from manual driving to partially automated driving to fully automated driving (Carsten et al., 2012), and drivers are up to 50% more likely to engage in NDRTs when using partial automation compared to no automation (Dunn et al., 2019). This body of evidence suggests that the driver's cognitive state must continue to be monitored and explored at each stage of shared responsibility between the driver and the vehicle.

However, other studies have failed to replicate these results. Naturalistic data obtained from vehicles instrumented with video cameras and eye trackers suggest that drivers continue to safely monitor the roadways even when Level 2 partial automation is engaged (Fridman et al., 2019; Hatfield et al., 2019), though they may do so with longer eye glances away from the forward roadway (Gaspar & Carney, 2019). McDonnell and colleagues (2021) collected electroencephalogram (EEG) data as participants drove in Level 0 manual mode and Level 2 partial automation mode on real world interstates—across different roadway conditions and vehicle types—and found no differences in level of automation on EEG measures of driver workload or visual engagement. Lohani and colleagues (2021) found a similar null effect of automation on physiological arousal, as measured by heart rate and heart rate variability. Furthermore, Weaver and colleagues (2022) periodically probed drivers and failed to find a significant effect of using partially

[1] Section contributors: Amy S. McDonnell, Kaedyn W. Crabtree, Joel M. Cooper, & David L. Strayer

automated technology on self-reported mind wandering, physiological arousal (heart rate and electrodermal activity), or driving performance. These results suggest there may be a notable difference between simulated driving and real driving, such that the potential for serious and adverse consequences on real roadways may lead drivers to remain engaged in such a way that is not captured in a driving simulator, where driving performance is less consequential. Regardless, the conflicting evidence regarding a driver's ability to remain engaged while driving under partial automation suggests the continued need for real-time assessment of driver cognitive states and how they change with new roles and responsibilities.

Current Experimental Study

It is possible that drivers who are new to automated technology and have no prior experience with Level 2 partial automation (such as those in Lohani et al., 2021; McDonnell et al., 2021; Weaver et al., 2022) remain aroused and engaged throughout testing due to the novelty of the technology. There is conjecture that increasing practice and familiarity with an automated system may influence workload and engagement over time, such that both wane as practice and comfort increase. Survey data collected from experienced Level 2 drivers found that trust increases with Level 2 experience and that increased familiarity with the technology leads to decreases in stress and increases in feelings of security (Endsley, 2017; Gaspar & Carney, 2019). The present study aims to experimentally manipulate familiarity with Level 2 automation. To do so, the experimental design of McDonnell et al. (2021) was replicated in which participants drove Level 2 vehicles on real interstates. However, in the present study, a novel, 6- to 8-week familiarization period was incorporated (described in Part 2), in which participants used the partially automated vehicle every day during their commute to work to gain practice and comfort with the vehicle. Participants then returned to the lab and completed a second experimental session at the end of the 6+ weeks. In the two experimental sessions, driver cognitive states were assessed from both a neural level and a behavioral level.

Electrophysiology

To understand cognition from a neural level, electrophysiological metrics of driver workload and driver engagement were collected, consistent with McDonnell et al. (2021). EEG has been used in past driving studies to assess driver cognitive states in real time (Lohani et al., 2019; Peng et al., 2022). Neurophysiological data is collected on a millisecond time scale from electrodes attached to the scalp, allowing for direct measurement of brain activity in response to the varying demands of the driving environment. Importantly, EEG is non-invasive and mobile, allowing it to be used in both simulators and real vehicles in on-road studies. EEG has become a relatively common method for assessing intelligent transport systems and in developing brain–computer interface frameworks for mental workload (Fan et al., 2022).

Importantly, the EEG signal can be decomposed into the frequency domain using a Fourier analysis (Cohen, 2014) and then categorized into different frequency bands (e.g., Theta ~4–8 Hz, Alpha ~8–12 Hz, Beta ~12–30 Hz) that index different cognitive processes. Driving studies that utilize EEG most commonly assess power in the frontally distributed theta frequency band from 4–8 Hz (referred to as frontal theta power) and power in the parietally distributed alpha frequency band from 8–12 Hz (referred to as parietal alpha power) to assess driver workload and visual engagement, respectively (for review see Borghini et al., 2014). Recent meta-analyses in the cognitive neuroscience literature verify that frontal theta power is sensitive to cognitive workload (Chikhi et al., 2022), such that frontal theta power increases with an increase in cognitive effort in response to task demands (Fairclough et al., 2005; Gevins & Smith, 2003; Puma et al., 2018). Parietal alpha power is inversely related to visual engagement (Goldman et al., 2002) such that, as engagement increases, parietal alpha power decreases (Foxe & Snyder, 2011; McDonnell et al., 2021).

Detection Response Task

In addition to EEG metrics, a behavioral metric of workload was employed: the Detection Response Task (DRT; ISO 17488, 2016). The DRT is a simple stimulus-response task in which participants are instructed to respond to a quasi-randomly presented stimulus with a button press against the steering wheel. The DRT is a commonly used metric in the driving literature that consistently shows that an increase in driving-related demands is associated with increased reaction times and decreased hit rate to the DRT (Cooper et al., 2016; Nilsson et al., 2018; Strayer et al., 2022; Young et al., 2013). Importantly, the DRT is not thought to interfere with performance on the driving task due to its simple nature (Castro et al., 2019; Strayer et al., 2015; Palada et al., 2019), making it an ideal candidate to safely assess driver workload while driving on real roadways at speed.

Research Questions and Hypotheses

The present study sought to answer three main questions related to driver cognitive states during partially automated driving:

1. How does partial automation influence driver workload (as measured with DRT reaction time, DRT accuracy (hit rate), and frontal theta power)?
2. How does partial automation influence drivers’ visual engagement (as measured with parietal alpha power)?
3. How does 6 to 8 weeks of familiarization and practice with a Level 2 partially automated vehicle influence driver workload and engagement?

It was hypothesized that if automation decreases the workload placed on the driver as it intends, lower frontal theta power would be observed, along with faster DRT reaction times and higher DRT hit rate while driving in Level 2 partial automation compared to Level 0. This would support the notion that since the vehicle removes some of the driving tasks from the responsibility of the driver, the driver then has more cognitive resources to allocate to the DRT. It was also hypothesized that if the safety concerns that partial automation may lead to under-arousal and subsequent disengagement from the driving environment are true, then an increase in parietal alpha power during Level 2 partial automation would be observed compared to Level 0. Lastly, it was hypothesized that 6 to 8 weeks gaining familiarity with the Level 2 vehicle would influence driver cognitive states. More specifically, it was predicted that if vehicle automation leads to a decrease in driver workload and engagement, there would be a decrease in DRT reaction time and frontal theta power and an increase in DRT hit rate and parietal alpha power during the partially automated driving in the second session.

Method

Participants

Participants (N = 30, 12 females, 18 males) between the ages of 18 and 55 (M = 35.7, SD = 9.3) were recruited via flyers, word of mouth, and online advertisements. Participants were compensated $20/hour on each of the two experimental testing days, for a total of ~10 hours across the two sessions. To be eligible for the study, participants had to have a valid U.S. driver's license, no at-fault accidents within the past two years (as confirmed by driving records acquired through the University of Utah Division of Risk Management), and no prior experience with Level 2 partial automation technology. To ensure substantial practice driving the vehicle over the 6- to 8-week familiarization phase (Part 2), participants were required to have a daily work commute of at least 40 minutes round-trip on a highway. Prior to testing, all participants were required to complete an online Defensive Driver Training as required by the Division of Risk Management. Additionally, for each session, participants were required to have a blood alcohol concentration of 0.0 (assessed with a BACtrack breathalyzer) and confirm that they had at least 6 hours of sleep the night before testing in order to be allowed in the vehicle. The study protocol was approved by the University of Utah Institutional Review Board (IRB).

Materials

Vehicles. Five, commercially available vehicles with Level 2 partial automation technology were utilized in this study:

- 2018 Tesla Model 3 AWD/Long Range with Autopilot
- 2017 Tesla Model S with Autopilot
- 2018 Cadillac CT6 Premium Luxury with Super Cruise
- 2018 Volvo XC90 Momentum with Pilot Assist
- 2018 Nissan Rogue SL Premium with ProPILOT Assist

All the vehicles were equipped with ACC and LCA (or LKA, Lane Keeping Assist), able to be used simultaneously (qualifying them as Level 2 partial automation). Six participants completed testing in the Tesla Model 3, eight tested in the Tesla Model S, one tested in the Cadillac CT6, nine tested in the Volvo XC90, and six tested in the Nissan Rogue. This array of five vehicles was included in testing in order to increase generalizability of the results across different vehicle makes, not to draw comparisons between vehicles (as the study was underpowered to draw such conclusions).

Highways. During the on-road, experimental sessions, participants drove on two interstate highways in the greater Salt Lake City area (see Figure 2). I-15 is a straight, high-traffic interstate that runs South to North, from Salt Lake City, UT, to Layton, UT, with five lanes of traffic in either direction and an average speed limit of 75 mph. I-80 is a curvy, mountain interstate running West to East, from Salt Lake City, UT, to Wanship, UT, with two to three lanes of traffic in either direction and an average speed limit of 60 mph. The intention behind including two interstates was to assess how different driving demands may differentially influence driver workload and engagement.

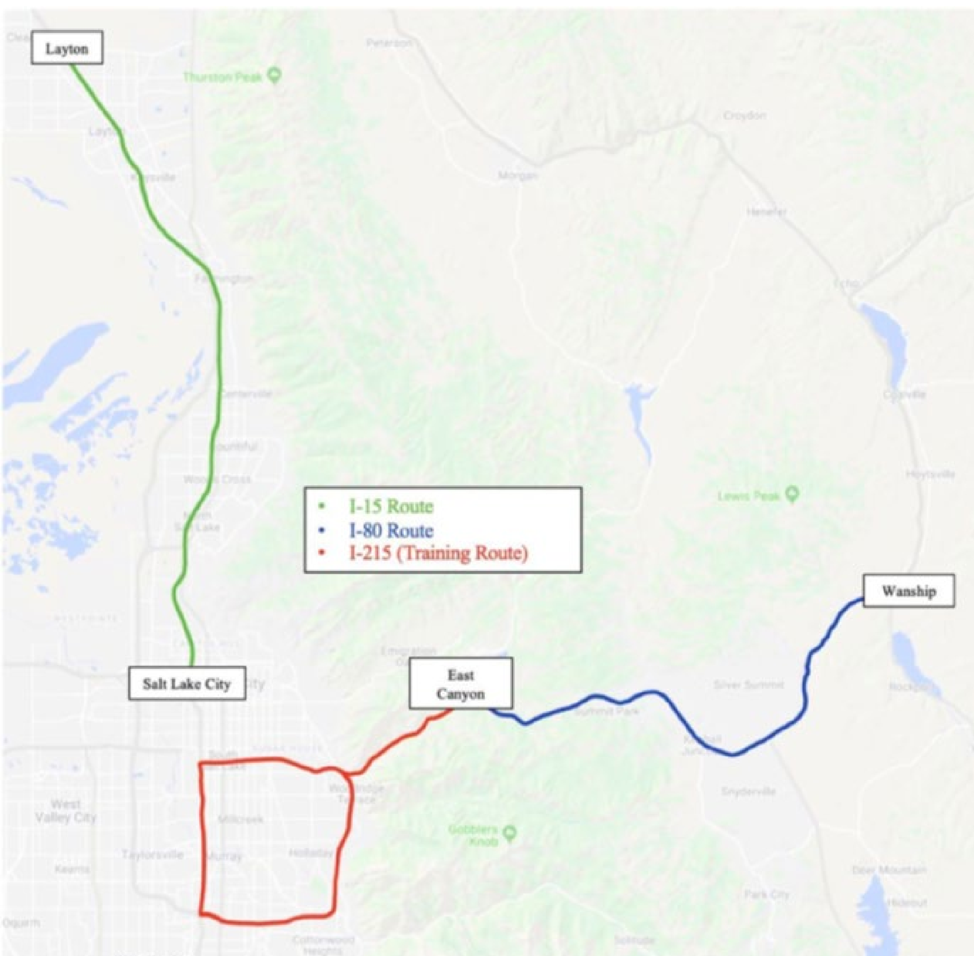

Figure 2. Map of the two experimental driving routes utilized in the study (I-15 in green and I-80 in blue), as well as the training route utilized in Session 1 (I-215 in red).

Detection Response Task. A vibrotactile DRT (Red Scientific Inc., Salt Lake City, UT, USA) was utilized as a behavioral metric of cognitive demand (ISO 17488, 2016). Participants were outfitted with a vibrotactile stimulus device taped to their right forearm and a microswitch response button attached to their right index finger (see Figure 3), consistent with previous driving research (Lohani et al., 2020; 2021; Strayer et al., 2022). For this simple stimulus–response task, a vibration stimulus was presented quasi-randomly every 3 to 5 seconds and participants were to respond by pressing the response button against the steering wheel when they felt the stimulus onset. The vibration lasted 1 second or until the participant pressed the response button. Response times (RTs) to the vibrotactile stimuli were recorded at millisecond resolution and stimuli that were not responded to were coded as misses. Average RT and Hit Rate to the DRT stimulus was calculated for each participant, on both highways, in both manual and automation conditions, for each experimental testing session. Any RTs that occurred sooner than 100 ms were removed before analyses, and RTs that were over 2500 ms were coded as misses (ISO 17488, 2016). Hit Rate was quantified as the proportion of hits (e.g. responses within 100–2500 ms of stimulus onset) out of the total number of stimuli presented in each condition.

Electrophysiology. EEG data were gathered using the gel-based, single-electrode BIOPAC system (BIOPAC Systems, Inc) and reusable electrodes (Ag/AgCl; NATUS Neurology). To best control the quality of the EEG recordings and prioritize quality over quantity (Luck, 2014), three passive electrodes were placed along three midline sites–frontal (Fz), central (Cz), and parietal (Pz)—whose locations followed the International 10-20 system (Jasper, 1958). A ground electrode was placed on the center of the forehead and a reference electrode was placed on the mastoid bone behind the right ear. Two more electrodes were placed above and below the right eye in line with the center of the pupil to record electrooculographic (EOG) activity from blinks and other eye movements for later data processing (see Figure 3). All impedances were kept below 10 kOhms and checked regularly throughout on-road testing to ensure quality data collection.

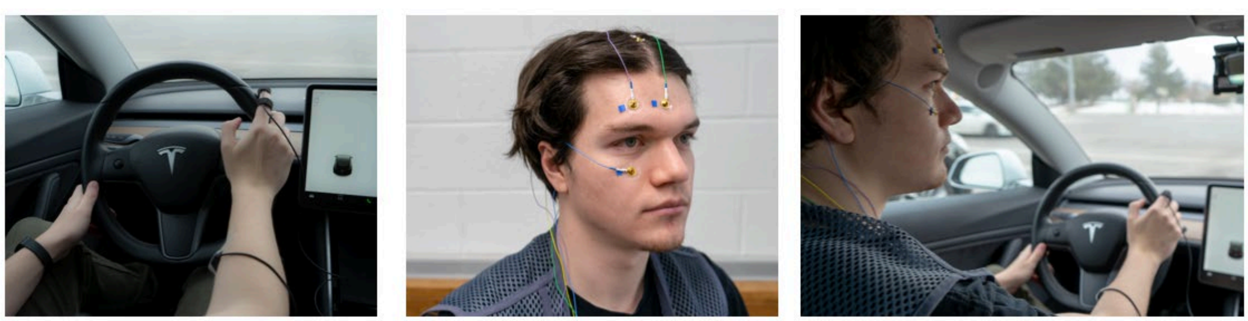

Figure 3. The DRT vibrotactile stimulus and microswitch response button setup (left), EEG electrode setup (middle), and the combination of the two while driving (right).

Design

The present study employed a 2 (Interstate) x 2 (Level of Automation) x 2 (Session) factorial design in which each participant completed the experimental testing on two interstates (I-15, I-80), in two levels of automation (Level 0 and Level 2), at two experimental sessions (before and after the 6- to 8-week familiarization).

During each ~5-hour testing session, participants completed three manipulation checks (Resting DRT, Eyes-closed parietal alpha manipulation check, Auditory N-back frontal theta manipulation check) and four experimental driving conditions (I-15 manual, I-15 partial automation, I-80 manual, I-80 partial automation). The Resting DRT manipulation check consisted of responding to four minutes of the DRT while sitting in the vehicle in the parking lot. This protocol allowed researchers to demonstrate that DRT behavioral metrics, frontal theta power, and parietal alpha power were all sensitive to driving demands such that RT and frontal theta power increase—and Hit Rate and parietal alpha power decrease—under dual-task, driving conditions compared to rest. Each participant also completed an eyes-closed, parietal alpha manipulation check in which they rested for 4 minutes with their eyes-closed while the vehicle was parked. An increase in parietal alpha power while the eyes are closed is one of the most reliable effects in the EEG literature (Goldman et al., 2002) and demonstrates that parietal alpha power is inversely related to visual engagement (because parietal alpha power increases when the eyes are closed). Lastly, each participant completed a frontal theta power manipulation check in which they completed an Auditory N-back counting task while EEG data were recorded and the vehicle was parked. The N-back is a task commonly used in the driving literature to induce cognitive load (Mehler et al., 2011; Strayer et al., 2019; Zhang et al., 2015) and provides a measure of frontal theta power during a standardized cognitive task to compare with driving.

Training

In order to be eligible for the study, it was essential that participants had no prior experience with Level 2 partial automation. Participants were trained for a total of 1.5 hours prior to the first experimental session. Prior to arriving at the lab for their first testing session, participants watched a 30-minute instructional video that explained the features of the specific vehicle they would be testing in. Then, before driving the car at their first testing session, they received an additional 30-minute familiarization period with the researcher in the parking lot in which they were able to adjust the seat and mirrors, identify the relevant sensors, and review how to activate and deactivate ACC and LCA. At this point, the researcher answered any questions and confirmed that the participant was ready to drive the training route.

Participants then completed an on-road training route (see Figure 1) for another 30 minutes. I-215, a medium-traffic interstate with a combination of straight and curving sections, was used for training. It forms a three-quarters loop around Salt Lake City, UT, with sections that run South, West, and North. It has three to four lanes of traffic in either direction and an average speed limit of 70 mph. On the training route, the researcher instructed the participant to engage and disengage the Level 2 partial automation features a few times until they reported feeling comfortable starting the experimental session. When both the researcher and the participant felt safe and comfortable traveling at the speed of traffic in both Level 0 and Level 2 partial automation, the experimental session could begin.

Procedure

The procedure was identical for both testing sessions, with the exception of reduced vehicle training during Session 2. Upon arrival to the research lab, participants signed the IRB approved consent document, were screened for alcohol using a BACtrack breathalyzer, and confirmed the amount of sleep they got the night before. They were then set up with the EEG electrodes on their face, right mastoid bone, and along the midline of their scalp. Once set up was complete, the participant and researcher walked to the nearby parking lot where the vehicle was located.

In the parked vehicle, the researcher then reviewed the technological features and controls/buttons in the vehicle with the participant and answered any questions. This familiarization period was often shorter in Session 2 because the participant was already very familiar with the vehicle. Following this brief vehicle familiarization phase, the participant completed the three manipulation checks (noted above). The DRT manipulation check consisted of responding to the DRT for four minutes while RTs and Hit Rate were recorded. The frontal theta manipulation check consisted of completing an auditory N-back for four minutes as EEG data were recorded. The parietal alpha manipulation check consisted of resting with their eyes-closed for four minutes as EEG data were recorded. Once the manipulation checks were complete, the participant started the vehicle, drove to the training route, and finished the training protocol (as described above).

Once the participant was comfortable with the vehicle and the researcher felt that the participant was able to drive safely in both manual and partial automation mode, the four experimental driving conditions began. The four experimental conditions were defined by driving on I-15 in manual mode, I-15 in partial automation mode, I-80 in manual mode, and I-80 in partial automation mode. The order of the experimental conditions was quasi-counterbalanced such that one interstate was completed before the other. For example, if one participant drove north on I-15 in partial automation mode first, they would then drive back south on I-15 in manual mode before they headed to I-80, in which they would then drive east on I-80 in manual mode and then west on I-80 in partial automation mode. The next participant would then start their experimental conditions on I-80 in partial automation mode, and so on. Counterbalancing in such a way controlled for potential practice effects or fatigue that may occur in the experimental conditions. Between each experimental condition, the participant was instructed to pull over for the researcher to check electrode impedances to ensure quality EEG data acquisition.

At the end of Session 1, the participant took the vehicle home with them and was instructed to drive it for 6+ weeks on their daily work commute. Part 2 describes the naturalistic driving component of the study. At the end of this 6- to 8-week familiarization phase, they were scheduled for their second experimental session, which employed the exact same experimental procedure as the first session.

EEG Recording and Processing

EEG data were recorded using wireless transmitters and amplified with the BioNomadix Smart Center (manufactured by BIOPAC Systems, Inc) at a 2000 Hz sampling rate. EEG data were observed online through AcqKnowledge (Version 5.0) and then processed offline in MATLAB using the EEGLAB toolbox (Delorme & Makeig, 2004). Data were down-sampled to 250 Hz, bandpass filtered from 0.1 Hz to 30 Hz, and then the continuous data were epoched into 1 second intervals with a Hanning window. Eye movement and blinks recorded by the EOG electrodes were corrected using Gratton's regression-based eye-movement correction procedure (EMCP; Gratton et al., 1983). To identify any additional artifacts that were not corrected by EMCP, a subsequent moving window artifact rejection was used to reject epochs containing flatlines or peak-to-peak activity greater than 200µV (Lopez-Calderon & Luck, 2014). This two-layer approach to eye movement correction and subsequent rejection preserved a majority of the EEG data with minimal data loss. The average percent of epochs lost due to blinks and eye movements after correction was 0.39% across all four experimental conditions (I-15 partial automation: 0.39%; I-15 manual: 0.42%; I-80 partial automation: 0.43%; I-80 manual: 0.33%) and is consistent with previous on-road EEG data loss (McDonnell et al., 2021). A Fast Fourier Transform was used to convert the artifact-free, clean EEG data from the time domain to the frequency domain (Cohen, 2014), and then the average power at each frequency from 1 Hz to 30 Hz was extracted for each participant at each level of automation, interstate, and session.

Statistical Analyses

DRT and EEG data were analyzed in R version 4.1.3 (R Core Team, 2022). Manipulation check results were analyzed with paired groups t-tests, comparing the manipulation check condition to the average of the driving conditions. These checks demonstrated that DRT metrics, frontal theta power, and parietal alpha power were sensitive to the various demands associated with driving. For the main analyses, linear mixed effects models were run using the lmer function in R's lme4 package (Bates et al., 2015) to account for sources of non-independence in the data (i.e., repeated measures within each participant) and any missing data. Participant ID was included in all models as a random intercept and DRT RT, DRT Hit Rate, frontal theta power, and parietal alpha power were all entered independently as the outcome variables. For each outcome variable, models were run to test the main effects of Level of Automation, Interstate, and Session, and then a model testing all two- and three-way interactions between these three predictors, with each predictor included as fixed effects in the models. Manipulation check conditions were not included in these models. Likelihood ratio tests were run using the ANOVA function in the stats package (R Core Team, 2022) to test the significance of all effects. The likelihood ratio test generated a chi-squared statistic that compared the model with the variable of interest (Level of Automation, Interstate, Session, or interaction between any of these three variables in the factorial design) entered as a fixed effect and Participant ID as a random effect, to a model with the fixed effect of interest removed. Effect sizes for significant effects were calculated as Cohen's *d*.

Results

Behavioral Results

Metrics derived from the vibrotactile DRT (RT and Hit Rate) offered insight into driver workload from a behavioral perspective. Each participant was presented with ~200 stimuli in each experimental condition at each of the two testing sessions, resulting in a final DRT dataset of 61,471 stimulus presentations across the entire study:

- Resting—3,062
- I-15 manual—13,741
- I-15 partial automation—14,334
- I-80 manual—15,036
- I-80 partial automation—15,298

In total, 59,200 of the stimuli were coded as hits and 2,271 were coded as misses.

The DRT manipulation check confirmed that DRT RT was sensitive to the workload associated with driving such that there were significantly slower RTs while driving (M = 450 ms, SD = 239, SE = 1.0) compared to while parked (M = 354 ms, SD = 195, SE = 3.5; $t(3558.8) = 26.0$, $p < .001$). Similarly, there was a statistically significantly lower Hit Rate while driving (M = 0.96, SD = 0.03, SE = 0.01) compared to while parked (M = 0.99, SD = 0.01, SE = 0.01; $t(31.8) = -6.80$, $p < .001$).

Table 1 presents the descriptive statistics of DRT RT and Hit Rate as a function of Level of Automation, Interstate, and Session, generated from the experimental driving conditions.

Table 1. Mean (standard deviation) of DRT RT and Hit Rate as a function of Level of Automation, Interstate, and Session.

		I-15		I-80	
Outcome	**Level of Automation**	**Session 1**	**Session 2**	**Session 1**	**Session 2**
RT (ms)	Manual	431 (220)	419 (201)	435 (231)	434 (228)
	Partial	476 (258)	439 (219)	483 (268)	479 (266)
Hit Rate	Manual	0.97 (0.07)	0.98 (0.04)	0.96 (0.05)	0.98 (0.02)
	Partial	0.95 (0.04)	0.97 (0.03)	0.94 (0.05)	0.94 (0.06)

Figure 4 shows the mean RT by driving condition (Resting DRT manipulation check, I-15 manual, I-15 partial automation, I-80 manual, I-80 partial automation) and Session. Based on the linear mixed effects models, there was a significant main effect of Level of Automation ($\chi^2(1) = 466.82$, $p < .001$) such that RTs were significantly slower when driving in Level 2 partial automation mode compared to Level 0 manual mode ($\beta = -39.82$, $p < .001$, $d = -0.17$). There was also a significant main effect of Interstate ($\chi^2(1) = 45.60$, $p < .001$) such that RTs were faster when driving on I-15 compared to on I-80 ($\beta = 12.44$, $p < .001$, $d = 0.05$). Furthermore, we found a significant main effect of Session ($\chi^2(1) = 60.18$, $p < .001$) such that RTs were significantly slower at Session 1 compared to Session 2 ($\beta = -14.40$, $p < .001$, $d = -0.06$). There was a significant interaction between Level of Automation and Interstate ($\chi^2(1) = 4.92$, $p = .03$), between Level of Automation and Session ($\chi^2(1) = 28.87$, $p < .001$), and between Interstate and Session ($\chi^2(1) = 26.93$, $p < .001$). Lastly, the three-way interaction between Level of Automation, Interstate, and Session was also significant ($\chi^2(1) = 5.20$, $p = .02$).

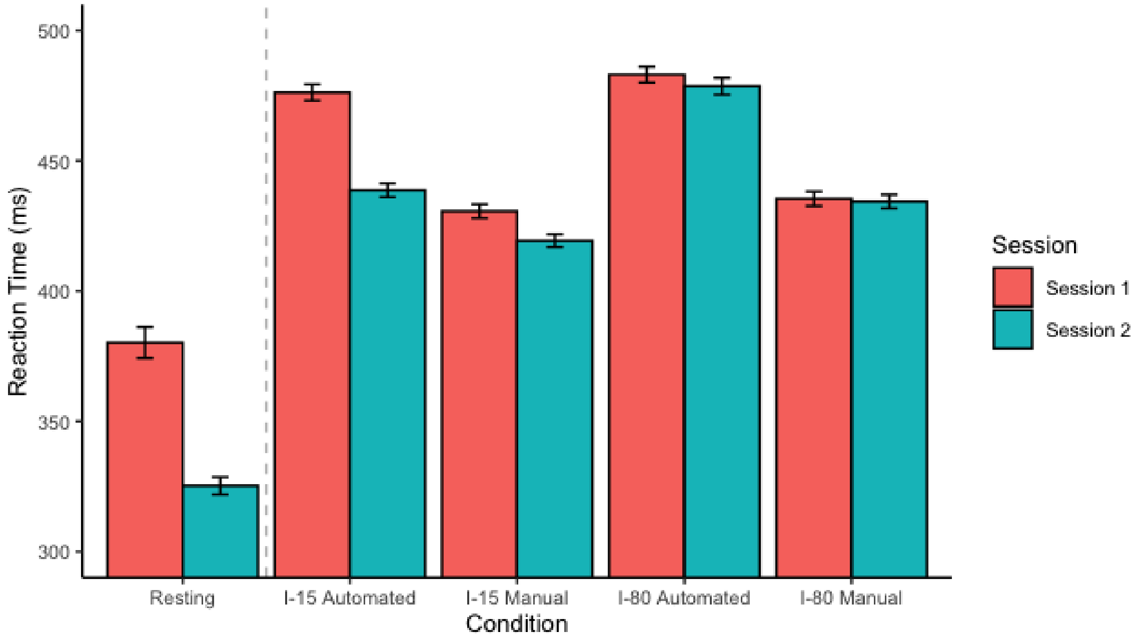

Figure 4. Average RT to the DRT in each of the four experimental driving conditions at Session 1 and Session 2. Error bars represent standard error of the mean. Results from the resting condition (manipulation check) are shown for reference.

Figure 5 shows the mean DRT Hit Rate broken down by driving condition (I-15 manual, I-15 partial automation, I-80 manual, I-80 partial automation) and Session. Based on the linear mixed effects models with Hit Rate as the outcome variable, there was a significant main effect of Level of Automation ($\chi^2(1) = 14.07$, $p < .001$) such that Hit Rate was higher when driving in Level 0 manual mode compared to Level 2 partial automation mode ($\beta = 2.03\text{e-}02$, $p < .001$, $d = 0.42$). A significant main effect of Interstate was also found ($\chi^2(1) = 4.71$, $p < .05$) such that Hit Rate was lower on I-80 than I-15 ($\beta = -0.01$, $p < .05$, $d = 0.25$). There was no significant main effect of Session ($\chi^2(1) = 3.30$, $p = .07$) on DRT Hit Rate, nor were there any significant two-way or three-way interactions (all *p*s = .07 to .65).

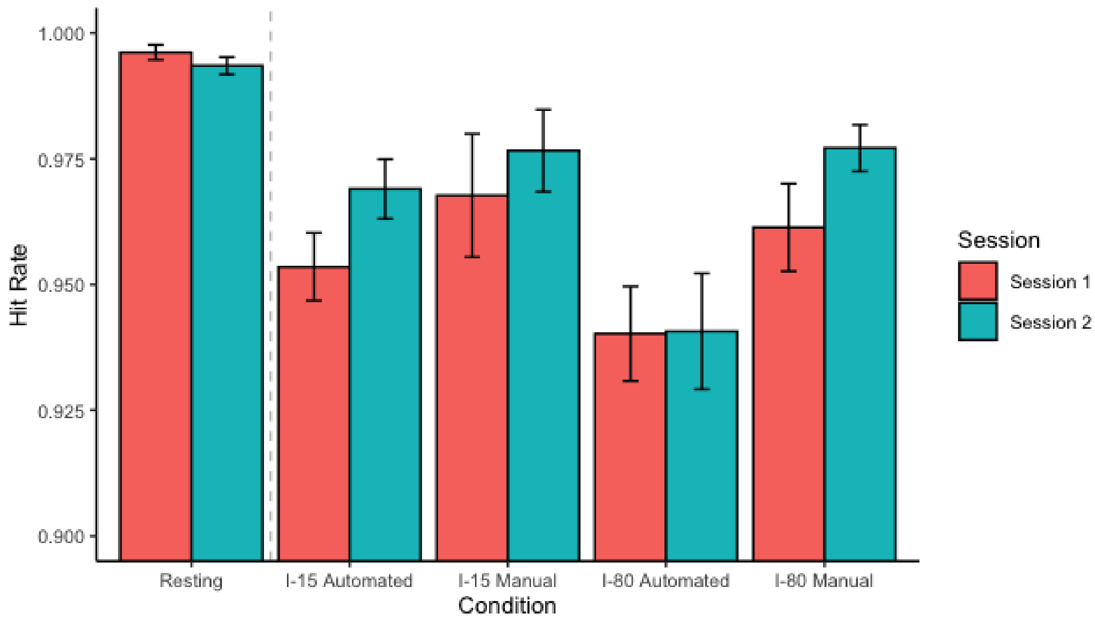

Figure 5. Average Hit Rate to the DRT in each of the four experimental driving conditions at Session 1 and Session 2. Error bars represent standard error of the mean. Results from the resting condition (manipulation check) are shown for reference.

Neurophysiological Results

Neurophysiological metrics derived from the EEG (frontal theta power and parietal alpha power) were assessed to examine covert changes in driver cognitive state (workload and engagement, respectively). The raw spectral curves from 1 Hz to 16 Hz at electrodes Fz and Pz for each condition are presented in Figure 6. Consistent with prior driving literature, frontal theta power is defined as the power in the frequency band between 4 Hz and 8 Hz at electrode Fz. Parietal alpha power is defined as the power in the frequency band between 8 Hz and 12 Hz at electrode Pz. The spectral plot at Fz includes the four experimental driving conditions as well as the Resting DRT and the Auditory N-back frontal theta manipulation checks. The spectral plot at Pz includes the four experimental driving conditions as well as the Resting DRT and Eyes-Closed (in which there is a large bump in alpha power between 8 Hz and 12 Hz) parietal alpha manipulation checks.

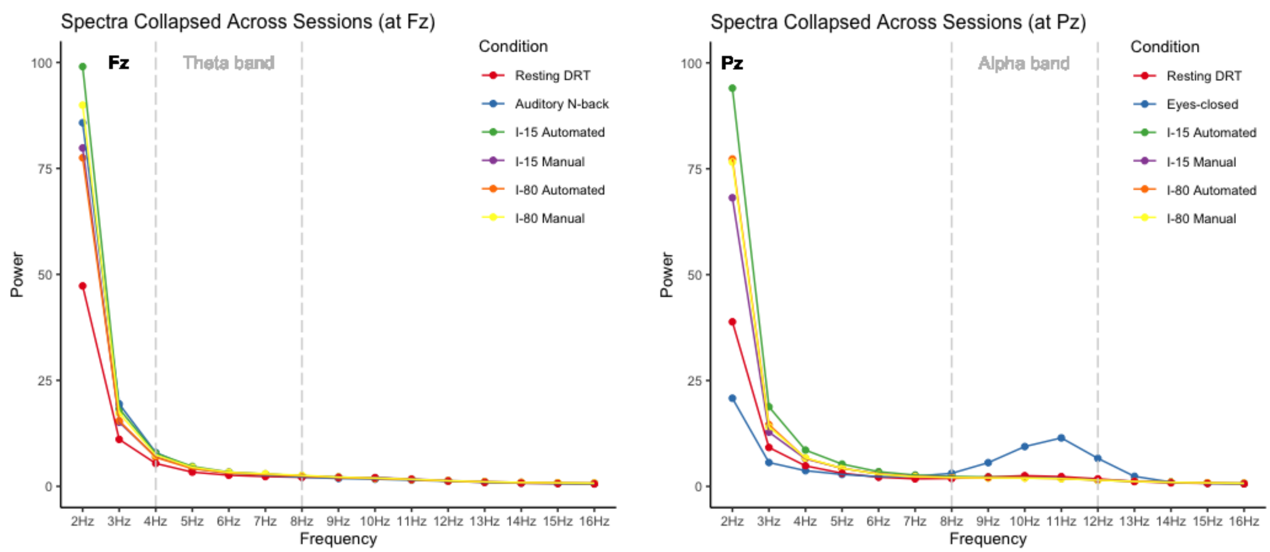

Figure 6. Spectral curves of the power at each frequency at electrodes Fz (left) and Pz (right).

With respect to the manipulation checks, the frontal theta power Auditory N-back manipulation check confirmed that frontal theta power derived from the EEG is sensitive to the cognitive demand associated with a driving task. Frontal theta power was nominally higher while driving (M = 3.05 μV^2/Hz, SD = 1.35, SE = 0.09) compared to the resting Auditory N-back (M = 2.80 μV^2/Hz, SD = 1.27, SE = 0.17), though this did not reach statistical significance (t(92.4) = -1.34, p = .18). Similarly, the Resting DRT manipulation check confirmed that frontal theta power while responding to the DRT during rest (M = 2.00 μV^2/Hz, SD = 1.23, SE = 0.16) was significantly lower than frontal theta power while responding to the DRT while driving (M = 3.05 μV^2/Hz, SD = 1.35, SE = 0.09; t(94.7) = 2.95, $p < .01$).

The eyes-closed parietal alpha power manipulation check confirmed that parietal alpha power derived from the EEG is a neural marker of visual engagement with the driving environment such that parietal alpha power while the eyes are closed at rest (M = 7.09 μV^2/Hz, SD = 10.2, SE = 1.34) is significantly higher than while driving (M = 1.74 μV^2/Hz, SD = 1.07, SE = 0.07; t(57.3) = 3.97 , $p < .001$). Lastly, the Resting DRT manipulation check indicated that parietal alpha power while responding to the DRT (M = 2.00 μV^2/Hz, SD = 1.54, SE = 0.20) was higher (indicative of less visual engagement) than parietal alpha power while responding to the DRT while driving (M = 1.74 μV^2/Hz, SD = 1.07, SE = 0.07), though this did not reach statistical significance (t(70.2) = -1.21, p = .23).

Table 2 presents the descriptive statistics of frontal theta power and parietal alpha power as a function of Level of Automation, Interstate, and Session, generated from the experimental driving conditions.

Table 2. Mean and standard deviation of each of the EEG metrics as a function of Level of Automation, Interstate, and Session.

		I-15		I-80	
Outcome	**Level of Automation**	**Session 1**	**Session 2**	**Session 1**	**Session 2**
Frontal theta power ($\mu V^2/Hz$)	Manual	3.11 (1.33)	2.98 (1.19)	2.99 (1.28)	3.24 (1.84)
	Partial	3.07 (1.19)	3.19 (1.53)	2.87 (1.13)	2.96 (1.29)
Parietal alpha power ($\mu V^2/Hz$)	Manual	1.85 (1.08)	1.71 (1.05)	1.65 (0.85)	1.73 (1.06)
	Partial	1.81 (1.17)	1.77 (1.35)	1.71 (1.02)	1.69 (1.04)

Frontal Theta Power*.* Across both testing sessions, 12 frontal theta power files were lost due to electrodes falling off during a driving condition, issues with the activation of the automated technology in the vehicle, or environmental conditions (such as rain or snow) that rendered the vehicle sensors unusable. Therefore, the final frontal theta power EEG analysis consisted of 58 data files in each of the I-15 manual, I-15 partial automation, and I-80 manual condition and 54 data files in the I-80 partial automation condition.

Figure 7 shows the mean frontal theta power as a function of Condition (I-15 manual, I-15 partial automation, I-80 manual, I-80 partial automation) and Session, along with the manipulation check benchmarks. The linear mixed effects models with frontal theta power as the outcome variable revealed that there were no significant main effects of Level of Automation, Interstate, or Session, and no significant two- or three-way interactions between each of the variables of interest (see Table 3).

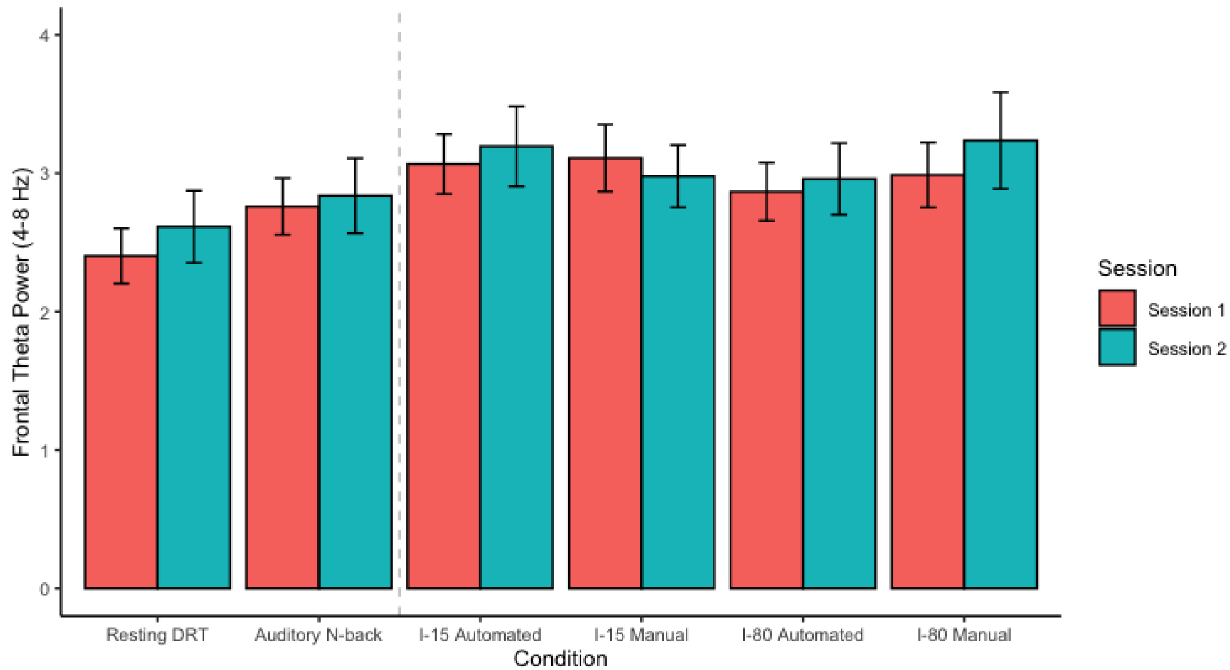

Figure 7. Average frontal theta power in the four experimental driving conditions at Session 1 and Session 2. Error bars represent standard error of the mean. Results from the resting DRT and auditory N-back conditions (manipulation checks) are shown for reference.

Table 3. Main effect and interactive results from linear mixed effect models predicting frontal theta power.

Predictor	χ^2	df	p
Level of Automation	0.0001	1	.99
Interstate	0.07	1	.79
Session	0.35	1	.55
Level of Automation x Interstate	1.11	1	.29
Level of Automation x Session	0.07	1	.80
Interstate x Session	1.05	1	.31
Level of Automation x Interstate x Session	1.92	1	.17

Parietal Alpha Power. Across both testing sessions, 14 parietal alpha power files were lost due to electrodes falling off during a driving condition, issues with the activation of the automated technology in the vehicle, or environmental conditions (such as rain or snow) that rendered the vehicle sensors unusable. Therefore, the final parietal alpha power EEG analysis consisted of 57 data files in the I-15 manual and I-15 partial automation conditions, 58 data files in the I-80 manual condition, and 54 data files in the I-80 partial automation condition.

Figure 8 shows the mean parietal alpha power broken down by Condition (I-15 manual, I-15 partial automation, I-80 manual, I-80 partial automation) and Session, along with the DRT and eyes-closed manipulation benchmarks. Results of the linear mixed effects models with parietal alpha power as the outcome variable are presented in Table 4. Mixed models revealed that there were no significant main effects of Level of Automation or Session. There was however a significant main effect of Interstate on parietal alpha power ($\chi^2(1) = 3.87$, $p < .05$) such that there was lower parietal alpha power when driving on I-80 compared to I-15 ($\beta = -0.10$, $p < .05$, $d = -0.10$), collapsed across Level of Automation and Session. There were no significant two- or three-way interactions between each of the variables in the factorial design.

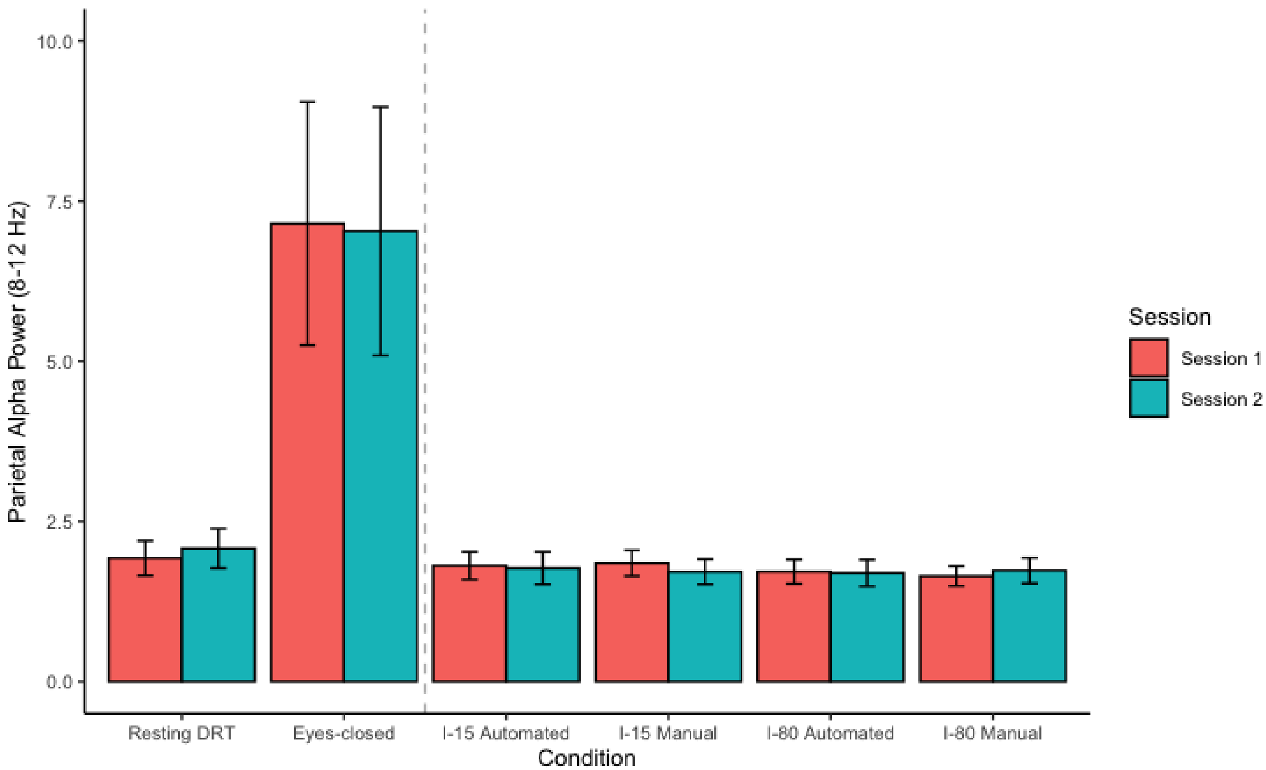

Figure 8. Average parietal alpha power in the four experimental driving conditions at Session 1 and Session 2. Error bars represent standard error of the mean. Results from the resting DRT and auditory N-back conditions (manipulation checks) are shown for reference.

Table 4. Main effect and interaction results from linear mixed effect models predicting parietal alpha power.

Predictor	χ^2	df	p
Level of Automation	0.24	1	.63
Interstate	3.87	1	.05
Session	0.30	1	.58
Level of Automation x Interstate	0.002	1	.97
Level of Automation x Session	0.045	1	.83
Interstate x Session	2.02	1	.16
Level of Automation x Interstate x Session	0.76	1	.38

Discussion

When driving a Level 2 partially automated vehicle, drivers are expected to remain vigilant and engaged with the driving task should the automated technology require human intervention (SAE, 2021). There is concern that because automation may decrease the workload of the driver, they therefore may disengage with the driving environment. Furthermore, drivers tend to perform poorly at monitoring tasks that require detection of rare events (Greenlee et al., 2018; Wolfe et al., 2005). The present study explores the effect of partial vehicle automation on driver workload and engagement and how these change after drivers gain practice and familiarity with automated systems. A multi-method, experimental approach was undertaken to measure driver cognitive states during real driving conditions from both a behavioral (DRT) and a neural (EEG) perspective. Each variable of interest was systematically manipulated in a 2 (Level of Automation: Level 0 and Level 2) x 2 (Interstate: I-15 and I-80) x 2 (Session: Session 1 and Session 2) factorial design. Manipulation checks for each of the outcome measures were also included to ensure that the measures were sensitive to the demands associated with driving. Overall, statistically significant effects were found for behavioral measures (DRT RT and Hit Rate) but not for neurophysiological metrics (frontal theta or parietal alpha power).

Behavioral Metrics

The DRT manipulation check (in which participants responded to four minutes of the DRT when the vehicle was parked) validated that RT and Hit Rate are sensitive to the cognitive demands associated with driving, providing confidence that these measures assess driver workload by using the behavioral metrics derived from the vibrotactile DRT.

RTs to the DRT demonstrated sufficient discriminability to detect differences in driver workload associated with varying levels of automation, interstate characteristics, and changes over time. The results revealed a main effect of automation such that participants had slower RTs when driving in partial automation mode compared to manual mode. This suggests that rather than decreasing driver workload as automation intends, driving under partial automation was associated with an increase in driver workload. Additionally, there was a significant main effect of Interstate such that driving on I-80 induced more cognitive load than driving on I-15, indicating that participants driving on the curvy section of I-80 paid more attention to the driving task regardless of the mode of automation. Lastly, there was a main effect of Session on RTs such that participants responded faster to the DRT at Session 2 compared to Session 1, suggesting evidence of a practice effect.

More pertinently, all two-way and three-way interactions were significant. Of particular interest was the interaction between Level of Automation and Session, which revealed that six weeks of practice and familiarity with vehicle automation influenced driver workload over time. The RT data tended to show a steeper decline (faster RTs) from Session 1 to Session 2 when driving in partial automation mode compared to manual mode. Interestingly, this was only the case when driving on I-15 and not I-80. These data suggest that practice with vehicle automation decreases driver workload over time, at least when driving on roads with relatively low demand. This further highlights the importance of including various highway characteristics in study designs.

DRT Hit Rate data support some of the conclusions drawn from the RT data. In particular, Hit Rate data show a significant main effect of Level of Automation such that Hit Rate was higher during manual driving compared to partially automated driving, consistent with the interpretations drawn from the RT data. Additionally, Hit Rate was higher on I-15 compared to I-80, further validating the interpretation that I-80 was more cognitively demanding. However, Hit Rate was less sensitive to variations in cognitive load than RT, possibly due to ceiling effects in Hit Rate.

Neurophysiological Metrics

The frontal theta and parietal alpha manipulation checks validated that these EEG metrics are sensitive to the demands (both cognitive and visual) associated with driving. In terms of frontal theta power, it was found that frontal theta power while responding to the DRT in the parking lot was significantly lower than frontal theta power while driving. Participants also completed a standardized Auditory N-back workload task while parked to assess frontal theta power in a cognitive task compared to while driving. Interestingly, frontal theta power was numerically higher in the conditions that involved driving compared to the N-back task alone, suggesting that driving is cognitively demanding. Additionally, parietal alpha power when the eyes are closed at rest was significantly larger than while driving, confirming well-validated EEG findings that parietal alpha power is inversely related to visual engagement (Goldman et al., 2002). It was also found that parietal alpha power was numerically lower (indicative of more visual engagement) when participants were driving and responding to the DRT than when they were responding to the DRT at rest.

The manipulation checks validate that these EEG metrics are sensitive to the demands of driving compared to rest; however, they were less sensitive to discriminating differences in workload and engagement between driving conditions. Unlike the behavioral results, our neurophysiological results do not show significant effects of Level of Automation or Session. There was a main effect of Interstate on parietal alpha power such that participants were more visually engaged on I-80 compared to I-15, likely due to the increased demand associated with the curvy section of I-80. In general, the null frontal theta and parietal alpha effects across a majority of the experimental conditions replicate the results of McDonnell et al. (2021), which had a similar experimental protocol, with the exception of a second testing session.

Conclusions

The DRT was more sensitive to experimental condition differences than the EEG measures, which is consistent with prior findings (e.g., Strayer et al., 2015). The data suggest that Level 2 partial automation may not decrease driver workload as expected, given our main effects of Level of Automation on RT and Hit Rate. In fact, the DRT data suggest that drivers in this experimental study may pay *more* attention to the driving environment under partial automation compared to manual mode. This conclusion is consistent with past work that utilized the DRT to test driver attention under partial vehicle automation (Strayer et al, 2020). However, after a 6-week familiarization period in which participants practiced driving in partial automation mode daily, there was a significant decrease in attention paid to the driving task under partial automation—at least in the simpler driving environment (I-15).

The observed difference across Interstate suggests that driver engagement is modulated by driving environment, consistent with prior work (Strayer et al., 2020). In particular, there was an increase in DRT RT, decrease in Hit Rate, and decrease in parietal alpha power when participants drove on I-80, a more winding, mountain-climbing section of the interstate, as compared to the straighter driving environment on I-15. Interestingly, this pattern held for both Level 0 manual driving as well as Level 2 partially automated driving and was observed even after participants had 6+ weeks of practice. This result has implications for future driving research in that it establishes the importance of including a variety of roadway conditions in the testing protocol, as the relationship between driver cognitive state and automation may be dependent on the specific demands of a given driving environment. Future research may also consider extending the length of familiarization and practice provided to a participant. This study allowed participants 6 to 8 weeks of practice with the Level 2 vehicle and found significant effects of Session on DRT RT. It is possible that with more time and practice, there would be an even greater decrease in workload and engagement associated with partial automation.

One limitation of the current experimental protocol is that an experimenter was present in the vehicle during testing in order to monitor data acquisition, the quality of the neurophysiological recordings, and electrode impedances. It is possible that the driver's behavior would change in a more naturalistic setting when the experimenter was not present. This hypothesis is directly examined in Part 2. Overall, the current study highlights the feasibility of collecting on-road behavioral and neurophysiological data across different driving environments and vehicle makes. As with all cognitive research, there is a desire for use-inspired basic research that utilizes traditional, well-validated psychological and cognitive neuroscience methods and constructs to test research questions with applied outcomes. The current experimental study taps into a timely and relevant topic in human factors research: as society is becoming more and more automated there is a need to continue to explore how humans engage and interact with automated technology when responsibilities are shared.

PART 2 (Naturalistic Study): Driver Behavior while using Level 2 Vehicle Automation, a Hybrid Naturalistic Study[2]

The naturalistic phase of this study aimed to address a number of research questions along several dimensions. Background and motivation for these dimensions are provided in the sub-sections below, followed by an overview of the main objectives and questions addressed in Part 2.

Effects of Practice on Usage

Research suggests that drivers' familiarity and experience with automation technologies, such as LCA or ACC, may influence usage patterns (Beggiato et al., 2015; Larsson, 2012). Initially, drivers may be hesitant to use automation due to lack of understanding or concerns about reliability. As they gain experience, they may become more comfortable and proficient. However, it is unclear how increased proficiency affects usage. Dunn et al. (2021) propose that experience with automation changes behavior through operational phases, but this has not been experimentally confirmed and likely depends on the driver's perception of control, usefulness, and reliability (Parasuraman & Riley, 1997). The current study aimed to better understand the relationship between practice and the use of vehicle automation.

Warnings and Driving Demand

Automation warnings occur for various reasons. In vehicles equipped with Level 2 automation, warnings related to driver state monitoring are common. Warnings arise when drivers fail to maintain sufficient steering torque or keep their eyes on the forward roadway. These warnings typically involve visual, auditory, and tactile cues such as vibrations through the steering wheel and seat. The specific types of warnings, their activation methods, and their intended message to drivers vary depending on the vehicle's automation system and capabilities.

However, research suggests that driver acceptance of system warnings is often low (Xu et al., 2021) and influenced by factors such as the driver's experience and familiarity with the technology, as well as the perceived reliability and usefulness of the automation (Abe & Richardson, 2004; Large et al., 2017). Changes in the frequency of system warnings may result from changes in a driver's understanding of the warning

---

[2] Section contributors: Joel M. Cooper, Kaedyn W. Crabtree, Amy S. McDonnell, Dominik May, Sean C. Strayer, Tushig Tsogtbaatar, Danielle R. Cook, Parker A. Alexander, David M. Sanbonmatsu, & David L. Strayer

cause, intent, and severity as well as their acceptance and use of the system. The current study allowed for the tracking of the frequency of system warnings across vehicles and over time.

Warnings are also occasionally issued to request that the driver take over steering control due to poor conditions. Although automated systems can sometimes function in challenging conditions, they are not currently intended for situations requiring extra driver caution and vigilance. The road-facing camera used in the current research allowed researchers to code the co-occurrence of warnings and various types of poor conditions. The frequency of system alarms and the continued use of vehicle automation in poor driving conditions reflect the automation control strategies and the extent to which drivers remain functionally vigilant to the driving task (Fridman et al., 2019).

Arousal: Fatigue and Fidgeting

The relationship between Level 2 automation and fatigue is complex and not fully understood. Several research studies using driving simulations have found that automation use led to an increase in driver passive fatigue, caused by under arousal and boredom (Ahlström, et al., 2021; Arefnezhad et al., 2022; Desmond & Hancock, 2000; Matthews et al., 2019). However, the controlled nature of these research designs often limits the types of natural countermeasures that drivers may employ to combat fatigue and under arousal. Indeed, research has shown that secondary task interactions may, in some cases, protect against fatigue that arises during the use of automation (Schömig et al., 2015; Feldhütter, et al., 2019), leading some to suggest secondary task use as a countermeasure for automation related fatigue (Vogelpohl, et al., 2019). However, complex secondary tasks can also distract from the driving task and result in slow resumption of vehicle control during a takeover request (Louw et al., 2015; Merat et al., 2014). Because research on driver fatigue during automation use has primarily been conducted in simulators and some on-road studies, outcomes have been inconclusive regarding the role and prevalence of drowsiness (Dunn et al., 2019). Thus, it remains uncertain if these findings can be extrapolated or generalized to real-world scenarios.

Fidgeting is defined by the Oxford Dictionary as making small movements, especially of the hands and feet, through nervousness or impatience. Research suggests that fidgeting is highly associated with mind wandering and inattention (Carriere et al., 2013) and is sometimes viewed as a distracting secondary task (Hasan et al., 2022). Fidgeting behaviors may therefore be indicative of a driver countermeasure to combat fatigue or boredom and a potential precursor to passive fatigue. Based on these definitions and findings, fidgeting behavior may serve as a useful indirect measure of driving task engagement, with lower rates of fidgeting suggesting higher driving engagement or potential fatigue, and higher rates of fidgeting suggesting lower driving engagement and possible mind wandering.

Secondary Task Engagement

Roadside observations of drivers suggest that they engage in non-driving related secondary tasks up to 32% of the time (Huisingh et al., 2015). With recent technological developments allowing for vehicle-phone pairing, voice control, and heads-up technology interactions, it is likely that this number is growing. Behavioral analyses using the SHRP2 naturalistic driving dataset suggest that observable distractions are prevalent in 52% of normal baseline driving (Dingus et al., 2016). While the prevalence of handheld phone use for talking by drivers has gradually decreased, the prevalence of handheld device manipulation for activities such as texting and internet use has increased (NHTSA, 2021).

Several studies have indicated that drivers are more likely to engage in secondary tasks when vehicle automation is active (Dunn et al., 2021; De Winter et al., 2014; Naujoks et al., 2016; Endsley, 2017; Reagan et al., 2021). Drivers are also able to more efficiently complete secondary tasks with automation than when manually driving (He & Donmez, 2019). The primary concern with secondary task engagements during automation use is that they reduce the driver's ability to safely monitor the automation through a diversion of visual and cognitive resources (Gaspar & Carney, 2019) and decrease a driver's ability to quickly resume full control of the vehicle (see Morales-Alvares et al., 2020, for a comprehensive review). A second concern is that the driver may develop automation-induced complacency over time, where drivers will over-rely on the automation and fail to monitor it appropriately. Contrasting results in two naturalistic driving studies analyzed by Dunn et al. (2021) suggest that driver complacency and willingness to engage in secondary tasks may develop through a series of phases. In the first phase, the learning phase, drivers begin to get acquainted with the automation, including learning about its potential uses and limitations. During this phase, drivers may not fully trust the automation and may be unwilling to engage in non-driving-related tasks. However, as experience with the automation grows, it has been suggested that drivers transition into an integration phase (Saad, 2004), indicated by an increased willingness to divert attention from the roadway toward secondary tasks. The existence of this type of phased learning has not, however, been demonstrated in a single study, and it is unclear whether this theory accurately characterizes the evolution of secondary task behaviors with automation use in the real world.

Naturalistic and Experimental Driving Approaches

The naturalistic driving approach, originally developed by the Virginia Tech Transportation Institute (Neale et al., 2005) and now used by researchers worldwide (Eenink, et al., 2014; Fridman et al., 2019), uses cameras placed in participant vehicles to passively collect video recordings of drivers during their normal use of the vehicle. This approach allows researchers to observe driving behavior as it occurs in real-world scenarios, while allowing drivers to act naturally.

Naturalistic driving research generates a continuous stream of video which can be challenging to transfer, catalog, and analyze. To help manage this complexity, several approaches have been developed to both identify events of interest and suitable sections of video to code for baseline behavior. In most cases, critical events are identified either through high-g (force) events or through some form of machine learning (e.g., Fridman et al., 2017). Baseline driving epochs are then selected to match as closely as possible to the event of interest with the exception that the event of interest is not found in the selected baseline video. The validity of these techniques hinges on the degree to which the baseline epochs match the event epochs, and great care is required to ensure that they match as closely as possible. However, because there are no true experimental controls in naturalistic driving research, it can be challenging to form causal relationships between events (Carsten et al., 2013). This is especially true with vehicle automation or secondary task interactions where an environmental factor, which is not easily matched in the baseline, may have a large influence on behavior.

Experimentally controlled evaluations of driver performance are commonly used to gain insights into the potential safety concerns that may arise with vehicle automation. Within the driving domain, these come in several variations that range from simple tracking tasks (Strayer & Johnston, 2001) to complex scenario mock-ups using multiple highly instrumented vehicles on climate-controlled test tracks (Gibson, 2015; Tan et al., 1998). The primary strength of tight experimental control is that it allows researchers to manipulate a single factor while holding all other factors constant. Unlike with the naturalistic driving approach, the performance baseline is often an identical or near identical scenario. This allows for confident statements about causality. The challenge with these types of studies is generalizability, as naturalism is often sacrificed for control.

Thus, another aim of this study was to deploy a hybrid naturalistic approach in order to contrast behaviors observed during Level 2 automation using different benchmarks: a naturalistic baseline (taken from video where drivers elected not to use automation) with an imposed experimental baseline (generated using a control condition in which the driver was instructed not to use vehicle automation).

Current Naturalistic Study

With the dimensions noted above as a backdrop, the current study expands on previous research in several key ways. First, all vehicles in the study have automated systems that meet the SAE definition of Level 2 automation. Prior research has often used a mixture of Level 1 and Level 2 vehicles (e.g., Dunn et al., 2021). Second, the study was designed to systematically control environmental differences that could influence automation use, such as varying road conditions, weather, traffic density, and infrastructure. This is a unique and important manipulation that, to the knowledge of the research team, has never been done before. Finally, novice users were followed over a 6- to 8-week period, longer than previous studies, which allowed for in-depth analysis of how driver behaviors change as drivers become more familiar with vehicle automation. Through this novel experimental design, the following questions are explored:

- Automation Usage (AU):
    - AU1 – Does experience *(based on number of weeks of system use)* with automation change the frequency with which drivers activate the automation?
    - AU2 – How does the time to re-engage the system (after a system disengagement) change with experience?
- Warning and Driving Demand (WD):
    - WD1 – Does the frequency of system warnings change over time?
    - WD2 – Does the frequency of automation use change during poor driving conditions?
- Fatigue and Fidgeting (FF):
    - FF1 – Do drivers show increased signs of fatigue when using automation and does this change over time?
    - FF2 – Do drivers show increased signs of fidgeting when using automation and does this change over time?
- Secondary Task Engagement (ST):
    - ST1 – How does the frequency of secondary task use (non-driving related) change during Level 2 use and over time?
    - ST2 – How does the frequency of task type, mode of interaction (voice vs manual), and interface (cell phone vs. In Vehicle Information System [IVIS]) change during Level 2 use and manual driving?

Method

As noted in previous sections, the video data analyzed and presented in this section form a subset of a larger research effort (see Figure 9), which included a 6- to 8-week naturalistic observation period, survey data collection (Part 3) and two 5-hour on-road performance evaluations (Part 1). Herein, methodological details that vary or are specific to the 6 to 8 weeks of naturalistic driving are described.

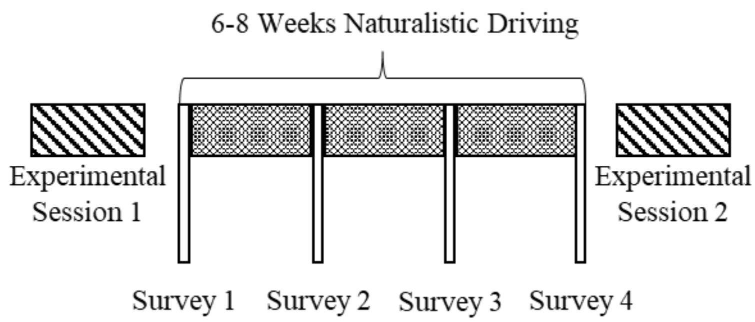

Figure 9. Research design overview.

Participants

Participants (*N*= 30, 12 females, 18 males) between the age of 18 and 55 (*M* = 35.7, *SD* = 9.3) were recruited through flyers, word of mouth, and online advertisements. For the 6- to 8-week naturalistic portion of the experiment, participants received an average compensation of $300. Eligibility criteria included having a valid U.S. driver's license, no at-fault accidents within the past two years (verified by driving records obtained through the University of Utah Division of Risk Management), and no prior experience with Level 2 automation. Participants needed to have a daily work commute of at least 20 minutes (40 minutes round trip) on a major local interstate (I-80, I-15, or I-215) and were instructed to use vehicle automation as often as they felt comfortable.

Materials

Vehicles. The naturalistic study used the same vehicles as the experimental study (Part 1): Tesla Model 3, Tesla Model S, Cadillac CT6, Volvo XC90, and Nissan Rogue. The distribution of participants testing in each vehicle was as follows: six participants in the Tesla Model 3, eight in the Tesla Model S, one in the Cadillac CT6, nine in the Volvo XC90, and six in the Nissan Rogue.

Cameras. Rosco-developed Dual-Vision XC4 cameras were installed under each vehicle's rear-view mirror. The cameras offered a view of both the forward roadway and the vehicle interior using a fish-eye lens. Additionally, an auxiliary camera captured either the screen (instrument panel) behind the steering wheel or the screen between the front seats, depending on the location of vehicle state icons indicating automation status. Video data were stored on Rosco and Transcend brand secure digital (SD) cards, and the cameras automatically started and stopped recording when the vehicle was turned on or off.

Video Coding. Videos were processed for analysis using BORIS (Behavioral Observation Research Interactive Software; Friard & Gamba, 2016). BORIS enabled coders to pre-specify activities of interest and then perform a frame-by-frame video playback to mark the beginning and end of each behavior. Summary results for each coded video were provided in a simple .csv format, with each output line containing details about individual observations, such as the behavior, location within the video, and start and stop times (Figure 10).

Observati	Observation date	Media file	Total length	FPS	Behavior	Behavioral	Modifiers	Behavior type	Start (s)	Stop (s)	Duration (s)
002_back_	6/14/2021 11:32	C:/Users/R	1739.2	10	Radio	Tasks	Listening	STATE	53.969	1407.718	1353.749
002_back_	6/14/2021 11:32	C:/Users/R	1739.2	10	L2 ON	L2		STATE	70.404	190.701	120.297
002_back_	6/14/2021 11:32	C:/Users/R	1739.2	10	Other	Other		STATE	78.455	83.956	5.501
002_back_	6/14/2021 11:32	C:/Users/R	1739.2	10	Other	Other		STATE	86.456	190.705	104.249
002_back_	6/14/2021 11:32	C:/Users/R	1739.2	10	L2 OFF	L2		STATE	190.702	1407.718	1217.016
002_back_	6/14/2021 11:32	C:/Users/R	1739.2	10	Poor Con	Other	Construction	STATE	195.956	256.804	60.848

Figure 10. Example BORIS video coding output file.

Procedure

After completing Experimental Session 1 (described in Part 1), participants received one of the five research vehicles, which they agreed to use on weekdays for commuting to and from work, not allowing other people inside the vehicle, and operating the vehicle according to the law. Participants were encouraged to use vehicle automation on interstate segments of their commute as often as they felt comfortable. They used the vehicle on workdays for 6 to 8 weeks (subject to scheduling constraints related to the final evaluation) before completing the final experimental session and returning the vehicle (see Part 1).

Experimental Control. Each week, one randomly selected day was designated as an experimental control day, during which participants were instructed *not* to use vehicle automation (Automation: NO). Notifications were sent to drivers the day prior. Control days were chosen at random and reassigned if they coincided with adverse weather unrepresentative of other drives that week. Videos from these days were coded and included in the analyses under the Experimental Control condition (see Figure 11).

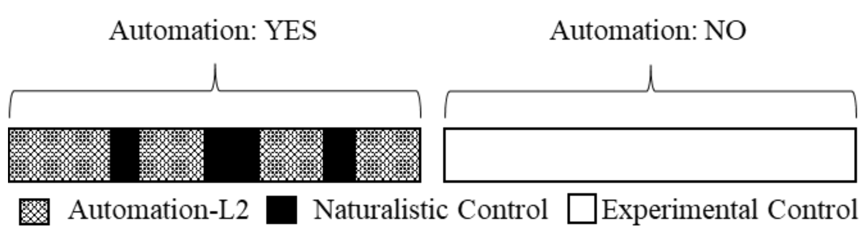

Figure 11. Automation and Control conditions. In the Automation: YES condition participants selected when to use automation (Automation-L2) and when to drive manually (Naturalistic Control). In the Automation: NO condition participants were instructed not to use automation.

Naturalistic Day. Due to the large volume of video collected during daily commutes, just one day each week was selected from the remaining days for coding (Automation: YES). This day was chosen at random, with the constraint that its weather closely matched that of the Experimental Control Day (e.g., if it was sunny in the control day then the Naturalistic Day was also sunny). Instances of automation use during this day were coded and analyzed under the Automation-L2 condition, while instances of manual driving were coded and analyzed under the Naturalistic Control condition (Figure 11).

Video Handling. After 6 to 8 weeks of naturalistic driving, participants completed the final experimental session and returned their vehicles. SD video cards were then removed from the vehicle cameras and processed for analysis. Prior to uploading and saving the videos, files were cleaned to eliminate all non-commute driving on the regional interstates. Thus, the length or duration of driving each day was roughly equivalent. Furthermore, video files were combined into AM and PM commutes for each day. Cleaned video files capturing highway driving during AM and PM commutes were uploaded to a secure server for analysis.

Video Blinding. To minimize potential bias among coders, several procedures were implemented to blind them to the experimental condition present in the videos. This primarily involved a two-pass approach to video coding, wherein all behaviors except the state of automation were coded during the first pass. Automation indicators were obscured during video playback using strips of painter's tape positioned on the monitor. During the second pass, the tape was removed, and the automation state was recorded. All other indicators of the experimental condition were eliminated, including file labels and other electronic data, until the final completion of each participant record, after which condition information was reintegrated into the record.

Video Coder Training. Video reduction took place over approximately 1.5 years, involving several different coders. To ensure coding consistency, new coders underwent a three-week peer-to-peer training focused on coding quality and consistency, established through redundant coding and regular checks of inter-rater reliability.

Additional steps were taken to further ensure coding consistency. First, with each new participant, reductionists group-coded video from at least one drive, allowing them to determine if any unique or challenging behavior was likely to arise from the participant and to reach a consensus on how to handle such behavior if observed. Second, at least one video was group-coded each week, regardless of whether it was from a new participant. This strategy led to a target of 40% of all videos being redundantly coded. Finally, inter-rater reliability was continuously assessed using an Excel-generated script and BORIS's kappa score generator. An acceptable kappa score on the unaggregated raw coding was set to 0.6, which when collapsed by coded task led to scores above 0.9. If significant differences were found between observations, coders would review the video as a group to identify and correct discrepancies.

Videos were generally coded in real-time, but coders often had to re-watch complex sections to accurately code the start and stop of overlapping behaviors. This demanding process required significant focused attention, so coders were encouraged to take breaks as needed to maintain high performance levels.

Video Coding Rubric. A comprehensive and systematic coding scheme was adapted from the task-modality distraction testing paradigm developed by Strayer et al. (2017). This rubric allowed for systematic differentiation of various participant behaviors, and it was used to create a video coding dictionary to guide video reduction. This dictionary included clear definitions of all behaviors of interest and examples of each behavior. To address the four sets of questions posed by this research, the following coding scheme was developed:

- *Level 2 Automation Usage* – Instances of automation engagement and disengagement were coded using an instrument-facing camera positioned in each vehicle to capture an image of the screen displaying automation state (i.e., when system was on/engaged or off/disengaged). The use of automation activation controls served as a redundant marker of automation use and helped to disambiguate system state when icon visibility was poor.
- *Warnings* – System warnings were marked as discrete events in the data file.
- *Driving Demand* – Driving demand was operationalized as the sum of concurrent poor conditions present, with low demand including no poor conditions, moderate demand including one poor condition, and high demand including two or more poor conditions. Poor conditions related to weather, traffic, construction, emergency vehicles, or other events that could adversely affect driving and automation system function.
- *Fatigue and Fidgeting* – Fatigue and fidgeting behavior were coded continuous events, meaning that the coders marked the start and stop times of each specific behavior. For fatigue, this included marking the beginning and end of visible signs of sleepiness, such as yawning, heavy eyelids, and head nodding. For fidgeting, this including identifying the beginning and end of body movements lasting more than 3 seconds, such as touching the face, neck, head/hair, or moving hands to and from the steering wheel. Additionally, reaching and grabbing, and eating and drinking behaviors were grouped into fidgeting.
- *Distraction and Inattention* – This was a comprehensive class of behaviors, and detailed data were collected on each instance. Five core distracting activities were defined: text messaging, calling and dialing, radio listening, navigation, and video interaction. Each of these activities was coded for modality of interaction, which included visual-manual or auditory-vocal, and interface, which included cell phone or in-vehicle-information-system (IVIS). Coders recorded the start and stop times of these distracting activities, capturing the frequency and duration of each behavior. This allowed for a detailed analysis of distraction and inattention on a trip-by-trip basis, as well as for the entire day's drive. Furthermore, the aggregate measure was used to

provide an overall assessment of secondary task engagement by summing all secondary task interactions across the various activities.

Appendix A provides additional information about the coding scheme for distraction and inattention behaviors, which includes definitions, examples, and coding instructions to ensure clarity and consistency among coders.

Statistical Analysis. BORIS provided a .csv file as output for each coded video, listing details for each task in separate columns with one row per task. To analyze this data, several R scripts were generated that converted outputs into a time-series format, with tasks organized in columns, time represented by each row, and a binary task state indicator listed in each column. This organization allowed for tasks to be combined and collapsed as required for various analyses. Transformations were primarily carried out using base R (R Core Team, 2022) and packages within the Tidyverse (Wickham et al., 2019).

To account for sources of non-independence in the data (i.e., repeated measures within each participant) and any missing data, data were analyzed with linear mixed-effects models using the lmer function found in the lmerTest library (Kuznetsova et al., 2017). Participant ID and the AM/PM drive indicator were included in all models as random intercepts and, where appropriate, Session and Condition were input as predictor variables (see bulleted list below). Outcome variables were dictated by the specific question and included fatigue, fidgeting, secondary task, etc., as described in the video coding rubric. Likelihood ratio tests were run using the ANOVA function in the stats package to test the significance of all effects, and pairwise comparisons were run using the contrasts function of the lmerTest library. Significance levels for all analyses were set at $p < .05$, $p < .01$, and $p < .001$.

Predictor variables of interest were as follows:

- ***Session*** – fixed continuous factor. This was the numerical indicator of week (e.g., 1, 2, 3, etc.). Session was handled as a continuous fixed factor for all relevant analyses but treated as discrete for plotting purposes.
- ***Condition*** – fixed discrete factor with 3 levels. The “Automation: Yes” day provided two levels of Condition, which were Automation-L2 and Naturalistic Control. The “Automation: No” day provided the third level of Condition which was Experimental Control. Condition was also entered as a discrete fixed effect in relevant models.
- ***Subject*** – random discrete factor. This was the simple subject identifier. Subject was modeled as a random intercept in all analyses.
- ***AM_PM*** – random discrete factor. Simple identifier of the AM or PM drives (e.g., the morning and evening commutes for each participant). AM_PM was also entered as a random slope in all analyses.

Results

Data Overview

Video Record. Video results were obtained from 30 participants, resulting in a total of 670 videos (353 Naturalistic; 317 Experimental Controls). Within the experimental control days, 26 videos contained instances of Level 2 automation, indicating a misunderstanding of the task for that day. These videos were excluded from the analysis leaving 291 baseline videos. For each week 1 through 8, data from the following number of participants were available to code: 26, 30, 30, 28, 27, 22, 17, and 12, respectively. Of the 670 available videos, 308 were double-coded and 80 were coded by three or more different reductionists. In total, 1060 coding records were entered into the analysis. Results from redundantly coded videos were averaged.

Automation Use. Coding only one Naturalistic and one Experimental Control Day per week resulted in 297 total hours of coded video with just over half collected during Naturalistic driving (161 hours). Overall, participants used automation between 25% and 99% of the time during the naturalistic observation period, resulting in 124 hours of video where participants engaged Level 2 automation.

Appendix B includes mean and standard error data for all of the analyses in the subsections below.

Level 2 Automation Usage

Two research questions related to operator trust and usage of Level 2 Automation:

- AU1 – Does experience (based on number of weeks of system use) with automation change the frequency with which drivers activate the automation?
- AU2 – How does the time to re-engage the system (after a system disengagement) change with experience?

To address these questions a mixed effects model was generated that treated frequency of usage and reengagement time as outcome measures with Week as a fixed effect and Subject and AM/PM drives as random effects. Re-engagement time was quantified as the amount of time between disengagement of automation and the participant actively re-engaging it, reflecting a difference score that would be expected to decrease over time if practice influenced reengagement.

Regarding the first question (AU1), results indicated that Week did not significantly predict usage frequency, $F(1,309) = 1.88$, $p = .17$ (see Figure 12). Similarly, there was no significant effect of Week on reengagement time, $F(1, 284) = 0.10$, $p = .75$ (AU2). Together these findings suggest that participants in this research maintained a similar level and interaction pattern of automation use throughout the 6 to 8 weeks of observation.

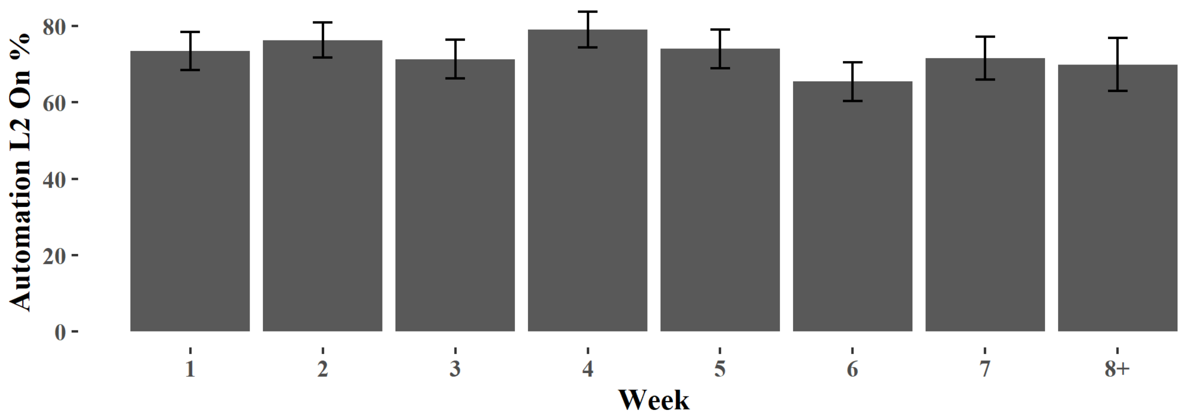

Figure 12. Automation-L2 Usage: Automation-L2 use by Week.

Warnings and Driving Demand

Two questions related to the misuse and unintended consequences of Level 2 automation use were identified:

- WD1 – Does the frequency of system warnings change over time?
- WD2 – Does the frequency of automation use change during poor driving conditions?

Regarding WD1, system warnings occurred when drivers either failed to apply sufficient tension to the steering wheel (Tesla, Nissan, Volvo), or failed to maintain their eyes on the forward roadway (Cadillac). A mixed effects model was generated that treated system warning frequency as the outcome measure with Week as a fixed effect and Subject and AM/PM drives as random effects. Of those participants that experienced warnings, the range of warning frequencies was 0.03 to 1.93 per minute. Results indicated that for these participants, warning frequencies increased during the observation period, $F(1, 423) = 9.84$, $p = .002$, suggesting that as drivers became more comfortable with automation they paid less attention to the driving task (see Figure 13, left panel).

To address the second question (WD2), we looked at the relationship between driving demand, as coded by the number of poor conditions that were present in the driving environment and the use of automation. The poor conditions analyzed included traffic impairing driving speed, weather (rain, snow, ice, or fog), road construction, emergency vehicles, and other outside influences affecting driving. Among these, traffic impairing driving speed was the most common poor condition observed. It is important to note that poor weather conditions can cause the system to disengage; however, in practice, this was rarely observed.

Results indicated that as the demand of the driving task increased, from Low to Moderate to High, the prevalence of automation use decreased ($F(2, 1072) = 9.93$, $p < .001$). These results suggest that drivers were aware of roadway demand and were less likely to use Level 2 automation when the roadway demands were higher (see Figure 13, right panel).

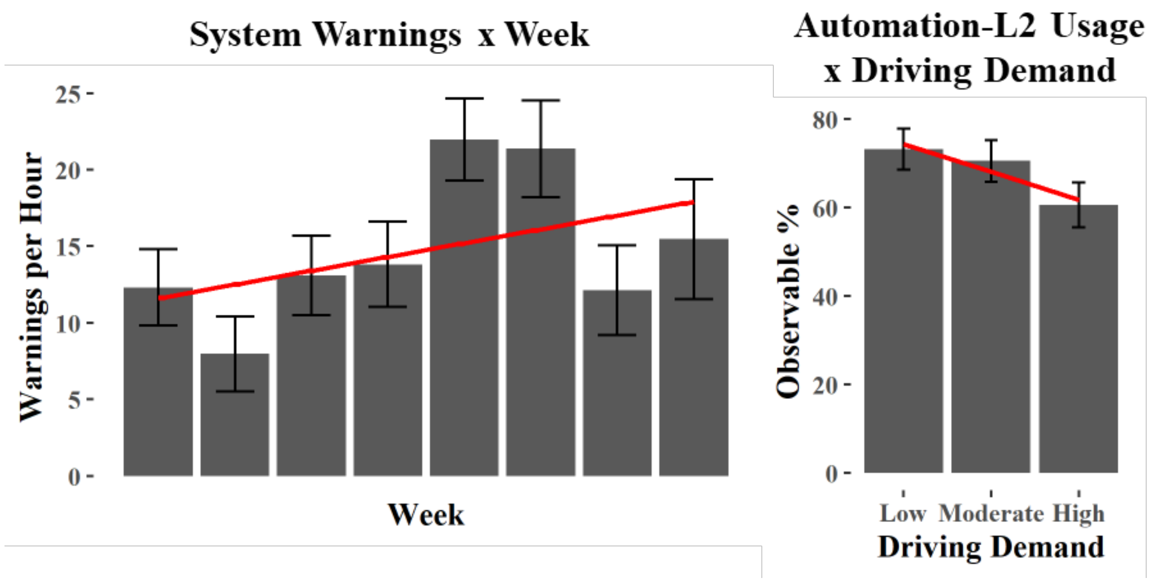

Figure 13. Misuse and Unintended Consequences: System Warnings by Week (left panel) and Automation use by Driving Demand (right panel).

Fatigue and Fidgeting

Two classes of observable behaviors related to arousal were coded in the video to address the following:

- *FF1* - Do drivers show increased signs of fatigue when using automation and does this change over time?
- *FF2* - Do drivers show increased signs of fidgeting when using automation and does this change over time?

For each question, three linear mixed effects models were generated with either Fatigue or Fidgeting behaviors treated as the outcome measure. Week, Condition, and Week by Condition were treated as fixed effects, while Subject and AM/PM drives were treated as random effects. Instead of using a single model with all the predictors included, three separate models were conducted for each predictor (Week, Condition, and Week X Condition) to reduce complexity and provide a clearer interpretation of the individual effects. Pairwise comparisons were completed on the effect of Condition (Automation-L2, Experimental Control, and Naturalistic Control) to determine how the different conditions affected fatigue and arousal.

Regarding the first question (FF1), there was a main effect of Condition on Fatigue ($F(2,882) = 3.84$, $p = .02$), but no significant effect of Week ($F(1,901) = 1.12$, $p = .29$) or interaction ($F(2, 878) = 0.27$, $p = .76$). As shown in Figure 14, pairwise comparisons indicated that Fatigue was higher in the Automation-L2 condition than in the Naturalistic Control condition. However, it did not differ statistically between the Automation-L2 condition and the Experimental Control condition, although it was nominally higher (see also Table 5).

With respect to the second question (FF2), there was a main effects of Condition ($F(2,868) = 11.8$, $p < .001$) and Week ($F(1,894) = 10.8$, $p = .001$) on fidgeting behaviors, but no interaction ($F(2, 865) = 0.54$, $p = .58$). As shown in Figure 14, pairwise comparisons indicated that the interpretation of fidgeting behavior depended on the type of control that was used. Drivers fidgeted relatively more during the Automation-L2 condition compared to the Naturalistic Control condition but showed no relative difference when compared to the Experimental Control condition (see also Table 5). The effect of fidgeting over week was relatively straightforward; fidgeting behaviors increased throughout the observation period (see Figure 15).

In summary, a relative increase in fatigue and fidgeting was observed in the Automation-L2 condition compared to the Naturalistic Control condition. Fidgeting behavior was also found to increase throughout the 6 to 8 weeks of study participation. However, when compared to the Experimental Control condition, neither fatigue nor fidgeting appeared to be affected by automation use.

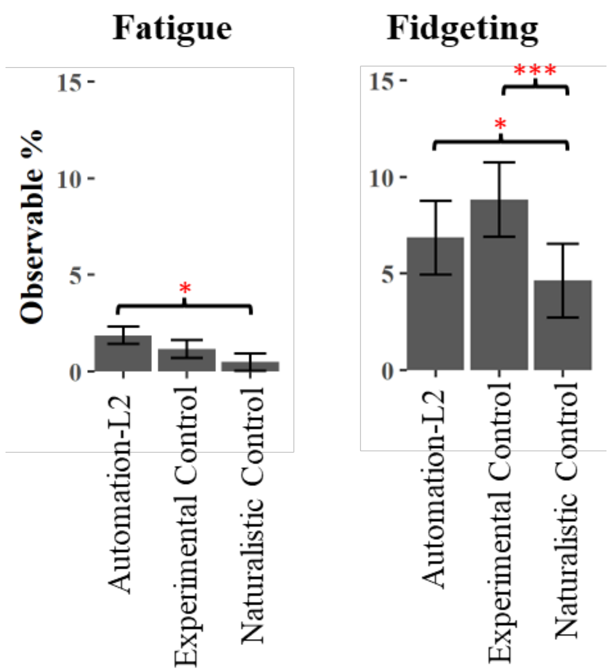

Figure 14. Fatigue and fidgeting by condition. Significant pairwise comparisons are indicated in the chart with asterisks.

Table 5. Pairwise comparisons between levels of condition for fatigue and fidgeting behavior.

Pairwise Comparisons		t ratio	df	p value
Fatigue	Automation-L2 vs. Experimental Control	-1.38	879	.35
	Automation-L2 vs. Naturalistic Control	-2.77	876	.02*
	Experimental Control vs. Naturalistic Control	1.35	878	.37
Fidgeting	Automation-L2 vs. Experimental Control	2.27	874	.06
	Automation-L2 vs. Naturalistic Control	-2.60	873	.03*
	Experimental Control vs. Naturalistic Control	4.84	874	<.001**

Note. $p < .05$*, $p < .01$**, $p < .001$***

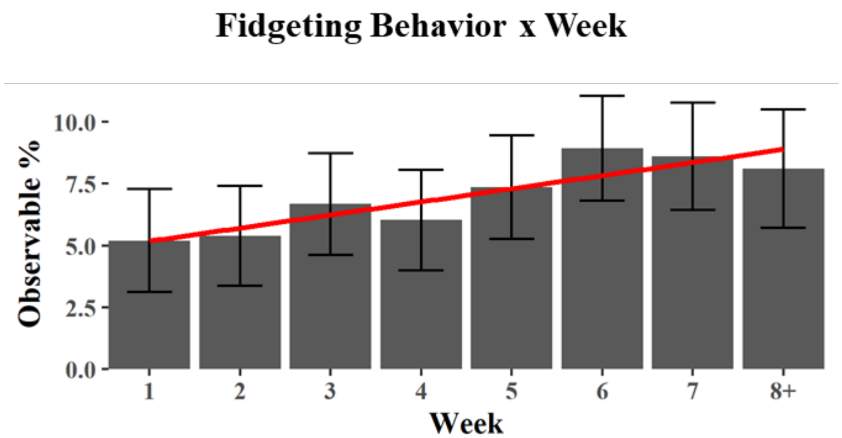

Figure 15. Observable percentage of fidgeting by Week.

Secondary Task Engagement

Two classes of observable behaviors related to secondary task engagement were coded in the videos. These were used to address two sets of questions on driver secondary task engagements during automation use.

- ST1 – How does the frequency of secondary task use (non-driving related) change during Level 2 use and over time?
- ST2 - How does the frequency of task type, mode of interaction (voice vs manual), and interface (cell phone vs. In Vehicle Information System [IVIS]) change during Level 2 use and manual driving?

To address these questions, several distinct secondary task behaviors were coded, including radio listening, text messaging, phone conversation, navigation, and video interaction. Each of these tasks was analyzed individually and in aggregate, which represents the sum of all secondary task interactions.

Regarding the first question (ST1), results indicated a main effect of Condition on the task aggregate, radio listening, and text messaging tasks while the effect of Week was significant on the task aggregate and text messaging tasks (see Figure 16 and Table 6). Pairwise comparisons of the task aggregate showed that greater secondary task engagement was observed in the Automation-L2 condition compared to the Naturalistic Control condition but not the Experimental Control condition; both controls differed from each other (See Figure 17 and Table 7). Pairwise comparisons of the radio listening task indicated that it was more common in the Automation-L2 condition than either of the control conditions. Finally, text messaging was found to be more common in the Automation-L2 condition than in the Naturalistic Control condition. A Condition x Week interaction was also observed on the navigation task; however, because neither of the main effects of Condition and Week were significant, the interpretation of this interaction is unclear.

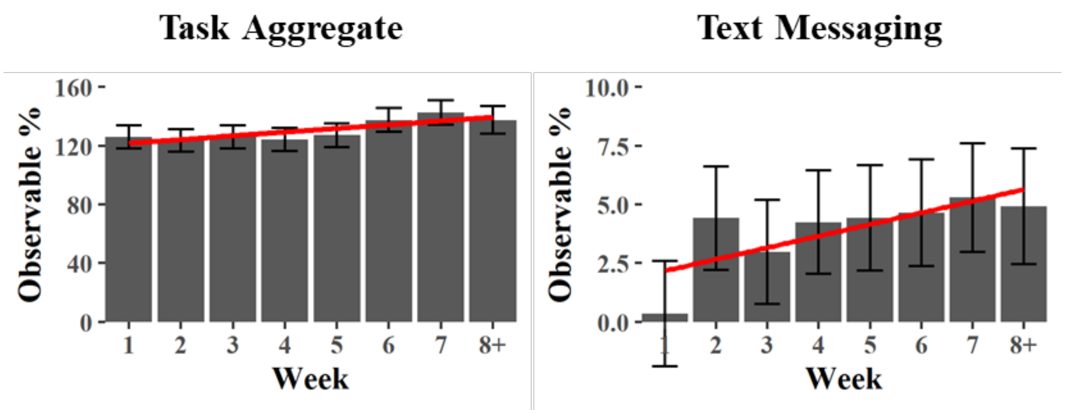

Figure 16. Secondary task engagement to task aggregate and text messaging. The y-axis indicates the percentage of time that any of the tasks were active. In the case of the task aggregate, the observable percentages over 100 indicate that that on average more than one task was active at any given time.

Table 6. Statistical tests for secondary tasks.

Main Effects	Condition	Week	Condition x Week
Task Aggregate	$F(2,873) = 15.6$, $p < .001$***	$F(1,883) = 12.8$, $p < .001$***	$F(1,869) = 1.18$, $p = .03$*
Radio Listening	$F(2,872) = 9.59$, $p < .001$***	$F(1,883) = 0.14$, $p = .71$	$F(2,868) = 0.24$, $p = .79$
Text Messaging	$F(2,871) = 3.27$, $p = .04$*	$F(1,878) = 10.7$, $p = .001$***	$F(2,868) = 0.23$, $p = .80$
Phone Conversation	$F(2,873) = 1.00$, $p = .33$	$F(1,894) = 2.86$, $p = .09$	$F(2,870) = 0.33$, $p = .72$
Navigation	$F(2,871) = 0.35$, $p = .71$	$F(1,880) = 0.84$, $p = .36$	$F(2,868) = 4.93$, $p = .007$**
Video Watching	$F(2,872) = 0.33$, $p = .72$	$F(1,882) = 0.05$, $p = .82$	$F(2,868) = 0.03$, $p = .97$

Note. $p < .05$, $p < .01$**, $p < .001$****

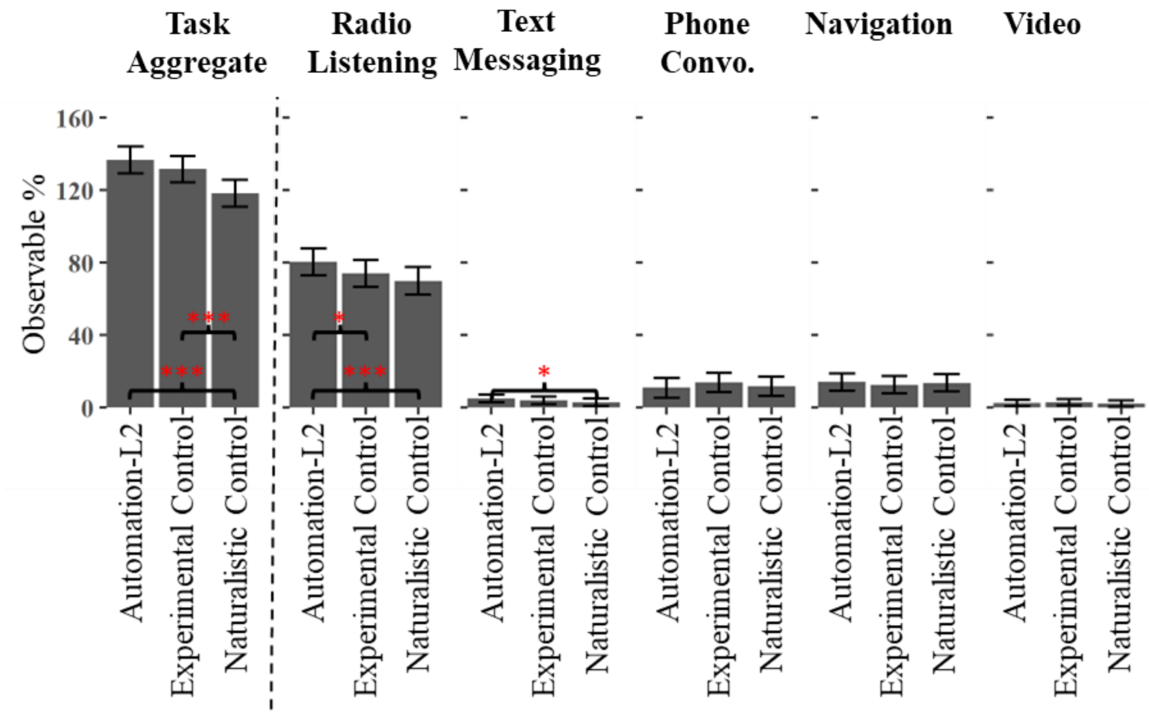

Figure 17. Secondary task engagement by condition, collapsed across Week.

Table 7. Pairwise comparisons between condition by different secondary tasks.

Pairwise Comparisons		t ratio	df	p value
Task Aggregate	Automation-L2 vs. Experimental Control	-1.47	872	.31
	Automation-L2 vs. Naturalistic Control	-5.41	872	<.001***
	Experimental Control vs. Naturalistic Control	-5.41	872	<.001***
Radio	Automation-L2 vs. Experimental Control	-2.61	872	.03*
	Automation-L2 vs. Naturalistic Control	-4.35	872	<.001***
	Experimental Control vs. Naturalistic Control	1.67	872	.22
Texting	Automation-L2 vs. Experimental Control	-1.14	872	.49
	Automation-L2 vs. Naturalistic Control	-2.55	871	.03*
	Experimental Control vs. Naturalistic Control	1.38	872	.35
Phone Conversation	Automation-L2 vs. Experimental Control	1.43	874	.33
	Automation-L2 vs. Naturalistic Control	0.37	873	.93
	Experimental Control vs. Naturalistic Control	1.07	873	.53
Navigation	Automation-L2 vs. Experimental Control	-0.82	872	.69
	Automation-L2 vs. Naturalistic Control	-0.28	872	.96
	Experimental Control vs. Naturalistic Control	-0.55	872	.85
Video	Automation-L2 vs. Experimental Control	0.26	872	.96
	Automation-L2 vs. Naturalistic Control	-0.54	872	.85
	Experimental Control vs. Naturalistic Control	0.80	872	.70

Note. p < .05, p < .01**, p < .001****

To address the second set of questions (ST2), all tasks were collapsed according to the modality of interaction (either auditory-vocal or visual-manual). This grouped all secondary task interactions that occurred either through the vehicle interface or through a secondary device such as a smartphone. Results indicated no significant effects of Condition but a main effect of Week on the aggregate of visual-manual tasks (see Figure 18 and Table 8). Thus, irrespective of the condition, drivers were more likely to engage in secondary visual-manual tasks with each week of vehicle use. Data were then reaggregated according to the interface (either smartphone or vehicle IVIS). Results indicated a main effect of Condition on vehicle IVIS use, but pairwise comparisons failed to indicate any significant contrasts (Table 9). Results also showed a main effect of Week on smartphone interactions, with smartphone interactions significantly increasing during each week of the study. Taken together, these findings indicate that drivers increased their visual-manual interactions with their smartphones with each week of the study, but the driving condition, either with or without automation, did not seem to affect the findings.

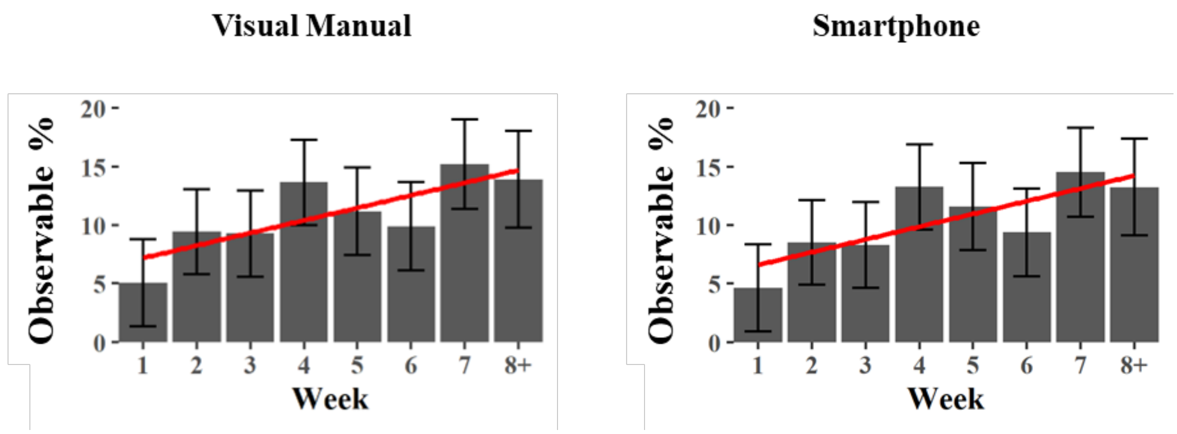

Figure 18. Frequency of engagement in visual-manual and smartphone interactions across week.

Table 8. Statistical tests for modality of interaction and interface type.

Main Effects	Condition	Week	Condition x Week
Auditory-Verbal	$F(2,878) = 1.71, p = .18$	$F(1,890) = 0.17, p = .68$	$F(2,874) = 0.37, p = .69$
Visual-Manual	$F(2,871) = 2.86, p = .06$	$F(1,880) = 15.4, p < .001$***	$F(2,868) = 0.28, p = .76$
Smartphone	$F(2,871) = 2.29, p = .10$	$F(1,880) = 16.3, p < .001$***	$F(2,868) = 0.10, p = .91$
Vehicle IVIS	$F(2,877) = 3.45, p = .03$*	$F(1,902) < .001, p = .98$	$F(2,873) = 0.44, p = .65$

Note. p < .05, p < .01**, p < .001****

Table 9. Pairwise comparisons for modality of interaction and interface by condition.

Pairwise Comparisons		t ratio	df	p value
Auditory-Verbal	Automation-L2 vs. Experimental Control	1.15	882	.48
	Automation-L2 vs. Naturalistic Control	-0.69	878	.77
	Experimental Control vs. Naturalistic Control	1.83	880	.16
Visual-Manual	Automation-L2 vs. Experimental Control	-0.72	872	.75
	Automation-L2 vs. Naturalistic Control	-2.34	872	.05
	Experimental Control vs. Naturalistic Control	1.58	872	.25
Smartphone	Automation-L2 vs. Experimental Control	-0.58	872	.83
	Automation-L2 vs. Naturalistic Control	-2.08	872	.11
	Experimental Control vs. Naturalistic Control	1.47	872	.31
Vehicle IVIS	Automation-L2 vs. Experimental Control	-0.58	872	.83
	Automation-L2 vs. Naturalistic Control	-2.08	872	.10
	Experimental Control vs. Naturalistic Control	1.47	872	.31

Discussion

The primary aim of this research was to better understand driver behavior when using Level 2 vehicle automation as driver exposure to and familiarity with the technology grows. Video data was collected and analyzed on 30 drivers, each of whom drove one of five partially automated (SAE Level 2), instrumented research vehicles for 6 to 8 weeks. Critically, participants were instructed not to use automation on one day each week (the experimental control day). This experimental control was compared with a more traditional naturalistic control condition where, for one reason or another, participants chose not to use automation even though it was available to them. Driver behavior in each of the two control conditions was then contrasted with behavior observed during automation use. Analyses center on four topical research areas: automation use, warnings and driving demand, fatigue and fidgeting, and secondary task engagement. Results from this hybrid research approach provide data that both bolster and challenge previous findings in each of these areas.

Use of Automation

Drivers in this research used Level 2 vehicle automation more than 70% of the time, an amount that stayed relatively consistent over the 6- to 8-week observation period. This high level of automation use was likely driven by the requirement that participants commute at least 40 minutes each day, to and from work, to be eligible for study participation, along with instructions provided to the participant to use the system as often as they were comfortable. Additionally, sections of each commute that were not on controlled access highways were not coded. These usage trends compare favorably with results obtained by Stapel et al. (2022), who observed highway use of Level 2 automation ranging from 57% to 63% with no reduction in use over a 12-week observation period. This steady use rate of automation suggests that participants remained comfortable with the automation performance, the monitoring requirements, and the potential benefits that it may have provided while driving.

Warnings and Driving Demand

Across the 6 to 8 weeks of automation use, an increase in the frequency of system warnings was observed as drivers become more experienced with the Level 2 vehicle automation. While the specific cause of this increase was not clear, the finding suggests an increased comfort with the automation and a tendency toward a more relaxed automation monitoring strategy over time. Warnings were found to vary widely between individuals. Some drivers rarely, if ever, experienced warnings while others received several warnings per minute and treated them as if they were simply a nuisance that could quickly be quieted through gentle pressure on the steering wheel.

Poor conditions related to weather, traffic, construction, emergency vehicles, or other events that would reasonably be expected to adversely affect driving were coded and aggregated to form a measure of driving demand. Results indicated that drivers were less likely to use vehicle automation when driving demands were higher. This suggests that drivers were aware of changes in roadway demand and were more likely to use automation when it was safer to do so. This is a key finding that helps to resolve differences that were observed in the experimental and naturalistic control conditions (discussed below).

Fatigue and Fidgeting

As discussed, a major safety concern with the use of Level 2 vehicle automation is that it may lead to an increase in driver fatigue. Findings were mixed in the current data. When contrasting the fatigue observed in the Automation-L2 condition with the Experimental Control condition, it was found that automation use did not increase either fatigue or fidgeting behaviors. However, an increase in fatigue *was* observed when comparing the Automation-L2 condition with the Naturalistic Control condition. Additionally, a decrease in fidgeting was also observed when comparing the Automation-L2 condition with the Naturalistic Control condition.

The finding that automation was or was not associated with fatigue and fidgeting depending on the benchmark used adds an interesting nuance to the literature and reinforces the importance of a strong and valid control condition. Automation is often singled out as the cause of fatigue in popular videos where drivers are seen to be sleeping as the vehicle drives itself. While this is clearly dangerous, it is not clear from such cases whether these drivers would have done the same under manual control and possibly driven off the road. If this were known, one might conclude that the automation prevented a fatigue-related crash. The answer to the question of whether automation does or does not lead to driver fatigue hinges on the question: *compared to what*? Compared to a strong experimental control, these data suggest that automation may not lead to levels of fatigue suggested by online videos and some prior research (ABC, 2023; Vogelpohl et al., 2019; Lu et al., 2021).

Secondary Task Engagement

One of the most frequently reported findings related to automation use is that it leads to an increase in the frequency of secondary task engagement. This is a significant safety concern for lower-level automated vehicles (Levels 1 and 2 and to a lesser extent Level 3), as secondary task use has been shown to reduce a driver's ability to take over vehicle control quickly and safely when required. Patterns of increased secondary task use with automation were found in the current study. Again, however, the nature and potential severity of these findings depended on which control condition is used for the comparison. When using the stronger Experimental Control, the results indicated that drivers were more likely to listen to the radio when automation was engaged, which, based on our prior work, is not a significant safety concern (Strayer et al., 2013). However, when compared to the Naturalistic Control condition, an increase in text messaging and the task aggregate, essentially the sum of all secondary task interactions, is also seen. Taken together, these findings indicate several notable secondary task trends, but again, they do *not* necessarily show the concerning increase in distracting behaviors that some have suggested occurs with vehicle automation.

Experimental versus Naturalistic Controls

The strength of naturalistic research is that it eschews experimental intervention in favor of naturalistic observation. However, two major limitations of the naturalistic method make it a poor approach to resolve the behavioral profile associated with automation use. The first limitation is that drivers may selectively choose when to engage in secondary tasks for reasons that are important but not, perhaps, obvious. This is especially problematic with automation, given the possibility that drivers use automation only when they feel it is appropriate. The selection of baseline events from the remaining drives (i.e., when automation is not engaged) is therefore confounded by the fact that drivers may feel that they are unsuitable for automation use. The second limitation of uncontrolled naturalistic designs for evaluation of automation use is that they often rely on the use of machine vision to automatically detect vehicle states. While these approaches have improved greatly, they require significant training data to implement and are sensitive to visual noise. The hybrid design implemented in this research helps ameliorate both issues.

Functional Vigilance

Drivers in the Experimental Control condition exhibited a behavioral profile that was markedly different from that observed in the Naturalistic Control condition. In most cases, behavior in the Experimental Control fell in between the Automation-L2 and Naturalistic Control conditions (e.g., fatigue, the secondary task aggregate, radio listening, and text messaging). In the case of fidgeting, it was shown that the most fidgeting was observed in the Experimental Control condition when drivers were not allowed to use automation. Considering the finding that automation use was lower when driving demands were higher, then one compelling and plausible explanation for these findings is that driving demand may mediate the relationship between automation use and secondary task engagements such that drivers may be less likely to use automation and less likely to engage in secondary tasks when driving demands are higher.

Additionally, the observation that results from the Experimental Control condition often fell between the Automation-L2 and Naturalistic Control conditions fits with the fact that driving demand was experimentally controlled to be comparable in the Automation: YES (Automation-L2 + Naturalistic Control) and the Automation: No

(Experimental Control) days. Overall, these results are consistent with the hypothesis that drivers maintain a level of functional vigilance when using automation that allows them to naturally slip in and out of automation as roadway demands change (Fridman et al., 2017).

Behavioral Adaptation

The current findings, when interpreted through the lens of the three-phase model of vehicle automation proposed by Dunn et al. (2019), suggest a trajectory of behavioral adaptation. A key indication of this adaptation is the observed increase in system warnings over time. This is consistent with the "post-novelty operational phase" of the model, where drivers, having developed a mental model of system operation, test the boundaries of the system's capabilities. The rising frequency of warnings suggests that drivers are growing more comfortable with the system, exploring its limitations, and responding to these cues as they navigate the automation. This trend might reflect a maturation of understanding the system, underscoring the notion of a learning curve associated with the use of advanced driver-assistance systems.

The increase in secondary task usage, reflected in the task aggregate (all secondary tasks combined) and the text messaging task, also corresponds with the behavioral adaptation concept. It indicates that as drivers grow more familiar with the system, they are more likely to engage in secondary tasks over time. However, the lack of interaction between secondary task engagements and Condition (Automation-L2, Experimental Control, Naturalistic Control) suggests that these behaviors may not be a direct consequence of over-reliance on the automation system. Instead, they might reflect a general trend of drivers becoming more comfortable with multitasking in these specific research vehicles over time, regardless of automation capabilities.

The rise in visual-manual task interactions with a smartphone over time further supports this trend. However, as with the task aggregate and text messaging behaviors, there was no interaction with Condition. This might have been expected if drivers were transitioning into overreliance within the "experienced user phase" of the model. The absence of this interaction in our data suggests that while drivers are engaging more in secondary tasks over time, they are not necessarily doing so as a direct consequence of using the automation system, nor are they doing so in a manner inconsistent with their driving behavior when automation is not available.

Overall, there is some evidence of behavioral adaptation in the current data, but it appears to be mostly functional in nature. That is, drivers in this research showed signs of adapting their behavior to automation but did not necessarily show strong signs of overreliance on the system to the point of compromising safety.

Limitations

While this study offers important insights into driver behavior during Level 2 automation, there are several limitations that should be acknowledged. These limitations pertain to the sample size, the duration of the study, the potential for Hawthorne effect, and the specific automation technology employed in the research vehicles, among other possible limitations.

- **Sample size**: The sample size of 30 participants may not be representative of the broader population of drivers. While the study aimed to recruit a diverse group of participants, a larger sample would allow for a more accurate representation of the general population, increased the statistical power, and increase the generalizability of the findings. Future studies should consider increasing the number of participants to better understand the effects of Level 2 automation on a wider range of drivers.
- **Study duration**: The study duration of 6 to 8 weeks may not be sufficient to fully understand the long-term effects of Level 2 automation on driver behavior. It is possible that drivers' behaviors may continue to evolve beyond the study period, as they become more familiar with and reliant on the technology. Future research should explore longer observation periods to better understand how drivers adapt to automation over time.
- **Hawthorne effect**: The potential for the Hawthorne effect should also be considered. Participants were aware that they were part of a study, which may have influenced their behavior during the observation period. It is possible that drivers may have behaved more cautiously or differently than they would have under normal, unobserved circumstances (e.g., if they were driving their own and not a study-owned vehicles). Also, due to IRB requirements, drivers in the current study received much more comprehensive a priori training regarding the vehicle automation compared to what real world owners of these vehicles might receive. These represent tricky issues that apply broadly to most driving research, however, future research should consider methods for mitigating the impact of the Hawthorne effect on driver behavior.

Future Directions

The findings presented in this study highlight several intriguing aspects that warrant further investigation. Firstly, understanding why drivers opt to use vehicle automation in certain scenarios and not others is crucial. Moreover, the factors that prompt drivers to disengage automation in favor of manual driving need to be explored. These questions hold considerable relevance not only for the design, marketing, and general uptake of vehicle automation but are also pivotal in the appropriate selection of baseline driving behavior in naturalistic driving research. These factors were touched upon in this study, but a comprehensive understanding is yet to be reached.

Secondly, future research should aim to clearly define the potential benefits and implications of vehicle automation for drivers. A crucial question is to determine what drivers should expect to gain from vehicle automation and how they should utilize their freed attention. In this study, for instance, there was an uptick in radio listening during automation use; while radio use is seemingly harmless, is engaging in activities like text messaging or video viewing appropriate when automation is handling most of the driving? These are important questions to address moving forward and as new forms of automation are developed and implemented.

Finally, apart from an increase in radio listening with Level 2 automation, these data suggest that drivers may be no more likely to engage in secondary tasks when using automation than they are when manually driving. This is a provocative finding that needs to be further investigated. Understanding the true impact of automation on driver behavior, particularly in terms of secondary task engagement, is crucial for designing effective policies and guidelines to ensure the safe integration of automated vehicles into our transportation system.

PART 3 (Survey Study): Automated Driving Experiences, Attention, and Intentions Following Extensive On-Road Usage of a Level 2 Automation Vehicle[3]

The survey phase of this study aimed to gauge driver perceptions and attitudes along a number of dimensions as they gained experience with the vehicles. Additional background and motivation are provided in the sub-sections below, followed by an overview of the main objectives of Part 3.

Research on Perceptions of Automated Driving Systems

Most of the research on perceptions of automated driving conducted during the last decade has focused on drivers' knowledge of and evaluations of advanced driver assistance systems (e.g., Beggiato & Krems, 2013; Gaspar & Carney, 2019; Körber et al., 2018; McDonald et al., 2018; Sanbonmatsu et al., 2018; Walker et al., 2018). Owners of partially automated vehicles generally report high levels of trust in the technology and perceive it to be useful and safe (e.g., McDonald et al., 2018). The acceptance of automated driving systems develops quickly with usage (e.g., Beggiato et al., 2015; Walker et al., 2018). Nevertheless, consumers sometimes lack understanding of the limitations of automated systems and over trust the technology (e.g., McDonald et al., 2018). Not surprisingly, the perceived trustworthiness, safety, and usability of automated systems diminish significantly when passengers experience on-road failures of the automation (e.g., Xu et al., 2021).

Both self-report and observational studies suggest that attention to driving is often lower and secondary activities are often greater during automated driving (e.g., Banks et al, 2018; Dunn et al., 2019; Gaspar & Carney, 2019; Körber et al., 2018). The perceived trustworthiness, usefulness, and safeness of the technologies have been linked to lower engagement during automated driving and delays in the manual take-over of vehicles, as well as stronger behavioral intentions to use automated systems (e.g., Körber et al., 2018; Morales-Alvarez et al., 2020; Payre et al., 2016; Xu et al., 2021).

Relatively few studies have examined the effects of automation on the driving experience. Limited research suggests that advanced assistance systems may reduce the stress and increase the comfort of driving. However, some of these studies examined the self-reports of purchasers of automated vehicles (e.g., Eichelberger & McCartt, 2016) who are likely to be favorably biased in their evaluations since they were willing to pay for the automation and may have sought to rationalize their expenditures. Other studies of

[3] Section contributors: David M. Sanbonmatsu, Kaedyn W. Crabtree, Amy S. McDonnell, Joel M. Cooper, and David L. Strayer

the experience of automated driving have been limited to the examination of simulated rather than actual on-road travel (Hartwich et al., 2018) or passengers of automated vehicles rather than actual drivers (Liu et al., 2021).

Some of these shortcomings were addressed by a previous on-road study of automated driving (Cooley et al., 2022). Participants who had never used advanced driving assistance systems drove a Tesla Model 3 sedan with Level 2 automation engaged or not engaged on a four-lane interstate freeway. After a single drive with the automation, the participants reported that driving was more enjoyable and less stressful during automated driving compared with manual driving. They also indicated that they were less anxious and nervous, and able to relax more with the automation. Participants did not report reduced attention to driving or greater engagement in secondary activities when the automated systems were operating. The positive experiences of the first-time users suggested that consumers may not need a great deal of persuading to develop an appreciation for partially automated vehicles.

The surprisingly favorable assessments led the researchers to speculate that an important contributor to the positive initial driving experience may have been the favorable road conditions of the test drive. The study was conducted during good weather on a section of a major freeway that was characterized by long straightaways and little traffic. The roadway is likely to have minimized concerns about the automation malfunctioning and causing an accident. However, the experience of automated driving may not be so favorable in normal travel conditions and over a longer period of time.

Researchers also believed that the effects of the automated systems on the attention and behavior of drivers may have been minimal because they were using it for the first time. As familiarity increases and drivers become even more comfortable with and reliant on automation, they may attend less to safely operating the vehicle and engage in more activities unrelated to driving (e.g., Banks et al, 2018; Dunn et al., 2019).

Current Survey Study

The current study was conducted to examine how perceptions of participants who had never used automated systems changed over the course of extensive on-road driving of a Level 2 vehicle. Participants were monitored and assessed continuously during 6 to 8 weeks of driving, which enabled researchers to examine the changes in driving experience, attention, and behavior that occur with increasing familiarity and usage of the automation. Participants' intentions to use and purchase automated systems in the future were also examined, along with how intentions interacted with perceived favorableness of automated driving experiences.

Finally, the study explored how pre-existing evaluations of and trust in automated systems may shape the experience of lengthy usage of automated systems. A large body of research has shown that pre-existing beliefs and hypotheses often bias the processing of information and personal experiences (e.g., Sanbonmatsu et al., 1998; Von Hippel & Tyre, 1995). As a result, people often interpret and remember events in a way that confirms their expectations. In this study, it was hypothesized that positive pre-driving beliefs about automation would be associated with greater enjoyment and less stress during automated driving, less attention to driving, and stronger intentions to use and purchase automated systems in the future. In particular, it was anticipated that participants who expressed trust in automated driving systems and trust in general technology prior to the study would report greater appreciation for driving with the automation. Before driving the Level 2 vehicle, participants completed the Propensity to Trust in General Technology Scale (McKnight et al., 2002), which measures faith in general technology (i.e., the belief that technology is usually reliable, functional, and helpful) and trusting stance (i.e., the belief that positive outcomes will result from relying on technology).

Method

The participants, vehicles, and general protocol are described in Parts 1 and 2. This section describes additional elements related to the surveys and their administration.

Procedure

Before the first test drive in the first experimental session (Part 1), participants completed questionnaires about their demographic backgrounds, health, personalities, knowledge of automated systems, and driving histories. They also filled out the Beliefs about Automated Systems Survey and the Propensity to Trust in General Technology Scale (a description of these measures is presented in “Survey Measures” below). The surveys were administered on a tablet using REDCap software.

After the initial experimental session, participants took the vehicle home and then drove the vehicle on their daily commute to and from work (described in Part 2). Shortly after taking possession of the vehicle, participants completed an online survey about their assessments of the vehicle’s automated systems and their automated driving experiences and behavior (see the “Bi-Weekly Survey” in “Survey Measures”). They filled out this survey every two weeks for a total of four times and received $20 for completing each survey.

Participants drove the vehicle each workday for 6 to 8 weeks before completing the second experimental session. The number of weeks varied because of difficulties that arose in scheduling the return of the vehicle. On the day participants returned the vehicle, they completed a final set of questionnaires about their beliefs and knowledge of automated systems on a tablet including measures of their automated driving experiences, attention, and intentions.

Survey Measures

Beliefs About Automated Driving Systems. The first set of questions measured participant's evaluations and expectations of the automated systems prior to driving a Level 2 automation vehicle in the study. Participants indicated the following:

- "How dependable are automated driving systems?" on a 5-point scale anchored by 1 = *not at all* and 5 = *highly dependable*
- "How well do automated driving systems work?" on a 5-point scale anchored by 1 = *do not work well at all* and 5 = *work extremely well*
- "How safe are automated driving systems?" on a 5-point scale anchored by 1 = *not at all safe* and 5 = *highly safe*
- "How much do you trust automated driving systems?" on a 5-point scale anchored by 1 = *do not trust at all* and 5 = *trust completely*
- "How useful are automated driving systems to you?" on a 5-point scale anchored by 1 = *not at all useful* and 5 = *highly useful*
- "What is your overall evaluation of automated driving systems?" on a 5-point scale anchored by 1 = *highly negative* and 5 = *highly positive*

Participants also conveyed their agreement with the following statements on 5-point scales anchored by 1 = *disagree completely* and 5 = *agree completely*:

- "Automated driving systems are susceptible to malfunctioning"
- "Automated driving systems handle a car well in terms of steering, acceleration, and braking"
- "Automated driving systems are prone to turning off unexpectedly"

Propensity to Trust in General Technology. Prior to driving, participants completed the Propensity to Trust in General Technology Scale (McKnight et al., 2011). The scale consists of two sub-scales with four items pertaining to "faith in general technology" and three items related to "trusting stance." Participants indicated their agreement with items that reflect the degree to which technology is assumed to be reliable and helpful (e.g., "I believe that most technologies are effective at what they are designed to do"). Other items, such as "I usually trust a technology until it gives me a reason not to trust it," measure the expectation that positive outcomes will result from relying on technology. The mean responses to the trusting stance and faith in general technology sub-scales were averaged to create a single index of the propensity to trust in general technology.

Bi-Weekly Survey. The first three questions of the Bi-Weekly Survey measured participants' trust, perceived utility, and favorability of automated systems. Specifically, participants reported the following:

- "How much do you trust the automated driving systems?" on a 5-point scale anchored by 1 = *do not trust at all* 5 = *trust completely*

- “How useful are the automated driving systems to you?” on a 5-point scale anchored 1 = *not at all useful* 5 = *highly useful*
- “What is your overall evaluation of the automated driving systems?” on a 5-point scale anchored 1 = *highly negative* 5 = *highly positive*

Subsequent questions addressed participants’ automated driving experience, attention during automated driving, and usage of the systems. Participants indicated their agreement with the following statements on 5-point scales anchored by 1 = *disagree completely* and 5 = *agree completely*:

- “The automated driving systems reduce the stress of driving”
- “I get bored and sleepy when the automated driving systems are on”
- “I am anxious and nervous when the automated driving systems are on”
- “I tend to allow my thoughts to wander while the automated driving systems are on”
- “I engage in more activities unrelated to driving while the car is driving in autonomous mode”
- “I am totally focused on the road and driving safely even when the automated systems are on”

Finally, participants conveyed their usage of automated driving systems by indicating their agreement with the following statements:

- “I do not use the automated systems in heavy traffic”
- “I do not use the automated systems on hilly or curvy roadways”
- “I utilize the automated driving systems in the vehicle as much as possible”

Driving Experiences and Attention during Automated Driving. At the end of the study, participants evaluated the automated driving systems by indicating their agreement with the following statements on 5-point scales anchored by 1 = *disagree completely* and 5 = *agree completely*:

- “The automated driving system made traveling safer”
- “I was concerned that the automated driving systems would shut off unexpectedly”

They also conveyed the favorableness of their experience of automated driving by indicating their agreement with the following statements:

- “I was able to relax when the automated driving systems were on”
- “The automated driving system made traveling boring for me”
- “The automated driving system made traveling safer”
- “I was anxious and nervous when the automated driving systems were on”
- “The automated driving system made traveling more enjoyable”
- “The automated driving system reduced the stress of driving”
- “The automated driving system took the fun out of driving”

Finally, participants rated their attentional state during automated driving by indicating their agreement with the following statements:

- "My mind would tend to wander when the automated driving systems were on"
- "I was able to engage in more activities unrelated to driving when the automated driving systems were on"
- "I was comfortable relinquishing control of the vehicle to the automated driving system"

Intentions to Use and Purchase Automated Vehicles. The final five items measured participants' intentions to use and purchase automated vehicles in the future. Participants indicated their agreement with the following statements on 5-point scales anchored by 1 = *disagree completely* and 5 = *agree completely*:

- "I would not feel comfortable using automated driving systems on most roads"
- "If I was tired or distracted, I would rely heavily on automated driving systems"
- "I would utilize the automated driving systems in a vehicle as much as possible
- "I would not feel comfortable using the automated driving systems in a vehicle without monitoring it closely"

Participants also indicated "What is the likelihood that you would buy automated driving systems for your car if you had the funds for them?" on a 5- point scale anchored by 1 = *definitely would not purchase* and 5 = *definitely would purchase*.

Results

Preliminary analyses indicated that there were no gender effects on almost all the measures. Because of the small sample size and lack of gender main effects and interactions, gender was not included in the reported analyses.

Bi-Weekly Survey Responses

Participants completed the bi-weekly surveys at the beginning and end of the 6- to 8-week driving period and two points of time in between. Linear trend analyses were performed to examine the changes in automated driving experiences and attention, and evaluations of the automated driving systems over the course of the study (see Table 10). The reported analyses were limited to the 20 participants who completed all four of the bi-weekly surveys.

Evaluations of the automation as reflected by participants' trust in automated systems, perceptions of the usefulness of automated systems, and overall evaluations increased in favorableness over the course of the study. There were no changes in the perceived safety, workings, or frequency of malfunction of the automation.

As their exposure (using week of the study as a proxy) to the vehicle increased, participants reported greater agreement that the automated systems reduce the stress of driving. There were no changes over time in participants' reports of sleepiness or anxiety while the automation was operating.

In terms of their attention during automated driving, participants reported increasing engagement in activities unrelated to driving while the automation was operating and less discomfort about using the automation without monitoring it closely over time. Nevertheless, they reported increased focus on the road and driving safely with the automation over the 6- to 8-week study period. These were the only linear trends that were significant. There were no changes over time in the reported usage of the automation.

Table 10. Responses to the bi-weekly survey across 6- to 8-week driving period

	Start	Week 2	Week 4	End	Linear Trend
Evaluations of automation					
How much do you trust the automated driving systems? (1 = *do not trust at all;* 5 = *trust completely)*	3.25 (0.64)	3.45 (0.76)	3.50 (0.61)	3.75 (0.55)	$F(1,19) = 6.00$, $p = .02$*
How useful are the automated driving systems to you? (1 = *not at all useful;* 5 = *highly useful)*	3.55 (0.89)	3.85 (0.81)	4.00 (0.80)	4.10 (0.79)	$F(1,19) = 5.83$, $p = .03$*
What is your overall evaluation of the automated driving systems? (1=*highly negative;* 5=*highly positive)*	3.95 (0.83)	4.15 (0.59)	4.20 (0.70)	4.35 (0.49)	$F(1,19) = 5.41$, $p = .03$*
The automated driving systems substantially increase the safeness of driving.	3.60 (0.75)	3.75 (0.79)	3.65 (0.81)	3.80 (0.77)	$F < 1$, $p = .46$
I am surprised by how well the automated driving systems work and what they can do.	3.90 (0.85)	3.80 (0.95)	3.85 (0.88)	4.05 (0.69)	$F(1,19) = 1.30$, $p = .27$
I am surprised by how often the automated driving systems malfunction or shut off unexpectedly.	3.15 (1.18)	2.85 (1.04)	3.15 (1.04)	2.95 (1.10)	$F < 1$, $p = .69$
Automated driving experience					
The automated driving systems reduce the stress of driving.	2.85 (1.14)	3.35 (1.04)	3.50 (0.95)	3.50 (1.00)	$F(1,19) = 5.48$, $p = .03$*
I get bored and sleepy when the automated driving systems are on.	2.25 (1.02)	2.15 (1.09)	2.10 (0.91)	2.25 (1.12)	$F < 1$, $p = .93$

	Start	Week 2	Week 4	End	Linear Trend
I am anxious and nervous when the automated driving systems are on.	3.15 (1.04)	3.00 (0.97)	3.05 (1.19)	2.70 (0.98)	$F(1,19) = 1.97$ $p = .18$
Attention during automated driving					
I engage in more activities unrelated to driving while the car is driving in autonomous mode.	2.60 (1.10)	3.00 (1.21)	2.85 (1.14)	3.25 (1.16)	$F(1,19) = 6.00$ $p = .02$*
I am totally focused on the road and driving safely even when the automated systems are on.	3.90 (0.91)	3.85 (1.23)	4.25 (0.91)	4.10 (0.91)	$F(1,19) = 6.55$ $p = .02$*
I tend to allow my thoughts to wander while the automated driving systems are on.	3.30 (1.03)	3.25 (0.85)	3.15 (1.09)	3.30 (1.03)	$F < 1$, $p = .87$
I do not feel comfortable using the automated systems in the vehicle without monitoring them closely.	3.70 (1.17)	3.40 (1.50)	3.10 (1.59)	3.30 (1.30)	$F(1,19) = 6.15$ $p = .02$*
Usage of automation					
I do not use the automated systems in heavy traffic.	3.20 (1.44)	3.20 (1.51)	3.35 (1.31)	3.35 (1.42)	$F < 1$, $p = .49$
I do not use the automated systems on hilly or curvy roadways.	2.90 (1.33)	3.25 (1.29)	2.90 (1.17)	2.70 (1.22)	$F(1,19) = 1.88$ $p = .19$
I utilize the automated driving systems in the vehicle as much as possible.	3.95 (1.15)	4.25 (1.16)	4.05 (1.15)	4.10 (1.07)	$F < 1$, $p = .68$

Notes. Scales anchored by 1 = disagree completely and 5 = agree completely, unless otherwise noted. N = 20.

Driving Experiences with the Automated Systems

At the completion of the study, participants conveyed their agreement with a series of statements about their automated driving experience, attention during automated driving, and usage of the automation. The mean levels of agreement were compared using one-sample t-tests against the scale midpoint of 3 to determine whether they generally agreed or disagreed with the statements (see Table 11).

Participants tended to agree that the automated systems made traveling safer. They also indicated that the automated systems reduced the stress of driving and made traveling more enjoyable, and that they were able to relax when the automated driving systems were on. Furthermore, participants reported that the automated systems did not make them anxious and nervous, did not make traveling boring, and did not take the fun out of driving.

Finally, participants appear to attend less to driving when the automation is operating. A large proportion of respondents reported that their minds would tend to wander during automated driving and they were comfortable relinquishing control of the vehicle to the automation. Moreover, there was a marginally significant tendency to engage in more secondary activities when the automated systems were operating.

Table 11. Experiences with automated driving systems, measured post-study.

	Percent Disagreement	Percent Agreement	Mean (St. Dev)	Comparison with Scale Midpoint
Evaluations of automation				
The automated driving system made traveling safer.	10.3%	72.4%	3.79 (0.86)	$t(28) = 4.96$ $p < .001$*
I was concerned that the automated driving systems would shut off unexpectedly.	55.1%	44.8%	2.59 (1.50)	$t(28) = 1.49$ $p = .15$
Experience of automated driving				
I was able to relax when the automated driving systems were on.	17.2%	82.8%	3.83 (1.04)	$t(28) = 4.30$ $p < .001$*
The automated driving system made traveling boring for me.	51.7%	27.6%	2.59 (1.09)	$t(28) = 2.05$ $p = .05$*
I was anxious and nervous when the automated driving systems were on.	51.7%	27.6%	2.48 (1.18)	$t(28) = 2.35$ $p = .03$*
The automated driving system made traveling more enjoyable.	17.2%	62.1%	3.66 (1.01)	$t(28) = 3.49$ $p = .002$*
The automated driving system reduced the stress of driving.	27.6%	58.6%	3.55 (1.15)	$t(28) = 2.58$ $p = .02$*
The automated driving system took the fun out of driving.	51.7%	20.6%	2.55 (1.09)	$t(28) = 2.22$ $p = .04$*
Attention during automated driving				
My mind would tend to wander when the automated driving systems were on.	24.1%	69.0%	3.48 (1.02)	$t(28) = 2.54$ $p = .02$*
I was able to engage in more activities unrelated to driving when the automated driving systems were on.	17.2%	62.0%	3.34 (1.11)	$t(28) = 1.67$ $p = .11$
I was comfortable relinquishing control of the vehicle to the automated driving system.	20.6%	69.0%	3.59 (1.05)	$t(28) = 3.00$ $p = .006$*

Notes. Scales anchored by 1 = disagree completely and 5 = agree completely. N = 29.

Intentions to Use and Purchase Automated Driving Systems

The mean intentions to use and purchase automated driving systems in the future are presented in Table 12. Once again, the mean levels of agreement were compared with the midpoint of 3 using one-sample t-tests to determine whether participants generally agreed or disagreed with the statements. Most participants reported that they would utilize the automated systems as much as possible while driving and few conveyed that they would not feel comfortable using automated driving systems on most roads. Most participants reported that they would purchase automated driving systems for their vehicles if they could afford it. Finally, almost 70% of participants reported that they monitored the car continuously when the vehicle was in automated mode. This is important because the safe operation of Level 2 vehicles requires motorists to take control when the automation disengages.

Table 12. Intentions toward automated driving systems (standard deviation in parentheses below means).

	Percent Disagreement	Percent Agreement	Mean (St. Dev.)	Comparison with Scale Midpoint
I would not feel comfortable using automated driving systems on most roads.	65.5%	17.2%	2.31 (1.20)	$t(28) = 3.10$, $p = .004$*
If I was tired or distracted, I would rely heavily on automated driving systems.	31.0%	55.1%	3.41 (1.45)	$t(28) = 1.54$ $p = .14$
I would utilize the automated driving systems in a vehicle as much as possible.	17.2%	75.0%	3.86 (1.09)	$t(28) = 4.25$ $p < .001$*
I would not feel comfortable using the automated driving systems in a vehicle without monitoring it closely.	20.7%	68.9%	3.72 (1.07)	$t(28) = 3.66$ $p = .001$*
If I can afford it, I am going to buy or lease a car with automated driving systems. (1 = *definitely would not purchase;* 5 = *definitely would purchase)*	17.2%	72.4%	3.90 (1.18)	$t(28) = 4.11$ $p < .001$*

Notes. Scales anchored by 1 = disagree completely and 5 = agree completely, unless otherwise noted. N = 29.

Relationship between Pre-Driving Beliefs about Automated Systems and Driving Experiences, Attention, and Intentions

Analyses were performed to explore the impact of pre-existing beliefs and expectations on automated driving experiences and attention, and evaluations of the automation (see Table 13). It was hypothesized that participants who expressed more positive pre-driving attitudes toward automation and more positive beliefs about technology would report more favorable automated driving experiences and less attention to the road. The correlational analyses focused on pre-driving overall evaluations of automated driving systems, pre-driving perceptions of the usefulness of automated systems, trust in automated systems, and the propensity to trust in general technology. The patterns of correlations that emerged for the evaluations and perceptions of the usefulness of automated systems were different from those that were observed for the trust in the automation. As a consequence, they are reported separately.

Participants who believed that the automated systems were useful prior to driving perceived that the automated systems tended to make driving safer. Otherwise, pre-driving evaluations and perceptions of usefulness did not appear to strongly influence post-driving evaluations of the automation. As expected, pre-driving beliefs appeared to influence the automated driving experience. As the favorableness of their pre-driving overall evaluations of automated systems increased, participants expressed less stress and anxiety during automated driving and were less apt to report that the automation took the fun out of driving. Similarly, as the perceived usefulness of automated systems pre-driving increased, the greater the enjoyment of automated driving and the lower the anxiety and stress reported. There was also a lower likelihood of reporting that the automation took the fun out of driving.

Pre-driving evaluations of automated systems were marginally significantly correlated with mind wandering and secondary task engagement during automated driving. Pre-driving perceptions of usefulness were marginally significantly correlated with the tendency to engage in secondary activities when the automation was on and comfort with relinquishing control of the vehicle to the automated driving systems.

In contrast, the pre-driving trust in automated systems and the pre-driving propensity to trust in general technology did not appear to influence evaluations of the automation or automated driving experiences and attention. These two pre-driving measures were significantly associated with only one of the eleven items. Note that separate correlational analyses for the “faith in general technology” and “trusting stance” sub-scales of the Propensity to Trust Technology Scale (McKnight et al., 2011) were similarly uncorrelated with evaluations of the automation, and automated driving experiences and attention.

Table 13. Correlations between pre-driving beliefs about automated systems and the Propensity to Trust Technology, and automated driving evaluations, experiences, and attention.

	Overall Evaluation[1]	Usefulness[2]	Trust[3]	Propensity to trust in general technology
Evaluations of automation				
The automated driving system made traveling safer.	.30 p = .12	.39 p = .04*	.31 p = .10	-.10 p = .61
I was concerned that the automated driving systems would shut off unexpectedly.	-.11 p = .59	-.02 p = .91	-.01 p = .97	.06 p = .76
Experience of automated driving				
I was able to relax when the automated driving systems were on.	.30 p = .11	.27 p = .15	.25 p = .19	-.20 p = .10
The automated driving system made traveling boring for me.	.16 p = .40	-.03 p = .88	-.12 p = .53	-.13 p = .52
I was anxious and nervous when the automated driving systems were on.	-.45 p = .02*	-.49 p = .01*	-.30 p = .11	.11 p = .56
The automated driving system made traveling more enjoyable.	.34 p = .07	.49 p = .01*	.28 p = .14	-.13 p = .50
The automated driving system reduced the stress of driving.	.47 p = .01*	.56 p = .002*	.26 p = .18	-.24 p = .21
The automated driving system took the fun out of driving.	-.38 p < .05*	-.39 p = .04*	-.50 p = .005*	-.15 p = .43
Attention during automated driving				
My mind would tend to wander when the automated driving systems were on.	.32 p = .09	.22 p = .25	.00 p = .99	-.25 p = .20
I was able to engage in more activities unrelated to driving when the automated driving systems were on.	.33 p = .08	.35 p = .06	.07 p = .71	-.08 p = .67
I was comfortable relinquishing control of the vehicle to the automated driving system.	.30 p = .11	.35 p = .06	.10 p = .59	-.27 p = .16

[1]*Pre-driving overall evaluation of automated systems (1=highly negative; 5=highly positive).*

[2]*Pre-driving beliefs about the usefulness of automated systems (1=not at all useful; 5=highly useful).*

[3]*Pre-driving trust in automated systems (1=do not trust at all; 5=trust completely); N = 30.*

Relation between Pre-Driving Beliefs about Automated Systems and Automated Driving Intentions

As shown in Table 14, the favorableness of pre-driving evaluations and perceptions of the usefulness of automated driving systems were associated with stronger intentions to use and purchase the automation in the future. Participants who expressed more favorable pre-driving evaluations of automated systems conveyed stronger intentions to utilize automated systems and rely on them when they were tired or distracted. Participants who perceived the automation to be useful expressed stronger intentions to utilize automated systems, rely on them when they were tired or distracted, and purchase them. In addition, they expressed greater comfort using automated systems on most roads.

In contrast, pre-driving trust in automated systems and the propensity to trust in general technology did not appear to influence automated driving intentions. Pre-driving trust in automated systems was not significantly correlated with any of the automated driving intention measures while the propensity to trust in general technology was significantly correlated only with the need to monitor the vehicle closely with the automated systems operating.

Table 14. Correlations between pre-driving beliefs about automated systems and the Propensity to Trust Technology, and intentions to use and purchase automated systems

	Overall Evaluation[1]	Usefulness[2]	Trust[3]	Propensity to trust in general technology
I would not feel comfortable using automated driving systems on most roads.	-.10 *p* = .61	-.37 *p* < .05*	.02 *p* = .92	-.02 *p* = .93
If I was tired or distracted, I would rely heavily on automated driving systems.	.67 *p* < .001*	.48 *p* = .01*	.21 *p* = .26	-.11 *p* = .56
I would utilize the automated driving systems in a vehicle as much as possible.	.65 *p* < .001*	.55 *p* = .002*	.35 *p* = .07	-.10 *p* = .60
I would not feel comfortable using the automated driving systems in a vehicle without monitoring it closely.	-.23 *p* = .23	-.49 *p* = .01*	-.03 *p* = .90	-.40 *p* = .03*
What is the likelihood that you would buy automated driving systems for your car if you had the funds for them?	.29 *p* = .13	.38 *p* < .05*	.11 *p* = .56	-.33 *p* = .08

[1]Pre-driving overall evaluation of automated systems (1=highly negative; 5=highly positive).

[2]Pre-driving beliefs about the usefulness of automated systems (1=not at all useful; 5=highly useful).

[3]Pre-driving trust in automated systems (1=do not trust at all; 5=trust completely); N = 29.

Relation Between Automated Driving Experiences and Attention, and Driving Intentions

Correlation analyses were performed to examine the extent to which automated driving intentions were associated with participants' evaluations of the automation, automated driving experiences, and attention and activities during automated driving. The analyses focused on the extent to which the favorableness of the automated driving experiences and behaviors were associated with intentions to utilize and purchase automated systems in the future (see Table 15).

Participants' intentions to use and purchase automated systems in the future appear to have been strongly related to the favorableness of their automated driving experiences, perceptions of the automation, and beliefs about the affordances of the automated systems. For example, the belief that automated systems make driving safer was strongly correlated with the intention to utilize the automated systems as much as possible and the likelihood of purchasing a vehicle with automated systems. Purchase and usage intentions were similarly stronger when participants believed that the automated systems make driving less stressful and traveling more enjoyable.

Participants expressed stronger intentions to use automated systems in the future when they let their minds wander during automated driving and were able to relinquish control of the vehicle. There also was a marginally significant tendency for usage intentions to be stronger when participants engaged in more secondary activities during automated driving. Participants also expressed a greater likelihood of purchasing automated systems if they felt comfortable relinquishing control of the vehicle to the automation.

Table 15. Correlations between automated driving evaluations, experiences, and attention, and intentions to use and purchase automated systems.

	Intention to Use[1]	Likelihood to Purchase[2]
Evaluations of automation		
The automated driving system made traveling safer.	.50* $p = .006$*	.54 $p = .002$*
I was concerned that the automated driving systems would shut off unexpectedly.	-.34 $p = .07$	-.49 $p = .007$*
Experience of automated driving		
I was able to relax when the automated driving systems were on.	.67 $p < .001$*	.63 $p < .001$*
The automated driving system made traveling boring for me.	-.11 $p = .57$	-.04 $p = .86$
I was anxious and nervous when the automated driving systems were on.	-.75 $p < .001$*	-.66 $p < .001$*
The automated driving system made traveling more enjoyable.	.70 $p < .001$*	.63 $p < .001$*
The automated driving system reduced the stress of driving.	.74 $p < .001$*	.70 $p < .001$*
The automated driving system took the fun out of driving.	-.29 $p = .12$	-.26 $p = .17$
Attention during automated driving		
My mind would tend to wander when the automated driving systems were on.	.38 $p = .04$*	.34 $p = .07$
I was able to engage in more activities unrelated to driving when the automated driving systems were on.	.31 $p = .11$	.30 $p = .11$
I was comfortable relinquishing control of the vehicle to the automated driving system.	.63 $p < .001$*	.74 $p < .001$*

[1]*I would utilize the automated driving systems in a vehicle as much as possible.*

[2]*What is the likelihood that you would buy automated driving systems for your car if you had the funds for them? (1 = definitely would not purchase; 5 = definitely would purchase); N = 30.*

Discussion

The study provides compelling evidence that experience with automated systems had a positive impact on drivers' perceptions and attitudes. Following several weeks of operating a Level 2 automation vehicle, participants reported that the automated systems reduce the stress of driving and make traveling more enjoyable. They further indicated that they were less anxious and nervous, and able to relax more during automated driving. Finally, most participants conveyed that the automation did not take the fun out of driving or make traveling boring.

The results provide strong support for previous findings (Cooley et al., 2022), which were based on a single automated drive. The current study goes beyond this previous work by suggesting that the effects of automation on the driving experience increase over time. The tendency for the automation to reduce the stress of driving was found to improve with greater familiarity with the technology. Moreover, participants felt increasingly comfortable driving with the automation without monitoring it closely.

Unlike the previous on-road study (Cooley et al., 2022), participants reported that their attention and activities during driving were affected by the automation. Most were reportedly comfortable relinquishing control of the vehicle to the automated systems. Following previous research (e.g., Lin et al, 2018), they also tended to report engaging in more activities unrelated to driving when the automated systems were operating. This study contributes to the literature on the attention and behavior of motorists by providing evidence that drivers are more apt to allow their minds to wander during automated driving. The study also showed that engagement in secondary activities during automated travel and the willingness to relinquish control of the vehicle becomes greater over time as familiarity with the automation increases.

The current study differed from most prior research in that participants' assessments of the experience of automated driving were based on actual on-road driving rather than simulator driving or traveling as a passenger. Prior studies of the automated driving experience have also been limited by the usage of potentially biased samples such as owners and purchasers of the technology. These concerns were substantially diminished in the current study by recruiting motorists who had never driven automated vehicles previously.

The favorableness of the automated driving experience and perceptions of the affordances of automated driving (i.e., what automation allows drivers to do) are important because they affect the usage and purchase of automated systems, and, hence, safety. Following previous research, participants expressed a strong willingness to use the automated systems in the future and purchase automated systems if they were affordable. As expected, the favorableness of driving experiences was associated with intentions to use the automation in different driving contexts and intentions to acquire automated systems. Drivers who had more positive experiences with the automation and used the automation to do things other than driving developed stronger intentions to use and purchase automated systems in the future.

Research has shown that people tend to gather and interpret information in a manner that is consistent with their beliefs and hypotheses (e.g., Von Hippel & Tyre, 1995). Consequently, their experiences and observations tend to be in line with their expectations. In the current study, the favorableness of experiences with the automated systems and future intentions were strongly correlated with pre-existing beliefs about the automation. For example, participants who believed that automated systems are useful prior to using the technology reported that the automation tended to increase the enjoyment and reduce the stress of driving. They also expressed stronger intentions to use and purchase automated driving systems in the future.

Unexpectedly, trust in automated systems and the propensity to trust in general technology did not appear to have influenced evaluations of the automation, automated driving experiences and attention, and usage and purchase intentions. This finding was very surprising because recent work on driving automation has assumed that trust is vital to the adoption and usage of the technology (e.g., Beggiato et al., 2015; Walker et al., 2018). Future research will need to closely examine the impact of trust in automated systems on driving behavior and consumer purchase decisions.

The survey findings are paralleled by the analyses of the videotaped driving behaviors of participants reported in Part 2. In line with the self-reports of participants, the behavioral coding and analysis found that secondary task engagement tended to be slightly higher during automated driving than manual driving (using a naturalistic control) and that secondary activities tended to increase as familiarity with the automation increased. Moreover, the behavioral analysis revealed that there were more system warnings about driver inattention as participants became more experienced with the Level 2 vehicle automation. This parallels participants' reports that they sometimes allowed their minds to wander while using the automated systems and became increasingly comfortable over time using the automation without monitoring it closely.

Some of the current findings may suggest that automated systems increase unsafe behavior by motorists. However, the analyses of the videotaped driving behaviors of our participants reported in Part 2 and other facets of the survey results suggest that most motorists are cognizant of the risks of automated driving and discreet in their usage of the automation. Almost 70% of our participants reported that they would not feel comfortable using the automated driving systems in a vehicle without monitoring them closely. In addition, as participants gained greater familiarity with the automation, they increasingly reported that they focus on the road and driving safely even when the automated systems are on. Finally, the majority of our participants reported that they would not use the automation in heavy traffic. Moreover, the behavioral analysis showed that participants were less likely to use the automation when driving demands were high, that is, when adverse weather, heavy traffic, and other potential hazards were present. Thus, while drivers may be less attentive when automated systems are operating, they are less likely to use the automation in roadway conditions in which the risk of driving are elevated.

GENERAL SUMMARY

To address critical research gaps in our understanding of vehicle automation, the current study recruited participants who had no prior experience with automated vehicle technology and examined their behaviors and interactions with the system over time. The study employed a hybrid research design that combined experimental, naturalistic, and survey components in a longitudinal design where drivers gained experience and familiarity with the vehicle technology over the course of several weeks. The design allowed for a comprehensive assessment of how drivers interact with and adapt to vehicle automation systems in real-world scenarios. By examining factors such as changes in system use over time, driver fatigue, and secondary task engagement, this study aimed to provide valuable insights into the safety concerns associated with automation use. This section highlights selected key outcomes gleaned from the different components of the study (described previously in the Parts 1, 2 and 3 above).

Experimental Study

- Results from the behavioral indices of workload (i.e., the Detection Response Task) suggested that driving under Level 2 automation was associated with an increase in driver workload compared to conditions without automation. That is, drivers in the experimental study may pay *more* attention to the driving environment under partial automation compared to manual mode.
- The experience of driving a vehicle for a period of 6 to 8 weeks impacted driver workload differentially across road environment, suggesting that practice with vehicle automation decreases driver workload while using the system over time, at least when driving on roads with relatively low demand. That is, after a 6-week familiarization period, there was a significant decrease in attention paid to the driving task under partial automation—at least in the simpler driving environment.
- The physiological measures of workload were less sensitive to discriminating differences in workload and engagement between driving conditions and did not show the same patterns as the behavioral indices.

Naturalistic Study

- Drivers used Level 2 vehicle automation more than 70% of the time, an amount that stayed relatively consistent over the 6- to 8-week observation period. Drivers were less likely to use vehicle automation when driving demands were higher.
- Over 6- to 8-week period of automation use, there was an increase in the frequency of system warnings as drivers become more experienced with the Level 2 vehicle automation, suggesting that drivers were becoming

increasingly comfortable with the automation and exhibited a tendency toward a more relaxed automation monitoring strategy over time.

- As drivers grow more familiar with the system, they are more likely to engage in secondary tasks over time; however, the lack of interaction between secondary task engagement and Condition suggests that these behaviors may not be a direct consequence of over-reliance on the automation system. Instead, they might reflect a general trend of drivers becoming more comfortable with multitasking in these specific research vehicles over time, regardless of automation capabilities.
- The Experimental Control condition yielded a behavioral profile that was markedly different from that observed in the Naturalistic Control condition. In most cases, behavior in the Experimental Control condition fell in between the Automation-L2 and Naturalistic Control conditions (e.g., fatigue, the secondary task aggregate, radio listening, and text messaging).
    - One plausible explanation for differences between the benchmarks is that driving demand may mediate the relationship between automation use and secondary task engagements such that drivers may be less likely to use automation and less likely to engage in secondary tasks when driving demands are higher.

Survey Study

- Experience with automated systems had a positive impact on drivers' perceptions and attitudes, including reduced stress and increased enjoyment.
- As their experience grew, drivers felt increasingly willing to relinquish control of the vehicle and comfortable driving with the automation without monitoring it closely—an outcome that aligns with the naturalistic study where more system warnings were observed as participants became more experienced with the Level 2 vehicle automation.
- Drivers were also more likely to report engaging in more activities unrelated to driving when the automated systems were operating as their experience with the systems grew—in line with outcomes from the naturalistic study.
- Unexpectedly, trust in automated systems did not appear to have influenced evaluations of the automation, automated driving experiences and attention, and usage and purchase intentions.
- While drivers may have reported being less attentive when automated systems are operating, they were also less likely to report using the automation in roadway conditions in which the risk of driving were elevated.

REFERENCES

ABC (2023). Woman thought Tesla driver caught sleeping at the wheel was having medical issue [Video]. Available at: https://www.youtube.com/watch?v=kQSA25RGjN4

Abe, G., & Richardson, J. (2004). The effect of alarm timing on driver behaviour: An investigation of differences in driver trust and response to alarms according to alarm timing. *Transportation Research Part F: Traffic Psychology and Behaviour*, *7*(4–5), 307–322. https://doi.org/10.1016/j.trf.2004.09.008

Ahlström, C., Zemblys, R., Jansson, H., Forsberg, C., Karlsson, J., & Anund, A. (2021). Effects of partially automated driving on the development of driver sleepiness. *Accident Analysis & Prevention*, *153*, 106058. https://doi.org/10.1016/j.aap.2021.106058

Alessandrini, A., Campagna, A., Delle Site, P., Filippi, F., & Persia, L. (2015). Automated vehicles and the rethinking of mobility and cities. *Transportation Research Procedia*, *5*, 145–160. https://doi.org/10.1016/j.trpro.2015.01.002

Arefnezhad, S., Eichberger, A., & Koglbauer, I. V. (2022). Effects of automation and fatigue on drivers from various age groups. *Safety*, *8*(2), 30. https://doi.org/10.3390/safety8020030

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*, *68*, 138–145. https://doi.org/10.1016/j.apergo.2017.11.010

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*, *18*, 47–57. https://doi.org/10.1016/j.trf.2012.12.006

Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *35*, 75–84. https://doi.org/10.1016/j.trf.2015.10.005

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, *44*, 58–75. https://doi.org/10.1016/j.neubiorev.2012.10.003

Carriere, J. S., Seli, P., & Smilek, D. (2013). Wandering in both mind and body: Individual differences in mind wandering and inattention predict fidgeting. *Canadian Journal of Experimental Psychology*, *67*(1), 19. https://doi.org/10.1037/a0031438

Carsten, O., Kircher, K., & Jamson, S. (2013). Vehicle-based studies of driving in the real world: The hard truth? *Accident Analysis & Prevention*, *58*, 162–174. https://doi.org/10.1016/j.aap.2013.06.006

Carsten, O., Lai, F. C., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control task substitution in semiautomated driving: Does it matter what aspects are automated? *Human Factors*, *54*(5), 747–761. https://doi.org/10.1177/0018720812460246

Casner, S. M., Hutchins, E. L., & Norman, D. (2016). The challenges of partially automated driving. *Communications of the ACM*, *59*(5), 70–77. https://doi.org/10.1145/2830565

Castro, S. C., Strayer, D. L., Matzke, D., & Heathcote, A. (2019). Cognitive workload measurement and modeling under divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 826–839. https://doi.org/10.1037/xhp0000638

Chikhi, S., Matton, N., & Blanchet, S. (2022). EEG power spectral measures of cognitive workload: A meta-analysis. *Psychophysiology*, *59*(6), e14009. https://doi.org/10.1111/psyp.14009

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. MIT press.

Cooley, E. H., Sanbonmatsu, D. M., Strayer, D. L., White, P. H., & Cooper, J. M. (2022). On-road vehicle study of the experience of automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *87*, 444–453. https://doi.org/10.1016/j.trf.2022.04.014

Cooper, J. M., Castro, S. C., & Strayer, D. L. (2016). Extending the detection response task to simultaneously measure cognitive and visual task demands. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 1962–1966. Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/1541931213601447

De Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196–217. https://doi.org/10.1016/j.trf.2014.06.016

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Desmond, P. A., & Hancock, P. A. (2000). Active and passive fatigue states. In *Stress, Workload, and Fatigue* (pp. 455–465). CRC Press.

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*, *113*(10), 2636–2641. https://doi.org/10.1073/pnas.1513271113

Dufour, A. (2014). Driving assistance technologies and vigilance: Impact of speed limiters and cruise control on drivers' vigilance. *Seminar on the Impact of Distracted Driving and Sleepiness on Road Safety, April*. Paris La Défense.

Dunn, N. J., Dingus, T. A., Soccolich, S., & Horrey, W. J. (2021). Investigating the impact of driving automation systems on distracted driving behaviors. *Accident Analysis & Prevention*, *156*, 106152. https://doi.org/10.1016/j.aap.2021.106152

Dunn, N., Dingus, T.A., & Soccolich, S. (2019). *Understanding the Impact of Technology: Do Advanced Driver Assistance and Semi-Automated Vehicle Systems Lead to Improper Driving Behavior?* (Technical Report). Washington, D.C.: AAA Foundation for Traffic Safety.

Eenink, R., Barnard, Y., Baumann, M., Augros, X., & Utesch, F. (2014). UDRIVE: the European naturalistic driving study. In *Proceedings of Transport Research Arena*. IFSTTAR.

Eichelberger, A. H., & McCartt, A. T. (2016). Toyota drivers' experiences with dynamic radar cruise control, pre-collision system, and lane-keeping assist. *Journal of Safety Research*, *56*, 67–73. https://doi.org/10.1016/j.jsr.2015.12.002

Endsley, M. R. (2017). Autonomous driving systems: A preliminary naturalistic study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making*, *11*(3), 225–238. https://doi.org/10.1177/1555343417695197

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, *37*(2), 381–394. https://doi.org/10.1518/001872095779064555

Fairclough, S. H., Venables, L., & Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology*, *56*(2), 171–184. https://doi.org/10.1016/j.ijpsycho.2004.11.003

Fan, C., Hu, J., Huang, S., Peng, Y., & Kwong, S. (2022). EEG-TNet: An end-to-end brain computer interface framework for mental workload estimation. *Frontiers in Neuroscience*, *16*. https://doi.org/10.3389/fnins.2022.869522

Feldhütter, A., Hecht, T., Kalb, L., & Bengler K. (2019). Effect of prolonged periods of conditionally automated driving on the development of fatigue: With and without non-driving-related activities. *Cognition, Technology & Work,* 21, 33–40. https://doi.org/10.1007/s10111-018-0524-9

Fisher, D. L., Lohrenz, M., Moore, D., Nadler, E. D., & Pollard, J. K. (2016). Humans and intelligent vehicles: The hope, the help, and the harm. *IEEE Transactions on Intelligent Vehicles*, *1*(1), 56–67. https://doi.org/10.1109/TIV.2016.2555626

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*, 154. https://doi.org/10.3389/fpsyg.2011.00154

Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Fridman, L., Brown, D. E., Glazer, M., Angell, W., Dodd, S., Jenik, B., ... & Reimer, B. (2019). MIT advanced vehicle technology study: Large-scale naturalistic driving study of driver behavior and interaction with automation. *IEEE Access*, *7*, 102021–102038. https://doi.org/10.48550/arXiv.1711.06976

Fridman, L., Brown, D. E., Glazer, M., Angell, W., Dodd, S., Jenik, B., ... & Reimer, B. (2017). MIT autonomous vehicle technology study: Large-scale deep learning based analysis of driver behavior and interaction with automation. *arXiv preprint arXiv:1711.06976*, *1*.

Gaspar, J., & Carney, C. (2019). The effect of partial automation on driver attention: A naturalistic driving study. *Human Factors*, *61*(8), 1261–1276. https://doi.org/10.1177/0018720819836310

Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, *4*(1–2), 113–131. https://doi.org/10.1080/14639220210159717

Gibson, T. (2015). Virginia's Smart Road: Where researchers make the extreme weather. *Weatherwise*, *68*(4), 20–27.

Goldman, R. I., Stern, J. M., Engel Jr, J., & Cohen, M. S. (2002). Simultaneous EEG and fMRI of the alpha rhythm. *Neuroreport*, *13*(18), 2487–2492. https://doi.org/10.1097/01.wnr.0000047685.08940.d0

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484. https://doi.org/10.1016/0013-4694(83)90135-9

Greenlee, E. T., DeLucia, P. R., & Newton, D. C. (2018). Driver vigilance in automated vehicles: Hazard detection failures are a matter of time. *Human Factors*, *60*(4), 465–476. https://doi.org/10.1177/0018720818761711

Hasan, A. S., Patel, D., Alfaris, R., & Jalayer, M. (2022). Identifying distracted-driving events from on-road observations using a moving vehicle: A case study in New Jersey. *Accident Analysis & Prevention*, *177*, 106827.

Hartwich, F., Beggiato, M., & Krems, J. F. (2018). Driving comfort, enjoyment and acceptance of automated driving–effects of drivers' age and driving style familiarity. *Ergonomics*, *61*(8), 1017–1032. https://doi.org/10.1080/00140139.2018.1441448

Hatfield, N., Yamani, Y., Palmer, D. B., Yahoodik, S., Vasquez, V., Horrey, W. J., & Samuel, S. (2019). Analysis of visual scanning patterns comparing drivers of simulated L2 and L0 systems. *Transportation Research Record*, *2673*(10), 755–761. https://doi.org/10.1177/0361198119852339

He, D., & Donmez, B. (2019). Influence of driving experience on distraction engagement in automated vehicles. *Transportation Research Record*, *2673*(9), 142–151. https://doi.org/10.1177/0361198119843476

Huisingh, C., Griffin, R., & McGwin Jr, G. (2015). The prevalence of distraction among passenger vehicle drivers: a roadside observational approach. *Traffic Injury Prevention*, *16*(2), 140–146. https://doi.org/10.1080/15389588.2014.916797

ISO 17488 (2016) Road vehicles — Transport information and control systems — Detection-response task (DRT) for assessing attentional effects of cognitive load in driving. Retrieved from: https://www.iso.org/standard/59887.html

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 370–375.

Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, *66*, 18–31. https://doi.org/10.1016/j.apergo.2017.07.006

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Large, D. R., Harvey, C., Burnett, G., Merenda, C., Leong, S., & Gabbard, J. (2017). Exploring the relationship between false alarms and driver acceptance of a pedestrian alert system during simulated driving. *Road Safety and Simulation Conference*.

Larsson, A. F. (2012). Driver usage and understanding of adaptive cruise control. *Applied Ergonomics*, *43*(3), 501–506. https://doi.org/10.1016/j.apergo.2011.08.005

Lin, R., Ma, L., & Zhang, W. (2018). An interview study exploring Tesla drivers' behavioural adaptation. *Applied Ergonomics*, *72*, 37–47. https://doi.org/10.1016/j.apergo.2018.04.006

Liu, P., Jiang, Z., Li, T., Wang, G., Wang, R., & Xu, Z. (2021). User experience and usability when the automated driving system fails: Findings from a field experiment. *Accident Analysis & Prevention*, *161*, 106383. https://doi.org/10.1016/j.aap.2021.106383

Lohani, M., Cooper, J. M., Erickson, G. G., Simmons, T. G., McDonnell, A. S., Carriero, A. E., Crabtree, K.W., & Strayer, D. L. (2021). No difference in arousal or cognitive demands between manual and partially automated driving: A multi-method on-road study. *Frontiers in Neuroscience*, 627. https://doi.org/10.3389/fnins.2021.577418

Lohani, M., Cooper, J. M., Erickson, G. G., Simmons, T. G., McDonnell, A. S., Carriero, A. E., ... & Strayer, D. L. (2020). Driver arousal and workload under partial vehicle automation: A pilot study. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 64(1), 1955–1959. https://doi.org/10.1177/1071181320641471

Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Frontiers in Human Neuroscience*, *13*, 57. https://doi.org/10.3389/fnhum.2019.00057

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

Louw, T., Kountouriotis, G., Carsten, O., & Merat, N. (2015). Driver inattention during vehicle automation: How does driver engagement affect resumption of control?. In *4th International Conference on Driver Distraction and Inattention (DDI2015), Sydney*. ARRB Group.

Lu, K., Karlsson, J., Dahlman, A. S., Sjöqvist, B. A., & Candefjord, S. (2021). Detecting driver sleepiness using consumer wearable devices in manual and partial automated real-road driving. *IEEE Transactions on Intelligent Transportation Systems*, *23*(5), 4801–4810. https://doi.org/10.1109/TITS.2021.3127944

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Makridis, M., Mattas, K., Ciuffo, B., Raposo, M. A., Toledo, T., & Thiel, C. (2018). Connected and automated vehicles on a freeway scenario. Effect on traffic congestion and network capacity. *7th Transport Research Arena*, *13*.

Matthews, G., Neubauer, C., Saxby, D. J., Wohleber, R. W., & Lin, J. (2019). Dangerous intersections? A review of studies of fatigue and distraction in the automated vehicle. *Accident Analysis & Prevention*, *126*, 85–94. https://doi.org/10.1016/j.aap.2018.04.004

McDonald, A., Carney, C., & McGehee, D. V. (2018). Vehicle Owners' Experiences with and Reactions to Advanced Driver Assistance Systems (Technical Report). Washington, D.C.: AAA Foundation for Traffic Safety.

McDonnell, A. S., Simmons, T. G., Erickson, G. G., Lohani, M., Cooper, J. M., & Strayer, D. L. (2021). This is your brain on Autopilot: Neural indices of driver workload and engagement during partial vehicle automation. *Human Factors*, 00187208211039091. https://doi.org/10.1177/00187208211039091

McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems (TMIS)*, *2*(2), 1–25. https://doi.org/10.1145/1985347.1985353

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

Mehler, B., Reimer, B., & Dusek, J. A. (2011). *MIT AgeLab Delayed Digit Recall Task (n-back)*. Cambridge, MA: Massachusetts Institute of Technology.

Merat, N., Jamson, A. H., Lai, F. C., Daly, M., & Carsten, O. M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 274–282. https://doi.org/10.1016/j.trf.2014.09.005

Morales-Alvarez, W., Sipele, O., Léberon, R., Tadjine, H. H., & Olaverri-Monreal, C. (2020). Automated driving: A literature review of the take over request in conditional automation. *Electronics*, *9*(12), 2087. https://doi.org/10.3390/electronics9122087

Musk, E. (2021). *Generalized self-driving is a hard problem, as it requires solving a large part of real-world AI. [...]. Nothing has more degrees of freedom than reality.* [Tweet]. https://twitter.com/elonmusk/status/1411280212470366213

National Highway Traffic Safety Administration (2017). *Automated Driving Systems 2.0: A Vision for Safety.* Washington, D.C.: U.S. Department of Transportation.

National Highway Traffic Safety Administration (2021). *Driver Electronic Device Use in 2021* (DOT HS 813 357). Washington, D.C.: National Highway Traffic Safety Administration.

Naujoks, F., Purucker, C., & Neukum, A. (2016). Secondary task engagement and vehicle automation–Comparing the effects of different automation levels in an on-road experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *38*, 67–82. https://doi.org/10.1016/j.trf.2016.01.011

Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J., & Goodman, M. (2005). An overview of the 100-car naturalistic study and findings. *National Highway Traffic Safety Administration, Paper*, *5*, 0400.

Nilsson, E. J., Aust, M. L., Engström, J., Svanberg, B., & Lindén, P. (2018). Effects of cognitive load on response time in an unexpected lead vehicle braking scenario and the detection response task (DRT). *Transportation Research Part F: Traffic Psychology and Behaviour*, *59*, 463–474. https://doi.org/10.1016/j.trf.2018.09.026

Palada, H., Neal, A., Strayer, D., Ballard, T., & Heathcote, A. (2019). Using response time modeling to understand the sources of dual-task interference in a dynamic environment. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 1331–1345. https://doi.org/10.1037/xhp0000672

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Payre, W., Cestac, J., & Delhomme, P. (2016). Fully automated driving: Impact of trust and practice on manual control recovery. *Human Factors*, *58*(2), 229–241. https://doi.org/10.1177/0018720815612319

Peng, Y., Xu, Q., Lin, S., Wang, X., Xiang, G., Huang, S., ... & Fan, C. (2022). The application of electroencephalogram in driving safety: Current status and future prospects. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.919695

Puma, S., Matton, N., Paubel, P. V., Raufaste, É., & El-Yagoubi, R. (2018). Using theta and alpha band power to assess cognitive workload in multitasking environments. *International Journal of Psychophysiology*, *123*, 111–120. https://doi.org/10.1016/j.ijpsycho.2017.10.004

R Core Team (2022). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria. URL* https://www.R-project.org/

Reagan, I. J., Teoh, E. R., Cicchino, J. B., Gershon, P., Reimer, B., Mehler, B., & Seppelt, B. (2021). Disengagement from driving when using automation during a 4-week field trial. *Transportation Research Part F: Traffic Psychology and Behaviour*, *82*, 400–411.

Saad, F. (2004). Behavioural adaptations to new driver support systems: Some critical issues. *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583), 1,* 288–293. https://doi.org/10.1109/ICSMC.2004.1398312

Sanbonmatsu, D. M., Posavac, S. S., Kardes, F. R., & Mantel, S. P. (1998). Selective hypothesis testing. *Psychonomic Bulletin & Review*, *5*, 197–220. https://doi.org/10.3758/BF03212944

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PloS One*, *8*(1), e54402. https://doi.org/10.1371/journal.pone.0054402

Sanbonmatsu, D. M., Strayer, D. L., Yu, Z., Biondi, F., & Cooper, J. M. (2018). Cognitive underpinnings of beliefs and confidence in beliefs about fully automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, *55*, 114–122. https://doi.org/10.1016/j.trf.2018.02.029

Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., & Othersen, I. (2015). The interaction between highly automated driving and the development of drowsiness. *Procedia Manufacturing*, *3*, 6652–6659. https://doi.org/10.1016/j.promfg.2015.11.005

Schroeter, R., Oxtoby, J., Johnson, D., & Steinberger, F. (2015). Exploring boredom proneness as a predictor of mobile phone use in the CAR. *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, 465–473. https://doi.org/10.1145/2838739.2838783

Sener, I. N., & Zmud, J. (2019). Chipping away at uncertainty: intent to use self-driving vehicles and the role of ride-hailing. *Transportation Planning and Technology*, *42*(7), 645–661. https://doi.org/10.1080/03081060.2019.1650423

Society for Automotive Engineering (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems* (Standard No. J3016). Warrendale, PA: SAE International. https://www.sae.org/standards/content/j3016_202104/

Stapel, J., Happee, R., Christoph, M., van Nes, N., & Martens, M. (2022). Exploring the usage of supervised driving automation in naturalistic conditions. *Transportation Research Part F: Traffic Psychology and Behaviour*, *90*, 397–411. https://doi.org/10.1016/j.trf.2022.08.013

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, *12*(6), 462–466. https://doi.org/10.1111/1467-9280.00386

Strayer, D. L., Castro, S. C., Turrill, J., & Cooper, J. M. (2022). The persistence of distraction: The hidden costs of intermittent multitasking. *Journal of Experimental Psychology: Applied, 28*(2), 262–282. https://doi.org/10.1037/xap0000388

Strayer, D. L., Cooper, J. M., Goethe, R. M., McCarty, M. M., Getty, D., & Biondi, F. (2017). *Visual and Cognitive Demands of Using In-Vehicle Infotainment Systems* (Technical Report). Washington, D.C.: AAA Foundation for Traffic Safety.

Strayer, D. L., Cooper, J. M., Goethe, R. M., McCarty, M. M., Getty, D. J., & Biondi, F. (2019). Assessing the visual and cognitive demands of in-vehicle information systems. *Cognitive Research: Principles and Implications, 4*(1), 1–22. https://doi.org/10.1186/s41235-019-0166-3

Strayer, D. L., Cooper, J. M., Sanbonmatsu, D. M., Erickson, G. G., Simmons, T. G., McDonnell, A. S., ... & Lohani, M. (2020). *Driver's Arousal and Workload Under Partial Vehicle Automation* (Technical Report). Washington, D.C.: AAA Foundation for Traffic Safety.

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J., Medeiros-Ward, N., & Biondi, F. (2013). *Measuring Cognitive Distraction in the Automobile* (Technical Report). Washington, D.C.: AAA Foundation for Traffic Safety.

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing cognitive distraction in the automobile. *Human Factors*, *57*, 1300–1324. https://doi.org/10.1177/0018720815575149

Tan, H. S., Rajamani, R., & Zhang, W. B. (1998). Demonstration of an automated highway platoon system. In *Proceedings of the 1998 American control conference. ACC (IEEE Cat. No. 98CH36207*), 3, 1823–1827. https://doi.org/10.1109/ACC.1998.707332

Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2019). Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. *Accident Analysis & Prevention*, *126*, 70–84. https://doi.org/10.1016/j.aap.2018.03.013

von Hippel, E., & Tyre, M. J. (1995). How learning by doing is done: problem identification in novel process equipment. *Research Policy*, *24*(1), 1–12. https://doi.org/10.1016/0048-7333(93)00747-H

Walker, F., Boelhouwer, A., Alkim, T., Verwey, W. B., & Martens, M. H. (2018). Changes in trust after driving level 2 automated cars. *Journal of Advanced Transportation*, *2018*. https://doi.org/10.1155/2018/1045186

Weaver, S. M., Roldan, S. M., Gonzalez, T. B., Balk, S. A., & Philips, B. H. (2022). The effects of vehicle automation on driver engagement: The case of adaptive cruise control and mind wandering. *Human Factors*, *64*(6), 1086–1098. https://doi.org/10.1177/0018720820974856

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., …, Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature*, *435*, 439–440. https://doi.org/10.1038/435439a

Xu, Y., Ye, Z., & Wang, C. (2021). Modeling commercial vehicle drivers' acceptance of advanced driving assistance system (ADAS). *Journal of Intelligent and Connected Vehicles*, *4*(3), 125–135. https://doi.org/10.1108/JICV-07-2021-0011

Young, R. A., Hsieh, L., & Seaman, S. (2013). The tactile detection response task: Preliminary validation for measuring the attentional effects of cognitive load. *Driving Assessment Conference*, 7. University of Iowa.

Zhang, Y., Angell, L., Pala, S., & Shimonomoto, I. (2015). Bench-marking drivers’ visual and cognitive demands: A feasibility study. *SAE International Journal of Passenger Cars–Mechanical Systems*, *8*(2), 584–593. https://doi.org/10.4271/2015-01-1389

APPENDIX A: DATA CODING DICTIONARY FOR OBSERVED BEHAVIORS

Adapted from Strayer et al. (2017).

Modes of Interaction	Definition
Texting	**Participant engages in cell-phone/IVIS* use for instant messaging.**
Verbal Cell	Participant begins to text through the voice functions available on cell-phone.
Manual Cell	Participant begins to text through manually inputting text into their phone.
IVIS Verbal	Participant begins to text through the voice function of the IVIS within the car.
IVIS Manual	Participant begins to text through IVIS preset text responses.
Unknown	Participant reaches for phone and coders are unable to see the function of the phone while in the participant's hand.
Listening	Participant listens to car's hands free system to read text messages received while driving.
Dialing	**Participant uses cell-phone/IVIS to dial and call.**
Verbal Dialing	Participant uses cell-phone voice functions to dial a number.
Manual Dialing	Participant manually begins a call on cell-phone.
IVIS Verbal	Participant uses IVIS to begin a call with a person through the voice function within the car.
IVIS Manual	Participant uses IVIS to begin a call through the steering wheel or center console of the vehicle.
Unknown	A phone call begins and coders are unable to percieve the mode by which it started.
Listening	Listening and speaking on the phone through any modality. Begins when dialing ends and the participant starts speaking.
Radio	**Participant is listening to music and makes adjustments via cell-phone or IVIS.**
Verbal Radio	Participant uses cell-phone voice functions to change music/volume etc.
Manual Radio	Participant manually uses cell-phone to change music/volume etc.
IVIS Verbal	Participant uses IVIS voice function to change music/volume etc.
IVIS Manual	Participant uses IVIS interface to change music/volume etc.
Unknown	Coders are unable to perceive how participant changes the music setting.
Listening	Participant is listening to music while driving freeway speeds.

Navigation	**Participant uses navigation features either through cell-phone or IVIS.**
Verbal Navigation	Participant starts navigation through the voice functions available on their cell-phone.
Manual Navigation	Participant starts navigation through manual input on cell-phone.
IVIS Verbal	Participant starts navigation through the voice function of the IVIS within the car.
IVIS Manual	Participant starts navigation by manually inputting locations through the IVIS.
Unknown	Coders are unable to perceive how participant activates navigation.
Listening	Participant has started navigation and is actively following directions through audio or visual instruction. Engaged for the entire duration of navigation, not just when the navigation is actively giving a new instruction at the time.
Unknown Tech	**Any action where the participant interacts with an object or tech that cannot be identified as Radio, Navigation, Dialing, or Texting.**
Verbal Cell	Participant speaks into their cell-phone to input a command, but audio is indiscernible.
Manual Cell	Participant manually inputs commands into their cell-phone that is not discernible as dialing, radio, texting, or navigation.
IVIS Verbal	Participant uses the IVIS voice function but purpose is indiscernible.
IVIS Manual	Participant manually inputs into IVIS but coders are unable to discern what they are doing. Accounts for temperature changes.
Unknown Verbal	Participant speaks command but audio is indiscernible and car functions do not seem to change.
Unknown Manual	Participant reaches for middle console area, it is unclear what they reach for, and their hands return to the steering area.
Listening	Participant is listening to indiscernible type of audio that cannot be categorized.
Fatigue	**Participant shows signs of tiredness/lack of awareness, such as yawning, heavy eyelids, drooping head, or any other behavior that is clearly a sign of drowsiness.**
L2 ON	**Participant has engaged steering assist and the vehicle now has control over acceleration/deceleration. A car-specific symbol is present on the dash or IVIS indicating that L2 is engaged.**
L2 OFF	**Participant disengages lane assist. Symbol on dash or IVIS disappears indicating that L2 is off. Does not include lane changes, where L2 is automatically turned on and off in some vehicles. (For Volvo/Nissan: If L2 remains off for more than 30 seconds after a lane change, mark as L2 off.) L2 is turned off at exit sign on freeway ramp but is kept on when switching freeways.**

Poor Conditions	**There are outside influences on the driving experience that are out of the driver's control/influence.**
Weather	Driving conditions are affected by rain, snow, ice, or fog, causing participant to use windshield wipers, fog light, reduce speed or turn off L2. Includes any time windshield wipers are on due to weather, with a 30 second offset interval between instances.
Traffic	Any traffic that impairs driving speed to under 50 miles per hour. If the participant slows to under 50 miles per hour on exit lanes, use this if there are any other cars in the exit lane. If they are just slowing down due to the steepness of the curve, do not use.
Construction	The presence of any road construction signs, workers, and/or vehicles. If no evidence of construction is present after 2 min and 30 seconds, go back to last cone or sign and stop construction duration 30 seconds after the last sign of construction. Do not count a single lone cone on the side of the road or orange stickers on HOV lane signs as construction.
Emergency Vehicles	Ambulances, police patrol vehicles, or roadside assistance have sirens and/or lights active and are driving on the road or are on the shoulder. Also includes snow plows and situations where people are waiting for emergency responders, such as an accident that has just happened, people pulled over on the side of the road with hazards on, or people outside their vehicle. Does not include: cars parked on the side of the road without flashers/people outside, construction vehicles in construction zones, or trucks with hazards on driving slowly up hills.
Other	Participant encounters any other outside influence that affects driving. Also used for marking unique/important observations such as FaceTime. If you think this may apply, it is best to discuss the observation with other coders first.
Audio Engagement	**Participant either sings or talks with a 10 second maximum interval between instances.**
Reaching/Grabbing	**Participant physically engages with object or pet inside of car other than IVIS or cell-phone, specifically reaching into center console, passenger seat, or back seat, for at least a period of 3 seconds with a 10 second offset interval between instances. Does not include minor car adjustment actions such as flipping down the sun visors or opening the sunroof.**
Food/Drink	**Participant is either actively drinking/eating or holding food or drink while driving, including holding food in lap or resting hand on drink.**

Self Grooming/Fidgeting	Participant engages in personal/hygene related tasks or fidgeting for at least a period of 3 seconds with a 10 second offset interval between instances. Includes touching the face, neck, head/hair, or moving hands while speaking on the phone as well as fidgeting fingers, readjustment in seat, and cleaning/wiping off hands/face. Hands moving from or returning to the steering are not a part of the duration. A single finger slightly moving is not enough to count. Start coding when participant is about to contact their face/body, not when hands are removed from the steering wheel.
System Warning	A system warning pops up on the HUD** or IVIS indicating that if the participant does not increase pressure on steering wheel, L2 will disengage. Tesla cars: add 1 system warning for text popping up, and another if the warning light begins to flash. You will usually add both.
Other considerations	A slight phone adjustment/check (including smart watches) does not count, and will be treated as an eyeglance behavior (i.e. see 003_code_1.am 20:30 m/s). When lane lines become solid before the actual exit (in the exit lane), new tasks should not be recorded.

** IVIS: In-Vehicle Information System*

*** HUD: Head Up Display*

APPENDIX B: SUPPLEMENTARY DATA FROM PART 2

The following tables provides the means and standard errors for the effects analyzed in Part 2. Where significant, these means and standard errors are used to generate the figures presented in the results section. Values are provided here for reference.

Automation-L2 Usage		
	Mean	**Std. Error**
Usage Frequency (%)		
Week 1	73.5	5.06
Week 2	76.3	4.60
Week 3	71.3	5.07
Week 4	79.1	4.68
Week 5	74.0	5.04
Week 6	65.5	5.08
Week 7	71.6	5.59
Week 8+	70.0	6.97
Reengagement Time (s)		
Week 1	131	78.3
Week 2	271	72.4
Week 3	199	78.4
Week 4	194	74.1
Week 5	187	80.0
Week 6	168	81.5
Week 7	148	89.6
Week 8+	424	125.6

Warnings and Driving Demand		
	Mean	**Std. Error**
Warnings Frequency (per minute)		
Week 1	0.205	0.04
Week 2	0.133	0.04
Week 3	0.218	0.04
Week 4	0.230	0.05
Week 5	0.366	0.05
Week 6	0.356	0.05
Week 7	0.202	0.05
Week 8+	0.258	0.07
Automation Use by Driving Demand (%)		
Low	0.732	0.05
Moderate	0.705	0.05
High	0.606	0.05

Fatigue and Fidgeting: Factor x Condition		
	Mean	**Std. Error**
Fatigue (%)		
Automation-L2	1.85	0.45
Experimental Control	1.15	0.45
Naturalistic Control	0.47	0.45
Fidgeting (%)		
Automation-L2	6.86	1.91
Experimental Control	8.83	1.92
Naturalistic Control	4.63	1.91

Fatigue and Fidgeting: Factor x Week		
	Mean	**Std. Error**
Fatigue (%)		
Week 1	1.14	0.65
Week 2	0.51	0.58
Week 3	1.21	0.62
Week 4	1.38	0.60
Week 5	0.86	0.65
Week 6	1.75	0.68
Week 7	1.52	0.74
Week 8+	1.31	0.94
Fidgeting (%)		
Week 1	5.20	2.09
Week 2	5.38	2.02
Week 3	6.68	2.06
Week 4	6.03	2.04
Week 5	7.37	2.09
Week 6	8.94	2.12
Week 7	8.61	2.18
Week 8+	8.11	2.40

Distraction and Attention: Factor x Condition		
	Mean	**Std. Error**
Task Aggregate (%)		
Automation-L2	137	7.34
Experimental Control	132	7.36
Naturalistic Control	118	7.33
Radio (%)		
Automation-L2	80.3	7.55
Experimental Control	73.9	7.56
Naturalistic Control	69.9	7.55
Texting (%)		
Automation-L2	4.69	2.12
Experimental Control	3.83	2.13
Naturalistic Control	2.79	2.12
Phone Conversation (%)		
Automation-L2	10.7	5.39
Experimental Control	13.7	5.40
Naturalistic Control	11.5	5.39
Navigation (%)		
Automation-L2	14.0	4.75
Experimental Control	12.3	4.76
Naturalistic Control	13.4	4.74
Video Watching (%)		
Automation-L2	2.40	1.70
Experimental Control	2.61	1.71
Naturalistic Control	1.98	1.70

Distraction and Attention: Factors x Week	Mean	Std. Error
Task Aggregate (%)		
Week 1	126	8.02
Week 2	124	7.73
Week 3	126	7.89
Week 4	124	7.80
Week 5	127	8.02
Week 6	138	8.13
Week 7	142	8.38
Week 8+	137	9.30
Radio (%)		
Week 1	74.3	7.91
Week 2	76.3	7.76
Week 3	74.1	7.84
Week 4	74.1	7.80
Week 5	73.6	7.91
Week 6	77.4	7.97
Week 7	74.1	8.10
Week 8+	71.2	8.59
Texting (%)		
Week 1	0.36	2.24
Week 2	4.41	2.20
Week 3	2.98	2.22
Week 4	4.24	2.21
Week 5	4.44	2.24
Week 6	4.64	2.26
Week 7	5.29	2.31
Week 8+	4.93	2.46
Phone Conversation (%)		
Week 1	10.5	5.75
Week 2	10.3	5.61
Week 3	11.7	5.69
Week 4	11.4	5.64
Week 5	13.5	5.75
Week 6	11.3	5.80
Week 7	15.0	5.93
Week 8+	16.4	6.39

	Mean	Std. Error
Navigation (%)		
Week 1	19.7	5.08
Week 2	13.7	4.94
Week 3	11.7	5.02
Week 4	9.36	4.97
Week 5	8.13	5.08
Week 6	17.4	5.13
Week 7	13.3	5.26
Week 8+	15.0	5.72
Video Watching (%)		
Week 1	1.29	1.86
Week 2	2.24	1.79
Week 3	2.67	1.83
Week 4	4.47	1.81
Week 5	2.09	1.86
Week 6	0.22	1.88
Week 7	2.55	1.94
Week 8+	2.18	2.14

Modality of Interaction: Factor x Condition		
	Mean	**Std. Error**
Auditory Verbal (%)		
Automation-L2	0.72	0.30
Experimental Control	1.08	0.31
Naturalistic Control	0.51	0.30
Visual Manual (%)		
Automation-L2	11.9	3.47
Experimental Control	10.9	3.47
Naturalistic Control	8.8	3.47

Hardware Interface: Factor x Condition		
	Mean	**Std. Error**
Smartphone (%)		
Automation-L2	11.2	3.44
Experimental Control	10.4	3.44
Naturalistic Control	8.41	3.43
Vehicle IVIS (%)		
Automation-L2	1.31	0.39
Experimental Control	1.52	0.39
Naturalistic Control	0.85	0.39

Modality of Interaction: Factor x Week		
	Mean	**Std. Error**
Auditory Verbal (%)		
Week 1	0.60	0.41
Week 2	0.79	0.37
Week 3	0.17	0.40
Week 4	0.89	0.38
Week 5	1.93	0.41
Week 6	0.44	0.43
Week 7	0.70	0.46
Week 8+	0.42	0.58
Visual Manual (%)		
Week 1	5.06	3.73
Week 2	9.45	3.64
Week 3	9.27	3.69
Week 4	13.64	3.66
Week 5	11.18	3.73
Week 6	9.89	3.76
Week 7	15.21	3.84
Week 8+	13.91	4.14

Hardware Interface: Factor x Week		
	Mean	**Std. Error**
Smartphone (%)		
Week 1	4.64	3.71
Week 2	8.45	3.62
Week 3	8.28	3.67
Week 4	13.3	3.64
Week 5	11.6	3.71
Week 6	9.38	3.75
Week 7	14.5	3.83
Week 8+	13.2	4.13
Vehicle IVIS (%)		
Week 1	0.99	0.46
Week 2	1.51	0.43
Week 3	1.06	0.44
Week 4	1.23	0.43
Week 5	1.40	0.46
Week 6	0.91	0.47
Week 7	1.45	0.49
Week 8+	1.10	0.57