Understanding the Impact of Technology: Do Advanced Driver Assistance and Semi-Automated Vehicle Systems Lead to Improper Driving Behavior?

December 2019

607 14th Street, NW, Suite 201 | Washington, DC 20005 | 202-638-5944

AAA Foundation for Traffic Safety

Title

Understanding the Impact of Technology: Do Advanced Driver Assistance and Semi-Automated Vehicle Systems Lead to Improper Driving Behavior?

(December 2019)

Authors

Naomi Dunn, Ph.D., Thomas Dingus, Ph.D., & Susan Soccolich, M.S.

Virginia Tech Transportation Institute

Foreword

Vehicle technologies can now actively control parts of the driving task for extended periods of time, potentially allowing drivers to disengage from their driving responsibilities. It is imperative that we continue to assess how these new features can affect driver behavior and performance. This report summarizes an analysis of data from two Naturalistic Driving Studies using vehicles equipped with advanced driver assistance systems (ADAS). The outcomes shed insight regarding driver behavior while using ADAS versus driving under normal conditions and the safety implications.

The results of this study should help researchers, automobile industry and government entities better understand driver-vehicle interactions and potential unintended consequences in vehicles with advanced technologies.

C. Y. David Yang, Ph.D.

Executive Director
AAA Foundation for Traffic Safety

About the Sponsor

AAA Foundation for Traffic Safety
607 14th Street, NW, Suite 201
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 education 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

List of Figures

List of Tables

Abbreviated Terms

AAAFTS	AAA Foundation for Traffic Safety
ACC	Adaptive Cruise Control
ADAS	Advanced Driver Assistance Systems
AEB	Automatic Emergency Braking
BSW	Blind Spot Warning
DAS	Data Acquisition System
EORT	Eyes-off-Road Time
FCW	Forward Collision Warning
HOW	Hands-on-Wheel
L1	Level 1
L2	Level 2
L2 MFA NDS	Level 2 Mixed Function Automation Naturalistic Driving Study
LCL	95% Lower Confidence Level
LDW	Lane Departure Warning
LKA	Lane-Keep Assist
MIT AVT	Massachusetts Institute of Technology Autonomous Vehicle Technology
NDS	Naturalistic Driving Study
PERCLOS	Percentage of Eye Closure
SCE	Safety-Critical Event
SHRP 2 NDS	Second Strategic Highway Research Program Naturalistic Driving Study
STE	Secondary Task Engagement
UCL	95% Upper Confidence Level
VCC L2 NDS	Virginia Connected Corridors Level 2 Naturalistic Driving Study
VTTI	Virginia Tech Transportation Institute

Executive Summary

Few studies exist to help assess whether drivers of vehicles equipped with advanced driver assistance systems (ADAS) experience a change in crash risk specific to the use of these automated features. The objectives of this study are to 1) Investigate driver behavior and the associated risks of ADAS use; 2) Fill a critical knowledge gap by providing information regarding the potential for changes in driver error, drowsiness, secondary task engagement (STE), and eye-glance behavior (e.g., surrogates for distracted driving behaviors) relative to ADAS use; and 3) Investigate changes in safety-critical event (SCE; i.e., crash or near-crash) risk related to the use of ADAS.

Two naturalistic driving studies (NDSs) — the Virginia Connected Corridors Level 2 NDS (VCC L2 NDS) and the Level 2 Mixed Function Automation NDS (L2 MFA NDS) — focusing on vehicles equipped with advanced automation features (i.e., adaptive cruise control [ACC], lane-keep assist [LKA], and Level 2 [L2, or ACC and LKA] automation), were used to inform the results of the investigation.

Drivers for both the VCC L2 NDS and the L2 MFA NDS were recruited from the Northern Virginia, Washington D.C., and Maryland areas, based on daily driving habits and driving history, among numerous other selection criteria. The VCC L2 NDS was a 12-month field evaluation of 50 participant-owned vehicles. Vehicle makes and models varied, but all VCC L2 NDS vehicles included in this study had a combination of ACC and LKA functionality. The L2 MFA NDS comprised 120 participants, each assigned to an L2-equipped study vehicle for four weeks, resulting in 120 months of driving data. Due to the large quantities of data collected during both NDSs, analyses conducted for the current study were based on sampled epochs of baseline driving, which may be defined as “ordinary” driving not tied to a specific event (e.g., a crash or near-crash). These baseline epochs were selected based on ADAS status (e.g., L2 active or none active).

The sampling strategy used varied between the two studies. The VCC L2 NDS used a matched baseline sampling strategy, whereby baseline epochs selected when ADAS were active were matched to a corresponding epoch where the same system was available but not active. For example, an L2 active epoch was matched to an epoch where L2 was available but not active. Baseline epochs were also matched by driver, day of the week, time of day, and vehicle speed to account for variability in driver behavior resulting from such factors. To avoid oversampling from any one trip, only one baseline epoch was sampled in the VCC L2 NDS from each trip. Sampling for the L2 MFA NDS was also based on ADAS status. However, the strategy was different, as the L2 MFA NDS did not use a matched sampling strategy. Rather, a baseline epoch was sampled from every period of ADAS activation during every trip when the vehicle speed was above 40 mph and lane markings were visible (i.e., verified by a trained data reductionist). This approach resulted in a much larger baseline data set. However, such sampling did not control for variability among drivers or driving scenarios. Due to the differences in sampling among the NDSs, separate analyses were conducted on the two data sets. For the VCC L2 NDS, comparisons were made between ADAS active and ADAS available but not active (e.g., L2 active was compared to L2 available; ACC active was compared to ACC available). For the L2 MFA NDS, comparisons were made between different levels of ADAS activation (i.e., L2 active was compared to L1 active and none active).

Driver behaviors were combined into two categories due to low numbers. Performance errors included a variety of vehicle operation and maneuvering errors, such as failing to signal or an improper turn. There were relatively few performance errors in either NDS data set. The prevalence of performance error occurring during a trip was 1.6% to 1.7% across both data sets, which was lower than the 4.8% found in overall analyses of the Second Strategic Highway Research Program (SHRP 2) NDS (Dingus et al., 2015). Conversely, the prevalence of judgment errors, which were operationally defined to include aspects of a momentary lapse of judgment by the driver (Dingus et al., 2016), was higher in the current study than that found in Dingus et al. (2016). However, this difference likely stems from a higher occurrence of exceeding the speed limit (≥ 10 mph over the posted speed) by drivers in both the VCC L2 NDS and L2 MFA NDS. When both ACC and LKA (i.e., L2) were available but not in use (i.e., active) in the VCC L2 NDS, the prevalence of a judgment error was 17.5%. However, this decreased to 11% with L2 active. Conversely, speeding in the L2 MFA NDS was significantly higher when L2 was active compared to when no systems were active (19% versus 16%, respectively), indicating that drivers exceeded the speed limit more frequently when using both lateral and longitudinal automation features simultaneously.

Observation of drowsy driving in the VCC L2 NDS was low, with only four trips (0.6%) indicating signs of driver drowsiness. Drowsy driving was present in a higher percentage of L2 MFA NDS baselines, most notably when both systems were active (5.4%), indicating a possible detrimental effect of automation use associated with driver underload. When looking specifically at possible drowsiness-related SCEs, the number of valid events was too low to form any conclusions based on the available data.

The results from the VCC L2 NDS indicate that the use of lateral and longitudinal ADAS (i.e., L2 automation) culminated in increased occurrence of distracted driving behaviors. Drivers with L2 active had 1.8 times the odds of engaging in a visual, manual, or visual-manual secondary task than when L2 was available but not active. When VCC L2 NDS drivers were involved in a secondary task with L2 active, they spent almost 30% of the time with their eyes off the forward roadway. In addition, drivers with L2 active took more frequent and longer duration non-driving-related task glances, subsequently spending less time with their eyes on driving-related tasks. Interestingly, drivers from the L2 MFA NDS did not display the same tendency toward distracted driving behaviors when automation systems were active. Instead, they were more likely to engage in a secondary task and take their eyes off the road when they were driving with no ADAS engaged (i.e., under periods of manual driving), possibly indicating a lack of trust in the systems.

These results — and the differences between the data sets used — may suggest the possibility of several phases of driver interaction with automated systems. The authors therefore propose a three-phase model of ADAS operation, comprising a "novelty" phase, a post-novelty operational phase, and an experienced user phase. The novelty phase includes learning and testing the ADAS in real time to understand limitations and capabilities. Trust in the ADAS may be lacking during this phase as the driver has little to no experience with the automation features. Once drivers move from the novelty phase to the post-novelty operational phase, behavioral adaptation may begin to occur, and overreliance and over-trust in the automation features may develop. The post-novelty phase is also when driver awareness of the limitations of the ADAS features ideally evolves to avoid an increase in SCEs associated with risky behaviors, such as distracted driving. Drivers then move to the experienced user phase, wherein overreliance and work underload may manifest in the form of drowsiness or inattention.

These results, ideally, will help raise awareness among drivers of the potential pitfalls they may experience throughout all phases of ADAS operation, from inexperience and initial system use to experienced use. To help mitigate such drawbacks of system use, comprehensive training for drivers purchasing ADAS-equipped vehicles may have a positive outcome. Additionally, every ADAS-equipped vehicle currently operates differently across functionality, ADAS activation, capabilities, and operating speeds, resulting in a learning curve each time a driver uses a different system. As such, standardization to a greater degree may also be beneficial in that it will limit the impact of novelty symptoms.

Introduction

The ultimate goal of advanced driver assistance systems (ADAS) is to increase road safety and driving comfort. These systems help with monitoring, braking, and steering tasks and alert the driver when corrective or evasive action is required. Several studies have estimated the impact of a range of ADAS on crashes, injuries, and deaths on U.S. roads, including forward collision warning (FCW), automatic emergency braking (AEB), lane departure warning (LDW), and blind spot warning (BSW) systems (Cicchino, 2017a; 2017b; 2017c; Benson et al., 2019).

The development of specific ADAS has aided in the evolution of semi-automated and automated vehicles. Generally speaking, adaptive cruise control (ACC) and lane-keep assist (LKA, or lane centering) are the two fundamental systems required for advancement of autonomous vehicles. As these systems that control the longitudinal and lateral movement of the vehicle become more advanced and refined, they can more reliably and consistently assume control of the driving task from the operator. However, new advances in vehicle technologies and rapid deployment of these advanced systems in vehicles already on the road leave unanswered questions about their impact on driver behavior. Specifically, as drivers are removed from the driving task for periods of time, their ability to retake control of the vehicle in the event of automation failure becomes questionable.

Recent automated-vehicle studies (e.g., Schoettle & Sivak, 2015; Blanco et al., 2016) have attempted to determine crash risks relative to semi-automated and autonomous vehicles. However, research conducted thus far is limited and provides only a narrow snapshot of what may occur when these vehicles are more widely deployed in the future. Few studies have been conducted to discover whether drivers of such vehicles will experience increased crash risk specific to the use of ADAS. But the evidence so far points to the need for more studies.

For example, one study measured drivers' willingness to engage in secondary tasks and their visual performance while operating an L2 automated system (Llaneras et al., 2013). The study found that, when operating under a vehicle capable of automated steering and maintaining speed and headway, engagement in secondary tasks increased. The length of drivers' off-road glances also increased, and they diverted their attention from the forward roadway more than 30% of the time. An earlier study by Rudin-Brown and Parker (2004) focusing on ACC found that drivers performed significantly better on a secondary task when using ACC, while their response time to hazard detection increased. The authors suggested that such results were due to behavioral adaptation to the system. A recent survey of owners of vehicles equipped with ADAS found that approximately 30% indicated they were comfortable engaging in non-driving-related tasks when ACC was activated (McDonald et al., 2018).

A key concern when considering the widespread adoption of these advanced automation features is that of behavioral adaptation. Evolutionarily speaking, humans adapt as they learn; thus, when presented with new technologies, the driver will learn and adapt to using these technologies over time and under changing conditions. Behavioral adaptation can implicate psychological factors, such as motivation, emotions, personality traits, and decision-making that inform how an individual will respond to the change (Vaa, 2013). The Organization for Economic Co-operation and Development (OECD, 1990) defines behavioral adaptation as “the collection of behaviors that occurs following a change to the road traffic system that were not intended by the initiators of the change”. Behavioral adaptation, including the related concept of risk compensation, has been postulated by numerous researchers as a contributing factor to changes in driver behavior following the introduction of road-safety measures, such as antilock brake systems (Vaa, 2013), airbags (Peterson & Hoffer, 1994), and seatbelts (Calkins & Zlatoper, 2001). The implementation of safety countermeasures does not occur in a vacuum, especially when humans are involved; thus, addressing one problem in a system has the potential to result in unintended consequences due to behavioral adaptation.

The development of behavioral adaptation involves two phases: the learning phase and the integration phase (Saad et al., 2004). In regard to the introduction of new technology, such as ADAS, the learning phase involves the driver initially becoming acquainted with the systems and learning their uses and limitations. Learning the limitations can be accomplished formally by instruction. However, learning to use the systems in real-world scenarios will largely occur organically as different driving scenarios are encountered. The integration phase occurs once the driver is past the initial learning phase and has incorporated the system use in everyday driving (Saad et al., 2004). Thus, behavioral adaptation takes time as drivers move through these phases of learning and integration. As a result, to better understand the process and development of behavioral adaptation to ADAS and the long-term impact on driver behavior, studies ideally should investigate ADAS use over an extended period of time.

A critical issue relative to behavioral adaptation and specific to ADAS use is that drivers may divert their attention away from the driving task to engage in secondary, non-driving-related tasks. This could have dire consequences should the ADAS fail or become inactive, meaning the driver needs to quickly redirect his/her attention back to the road and resume manual control of the vehicle (Sullivan et al., 2016). If the driver is distracted and has his/her attention on a non-driving-related task, then the operator will not be prepared to intervene in the event of automation failure. For example, numerous studies have indicated that the combination of ACC and active steering results in longer response times, increased hard braking, and more collisions or near collisions when automation fails, as compared to manual driving (Strand et al., 2014; Merat et al., 2014; de Winter et al., 2014). Conversely, a recent study focusing on Tesla drivers' behavioral adaptation by Lin et al. (2018) found that, while drivers universally engaged in secondary tasks under automated driving conditions, these drivers also reported learning from their experiences with the vehicles to identify "safe" scenarios, as well as adopting a margin of safety to avoid excessive risk. It should be noted that this was behavior self-reported in a semi-structured interview, so real-world behavior may be different.

The Massachusetts Institute of Technology (MIT) is currently undertaking a large-scale real-world driving data collection effort called the MIT Autonomous Vehicle Technology (MIT AVT) study, which focuses on driver behavior and interaction with advanced vehicle automation (Fridman et al., 2017). While this research is ongoing, initial results from a subset of the MIT AVT data examining the Tesla Autopilot feature relative to driver vigilance indicated that drivers did not over-trust the system and remained functionally vigilant during use of the Autopilot feature (Fridman et al., 2019). Drivers either performed anticipatory actions or responded immediately to what the researchers categorized as "tricky situations," meaning challenging scenarios that required input from the driver to maintain safe operation of the vehicle, resulting in disengagement of the Autopilot feature. These results, though surprising and contrary to what would be expected if behavioral adaptation occurred among this subset of drivers, are nonetheless preliminary. It therefore remains imperative in transportation research to investigate potential over-trust and overreliance on ADAS as these issues could mitigate some of the safety benefits of these advanced automation features (Inagaki & Itoh, 2013).

In addition to behavioral adaptation, the use of automation features that remove sources of workload from the driver may create issues associated with mental underload. Reducing mental workload may be helpful, to a point. However, reducing it too much has been shown to negatively impact driver performance (Nachreiner, 1995; Young & Stanton, 2002). Underload occurs when the task demands are relatively low and the driver does not need to mobilize too many cognitive resources, or mental effort, to maintain performance (Gimeno et al., 2006). Young and Stanton (2004; 2006) conducted a series of studies on the impact of automation on workload and driver performance and found higher levels of automation (e.g., ACC) were associated with significantly lower workload and poorer response times (1-1.5 s slower) to unexpected automation failure. Underload has also been linked to monotony and fatigue. In a separate simulator study investigating the detrimental effects of passive task-related fatigue (i.e., fatigue associated with monotonous, low-demand driving), Matthews et al. (2009) found drivers were slower to respond to an unexpected hazard and were more likely to collide with the hazard. Thus, while there are safety benefits associated with ADAS, there may also be associated underload and fatigue-related declines in driver performance that need to be investigated.

Finally, many of these advanced safety systems have limitations and specific conditions under which they are designed to operate optimally. Unfortunately, owners of vehicles equipped with these advanced safety systems typically lack awareness of the key limitations of these technologies. A recent survey of registered owners of vehicles equipped with various ADAS found more than three-quarters of owners of vehicles with blind spot monitoring systems had misconceptions about its function or were unsure of the system limitations. In addition, one-third of respondents whose vehicles had AEB systems did not realize the sensors and cameras on which the system relies could be blocked or obstructed by dirt, snow, or ice (McDonald et al., 2018).

The majority of studies investigating various safety factors related to the use of ADAS have been simulator or test-track studies with instrumented vehicles. These studies have the advantage of keeping participants safe, but their external validity may be questioned because of the controlled environment. To help address the disadvantages of such studies, the Virginia Tech Transportation Institute (VTTI) created the naturalistic driving study (NDS) method to continuously capture video and other in-depth data from key-on to key-off in drivers' personal vehicles, or in leased vehicles provided to participants, without an experimenter present. The outcomes provided robust information about real-world driver behavior and performance leading to safety-critical events (SCEs; e.g., Dingus et al., 2006; 2015; Fitch et al., 2013; Klauer et al., 2010). SCEs include crashes, wherein the vehicle makes contact with another object, and near-crashes, wherein a crash is avoided by an evasive maneuver (Guo & Fang, 2013).

Naturalistic driving data also comprise ordinary, or baseline, driving during which no SCEs occur. Because detailed information about driver behavior and the surrounding environment is available before, during and after a crash – and during ordinary driving, naturalistic data can help overcome significant pre-crash behavioral limitations of epidemiological studies. For example, VTTI recently investigated the crash risk associated with cognitive distraction and drowsy driving (Dingus et al., 2019). In this study, the authors examined a sample of data from more than 3,500 drivers in the Second Strategic Highway Research Program (SHRP 2) NDS to determine the prevalence of engagement in tasks involving cognitive distraction (i.e., tasks with limited visual-manual demands, such as interacting with a passenger) across events in which crashes occurred and in comparison to baseline (non-crash, ordinary) driving. Results of the study indicated that engagement in some secondary tasks associated with cognitive distraction had an increase in crash risk when compared to baseline driving. The authors also found a significant increase in crash risk relative to drowsy driving, which was identified using percentage of eye closure (PERCLOS), or the percentage of time the driver's eyes were closed more than 80% across one- or three-minute periods.

Additionally, naturalistic driving data can be used to assess various eye-glance metrics that act as surrogates for driver distraction. Eye-glance locations are coded according to whether they are on-road or off-road and related to a driving task (e.g., checking mirrors) or a non-driving-related task (e.g., looking at a cell phone). Numerous eye-glance metrics have been shown in past studies to be associated with crash risk. For example, Klauer et al. (2006)

found that, when the driver's eyes were off the forward roadway for longer than a total of 2 seconds (i.e., total cumulative eyes-off-road duration > 2 seconds), the driver's crash/near-crash risk doubled. Additionally, the authors found the number of glances away from the forward roadway and the duration of the longest glance increased as the severity of the SCE increased. Dingus et al. (2016) also found evidence that activities requiring eyes-off-road time resulted in a higher crash risk. Distracting visual-manual activities — such as texting on a handheld cell phone, dialing a number on a handheld cell phone, and reaching for an object — were all associated with increased crash risk.

Using the NDS method, VTTI collected naturalistic driving data from ADAS and Level 2 (L2) automation-equipped vehicles. In the current study, the resulting database of continuous naturalistic driving data was mined and analyzed to determine if ADAS have unintended consequences on driver safety. This offers a chance to investigate how drivers use ADAS in real-world conditions and whether these systems have a detrimental effect on driver behavior. This study is exploratory in nature and aims to fill a critical knowledge gap regarding driver risk in relation to ADAS use. As a result, the research questions cover a broad range of topics, including driver behavior, secondary task engagement (STE), and driver drowsiness. The goal is to provide a better understanding of possible changes in driver risk due to distraction, fatigue, and other factors.

Research Questions

The research questions for this study are listed below. Each research question is addressed using the two NDS data sets, which are described in the Methods chapter.

1. What driver behaviors are observed when ADAS are active?
2. Do unsafe driver behaviors occur more frequently when ADAS are active?
3. Does STE occur more frequently when ADAS are active?
4. Do the characteristics of SCEs change when ADAS are active?
5. Do SCE rates differ when ADAS are active?
6. Is there an increased prevalence of STE during SCEs that occur when ADAS are active?

7. Do drivers spend more time with their eyes off the forward roadway when ADAS are active?
8. When engaged in a secondary task, do drivers take longer glances away from the roadway when ADAS are active?
9. In general, do drivers engage in less scanning of the roadway environment when ADAS are active?
10. Is driver drowsiness observed more often when ADAS are active?
11. Is driver drowsiness more prevalent during SCEs that occur when ADAS are active?
12. How do drivers respond to ADAS alerts?
13. How long does it take drivers to respond to ADAS alerts?
14. Is driver drowsiness more prevalent when drivers receive ADAS alerts?

Methods

Vehicles and Equipped ADAS

Data for this study were derived from two naturalistic driving data collection efforts using vehicles equipped with longitudinal and lateral ADAS: 1) The L2 Mixed Function Automation (MFA) NDS and 2) the Virginia Connected Corridors (VCC) L2 NDS. Vehicles used in both studies were equipped with varying driving automation systems, which were reflective of market availability at the time of data collection. Individual systems differed in terms of their functionality and capabilities (across vehicle make and manufacturer). However, all vehicles in the study allowed drivers to simultaneously activate longitudinal and lateral automation systems (i.e., L2).

Longitudinal automation features are commonly referred to as Adaptive Cruise Control (ACC) systems. When activated, ACC automatically adjusts the speed of the vehicle to maintain a safe following distance from a lead vehicle. Following distance, also called headway, can usually be set manually, from one car length to three or four car lengths. Lateral automation (i.e., lane-keeping) features differ across vehicle make and model and are largely based on whether the lane-keeping system is proactive or reactive. Lane-Keep Assist (LKA) is a reactive system that will steer the vehicle back into the designated lane if it begins to drift over the road's lane markings. LKA systems also typically include Lane Departure Warning (LDW) features, whereby the driver is provided with alerts (i.e., audible, visual, and/or tactile) if the vehicle drifts over lane markings. Conversely, lane centering is a proactive lane-keeping system that monitors lane markings and provides steering input to actively keep the vehicle in the center of the detected lane (Motor1.com, 2019).

Due to variety in the study vehicles equipped with ADAS, longitudinal automation features will generically be referred to herein as ACC. Given the broad array of vehicles included and the exploratory nature of the current study, both proactive and reactive lane-keeping (lateral automation) systems are grouped together and referred to collectively as LKA. Functionality achieved when ACC and LKA are activated simultaneously will be referred to as L2 active. It should be noted that true L2 automation, as defined by SAE International (2018), requires proactive lane-keeping systems, such as lane-centering, which perform part of the dynamic driving task on a sustained basis (i.e., the driver is relieved of his/her role in steering the vehicle). Reactive LKA systems, with which several of the vehicles in the VCC L2 NDS were equipped, only provide momentary action as a response to potentially hazardous situations (i.e., the driver is not relieved of his/her role in steering the vehicle). Thus, while these are active safety systems, they are not considered to be driving automation (SAE International, 2018).

Regardless of individual vehicle capability, all features available on each vehicle included in this study required active monitoring from the driver and intervention when and where necessary. The lateral automation features, regardless of whether they were proactive or reactive LKA systems, required the driver's hands to be on the wheel. There were also conditions in which the ADAS were not intended to be used, which were consistent across all vehicles. These typically included conditions that would result in poor visibility, such as fog, heavy rain, or snow. A main difference between study vehicles and equipped ADAS features was whether LKA could be activated independently of ACC. Some vehicles, such as Tesla, require ACC to be activated prior to the engagement of LKA, meaning the driver can only have ACC active or both systems active, but not LKA alone. Other vehicles allow ACC or LKA or both to be activated. Table 1 provides a breakdown of the vehicles and equipped ADAS features included in the VCC L2 NDS and L2 MFA NDS.

Table 1. VCC L2 NDS and L2 MFA NDS vehicles and breakdown of ADAS features.

						ACC		LKA	
						Operating Speed (mph)		Operating Speed (mph)	
	Vehicle Make/Model	**N**	**L2**	**ACC**	**LKA**	**Min**	**Max**	**Min**	**Max**
VCC L2 NDS	2015 Tesla Model S	3	X	X	X	18	90	30	85
	2016 Tesla Model S	8	X	X	X	18	90	30	85
	2017 Tesla Model X	1	X	X	X	18	90	35	90
	2014 Acura MDX	1	X	X	X	25	--	45	90
	2015 Acura TLX	2	X	X	X	25	--	45	90
	2016 Acura RDX	2	X	X	X	25	--	45	90
	2015 Ford Fusion	1	X	X	X	16	--	40	--
	2017 Ford Fusion	1	X	X	X	16	--	40	--
	2016 Honda Accord	2	X	X	X	25	--	45	90
	2017 Honda Accord	2	X	X	X	25	--	45	90
	2015 Hyundai Genesis	2	X	X	X	20	110	40	110
	2016 Hyundai Genesis	1	X	X	X	20	110	40	110
	2017 Chrysler Pacifica	1	X	X	X	20	--	37	112
	2014 Jeep Cherokee	1	X	X	X	20	--	37	112
	2016 Hyundai Sonata	2	--	X	LDW	20	110	N/A	N/A
L2 MFA NDS	2015 Tesla Model S	2	X	X	X	18	90	30	85
	2017 Audi Q7	2	X	X	X	20	95	40	--
	2015 Infiniti Q50	2	X	X	X	20	90	45	--
	2016 Mercedes E350	2	X	X	X	20	120	37	120
	2016 Volvo XC90	2	X	X	X	20	125	40	--

Note: Dashes indicate there is no maximum operating speed for the ADAS feature.

Study Sample

VCC L2 NDS

For the VCC L2 NDS, 50 participants were recruited from the Northern Virginia and Washington, D.C., areas. Due to vehicle variability and dash icon arrangement or design, a subset of 30 participants was used for this study. Participants were required to own or lease a vehicle with at least ACC functionality, as well as meeting the following criteria:

- Must be at least 18 years of age and have at least two years of driving experience
- Must regularly drive (i.e., two to three days per week) on Interstate 66, U.S. 29, U.S. 50, or Interstate 495 in the Northern Virginia or Washington, D.C., area

Data collection for the study took a total of 20 months, from November 2016 to June 2018, with approximately one year of data collected for each participant. Of the 30 participants used in the current study, the majority were male (22 males and eight females). There was a wide age range, with eight participants between 25 and 39 years of age, 15 participants between 40 and 54 years old, and seven participants between 55 and 77 years of age.

Since participants used their own vehicles during the VCC L2 NDS, there was no specific training on how and when to use the ADAS. It was assumed that participants would drive as they normally would and used the systems if, and when, they chose. Responses to a questionnaire completed upon enrollment in the study indicated these participants were familiar with the ADAS and trusted the systems in their vehicles.

Participants were compensated at a rate of $300 per month of participation up to a total of 12 months per participant. An additional payment of $125 was made to each participant following installation of a data acquisition system (DAS) in their vehicles and again once the DAS equipment was removed from their vehicles. Thus, a participant enrolled in the study for a full year was compensated $3,850 for his/her participation.

L2 MFA NDS

Participants in the L2 MFA NDS were recruited from the Washington, D.C., area, which included both Northern Virginia and Maryland suburbs. For the L2 MFA NDS, participants were provided a study vehicle equipped with ADAS. Participants were selected based on the following criteria:

- Must drive at least 60 miles per weekday
- Must not have driver’s license suspension within the last seven years and must agree to submit to a driving history check

- Must not have been convicted of more than two driving violations or been involved in an at-fault crash causing injury within the past three years
- Must own and use a smartphone with Bluetooth capability

A total of 120 participants completed the study. Each participant was assigned to one of 10 study vehicles, resulting in 12 participants per vehicle. This study was designed to balance participants across age and gender classifications for each vehicle. The age group classifications were 25 to 39 years old and 40 to 54 years old. Of the 12 participants per vehicle, six were in the 25- to 39-year-old age group (three males and three females), and six were in the 40- to 54-year-old age group (three males and three females).

Data were collected for a total of 16 months, from September 2016 to December 2017, with each participant enrolled for a period of four weeks. Since participants were assigned a study vehicle, it was not assumed they had prior experience using a vehicle equipped with any ADAS features (e.g., ACC or LKA). Thus, all participants received an orientation to their assigned vehicle, as well as training regarding the use of the driving automation system features. Training comprised a static orientation of all basic vehicle controls (e.g., seat adjustment, windshield wipers), comfort features (e.g., seat warmers, navigation and entertainment systems), and driving automation system features (e.g., ACC and LKA). This included verbal instruction on the location and operation of all associated buttons, levers, and alerts (i.e., auditory, visual, and haptic). After the static orientation, the participant was taken on a two-part test drive, whereby the researcher first drove and demonstrated the vehicle automation features, then the participant took over driving. The test drive was considered complete once the researcher had answered all of the participant's questions and the participant indicated he/she felt comfortable driving the study vehicle with the driving automation features. The orientation and training session took approximately 1.5 hours. Participants also completed several self-report questionnaires, one of which asked about their previous experiences with ADAS. The majority of participants reported they had little to no direct experience with any driving automation systems. Survey results showed that 63% had heard of an "automated vehicle system" of some type.

Participants were compensated up to $500 based on total mileage driven and questionnaire completion. No monetary incentive was provided to use the driving automation systems.

Naturalistic Driving Data Sets

The two NDS data sets used in this study were produced by equipping each study vehicle with the VTTI NextGen DAS – the same system used in the SHRP 2 NDS (Dingus et al., 2015). The DAS used unobtrusive video cameras and vehicle sensors to continuously collect and store driving-related data. Five cameras installed in the vehicle provided views of the forward roadway, the driver's face, an over-the-shoulder view of the driver's hands and lap, the foot well (i.e., the accelerator and brake), and a rear view (Figure 1). The DAS also recorded vehicle data, including speed, acceleration, brake application, lane position, and GPS coordinates. A sixth camera was installed specifically to capture a view of the instrument panel (Figure 2), which provided information related to ADAS status.

Figure 1. Data Acquisition System camera views

Figure 2. Example of an instrument panel camera view

A VTTI-developed machine-learning algorithm was used to determine the ADAS status of each study vehicle at any given point during a trip. This algorithm was based on image classification using deep neural networks, specifically ResNet-18 architecture (He et al., 2015). A training set of images was created for each vehicle make and model used in the current study. The images depicted dash icons representing each different ADAS status assessed in the current study (i.e., ACC active, ACC available, LKA active, LKA available, L2 active, L2 available). Six vehicles from the VCC L2 NDS were excluded from the study as a result of poor post-training accuracy when tested on the validation data sets. This was typically due to multiple dash icons appearing in the same location or the style or size of the icons, all of which contributed to higher false-positive rates or a misclassification of ADAS status. The resulting trained model was then applied to video collected by the instrument panel camera, which in turn produced time-series data for each vehicle that indicated ADAS status.

Data Sampling and Reduction

NDSs provide large amounts of continuous data recorded while study participants are driving. To create a manageable data set that can be investigated and analyzed, these data need to be sampled to identify epochs of interest. These epochs are then inspected by trained data reductionists and coded to allow for comparison of driver-, environment-, roadway-, and vehicle-related variables. Three groups of epochs were gathered to inform different research questions: baseline epochs, safety critical event (SCE) epochs, and alert epochs. Sampling of baselines varied between the two NDS data sets.

VCC L2 NDS Sampling

For the VCC L2 NDS, driving epochs of 10 seconds were identified based on driver and ADAS activation, then matched with a corresponding epoch from the same driver where the same system was available but not active. For example, if Driver #1 had ACC active, a corresponding epoch from Driver #1 was identified where ACC was available but not active. Similarly, if Driver #5 had both ACC and LKA systems (i.e., L2) active, this was matched to a corresponding epoch from Driver #5 where both systems were available but not active. Baseline epochs were also matched by:

- Driver;
- Day of the week (i.e., weekday or weekend);
- Time of day (i.e., 6 am-9 am, 9 am-4 pm, 4 pm-7 pm, 7 pm-11 pm, 11 pm-6 am); and
- Vehicle speed above 20 mph but within +/- 5 mph.

The initial study plan was to randomly sample an equal number of baseline epochs from each driver, which were matched based on ADAS status. The sampling approach was as follows:

- 16 matched pairs for L2 active (i.e., eight epochs of L2 active matched with eight epochs of L2 available but not active);
- 8 matched pairs for ACC active (i.e., four epochs of ACC active matched with four epochs of ACC available but not active);
- 8 matched pairs for LKA active (i.e., four epochs of LKA active matched with four epochs of LKA available but not active).

In practice, these numbers were not observed in all cases for all vehicles in the VCC L2 NDS. The LKA active group was decidedly lower than originally planned as some of the vehicles (e.g., Tesla) do not allow for activation of LKA independently of ACC (i.e., ACC must be active before LKA can be activated). In these instances, the sampling approach was adjusted, and the number of ACC active epochs was doubled to 16 matched pairs when available (i.e., instead of eight). Table 2 provides a summary of the planned sampling approach versus the total epochs sampled in the VCC L2 NDS.

Table 2. Summary of planned sampling approach and final total number of epochs in the VCC L2 NDS.

ADAS Status	Planned Number of Epochs	Actual Number
L2 Active	240	200
L2 Available	240	200
ACC Active	120	133
ACC Available	120	133
LKA Active	120	71
LKA Available	120	71
Total	960	808

The matched baseline sampling strategy was designed to obtain data from each driver engaged in each ADAS status—regardless of their time in any one ADAS status—while controlling for potentially confounding factors by matching samples of certain conditions. This sampling plan was not designed to be representative of all possible driving scenarios or representative of time spent in each ADAS status. This study and sampling method were designed instead to give all drivers equal weight, allowing equal influence among different user types on the understanding of how ADAS status may affect driving behaviors. This matched sampling strategy accounted for variability between drivers, in general, and between drivers in different scenarios. As such, driver behavior varied between individual drivers in different types of traffic (e.g., commuting to work versus weekend driving), at different times of the day (e.g., driving late at night versus in the middle of the day), and at different vehicle speeds (e.g., driving at 70 mph on an interstate versus driving 30 mph on a residential street).

In addition to the matching criteria above, only one baseline epoch was sampled from each trip in the VCC L2 NDS. This avoided potential pitfalls associated with oversampling from any one particular trip. For example, if a driver was highly fatigued during one particular trip and 10 baseline epochs were sampled from that trip, it may give an overrepresentation of the prevalence of driver drowsiness.

Established kinematic algorithms developed in previous NDSs (e.g., Blanco et al., 2016; Dingus et al., 2006; Fitch et al., 2012; Hanowski et al., 2008; Simons-Morton et al., 2011) were used to identify potential SCEs. The algorithms identified rapid longitudinal decelerations, rapid lateral accelerations, short time-to-collisions, and substantial swerving. Trained data reductionists then inspected the videos associated with these events to verify the occurrence of an SCE. All verified SCEs (i.e., crashes and near-crashes) from each data set were included in the analysis (see Table 3).

Along with SCEs and matched baselines, three different kinds of alerts—FCW, immediate takeover, and hands-on-wheel (HOW)—were sampled from the Tesla vehicles in the VCC L2 NDS (see Figure 3 for an example of each alert type). The Tesla was chosen because the dash icons for each alert were the most easily identified by the machine vision algorithm. In addition, these types of alerts may not occur frequently; thus, since there were 12 Teslas in the VCC L2 NDS, using the Tesla maximized the potential to identify these alerts. Different vehicle makes and models use different criteria to determine what constitutes an event requiring an alert; hence, using alert data from multiple different vehicular makes and models would have produced mixed results.

Figure 3. Examples (from left to right) of Forward Collision Warning (FCW), Immediate Takeover, and Hands-on-Wheel (HOW) alerts in the VCC L2 NDS

FCW relates to longitudinal vehicle control and warns the driver when he/she is too close or closing too quickly to the lead vehicle. Immediate takeover alerts relate to lateral vehicle control and warn the driver when he/she needs to assume steering of the vehicle. HOW prompts occur when the vehicle senses the driver does not have his/her hands on the steering wheel and prompts the driver to place his/her hands on the wheel (see Table 3 for summary of all alert epochs).

Table 3. Total epochs (SCEs and sampled alerts) for the VCC L2 NDS.

Epoch		Total
SCEs		159
Alerts	**FCW**	63
	Immediate Takeover	61
	HOW Prompt	391

L2 MFA NDS Sampling

A different strategy was used for baseline sampling of the L2 MFA NDS. All periods in which ADAS were active and ADAS were available, but not active, were identified using vehicle network information or the VTTI-developed machine-learning process. Baseline epochs were only sampled when the vehicle was traveling above the speed required for activation of the driving automation features, which was designated as 40 mph in the L2 MFA NDS, and the vehicle was traveling on a road with visible lane markings (i.e., verified by trained data reductionists). Using these criteria, baseline epochs of 15 s were selected from every period of system activation or availability within every trip, based on whether both ACC and LKA systems were active (i.e., L2 active), one system was active but L2 was available (i.e., ACC or LKA active), and no systems were active but L2 was available. Baseline epochs were also stratified by each week of study participation. In addition to baselines, all SCEs were identified, along with a subset of alerts. The alerts included in the L2 MFA NDS were all generated as part of the lateral driving automation feature. Thus, alerts were based on lack of detected steering input from the driver and/or crossing a detected lane marking. FCW alerts and HOW prompts were not sampled in the L2 MFA

NDS. Table 4 shows the number of sampled baseline epochs for each ADAS level in the L2 MFA NDS, as well as SCEs and alerts.

Table 4. Total epochs (baselines sampled for each ADAS status, SCEs, and alerts) for the L2 MFA NDS.

Epoch	Total
L2 Active	1,388
L1 Active (L2 available)	1,139
None Active (L2 available)	1,228
SCEs	71
Alerts	450

For each epoch, trained data reductionists used recorded video and kinematic data to annotate the driver, vehicle, and environmental factors present during each of the ADAS activation levels, SCEs, or alerts. All data reductionists for both the VCC L2 NDS and L2 MFA NDS underwent identical training procedures, including a coding and feedback loop wherein they were required to reach an accuracy rate of at least 90% (i.e., when compared to an expertly reduced set of events) before they could progress to independently reducing new events. All data reduction completed by new reductionists was initially quality checked by a senior reductionist or reduction coordinator for a 100% rate of quality-control checks. The rate of quality checks was gradually reduced to 75%, then to 50%, if the reductionist's accuracy remained consistent for at least one week. The VTTI data dictionary (VTTI, 2015) was used for data reduction in both studies. This dictionary includes definitions used in the data reduction process, descriptions of all reduction variables, and examples of secondary tasks. Environmental variables included weather, lighting, roadway type, and traffic density, all of which may impact ADAS usage. In addition to epoch reduction, eye-glance analysis was performed on all baseline epochs, SCEs, and alerts for both studies. Eye-glance locations were then classified as on-road/off-road and driving-related/non-driving-related for the purposes of analysis. Table 5 includes eye-glance locations and on-road or driving-related classifications used in both data sets.

Table 5. Eye-glance location classifications used in both NDS data sets.

Eye-Glance Location	On-Road Classification	Driving-Related Classification
Center Dashboard Console	No	No
Interior Object	No	No
Cell Phone	No	No
Passenger	No	No
Over-the-Shoulder (left or right)	No	Yes
Eyes Closed	No	No
Instrument Panel	No	Yes
Other	No	No
No Eyes Visible - Location Unknown	Unknown	Unknown
Forward	Yes	Yes
Right Windshield	Yes	Yes
Rearview Mirror	No	Yes
Left Windshield	Yes	Yes
Left Window/Mirror	No	Yes
Right Window/Mirror	No	Yes

Driver drowsiness was also assessed on both NDS data sets using PERCLOS, which uses video of the driver's face to determine the percentage of time a driver's eyes are closed, here across a one-minute period (i.e., PERCLOS 1). PERCLOS 1 was performed on epochs where the driver's eyes were visible (i.e., not occluded by an object, sunglasses, glare, or poor video quality) for at least 80% of the epoch. Trained data reductionists viewed the video frame-by-frame (15 Hz capture rate) and selected one of three classification options for each frame: *Eyes Open*, *Eyes Closed*, or *Unknown*. Consistent with the PERCLOS definition, *Eyes Closed* was operationally defined as the eyelid being at least 80% closed and covering the pupil. After completion of an event, reductionists reviewed their work by watching the video at half-time playback speed to ensure accuracy. If the driver's eyes were coded *Eyes Closed* for more than 12% of the valid video frames in an epoch, the driver was considered drowsy. PERCLOS was performed on all baseline epochs, SCEs, and alerts for the VCC L2 NDS and on all SCEs and alerts for the L2 MFA NDS. Due to the large number of baseline epochs in the L2 MFA NDS, a subset of baselines was sampled for PERCLOS. The plan was to randomly sample three epochs for each ADAS status (i.e., three ADAS levels) per participant, which would have yielded 1,080 PERCLOS epochs. However, availability of valid video data resulted in 919 epochs sampled from the L2 MFA NDS for PERCLOS 1 analysis. These epochs were randomly sampled from available data for each participant and were not matched across ADAS status.

Data Analysis

In the current study, descriptive statistics used to understand the distribution of variables included calculations for average (mean), standard deviation, median, first and third quartile (25^{th} and 75^{th} percentiles, respectively), and minimum and maximum values. These metrics were particularly useful for research questions with data limitations, allowing for meaningful comparisons even without inferential statistics.

A mixed-effect logistic regression model was used to evaluate the relationship between engagement in a secondary task and ADAS status (i.e., active versus available). For each baseline, STE was treated as an indicator variable (1 = yes, STE was observed during the baseline epoch; 2 = no, STE was not observed during the baseline epoch). The model predicted the log odds of STE from system status; a driver-level random effect was included to account for correlations in data from the same driver. Odds ratios and 95% confidence intervals calculated from the logistic regression model compared the estimated odds of STE in baseline epochs for active and available ADAS. If the confidence interval included "1," the two systems did not have a statistically significant difference in odds of STE. A confidence interval that was fully above or below "1" indicated a statistically significant difference in odds of STE by system status. Individual models were built for each system. If a model did not converge, a fixed-effects model (i.e., without a random effect for driver) was used. This same methodology was used to assess differences in STE during SCEs by system status.

A Poisson mixed-regression model was used to assess the impact of ADAS status on SCE rate. This model estimates how counts over time change with predictor variables. The number of SCEs and driving minutes in each system and status level were tabulated for each participant. The SCE counts were modeled by system status, using log of total minutes as an offset. As in the mixed-effect logistic regression model, a driver-level random effect was included to account for correlations in data from the same driver.

Several analysis techniques were used to understand eye-glance behaviors in the different ADAS levels. Generalized linear mixed models were used to model total eyes-off-road time (EORT) and duration of the longest glance. In these models, the eye-glance metric was the response variable, with system status and a driver-level random effect as predictor variables. Two eye-glance metrics were bound by 0 (0%) and 1 (100%) lower and upper limits: 1) percent total EORT and 2) percent of glances greater than 2 seconds. For these eye-glance metrics, Beta regression models were used to assess the impact of ADAS status. Data for Beta regression models must be within the bounds of (0, 1), not inclusive of 0 or 1. Because the current study data could include these bounds, data with these values were transformed using the method outlined by Smithson and Verkuilen (2006) and tested by Blanco et al. (2015). Beta regression models for each eye-glance metric used system status as a predictor variable.

An analysis of drivers' reaction times to alerts used the time from alert to first response, measured in seconds, from sampled alerts wherein the driver was observed reacting to the alert (not to the precipitating event). Epochs without a visible driver response to the alert were removed from the analysis of reaction time. The differences in reaction time between alert types were modeled using generalized linear models, with alert type as the predictor variable.

All analyses were completed using SAS software, version SAS 9.4[1].

[1] For an in-depth discussion of naturalistic driving analysis models, please see Guo (2019).

Results

This section provides an initial breakdown of the two NDS data sets, including a summary of ADAS usage. The research questions are then presented individually, followed by a description of the results of the analyses. Each research question was examined via individual tests, either descriptive or inferential.

Summary of Data Sets

This section summarizes the trip and ADAS status or activation information for both NDS data sets (i.e., the VCC L2 NDS and L2 MFA NDS). In addition, environmental and roadway scenarios were assessed and categorized where drivers had L2 active, L1 active, or no systems active (i.e., systems available but not active). Factors such as locality, traffic density, weather, and traffic flow have previously been shown to predict whether automation features were active at the time of driving (Russell et al., 2018).

Across the 30 VCC L2 NDS participants used in the current study, a total of 359,155 miles were driven (11,972 average miles per participant), which equated to 12,619 total hours of driving (421 average hours per participant). In the L2 MFA NDS, 120 participants drove a total of 217,207 miles (1,810 average miles per participant), which equated to a total of 6,957 hours of driving (58 average hours per participant). Relative to ADAS activation, Table 6 provides a summary of total and average hours per driver of system activation for both the VCC L2 NDS and L2 MFA NDS, as well as the duration of ADAS use as a percentage of the total hours of driving. Due to the different sampling strategies used in the two studies, the ADAS activation categories are slightly different. For example, in the VCC L2 NDS, the "None Active" category includes instances when ADAS were available (but not active) and when the ADAS were completely unavailable. Conversely, in the L2 MFA NDS, the epochs selected for inclusion in the data set had to meet specific criteria other than L1 active or none active. In both cases, L2 needed to be available (but not active), the minimum speed of the vehicle had to be 40 mph, and there needed to be verified visible lane markings. Thus, all other epochs that did not meet these criteria were combined under "All Other" in Table 6. This category includes epochs wherein ADAS were unavailable or ADAS were available but the epoch did not meet the set criteria. Interestingly, the percentage of ADAS use was fairly similar across the two studies when similar categories were combined, with L2 active being approximately 5% higher in the L2 MFA NDS than in the VCC L2

NDS (approximately 17% versus 12%, respectively). L1 was active (i.e., ACC active or LKA active in the VCC L2 NDS, when collectively tallied) approximately 12% of the time, and the majority of driving (i.e., approximately three-quarters of the total duration) occurred under manual driving conditions (i.e., "None Active" in the VCC L2 NDS and "None Active" plus "All Other" in the L2 MFA NDS).

Table 6. Summary of ADAS activation for the VCC L2 NDS and L2 MFA NDS.

	ADAS Activation	**N**	**Total Duration of ADAS Use (hrs)**	**% of ADAS Use**	**Average ADAS Use per Driver (hrs)**
VCC L2 NDS	L2 Active	30	1,500.52	11.89%	50.02
	ACC Active		880.90	6.98%	29.36
	LKA Active		635.15	5.03%	21.17
	None Active (includes ADAS unavailable)		9,602.18	76.09%	320.07
L2 MFA NDS	L2 Active	120	1,177.73	16.93%	9.81
	L1 Active (L2 Available)		841.07	12.09%	7.01
	None Active (L2 Available)		1,601.32	23.02%	13.34
	All Other (i.e., did not meet inclusion criteria)		3,336.88	47.96%	27.81

In both the VCC L2 NDS and L2 MFA NDS, environmental and roadway scenarios differed somewhat depending on ADAS status. Figure 4 and Figure 5 show that the majority of baseline epochs for both NDSs occurred during clear/partly cloudy weather, regardless of ADAS status. Participants in the VCC L2 NDS were slightly more likely to have L2 active on overcast days. There was also a similar percentage of baseline epochs in both NDSs where drivers had L2 active, despite rainy weather. This is typically against the recommendations of the vehicle manufacturers, who strongly caution drivers not to use ADAS in inclement weather.

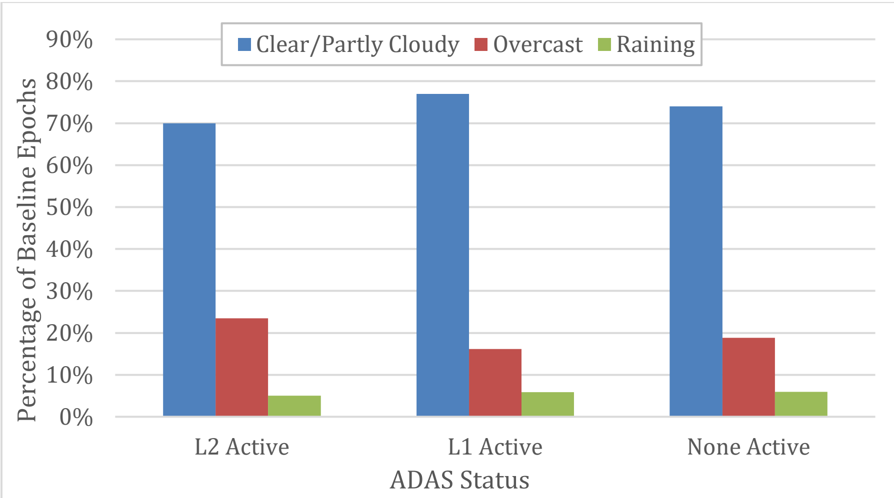

Figure 4. ADAS status based on weather conditions for the VCC L2 NDS

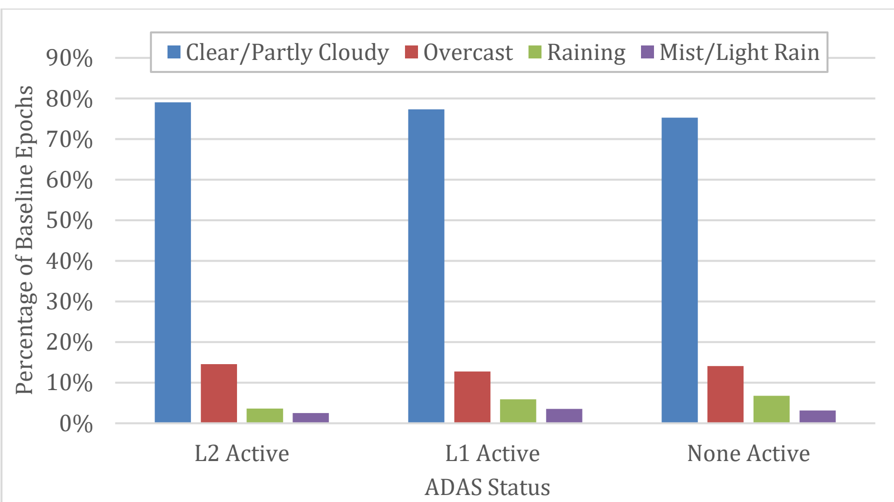

Figure 5. ADAS status based on weather conditions for the L2 MFA NDS

Figure 6 and Figure 7 show ADAS activation across both NDS data sets as based on road type. Participants in both NDSs drove with L2 active far more frequently on two-way, divided roads compared to other two-way or one-way road types. In the VCC L2 NDS, participants drove with one system active more frequently on two-way (divided and undivided) roads compared to roads with one-way traffic. In both NDSs, no systems active was more frequent on roads with one-way traffic compared to other ADAS use (i.e., L2 or L1 active).

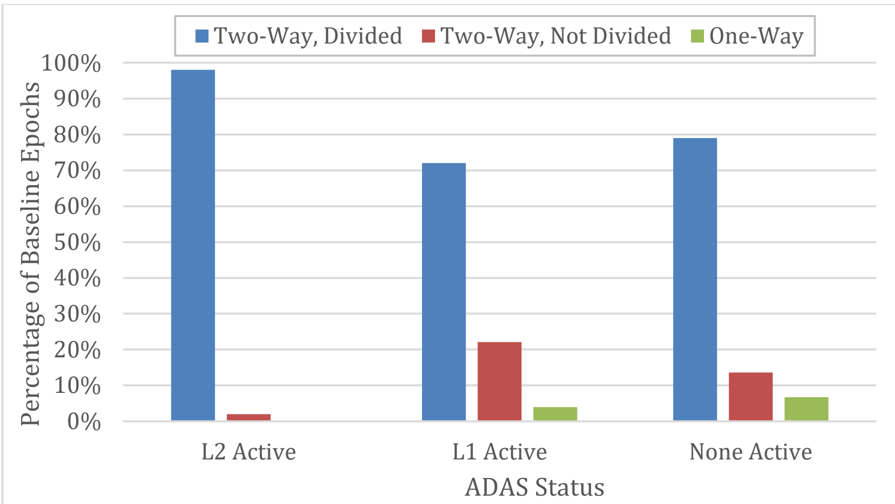

Figure 6. ADAS status based on road type for the VCC L2 NDS

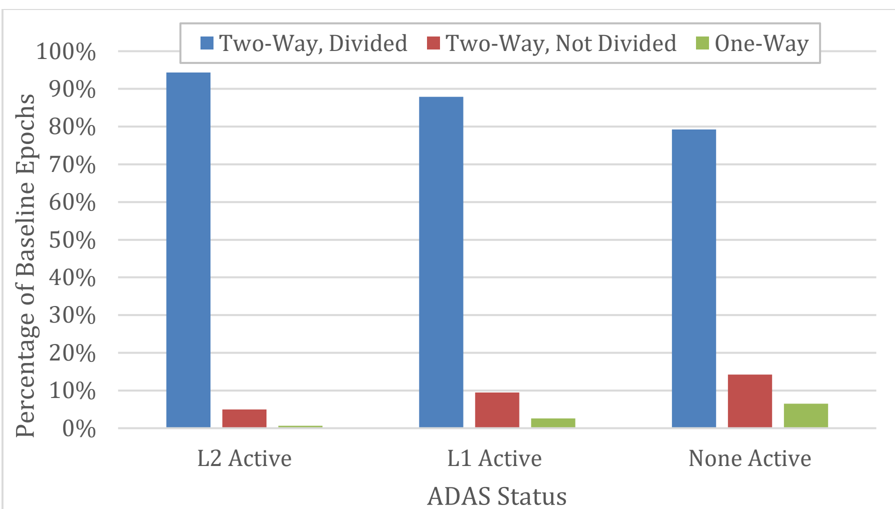

Figure 7. ADAS status based on road type for the L2 MFA NDS

Figure 8 and Figure 9 show the percentage of ADAS activations across traffic density types in both NDSs. Participants in both NDSs were much more likely to drive with ADAS active under free or stable maneuvering conditions than during unstable maneuvering conditions. Participants in the VCC L2 NDS were also more likely to have L2 active under stable maneuvering conditions than participants in the L2 MFA NDS (73% versus 67%, respectively).

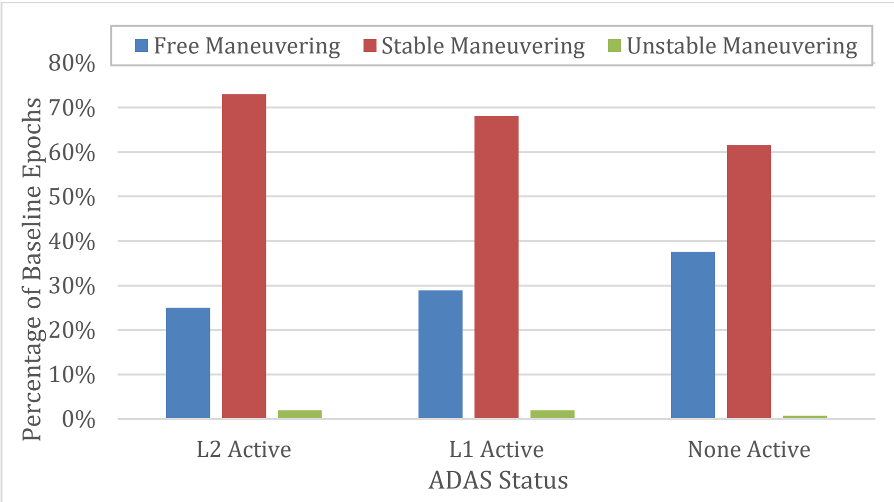

Figure 8. ADAS status based on traffic density types for the VCC L2 NDS

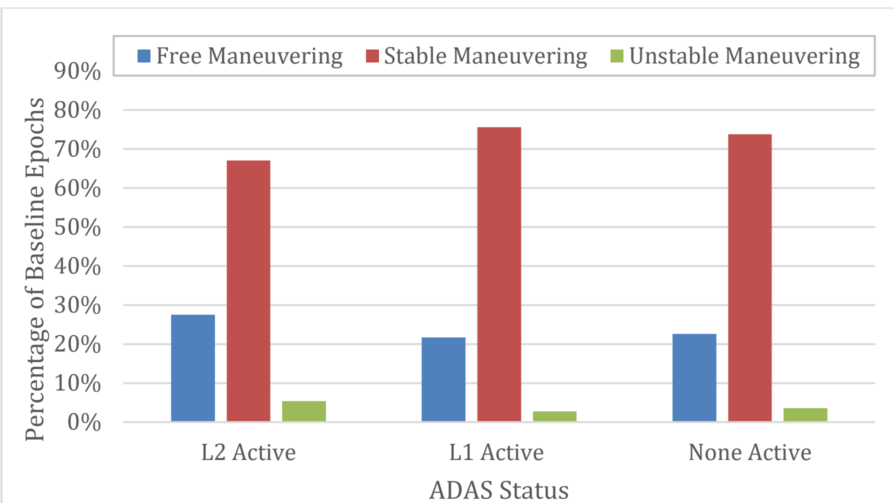

Figure 9. ADAS status based on traffic density types for the L2 MFA NDS

Figure 10 and Figure 11 show the percentage of ADAS activations across locality types in both NDSs. Participants in both NDSs activated both systems more frequently on interstate roads (i.e., controlled highways) compared to all other localities. In both studies, driving with one or no systems active was frequent on not only controlled highways but in business/industrial localities.

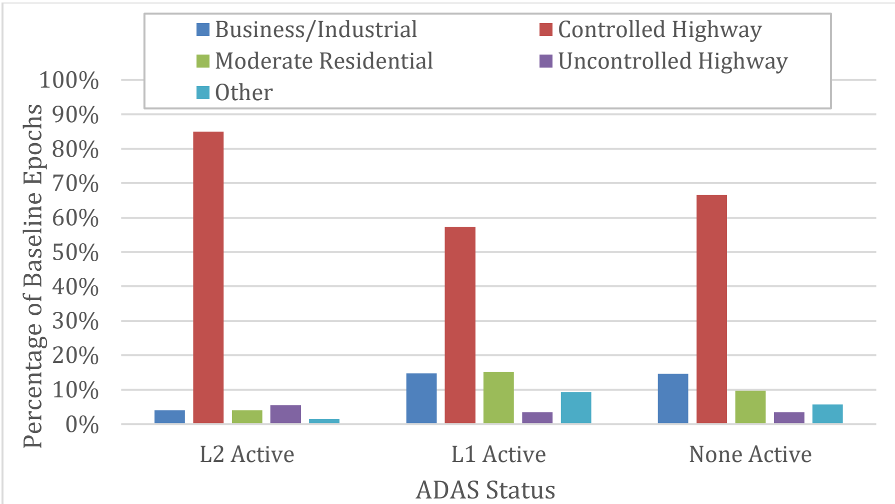

Figure 10. ADAS status based on locality types for the VCC L2 NDS

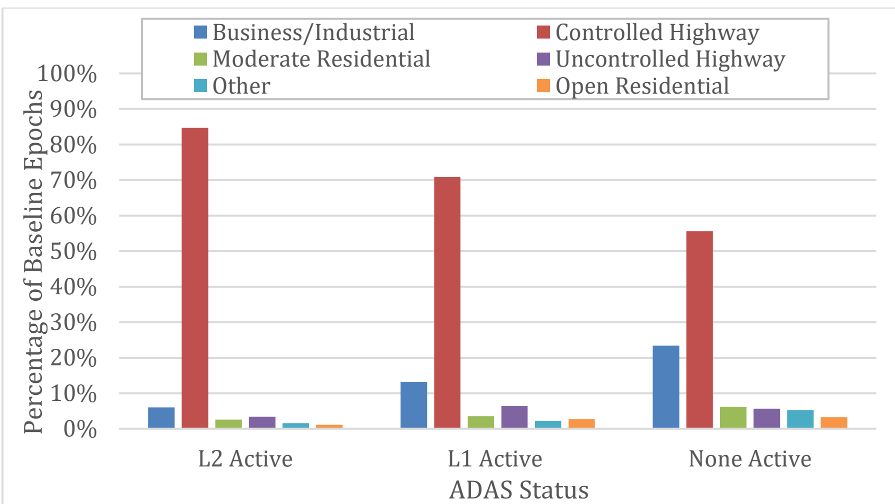

Figure 11. ADAS status based on locality types for the L2 MFA NDS

Research Question 1: What driver behaviors are observed when ADAS are active?

The occurrence of driver behaviors that one could consider meaningful to the presence of automation was relatively low in the VCC L2 NDS. Therefore, individual driver behaviors were combined into two broader categories for the purposes of reporting performance errors and judgment errors. *Driver performance errors* included a variety of vehicle operation and maneuver errors. This category comprised the following driver behaviors:

- Apparent Inexperience: Driver behaved in an unsafe manner, apparently due to lack of experience with the driving task (e.g., hyper-focused driving, overly cautious maneuvers).
- Blind Spot: Driver was traveling close to another vehicle in such a way that the driver of the other vehicle was not expected to be able to see him/her (note that the study driver must have maintained this position for at least 5 seconds).
- Improper Turn: Driver turned left or right from the initial travel path, unnecessarily encroaching into the adjacent lane, median, or shoulder/curb.
- Right-of-Way: Driver made the incorrect decision regarding who had the right-of-way (i.e., his/her own vehicle or another vehicle or pedestrian) due to a misunderstanding or improper analysis of the situation.
- Signal Violation: Driver did not notice a traffic signal (i.e., this error was not intentional).
- Stop-Yield Violation: Driver did not notice an intersection with either a stop sign or a yield sign (i.e., this error was not intentional).
- Wrong Side of Road: Driver was traveling on the wrong side of the road with no intent of passing or overtaking another vehicle.
- Driving Too Slow: Driver was traveling at a speed much lower than the posted speed limit (i.e., $\geq$ 10 mph under posted speed limit) when higher speeds were appropriate.
- Sudden or Improper Braking: Driver braked suddenly, in an unsafe manner, or at an unsafe time in the roadway, but did not come to a complete stop.
- Failed to Signal: Driver failed to properly signal his/her intent by not signaling at all (e.g., changed lanes or made a turn without signaling).

As noted, there were very few performance errors made, and most of the driver behaviors listed above did not appear in either data set. In the VCC L2 NDS, there were 13 performance errors (approximately 1.6% prevalence) across all ADAS levels from the baseline data set. When systems were available but not active, 12 drivers failed to signal. One driver showed signs of apparent inexperience when ACC was active. In the L2 MFA NDS, there were 65 performance errors from the baseline data set (approximately 1.7% prevalence), the majority of which occurred when drivers had no systems active (i.e., under manual driving conditions). Of these performance errors, the most frequent was failing to signal, which occurred 34 times when drivers had no systems active. Five drivers were driving too slowly when they had L2 active, and six drivers showed signs of apparent inexperience with the vehicles.

The remaining driver behaviors were classified as *judgment errors*, which were operationally defined to include aspects of a momentary lapse of judgment by the driver (Dingus et al., 2016). Judgment errors included speeding, driving too fast for the conditions, and other forms of aggressive driving (e.g., illegal passing or following too closely). As shown in Table 7, exceeding the speed limit was the most frequent judgment error made in both the VCC L2 NDS and the L2 MFA NDS, followed by illegal passing. Speeding in a work zone was a more frequent judgment error made among the L2 MFA NDS drivers than among drivers in the VCC L2 NDS.

Table 7. Judgment errors across all ADAS statuses in the VCC L2 NDS and L2 MFA NDS.

		Aggressive Driving	Exceed Speed Limit	Speed in Work Zone	Illegal Passing	Follow too Closely	Intentional Signal Violation	Intentional Stop-Yield Violation
VCC L2 NDS	L2 Available	0	35	2	3	1	0	0
	L2 Active	0	22	0	2	0	0	0
	ACC Available	0	14	0	3	0	0	0
	ACC Active	0	13	0	2	0	0	0
	LKA Available	0	4	1	2	0	0	0
	LKA Active	1	6	0	3	1	0	0
L2 MFA NDS	L2 Active	0	263	5	1	0	0	0
	L1 Active	0	181	2	4	1	2	0
	None Active	4	186	3	9	1	3	0

Research Question 2: Do unsafe driver behaviors occur more frequently when ADAS are active?

When assessing the unsafe driver behaviors that compose the group of judgment errors relative to the VCC L2 NDS, exceeding the speed limit was the only unsafe driver behavior with enough data to investigate further (Table 7). Exceeding the speed limit refers to instances wherein the driver was traveling greater than 10 mph above the posted speed limit and was not in a work zone. The data indicate that exceeding the speed limit occurred more often when VCC L2 NDS drivers had L2 available compared to when they had L2 active (17.5% versus 11%, respectively). Thus, in the VCC L2 NDS, the simultaneous use of both LKA and ACC resulted in a reduction in the occurrence of speeding (i.e., $\geq$ 10 mph over the posted speed). Further testing revealed this difference was significant (OR = 0.53, 95% LCL = 0.29, 95% UCL = 0.95).

In the L2 MFA NDS, the occurrence of exceeding the speed limit was greater when L2 was active. Further testing showed drivers with L2 active had significantly higher odds of exceeding the posted speed limit by at least 10 mph (OR = 1.27, 95% LCL = 1.01, 95% UCL = 1.60) when compared to no systems active. This is opposite the trend demonstrated in the VCC L2 NDS, which showed a reduction in the occurrence of exceeding the speed limit when L2 was active.

Research Question 3: Does STE occur more frequently when ADAS are active?

In the VCC L2 NDS, engagement in a secondary task was assessed when L2 was active and when it was available, when ACC only was active and when it was available, and when LKA only was active and when it was available. The initial approach assessed the overall presence of any secondary task (i.e., versus just driving [no STE]). Overall, drivers were observed engaging in a secondary task during 45% to 58% of baseline epochs (see Table 8). A mixed-effect logistic regression model compared STE under active and available statuses for all ADAS levels. The results revealed no significant differences for ACC and LKA individually active compared to available. For L2 active, the odds of STE when L2 was active were significantly higher than when L2 was available but not active (OR = 1.54, 95% LCL = 1.03, 95% UCL = 2.30).

In the L2 MFA NDS, engagement in a secondary task was assessed when L2 was active, when L1 was active, and when no systems were active (i.e., under manual driving conditions). Interestingly, the results of STE tended in the opposite direction to that found in the VCC L2 NDS. Overall, drivers were observed engaging in a secondary task during 60% of baseline epochs with L2 active compared to 69% with no systems active. A mixed-effect logistic regression model revealed drivers who had no systems active had significantly higher odds of STE compared to both baselines with L1 active (OR = 1.23, 95% LCL = 1.02, 95% UCL = 1.49) and baselines with L2 active (OR = 1.38, 95% LCL = 1.16, 95% UCL = 1.64).

Table 8. Overall STE during baselines for all ADAS statuses in the VCC L2 NDS and L2 MFA NDS.

	ADAS Status	Total Number of Baselines	% with STE
VCC L2 NDS	**L2 Available**	200	47%
	L2 Active	200	58%
	ACC Available	133	53%
	ACC Active	133	52%
	LKA Available	71	45%
	LKA Active	71	46%
L2 MFA NDS	**L2 Active**	1,388	60%
	L1 Active	1,139	63%
	None Active	1,228	69%

To further investigate the types of secondary tasks performed, secondary tasks were categorized as visual-manual, visual, manual, or cognitive based on the primary demands of the task. For example, a visual task required the driver to look away from the forward road to visually obtain information but did not require the driver to take his/her hands off the steering wheel (e.g., looking at a pedestrian). A manual task required the driver to take one or both hands off the steering wheel, for example, to manipulate a device or control. However, the driver was not required to look away from the forward roadway (e.g., holding a cell phone, smoking a cigarette). A visual-manual task required a combination of hands-off the steering wheel and eyes-off the forward roadway (e.g., texting, adjusting/monitoring climate control). Finally, a cognitive task required the driver to divert his/her mental attention away from the driving task; examples included interacting with passengers and talking/listening on a hands-free cell phone. Figure 12 and Figure 13 show the prevalence for each secondary task category in the VCC L2 NDS and the L2 MFA NDS data sets, respectively. Note that drivers may have engaged in more than one type of secondary task during the epoch; thus, percentages within each ADAS level did not total 100%.

When drivers performed a secondary task in the VCC L2 NDS data set, it was most often a visual-manual task (Figure 12). Interestingly, the occurrence of engagement in a visual-manual secondary task increased from 24% with L2 available to 32% with L2 active. The occurrence of manual STE also increased for all ADAS statuses when comparing systems available to systems active. For example, when ACC was available, manual task engagement was 5%; this increased to 11% with ACC active. Cognitive STE showed the opposite trend, with occurrence decreasing slightly when switching from available to active for L2 (~20% and 17%, respectively) and ACC (23% and ~20%, respectively).

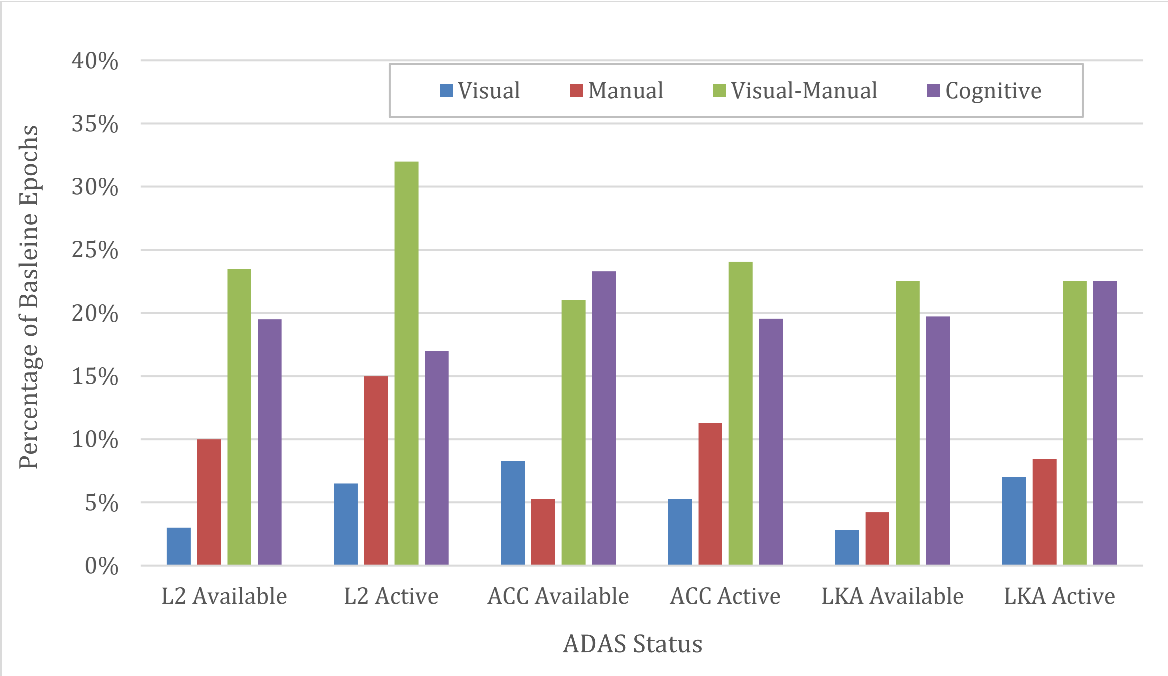

Figure 12. Occurrence of STE by ADAS status in the VCC L2 NDS

As seen in Figure 13, the above trends did not hold true for the L2 MFA NDS data set. The occurrence of cognitive, visual-manual, and visual STE was higher when drivers had no systems active (i.e., manual driving) than when L1 was active or when L2 was active. The exception was manual STE, which was approximately the same with no active systems as when L2 was active.

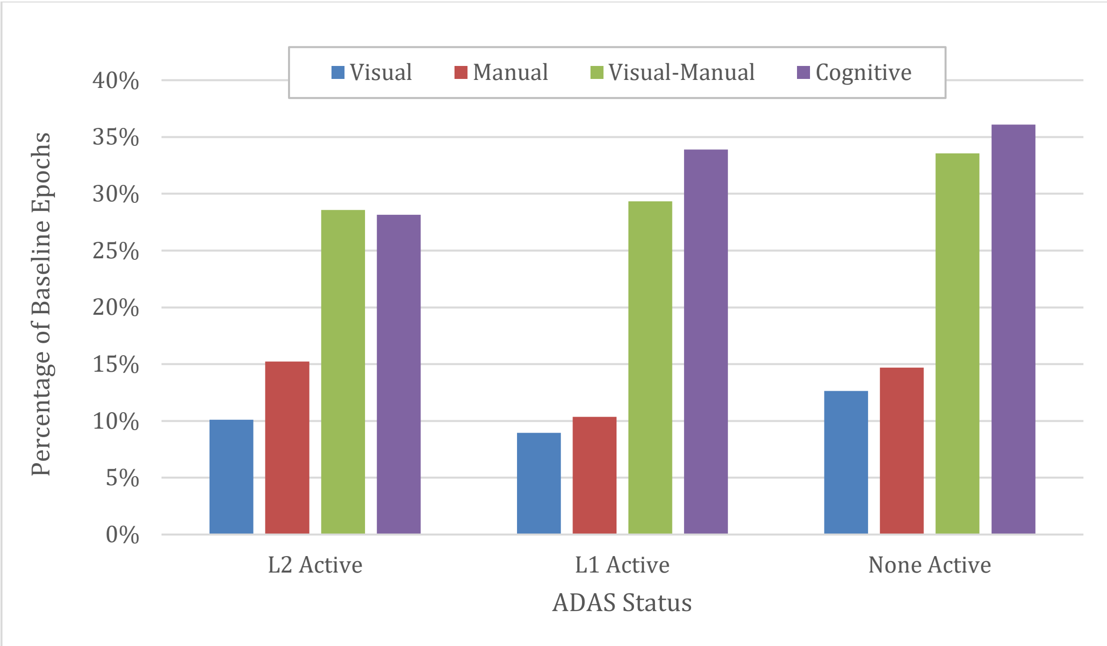

Figure 13. Occurrence of STE by ADAS status in the L2 MFA NDS

To conduct inferential testing on secondary task types across ADAS statuses, the visual, manual, and visual-manual tasks were combined into one group. Figure 14 shows the occurrence of cognitive STE compared to the occurrence of the visual, manual, or visual-manual combined STE for all ADAS statuses in the VCC L2 NDS. A mixed-effect logistic regression model comparing systems active to systems available was not significant for ACC or LKA across either the cognitive secondary tasks or the combined visual, manual, or visual-manual tasks. When L2 was active, however, the odds of a driver performing a visual, manual, or visual-manual task were significantly higher than when both systems were available (OR = 1.81, 95% LCL = 1.20, 95% UCL = 2.74). The odds of a driver performing a cognitive task were not significantly different when L2 was active compared to when L2 was available.

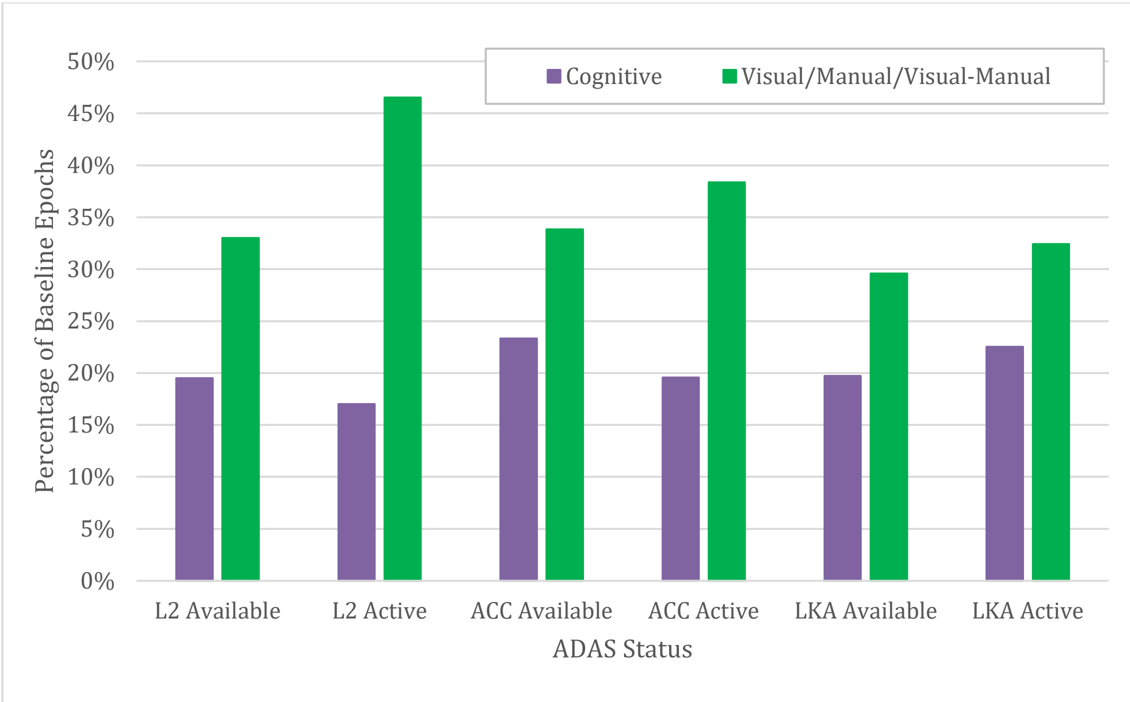

Figure 14. Occurrence of cognitive STE and combined visual, manual, and visual-manual STE compared across all ADAS statuses in the VCC L2 NDS

Figure 15 shows the occurrence of cognitive STE compared to the combined visual, manual, or visual-manual STE for all ADAS statuses in the L2 MFA NDS. A mixed-effect logistic regression model comparing the three ADAS statuses found drivers with no systems active had significantly higher odds of engaging in a visual, manual, or visual-manual task compared to drivers with L1 active (OR = 1.27, 95% LCL = 1.06, 95% UCL = 1.52), but did not differ significantly from drivers with L2 active. The model also found drivers who had L1 active and drivers with no systems active both had significantly higher odds of engaging in a cognitive task than drivers with L2 active (OR = 1.32, 95% LCL = 1.10, 95% UCL = 1.59; OR = 1.51, 95% LCL = 1.26, 95% UCL = 1.80, respectively). No other comparisons were significant.

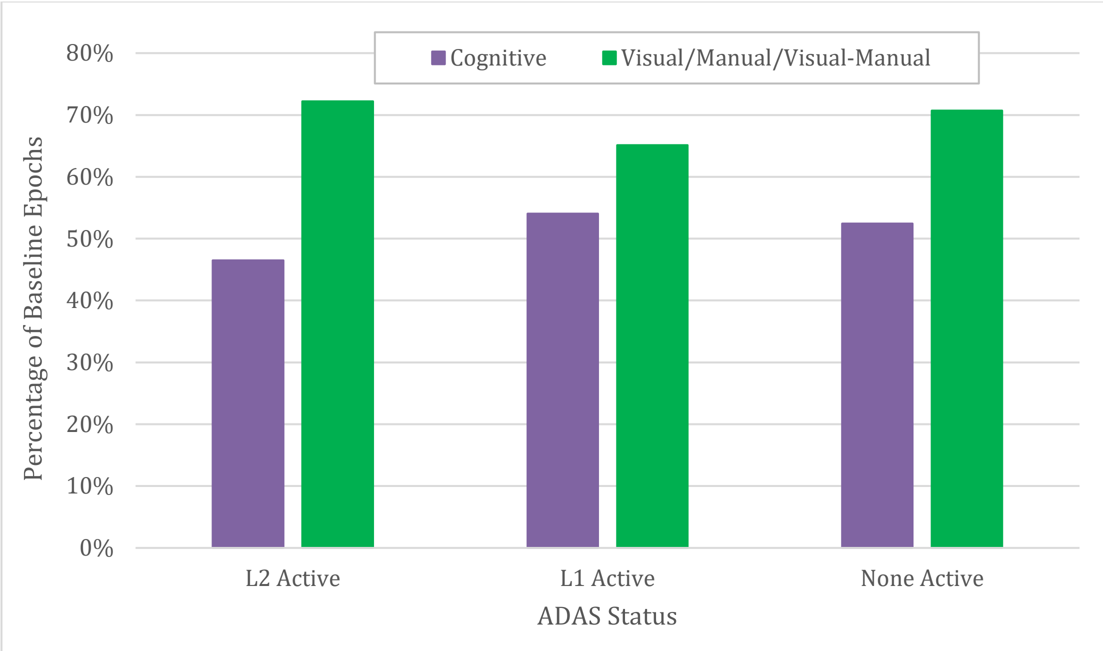

Figure 15. Occurrence of cognitive STE and combined visual, manual, or visual-manual STE compared across all ADAS statuses in the L2 MFA NDS

Research Question 4: Do the characteristics of SCEs change when ADAS are active?

In the VCC L2 NDS, there were 159 safety critical events (SCEs), four of which were crashes (or minor collisions) and 155 that were near-crashes. Twelve of the SCEs were excluded from the following analysis due to an inability to determine ADAS status. Additionally, due to low numbers of some ADAS statuses, SCEs were combined into the following categories:

- ADAS Active, which meant one or both systems were active at the time of the SCE;
- ADAS Available, which meant one or both systems were available but were not active at the time of the SCE; and
- ADAS Unavailable, which meant none of the systems were active or available at the time of the SCE.

There were 45 SCEs with ADAS Active, 92 with ADAS Available, and 10 with ADAS Unavailable.

The majority of the SCEs in the VCC L2 NDS were rear-end striking incidents (*n* = 77), defined as "the subject vehicle (V1) makes contact or nearly makes contact with any portion of the back of the vehicle in front (V2)" (VTTI, 2015). Sixty-two percent of the rear-end striking SCEs occurred with systems available; 38% occurred with systems active. The second most frequent incident type was sideswipe, same direction (left or right), defined as the "subject vehicle (V1) is struck or nearly struck by another vehicle (V2) or strikes or nearly strikes another vehicle (V2) on either the driver or passenger side of the vehicle (V1 or V2) when the vehicles were traveling in the same direction" (VTTI, 2015). Of the 30 sideswipe SCEs, 60% occurred with systems available. Only 40% occurred when systems were active.

SCEs in the VCC L2 NDS mostly occurred during daylight hours (*n* = 119) in clear/partly cloudy (*n* = 105) or overcast (*n* = 35) weather, regardless of ADAS status. A higher percentage of the SCEs with ADAS Active occurred on divided (median strip or barrier) roads than with ADAS Available (98% versus 75%, respectively). Traffic density when SCEs occurred was similar regardless of ADAS status, ranging from level-of-service B (flow with some restrictions) to level-of-service E (flow is unstable, vehicles are unable to pass, temporary stoppages, and all levels in between). There were some differences between the localities of SCEs depending on ADAS status. A higher percentage of SCEs with ADAS Active occurred on interstates/bypasses/divided highways (controlled access) compared to ADAS Available (78% versus 53%, respectively). SCEs with ADAS Available occurred more frequently in business/industrial localities compared to when systems were active (29% versus 9%, respectively).

In the L2 MFA NDS, there were 71 SCEs, five of which were crashes and 66 that were near-crashes. Seventeen of these were excluded from the following analysis due to indeterminable ADAS status at the time of the SCE. Of the remaining 54 SCEs, there were 13 SCEs with L2 active, 11 with L1 active, and 30 with no systems active.

Similar to the VCC L2 NDS, the two most frequent types of SCEs occurring in the L2 MFA NDS were rear-end collisions (*n* = 35) and sideswipe, same direction (*n* = 13). For rear-end SCEs, approximately 26% occurred with L2 active, 14% occurred with L1 active, and 60% occurred with no systems active. For sideswipe SCEs, 23% occurred with L2 active, 38% occurred with L1 active, and 39% occurred with no systems active. The SCEs typically occurred during daylight hours (n = 41) in clear/partly cloudy weather (n = 44). Unlike the VCC L2 NDS, four of the SCEs in the L2 MFA NDS occurred when it was raining (i.e., when ADAS are not recommended for use). Most of the SCEs occurred on divided (median strip or barrier) roads, with all 13 SCEs with L2 active occurring on these road types. Level-of-service B was associated with the most SCEs, regardless of ADAS status. Levels-of-service C (stable flow, maneuverability, and speed are more restricted) and D (unstable flow – temporary restrictions substantially slow driver) were also applicable to a smaller number of SCEs. Also similar to the VCC L2 NDS, the most frequent locality for SCEs occurring in the L2 MFA NDS, regardless of ADAS status, was interstate/bypass/divided highway (controlled access), which accounted for all SCEs with both systems active. SCEs with one system active and no systems active also frequently occurred in business/industrial localities (27% and 37%, respectively).

Research Question 5: Do SCE rates differ when ADAS are active?

Due to the low numbers of safety critical events (SCEs) across multiple ADAS levels (e.g., 13 SCEs had ACC active and 9 SCEs had LKA active) in the VCC L2 NDS, ADAS levels were combined into two categories for the purposes of analysis: 1) ADAS Active and 2) ADAS Available/Inactive. Table 9 shows a summary of the total SCE counts across both categories, the total minutes of driving, and the averages of both SCE counts and minutes per driver for the VCC L2 NDS.

Table 9. Summary of SCE counts and minutes of driving for each ADAS status in the VCC L2 NDS.

ADAS Status	Total SCE Count	Average (SD) SCE Count per Driver	Total Minutes	Average (SD) Minutes per Driver
Active	45	1.50 (2.60)	166,022.52	5,534.08 (5,098.57)
Available/Inactive	102	3.40 (3.33)	531,643.93	17,721.46 (8,116.02)

Figure 16 shows drivers' SCE rates per 1,000 driving minutes with ADAS Active and ADAS Available/Inactive in the VCC L2 NDS. As can be seen in the plot, there was a small number of drivers whose SCE rates per 1,000 minutes of driving were noticeably higher when ADAS were active compared to when ADAS were available or inactive. Participant 1 in the figure, for example, had an SCE rate of approximately 1.3 per 1,000 driving minutes when ADAS were active compared to approximately 0.1 per 1,000 driving minutes when ADAS were available/inactive.

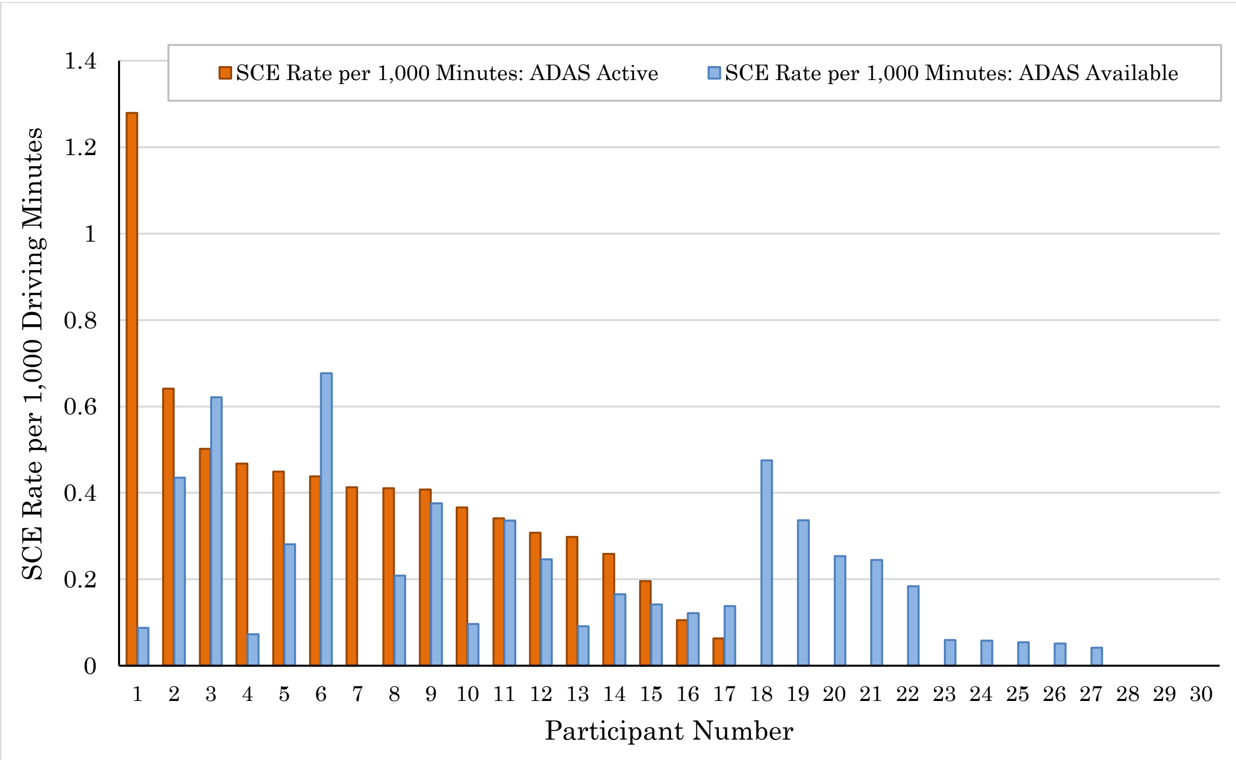

Figure 16. Drivers' SCE rates per 1,000 minutes of driving with ADAS Active and ADAS Available/Inactive in the VCC L2 NDS

To consider the differences between drivers' SCE rates in the VCC L2 NDS, each drivers' SCE rate was calculated for each ADAS status based on the sum of their SCEs (i.e., with systems active and with systems available/inactive) divided by their total minutes of driving (i.e., minutes of driving with systems active and minutes of driving with systems available/inactive). Thus, rather than an overall average SCE rate, this method provided a per-driver average SCE rate per driving minutes. Table 10 shows the per-driver average SCE rates per 1,000 minutes of driving for ADAS Active and for ADAS Available/Inactive. A Poisson mixed-regression model was used on the SCE rates, with a log of total minutes as an offset. The model found no significant differences in the SCE rates with ADAS Active and ADAS Available/Inactive (F value = 1.11, p = 0.30). The similarity in SCE rates between ADAS Active and ADAS Available/Inactive may seem surprising in light of outcomes for Research Question 4 above. However, Table 9 shows that, despite a greater SCE count across the available/inactive category, that category of driving also accounted for approximately three times as many minutes. Thus, the SCE rates per 1,000 minutes across the two groups were similar.

Table 10. Per-driver average SCE rates for ADAS Active and ADAS Available/Inactive in the VCC L2 NDS.

ADAS Status	Average (SD) SCE Rate per 1,000 Driving Minutes
Active	0.23 (0.29)
Available/Inactive	0.20 (0.18)

The approaches used above were also applied to the L2 MFA NDS, with ADAS status grouped as active or available/inactive. Table 11 shows a summary of the total SCE counts, total minutes of driving, and the averages of both counts and minutes per driver for the L2 MFA NDS.

Table 11. Summary of SCE counts and minutes of driving for each ADAS status in the L2 MFA NDS.

ADAS Status	Total SCE Count	Average (SD) SCE Count per Driver	Total Minutes	Average (SD) Minutes per Driver
Active	24	0.20 (0.51)	116,295.25	961.12 (514.34)
Available/Inactive	30	0.25 (0.71)	301,136.22	2,488.73 (1,275.56)

Figure 17 shows drivers' SCE rates per 1,000 driving minutes with ADAS Active and ADAS Available/Inactive in the L2 MFA NDS. The plot does not show the complete study sample of 120 participants, as the majority of participants were not involved in any SCEs, regardless of ADAS status. Interestingly, many drivers in Figure 17 appear to be involved in SCEs with ADAS Active but not with ADAS Available/Inactive and vice versa.

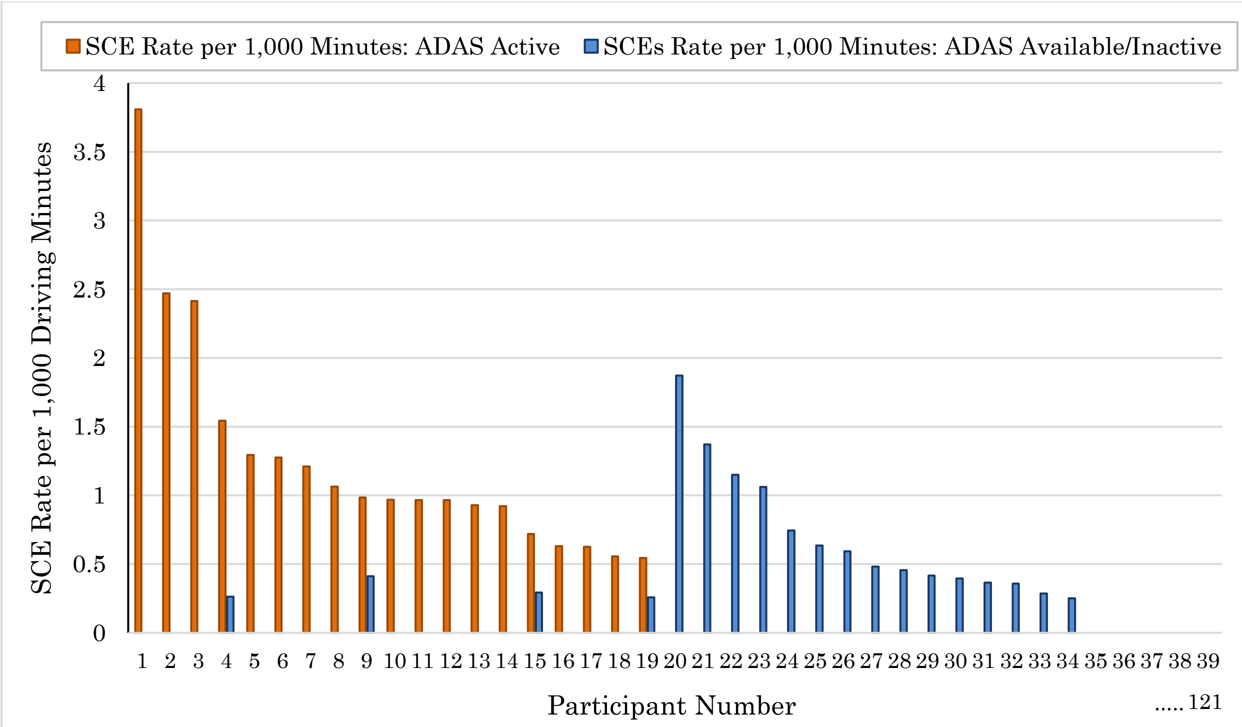

Figure 17. Drivers' SCE rates per 1,000 minutes of driving with ADAS Active and ADAS Available/Inactive in the L2 MFA NDS

As in the VCC L2 NDS, drivers' SCE rates in the L2 MFA NDS were calculated based on each drivers' SCE count and their total minutes of driving for each ADAS level. Table 12 shows the per-driver average SCE rates per 1,000 minutes of driving for ADAS Active and for ADAS Available/Inactive. A Poisson mixed-regression model was used on the SCE rates, with a log of total driving minutes in each ADAS status as an offset. The model found SCE rates to be significantly higher than expected with ADAS Active compared to ADAS Available/Inactive (F value = 4.13, p = 0.04).

Table 12. Per-driver average SCE rates for ADAS Active and ADAS Available/Inactive in the L2 MFA NDS.

ADAS Status	Average (SD) SCE Rate per 1,000 Driving Minutes
Active	0.20 (0.56)
Available/Inactive	0.10 (0.28)

Research Question 6: Is there an increased prevalence of STE during SCEs that occur when ADAS are active?

As in Research Question 5, due to low numbers of SCEs in the VCC L2 NDS, ADAS levels were combined into two broad categories of ADAS Active and ADAS Available/Inactive. Table 13 shows a summary of the secondary tasks present during SCEs in the VCC L2 NDS data set, based on whether ADAS were active or available/inactive.

Table 13. Summary of secondary tasks present during SCEs where ADAS were active or available/inactive in the VCC L2 NDS.

Secondary Task	**ADAS Active**	**ADAS Available/Inactive**
Adjusting/monitoring in-vehicle device, instrument panel, or radio	3	9
Cell phone-related tasks (i.e., browsing, holding, texting, talking/listening on handheld phone)	6	16
Passenger interactions	0	9
Personal hygiene-related tasks	0	3
Interacting with object or pet in vehicle	0	11
Food- or drink-related task	0	2
Smoking cigar/cigarette	1	0
Other external distraction	1	8
Talking/singing, audience unknown	4	7

As seen in Table 13, cell phone-related tasks were the most prevalent secondary tasks performed in SCEs that occurred when ADAS were active, as well as when ADAS were available/inactive. Talking/singing and adjusting/monitoring in-vehicle devices were the next most frequent secondary tasks performed during ADAS Active SCEs. Interacting with an object or a pet in the vehicle, adjusting/monitoring in-vehicle devices, and passenger interaction were all common secondary tasks performed during SCEs when ADAS were available/inactive.

To determine if there was an increased prevalence of STE during SCEs in the VCC L2 NDS, the proportion of SCEs that included STE for each ADAS status was calculated, along with the average SCE rate per 1,000 minutes of driving. A mixed-effect logistic regression model predicted STE involvement during SCEs by ADAS status using STE as a binary (yes/no) variable for all SCEs. The model found no statistically significant difference in the odds of STE in SCEs by ADAS status (OR = 2.15, 95% LCL = 0.96, 95% UCL = 4.79; Table 14).

Table 14. Summary of SCEs occurring when drivers engaged in a secondary task for ADAS Active or ADAS Available/Inactive in the VCC L2 NDS.

ADAS Status	**Total SCEs with STE**	**Average (SD) SCE with STE Count per Driver**	**Proportion of SCEs with STE**	**Average (SD) SCE with STE Rate per 1,000 Minutes**
Active	14	0.47 (0.86)	31.11%	0.07 (0.14)
Available/Inactive	53	1.77 (1.55)	46.49%	0.11 (0.11)

Table 15 shows a summary of secondary tasks present during SCEs that occurred when ADAS were active and when ADAS were available/inactive in the L2 MFA NDS. The most frequent secondary tasks present during SCEs were the same when ADAS were active and when ADAS were available/inactive: cell phone-related tasks, adjusting/monitoring in-vehicle devices, and passenger interactions.

Table 15. Summary of secondary tasks present during SCEs where ADAS were active or available/inactive in the L2 MFA NDS.

Secondary Task	**ADAS Active**	**ADAS Available/Inactive**
Adjusting/monitoring in-vehicle device, instrument panel, or radio	7	6
Cell phone-related tasks (i.e., browsing, holding, locating/reaching/answering, talking/listening hands-free phone)	21	15
Passenger interactions	5	5
Personal hygiene-related tasks	1	2
Interacting with object or pet in vehicle	3	1
Food- or drink-related task	1	0
Smoking cigar/cigarette	0	1
Other external distraction	4	1
Dancing, Talking/singing, audience unknown	4	3

To determine if there was an increased prevalence of STE during SCEs in the L2 MFA NDS, the proportion of SCEs that included STE for each ADAS status was calculated, along with the average SCE rate per 1,000 minutes of driving (Table 16). A mixed-effect logistic regression model predicted STE as a binary (yes/no) variable for all SCEs by ADAS status. The model found no statistically significant difference in the odds of STE in SCEs by ADAS status (OR = 1.54, 95% LCL = 0.38, 95% UCL = 6.24).

Table 16. Summary of SCEs occurring when drivers engaged in a secondary task for ADAS Active or ADAS Available/Inactive in the L2 MFA NDS.

ADAS Status	Total SCEs with STE	Average (SD) SCE with STE Count per Driver	Proportion of SCEs with STE	Average (SD) SCE with STE Rate per 1,000 Minutes
Active	17	0.14 (0.39)	70.83%	0.12 (0.33)
Available/Inactive	20	0.17 (0.62)	66.67%	0.06 (0.24)

The proportion of SCEs with a secondary task present appears to be much higher in the L2 MFA NDS compared to the VCC L2 NDS. In the former, approximately two-thirds of SCEs with ADAS Available/Inactive had a secondary task present, compared to less than one-half with ADAS Available/Inactive in the VCC L2 NDS. This proportion increased to more than 70% with ADAS Active in the L2 MFA NDS, compared to 31% with ADAS Active in the VCC L2 NDS.

Research Question 7: Do drivers spend more time with their eyes off the roadway when ADAS are active?

On-Road vs. Off-Road Glances

There are several eye-glance metrics that can be used as surrogates for distraction. In the current study, these metrics included:

- Total EORT: Cumulative total number of seconds a driver's eyes were off the forward roadway within the designated baseline epoch (i.e., 10 seconds in the VCC L2 NDS, 15 seconds in the L2 MFA NDS);
- Percent of EORT: Total EORT divided by the length of the baseline epoch;
- Percentage of glances > 2 seconds: The percentage of individual glances within the designated baseline epoch that were greater than 2 seconds;
- Number of off-road glances: The number of off-road glances within the designated baseline epoch; and
- Longest single glance: The duration (in seconds) of the longest single off-road glance within the designated baseline epoch.

Baseline epochs with valid eye-glance data for each ADAS status across both NDSs were initially assessed to determine the percentage of epochs comprising no off-road glances (i.e., drivers' eyes were on the forward roadway the entire time). An important difference between the VCC L2 NDS and L2 MFA NDS relative to eye-glance metrics is that of the duration of the baseline epochs (i.e., baseline epochs were 10 seconds in the VCC L2 NDS and 15 seconds in the L2 MFA NDS). Table 17 provides a summary of the on-road-only eye-glance data for each ADAS level in both the VCC L2 NDS and L2 MFA NDS.

Table 17. On-road-only eye-glance data for all ADAS statuses in the VCC L2 NDS and L2 MFA NDS.

	ADAS Status	Total Epochs with Valid Eye-Glance	Epochs with On-Road-Only Eye-Glance	% with On-Road-Only Eye-Glance
VCC L2 NDS	**L2 Available**	195	66	34%
	L2 Active	195	51	26%
	ACC Available	128	41	32%
	ACC Active	132	39	30%
	LKA Available	69	22	32%
	LKA Active	70	27	39%
L2 MFA NDS	**L2 Active**	1,321	200	15%
	L1 Active	1,097	212	19%
	None Active	1,167	185	16%

As seen in Table 17, approximately one-third of all valid eye-glance epochs in the VCC L2 NDS comprised on-road-only eye-glance data for all ADAS statuses. The exception was that, when L2 was active, the percent of on-road-only eye-glance decreased to one-quarter of epochs, meaning drivers less frequently kept their eyes on-road the entire time when both lateral and longitudinal systems (i.e., ACC and LKA) were active. When LKA was active, however, nearly 40% of drivers kept their eyes on-road the entire time.

Results from the L2 MFA NDS were similar across ADAS statuses (Table 17). When L2 was active, 15% of drivers kept their eyes on the forward road for the entire epoch; this was slightly lower than when one or no systems active, but the difference was small. Interestingly, the L2 MFA NDS percentages were lower across the board compared to the VCC L2 NDS, suggesting drivers in the L2 MFA NDS spent less time with their eyes consistently on the road during the sampled epochs. However, the baseline epochs in the L2 MFA NDS were 5 seconds longer, which could account for some of the differences between the two studies.

When looking specifically at Total EORT (Table 18), the average total time drivers spent with their eyes-off-road in the VCC L2 NDS was less than 1.5 seconds across all ADAS statuses, with the exception of L2 active. When drivers had L2 active, their total EORT increased by more than one-half of a second (to 2.02 seconds). Mixed-model results indicated this difference was significant ($p = 0.0003$). Thus, drivers with L2 active had a significantly higher Total EORT compared to when L2 was available. None of the other ADAS statuses showed significant differences between system active compared to system available.

Total EORT in the L2 MFA NDS was higher than in the VCC L2 NDS, averaging close to 2.5 seconds across all ADAS statuses (Table 18). However, as mentioned previously, the baseline epochs for the L2 MFA NDS were 5 seconds longer than the VCC L2 NDS. Interestingly, the highest average Total EORT was found when drivers had no systems active (2.70 seconds), followed by L2 Active (2.51 seconds), then L1 Active (2.20 seconds). Mixed-model results indicated significant differences between L1 Active and L2 Active ($p = 0.0018$) and L1 Active and None Active ($p = 0.0004$). The total EORT was not significantly different between L2 Active and no systems active.

Table 18. Total EORT and Percent EORT for all ADAS statuses in VCC L2 NDS and L2 MFA NDS (including on-road-only eye-glance epochs).

	ADAS Status	N	Total EORT (s)	% EORT
VCC L2 NDS	**L2 Available**	195	1.29	13.0%
	L2 Active	195	2.02	20.3%
	ACC Available	128	1.29	13.0%
	ACC Active	132	1.43	14.8%
	LKA Available	69	1.43	14.4%
	LKA Active	70	1.17	11.8%
L2 MFA NDS	**L2 Active**	1,321	2.51	16.8%
	L1 Active	1,097	2.20	14.7%
	None Active	1,167	2.70	18.1%

In terms of glances, the number of glances and the longest single glance were also important indicators of distraction. In the VCC L2 NDS, the average longest single glance and the average number of glances were greater when L2 was active. The model indicated these results were significant compared to when L2 was available ($p < 0.0001$ and $p = 0.048$, respectively; see Table 19). None of the other ADAS statuses showed significant differences between systems active versus system available.

In the L2 MFA NDS, the average longest single glance when no systems were active was significantly longer than when L1 was active ($p < 0.0001$) or when L2 was active ($p = 0.0045$; see Table 19). Having L2 active also resulted in a significantly longer single glance compared to when L1 was active ($p = 0.013$).

Table 19. Summary of off-road glance metrics for all ADAS statuses in the VCC L2 NDS and L2 MFA NDS in epochs with at least one off-road glance.

	ADAS Status	N	Longest Single Glance (s)	Number of Glances per Epoch	% of Glances > 2 s
VCC L2 NDS	**L2 Available**	129	0.94	2.53	0.2%
	L2 Active	144	1.35	2.88	4.4%
	ACC Available	87	0.93	2.38	1.9%
	ACC Active	93	1.08	2.43	1.7%
	LKA Available	47	1.11	2.43	3.2%
	LKA Active	43	1.14	2.23	2.5%
L2 MFA NDS	**L2 Active**	1,121	1.22	3.28	3.4%
	L1 Active	885	1.10	3.35	1.7%
	None Active	982	1.40	3.35	3.3%

The percentage of glances longer than 2 seconds provides further insight into potential distraction that drivers may experience with ADAS activation. As shown in Table 19, when L2 was active in the VCC L2 NDS, a higher percentage of glances were longer in duration compared to when L2 was available (4.4% versus 0.2%). A mixed model showed this difference to be significant ($F = 12.18$, $p = 0.0006$). None of the other comparisons in the model were significant. The percentage of glances greater than 2 seconds was also highest for L2 Active in the L2 MFA NDS when compared to L1 Active and None Active. However, the model showed no significant differences between each ADAS status.

Driving-Related vs. Non-Driving-Related Task Glances

An additional way to assess eye-glance behavior is to break glance locations down into those associated with a driving-related task or a non-driving-related task. Driving-related glances incorporate the front windshield (i.e., left, right, forward), left- and right-side windows/mirrors, rearview mirror, instrument panel, and over the shoulder (i.e., left and right blind spots). Non-driving-related glances encompass the center dashboard console, cell phone, passenger, and any other interior object. Baseline epochs with valid eye-glance data for all ADAS statuses were initially assessed to determine the percentage of epochs comprising driving-related-only glances (i.e., drivers' eyes were on driving-related tasks the entire time). Table 20 provides a summary of the driving-related-only eye-glance data for each ADAS status in the VCC L2 NDS and L2 MFA NDS.

Table 20. Driving-related-only eye glances for all ADAS statuses in the VCC L2 NDS and L2 MFA NDS.

	ADAS Status	Total with Valid Eye-Glance	Driving-Related-Only Eye-Glance	% with Driving-Related-Only Eye-Glance
VCC L2 NDS	**L2 Available**	195	142	73%
	L2 Active	195	113	58%
	ACC Available	128	89	70%
	ACC Active	132	92	70%
	LKA Available	69	51	74%
	LKA Active	70	53	76%
L2 MFA NDS	**L2 Active**	1,321	900	68%
	L1 Active	1,097	766	70%
	None Active	1,167	735	63%

As seen in Table 20, approximately 70% or more of the baseline epochs in the VCC L2 NDS comprised driving-related-only glances, with the exception of L2 Active, during which driving-related-only glances decreased to 58%. As such, it may be inferred that when drivers had both longitudinal and lateral automation systems active (i.e., L2), they spent less time with their eyes consistently on driving-related tasks. In the L2 MFA NDS data set, the trend appeared to be in the opposite direction. When drivers had no systems active (i.e., manual driving), only 63% of the baseline epochs comprised driving-related-only glances, with that percentage rising to 70% when L1 was active and 68% when L2 was active.

Relative to the time drivers spent with their eyes not on driving-related tasks (Table 21), the average total time drivers' eyes were on non-driving-related tasks in the VCC L2 NDS ranged from nearly half a second when L2 was available (but not active) to just over 1 second when L2 was active. A mixed model indicated that drivers spent significantly more time with their eyes on non-driving-related tasks when L2 was active compared to when L2 was available ($p = 0.0003$). None of the other ADAS statuses showed significant differences between systems active and systems available.

As with previous metrics, the L2 MFA NDS showed somewhat different results to the VCC L2 NDS (Table 21). When drivers had no systems active, they spent significantly more time with their eyes on non-driving-related tasks compared to when they had L2 active ($p = 0.01$) or L1 active ($p = 0.0001$).

Table 21. Total time and percent of time eyes on non-driving-related tasks for all ADAS statuses in the VCC L2 NDS and L2 MFA NDS (including on-road-only epochs).

	ADAS Status	N	Total Time (s)	% of Time
VCC L2 NDS	**L2 Available**	195	0.43	4.4%
	L2 Active	195	1.06	10.7%
	ACC Available	128	0.56	5.7%
	ACC Active	132	0.57	5.7%
	LKA Available	69	0.51	5.1%
	LKA Active	70	0.42	4.2%
L2 MFA NDS	**L2 Active**	1,321	0.84	5.7%
	L1 Active	1,097	0.71	4.8%
	None Active	1,167	1.16	7.8%

The percent of time drivers spent with their eyes on non-driving-related tasks was also assessed (Table 21). In the VCC L2 NDS, the analysis revealed that, when drivers had L2 active, they spent roughly 11% of the time with their eyes on non-driving-related tasks compared to when L2 was available (~4%). A mixed model showed this difference to be significant ($F = 19.62$, $p < 0.0001$). None of the other comparisons between ADAS active and ADAS available were significant.

The results were different for the L2 MFA NDS, with drivers who had no systems active spending approximately 8% of time with their eyes on non-driving-related tasks compared to nearly 5% with L1 Active and approximately 6% with L2 Active. A mixed model showed significant differences between L1 Active and None Active ($t = -4.34$, $p < 0.0001$) and L2

Active and None Active (t = -2.70, p = 0.007), with no systems active being higher in both comparisons.

The glance data (Table 22) showed a similar pattern for non-driving-related tasks as for glances off-road (Table 19). For the VCC L2 NDS, the average longest single glance and average number of glances per epoch were both significantly higher when L2 was active compared to when L2 was available (p = 0.012 and p = 0.017, respectively). As before, these results tended in the opposite direction for the L2 MFA NDS, with the average longest single glance being significantly higher under manual driving conditions (i.e., no systems active) compared to L2 Active (p = 0.01) and L1 Active (p = 0.0002). However, the number of non-driving-related glances per epoch was not significantly different across ADAS statuses for the L2 MFA NDS.

Table 22. Summary of non-driving-related task glance metrics for all ADAS statuses in the VCC L2 NDS and L2 MFA NDS in epochs with at least one non-driving-related task glance.

	ADAS Status	N	Longest Single Glance (s)	Number of Glances per Epoch	% of Glances > 2 s
VCC L2 NDS	**L2 Available**	53	0.95	1.77	0.5%
	L2 Active	82	1.39	2.37	6.7%
	ACC Available	39	0.97	1.97	0.0%
	ACC Active	40	1.16	1.83	1.3%
	LKA Available	18	1.16	1.94	2.8%
	LKA Active	17	1.10	1.88	4.9%
L2 MFA NDS	**L2 Active**	421	1.30	2.47	4.7%
	L1 Active	331	1.16	2.30	2.0%
	None Active	432	1.61	2.52	4.7%

When glances greater than 2 seconds were assessed for non-driving-related tasks, drivers in the VCC L2 NDS with L2 active were again shown to have a significantly higher percentage of non-driving-related glances greater than 2 seconds compared to when L2 was available ($F = 7.35$, $p = 0.008$; Table 22). When LKA was active in the VCC L2 NDS, the percentage of long glances (i.e., > 2 seconds) was also higher relative to when LKA was available; however, this difference was not significant (Table 22). Drivers in the L2 MFA NDS who had no systems active and those who had L2 active had the highest proportion of long glance durations greater than 2 seconds compared to L1 Active. However, the model showed no significant differences.

Research Question 8: When engaged in a secondary task, do drivers take longer glances away from the roadway when ADAS are active?

To investigate glance behavior when drivers were engaged in a secondary task, the percentage of time drivers' eyes were off the forward roadway and the number of long glances greater than 2 seconds were assessed. Table 23 summarizes relevant eye-glance metrics during STE across all ADAS statuses in both NDS data sets. The percent of EORT included epochs wherein drivers did not glance away from the forward roadway (i.e., including zero EORT). However, the percent of long glances (i.e., > 2 seconds) excluded epochs with zero EORT.

During secondary tasks completed when L2 was active in the VCC L2 NDS, drivers spent approximately 29% of the time with their eyes off the forward roadway, compared to just over 18% when L2 was available. A mixed model showed this difference was significant (F = 16.31, $p < 0.0001$). This was the only significant difference between systems active and systems available in the VCC L2 NDS.

In terms of glances longer than 2 seconds, LKA active, ACC active, and L2 active all had higher percentages compared to when these systems were available, but not active. However, further testing showed none of these differences were statistically significant. In addition, comparing STE eye-glances to baseline eye-glances in Research Question 7 (Table 18 and Table 19) revealed the percentages of EORT and long glances were higher during STE than during baseline driving across all levels of ADAS activation (i.e., ACC, LKA, and L2).

Table 23. Eye-glance metrics during STE across all ADAS statuses for the VCC L2 NDS and L2 MFA NDS.

	ADAS Status	N (incl. zero EORT)	% EORT	N (excl. zero EORT)	% of Glances > 2 s
VCC L2 NDS	**L2 Available**	93	18.4%	75	0.3%
	L2 Active	114	28.8%	109	5.8%
	ACC Available	68	18.9%	59	1.4%
	ACC Active	68	21.1%	57	2.7%
	LKA Available	31	19.8%	26	1.9%
	LKA Active	33	18.7%	28	3.9%
L2 MFA NDS	**L2 Active**	791	21.1%	727	4.4%
	L1 Active	681	18.6%	600	1.9%
	None Active	792	22.3%	711	4.4%

In the L2 MFA NDS, drivers with no systems active spent about 22% of the time with their eyes off the forward roadway during STE, compared to approximately 21% with L2 active and 19% with L1 active (Table 23). The model revealed the difference between no systems active and L1 Active was significant (t = -2.97, p = 0.003), and the difference between L1 Active and L2 Active was significant (t = -3.18, p = 0.002). Comparing STE eye-glances to baseline eye-glances (Table 18 and Table 19) showed the percentages of EORT and longer glances (> 2 seconds) were higher during STE across all three ADAS statuses in the L2 MFA NDS.

To evaluate eye-glance patterns associated with STE across both NDS data sets, secondary tasks were grouped into two broad categories: cognitive secondary tasks and visual, manual, or visual-manual tasks. As seen in Figure 18, the percent of time drivers spent with their eyes off-road was higher in the VCC L2 NDS when ADAS were active compared to when ADAS were available for both cognitive tasks and visual, manual, or visual-manual tasks. The exception was when drivers had either LKA active or ACC active, during which time they spent less time with their eyes off-road when engaged in other tasks than when the corresponding system was available. When drivers had L2 active, they spent nearly one-third of the time with their eyes off-road when engaged in a visual, manual, or visual-manual task compared to nearly one-quarter of the time when L2 was available when engaged in a visual, manual, or visual-manual task. Further model testing revealed this to be the only significant difference when comparing systems active to systems available (F = 10.65, p = 0.001). Similarly, for cognitive tasks, drivers spent significantly more time with their eyes off-road with L2 active compared to L2 available (F = 6.09, p = 0.017).

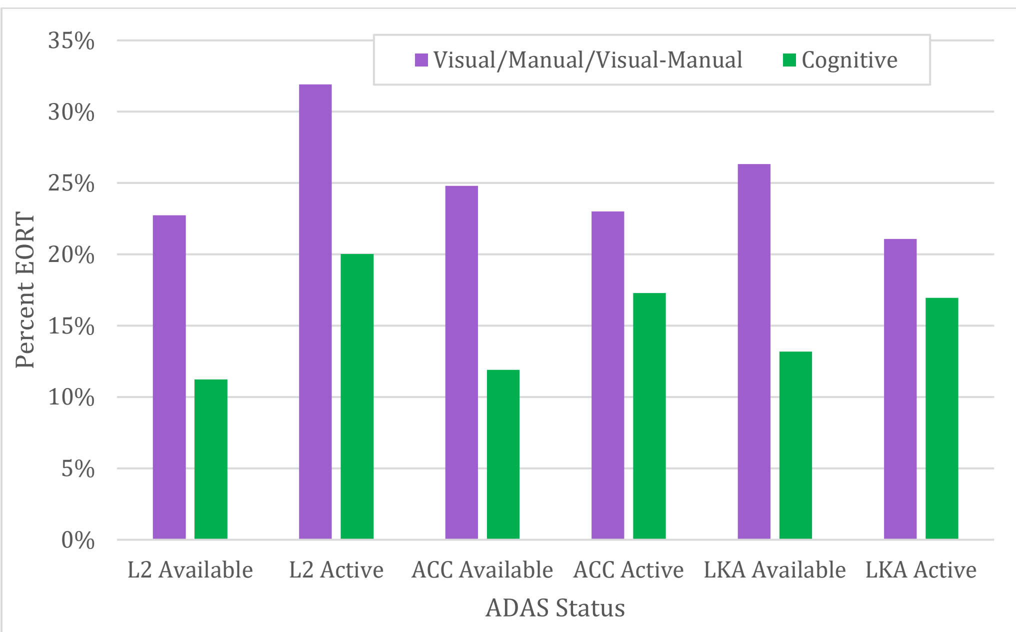

Figure 18. Percent EORT during cognitive and visual, manual, or visual-manual tasks for the VCC L2 NDS

As seen in Figure 19, the percent of time drivers spent with their eyes off-road in the L2 MFA NDS was similar for L2 active and L1 active when engaged in a visual, manual, or visual-manual task. The percent of EORT across both secondary task categories was slightly higher when drivers had no systems active. When engaged in a cognitive task, however, drivers spent a significantly higher proportion of time with their eyes off-road when L2 was active compared to when L1 was active (t = -2.66, p = 0.008). L2 Active was also higher relative to cognitive tasks than no systems active, but the difference was not significant.

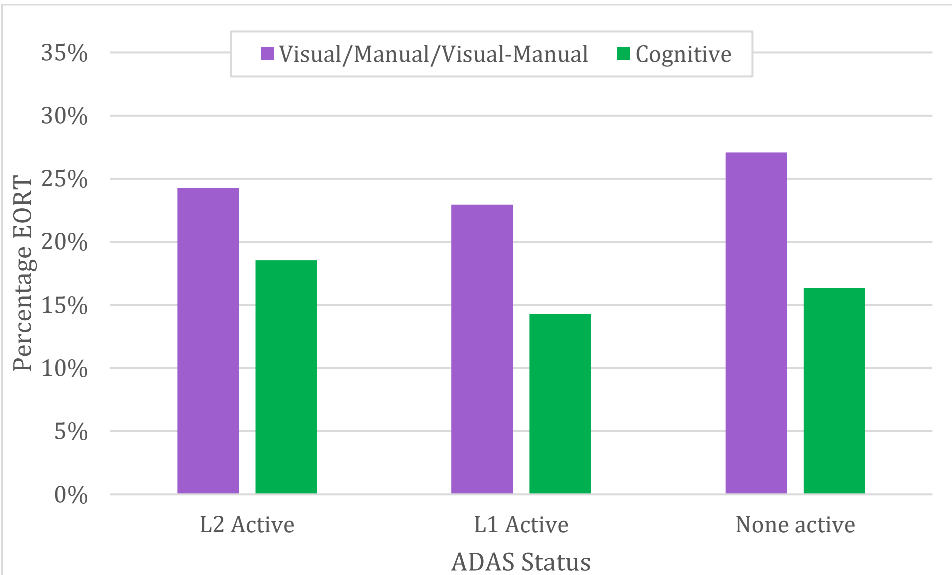

Figure 19. Percent EORT during cognitive and visual, manual, or visual-manual tasks for the L2 MFA NDS

In terms of long glances (i.e., > 2 seconds), a higher proportion of long, off-road glances occurred in the VCC L2 NDS when L2 was active or LKA was active as drivers engaged in a visual, manual, or visual-manual secondary task, compared to when those corresponding systems were available (see Table 24). Of these two differences, only L2 Active compared to L2 Available was significant (F = 9.68, p = 0.002). When engaged in cognitive tasks, the majority of drivers in the VCC L2 NDS did not take long off-road glances, the exceptions being when L2 was active and when ACC was active. L2 Active compared to L2 Available was the only significant comparison (F = 8.60, p = 0.006).

Table 24. Percent of glances greater than 2 seconds for visual, manual, or visual-manual and cognitive tasks across all ADAS statuses in the VCC L2 NDS and L2 MFA NDS in epochs with at least one off-road glance.

		Visual/Manual/Visual-Manual Tasks		Cognitive	
	ADAS Status	**N**	**% of Glances > 2 s**	**N**	**% of Glances > 2 s**
VCC L2 NDS	**L2 Available**	60	0.4%	26	0%
	L2 Active	91	7.0%	29	4.6%
	ACC Available	43	1.9%	22	0%
	ACC Active	44	1.6%	21	4.0%
	LKA Available	19	2.6%	10	0%
	LKA Active	22	4.9%	12	0%
L2 MFA NDS	**L2 Active**	554	5.2%	328	3.3%
	L1 Active	419	2.6%	307	0.7%
	None Active	535	5.5%	362	2.4%

When drivers in the L2 MFA NDS engaged in a visual, manual or visual-manual secondary task, they took significantly less longer glances (i.e., > 2 seconds) off-road when L1 was active compared to L2 active (t = -4.13, p < 0.0001) and no systems active (t = -3.51, p = 0.0005; Table 23). During cognitive tasks, L2 MFA NDS drivers took longer glances off-road when L2 was active and when no systems were active. However, neither of these differences were significant.

Research Question 9: In general, do drivers engage in less scanning of the roadway environment when ADAS are active?

Scanning the driving environment is an important element of safe driving behavior, as long as the off-road glances are short (i.e., less than 2 seconds off the forward roadway) and the glances are driving-related (e.g., checking mirrors or blind spots). The longer the driver has his/her eyes off the forward roadway, the riskier the secondary activity becomes (Klauer et al., 2006). To assess eye-scanning behavior by ADAS status, off-road glances, driving-related glances, and non-driving-related glances were assessed. Total EORT, percent of time eyes were off-road, average single glance duration, and percent of glances greater than 2 seconds were considered.

In the VCC L2 NDS, it appears drivers with ACC available or L2 available (but not active) engaged in the safest eye-scanning behavior of the roadway environment. The percent of long glances on non-driving-related tasks when these two systems were available were lower than any other ADAS status (ACC available = 0.0%; L2 available = 0.5%; see Table 22 under Research Question 7). Similarly, the longest single non-driving-related glance when ACC or L2 was available was lower than all other ADAS levels (ACC = 0.97 seconds; L2 = 0.95 seconds; see Table 22). In addition, the total time and percent of time drivers' eyes were off the forward roadway were lower when ACC or L2 was available than the majority of the other ADAS statuses (ACC Available: 1.29 seconds, 13.0%; L2 Available: 1.29 seconds, 13.0%, respectively), with the exception of LKA Active (1.2 seconds, 11.8%; Table 18 under Research Question 7).

Conversely, drivers who had L2 active in the VCC L2 NDS engaged in poor eye-scanning behavior of the roadway environment. When L2 was active, drivers had the highest percent of non-driving-related glances greater than 2 seconds (6.7%), the greatest number of non-driving-related glances per epoch (~2.4), and the longest single glance at non-driving-related tasks (~1.4 seconds; see Table 22 under Research Question 7).

When L2 was active, the percent of time drivers spent with their eyes on driving-related tasks (Table 25) was significantly lower compared to when L2 was available (89.4% versus 95.6%, respectively); the average total time with eyes on driving-related tasks was also significantly lower when comparing L2 Active and L2 Available (8.9 seconds versus 9.5 seconds, respectively). Thus, when L2 was active, VCC L2 NDS drivers looked away from the forward roadway more frequently and for longer periods of time and spent more total time with their eyes not on driving-related tasks.

Table 25. Total time and percent of time drivers' eyes were on driving-related tasks in the VCC L2 NDS and L2 MFA NDS (including epochs with only driving-related glances)

	ADAS Status	N	Average Total Time (s)	% of Total Time
VCC L2 NDS	**L2 Available**	195	9.5	95.6%
	L2 Active	195	8.9	89.4%
	ACC Available	128	9.4	94.3%
	ACC Active	132	9.4	94.3%
	LKA Available	69	9.4	94.9%
	LKA Active	70	9.5	95.8%
L2 MFA NDS	**L2 Active**	1,321	14.1	94.3%
	L1 Active	1,097	14.2	95.2%
	None Active	1,167	13.8	92.2%

In the L2 MFA NDS, it appears drivers with L1 active engaged in the safest eye-scanning behavior of the roadway environment. These drivers had the lowest percent of long glances on non-driving-related tasks (2.0%), the shortest single non-driving-related glance (~1.2 seconds), the lowest EORT (2.2 seconds), and the lowest percent of time drivers' eyes were off-road (14.7%) compared to drivers with L2 Active or no systems active (Table 18 and Table 22 under Research Question 7).

L2 MFA NDS drivers with no systems active appeared to engage in the poorest eye-scanning behavior. These drivers spent significantly less time with their eyes on driving-related tasks compared to L2 active ($t = 1.97$, $p = 0.05$) and L1 active ($t = 2.56$, $p = 0.01$; Table 25). When no systems were active, drivers had the highest number of non-driving-related glances (~2.5) and the longest single glance at non-driving-related tasks (~1.6 seconds). The percent of long non-driving-related glances was highest with L2 Active (4.7%) and no systems active (4.7%), both of which were higher than when L1 was active (2.0%; Table 22 in Research Question 7).

Research Question 10: Is driver drowsiness observed more often when ADAS are active?

PERCLOS 1 was performed on all baseline epochs from the VCC L2 NDS data set and a subset of baseline epochs from the L2 MFA NDS data set. The number of valid events (i.e., having one minute of data available with at least 80% of video frames usable for analysis), the number of events rated as "drowsy," and the prevalence of drowsy driving for each ADAS status in both NDSs are presented in Table 26.

Table 26. Prevalence of drowsy driving across all ADAS statuses in the VCC L2 NDS and L2 MFA NDS using PERCLOS 1.

	ADAS Status	# of Valid PERCLOS 1 Events	Number of Events Rated "Drowsy" (>12% PERCLOS 1)	Drowsy Driving Prevalence
VCC L2 NDS	**L2 Available**	174	0	0%
	L2 Active	179	1	0.6%
	ACC Available	110	2	1.8%
	ACC Active	117	0	0%
	LKA Available	57	0	0%
	LKA Active	58	1	1.7%
L2 MFA NDS	**L2 Active**	317	17	5.4%
	L1 Active	265	10	3.8%
	None Active	290	10	3.4%

Drowsy driving was present in a small percentage of baseline epochs for the VCC L2 NDS, with the highest prevalence being 1.8% when ACC was available, followed by 1.7% when LKA was active. In the L2 MFS NDS, drowsy driving was present in a higher percentage of baselines. Notably, when L2 was active, drowsy driving was present in 5.4% of baselines compared to 3.8% when L1 was active and 3.4% when no systems were active.

Due to the nature of driver drowsiness, there is a possibility that drowsy driving events may be a function of trip duration, with drowsiness being more likely to occur during trips of longer duration. In addition, time-on-task is an important consideration, which refers herein to the time elapsed from the start of the trip to the sampled PERCLOS epoch. To investigate this, the duration of trips containing a drowsy driving event were compared to trips with no drowsy driving events, and the time from trip start to PERCLOS epoch start was calculated for both drowsy and non-drowsy trips. In the VCC L2 NDS, the mean trip duration of four drowsy driving trips was 31 minutes (*SD* = 16 min), and time-on-task was

17 minutes (*SD* = 10.4 min). The mean duration of 804 non-drowsy driving trips was 42.3 minutes (*SD* = 29.8 min), and time-on-task was 21 minutes (*SD* = 22.5 min). Thus, the drowsy driving trips were shorter in duration; time-on-task prior to the occurrence of a drowsy driving event was just over four minutes less than a non-drowsy driving event. Figure 20 shows the distribution of all VCC L2 NDS trip durations as a function of whether or not the trip contained a drowsy driving event (as determined by the PERCLOS 1 analysis).

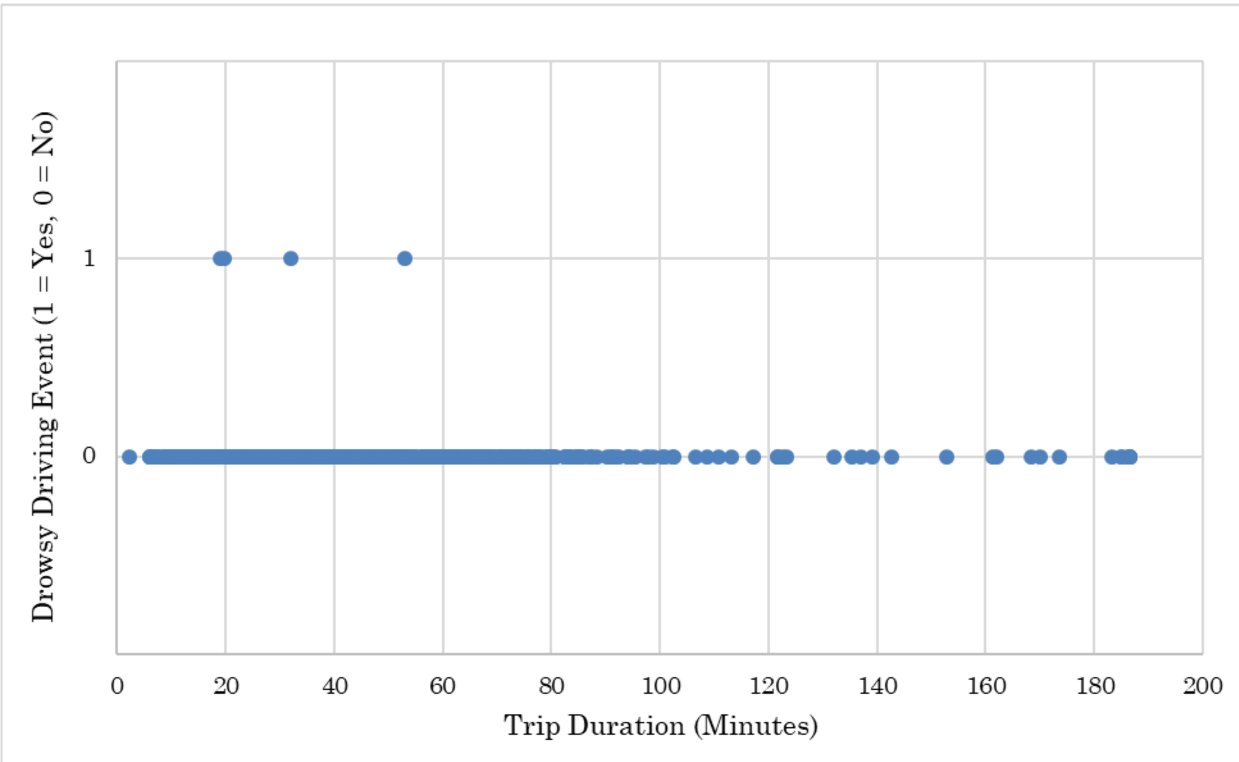

Figure 20. Distribution of trip durations with and without a drowsy driving event for the VCC L2 NDS

In the L2 MFA NDS, the durations of the drowsy and non-drowsy trips were similar, as were time-on-task prior to a drowsy and non-drowsy driving event. The mean trip duration for 37 drowsy driving trips was 53.9 minutes (*SD* = 32.9 min), with a time-on-task of 31 minutes (*SD* = 31.4 min). For the 835 non-drowsy driving trips, the mean duration was 57.5 minutes (*SD* = 38.1 min), with a time-on-task of 27 minutes (*SD* = 25.0 min). Thus, the time-on-task to a drowsy driving event was roughly four minutes longer than a non-drowsy driving event. Figure 21 shows the distribution of all L2 MFA NDS trip durations as a function of whether or not the trip contained a drowsy driving event (as determined by the PERCLOS 1 analysis).

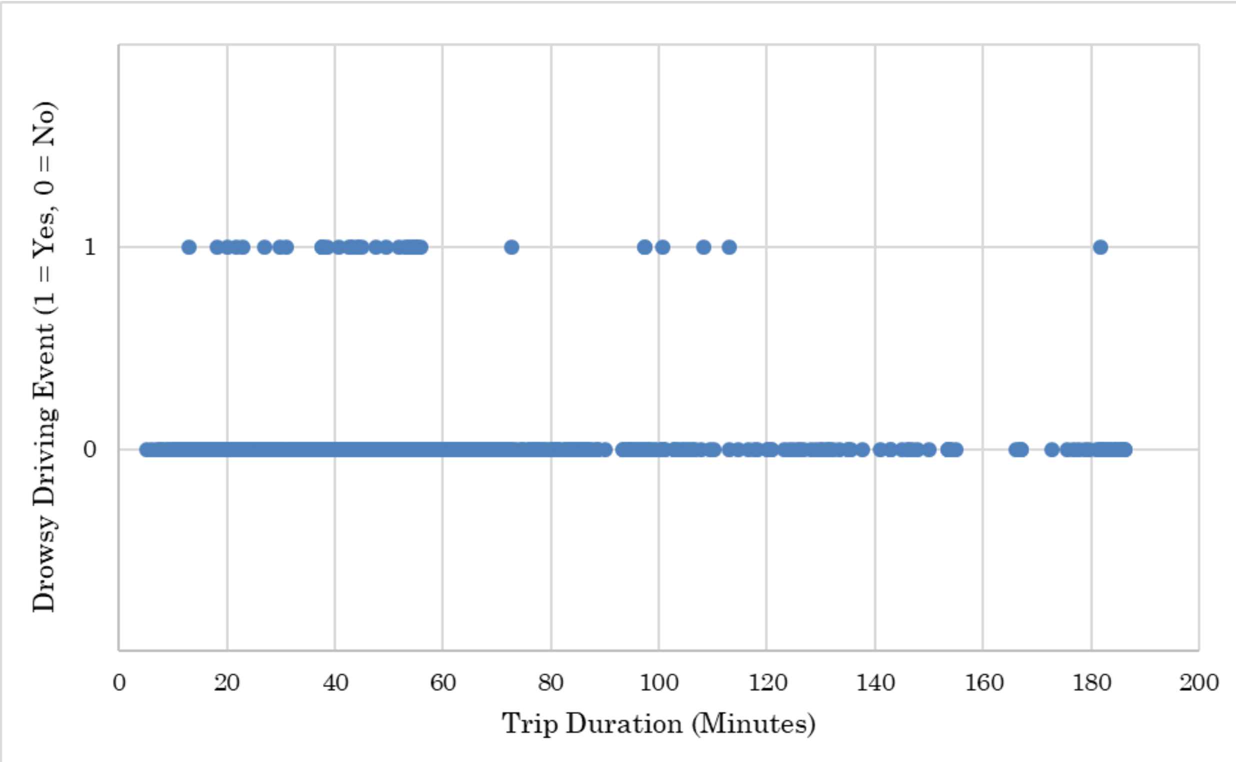

Figure 21. Distribution of trip durations with and without a drowsy driving event for the L2 MFA NDS

As can be seen in Figure 20 and Figure 21, all trips in the VCC L2 NDS and the majority of trips in the L2 MFA NDS containing a drowsy driving event were under 60 minutes in duration; time-on-task prior to a drowsy driving event was not significantly longer than for a non-drowsy driving event. Of the 37 trips in the L2 MFA NDS that included a drowsy driving event, 10 occurred under manual driving conditions (i.e., no ADAS active), and 10 occurred with L1 active. The remaining 17 drowsy driving events occurred under L2 activation, where both ACC and LKA were activated simultaneously, with one of the trips being just over three hours long (181 minutes). Although preliminary, these results indicate there is not substantial workload underload that leads to bouts of unusual drowsiness when operating with L2 automation relative to manual driving.

Research Question 11: Is driver drowsiness more prevalent during SCEs that occur when ADAS are active?

PERCLOS 1 was performed on all SCEs from the VCC L2 NDS and the L2 MFA NDS. The number of valid events (i.e., having one minute of data available with at least 80% of video frames usable for analysis), the number of events rated "drowsy," and the prevalence of drowsy driving for each ADAS status in both NDSs are presented in Table 27.

Table 27. Prevalence of drowsy driving during SCEs across all ADAS statuses in the VCC L2 NDS and L2 MFA NDS.

	ADAS Status	# of Valid PERCLOS Events	Number of Events Rated "Drowsy" (>12% PERCLOS 1)	Drowsy Driving Prevalence
VCC L2 NDS	**L2 Available**	57	1	1.8%
	L2 Active	21	0	0%
	ACC Available	9	0	0%
	ACC Active	13	0	0%
	LKA Available	20	1	5%
	LKA Active	9	0	0%
L2 MFA NDS	**L2 Active**	13	0	0%
	L1 Active	11	0	0%
	None Active	30	0	0%

Drowsy driving was present in a very small number of SCEs for the VCC L2 NDS, with the highest prevalence being 5% when LKA was available, followed by 1.8% when L2 was available. Conversely, in the L2 MFA NDS, drowsy driving was not present in any SCEs. The number of SCEs in each data set, particularly in the L2 MFA NDS, was also relatively small.

Research Question 12: How do drivers respond to ADAS alerts?

In the VCC L2 NDS data set, there were three types of alerts that were identified by the machine-vision process (applied only in the Tesla vehicles): FCW and immediate takeover events requested that drivers react by putting their hands on the steering wheel. HOW prompts – where there was not immediate danger – were analyzed separately from those alerts because they lacked the same urgency. Alerts in the L2 MFA NDS were also analyzed separately, as they did not include FCW alerts and were specific to lateral automation features.

Response Types to FCW and Immediate Takeover Alerts in the VCC L2 NDS

Overall, response types varied based on alert types in the VCC L2 NDS. There were no instances of alerts occurring when the driver had his/her hands off the wheel. All drivers had at least several fingers, one hand, or both hands on the steering wheel just prior to the onset of the alert. Table 28 shows the response types for FCW and immediate takeover alerts, including response combinations wherein drivers simultaneously engaged in two different response types.

Table 28. Response types to FCW and immediate takeover alerts in the VCC L2 NDS.

Response Type	% Response to FCW	% Response to Immediate Takeover
Engage brakes	50.8%	3.3%
Instrument cluster glance	0%	36.1%
Reach for steering wheel	0%	4.9%
Prepare to brake	3.2%	0%
Accelerate	1.6%	0%
Engage brakes + Instrument cluster glance	4.8%	1.6%
Engage brakes + Reach for steering wheel	22.2%	0%
Instrument cluster glance + Reach for steering wheel	0%	19.7%
Instrument cluster glance + Prepare to brake	0%	4.9%
Other	4.8%	6.6%
No Response	12.7%	23%

As seen in Table 28, the majority of VCC L2 NDS drivers responded to FCW alerts by engaging the brakes, either as their only response or in combination with a second response type. Responses to immediate takeover alerts tended to involve glancing at the instrument cluster, either alone or simultaneously with a second response. Given that none of the drivers had their hands completely off the steering wheel, the response type of "reach for steering wheel" is low. When such action did occur, the drivers had only several fingers or one hand on the steering wheel; thus, placing both hands on the steering wheel gave them additional control over the vehicle.

Response Types to HOW Prompts in the VCC L2 NDS

Of the 391 hands-on-wheel (HOW) prompts (see Table 3), all occurred when drivers' hands were completely off the steering wheel. As shown in Table 29, the two most frequent response types were an instrument cluster glance or reaching for the steering wheel. These two responses also happened simultaneously for approximately one-quarter of the HOW prompts. Interestingly, approximately 15% of HOW prompts resulted in no response from the driver, meaning the driver chose to ignore the prompt and to continue driving with his/her hands off the steering wheel.

Table 29. Response types to HOW prompts in the VCC L2 NDS.

Response Type	**Percent**
Engage brakes	0.8%
Instrument cluster glance	30.4%
Reach for steering wheel	24.8%
Adjust grip on steering wheel	0.5%
Instrument cluster glance + Reach for steering wheel	24.6%
Other	3.3%
No Response	15.6%

Response Types to Alerts in the L2 MFA NDS

Alerts in the L2 MFA NDS were specific to lateral automation features; thus, these alerts were activated when the vehicle required the driver to resume control of the vehicle. Of the 450 alerts in the L2 MFA NDS (see Table 4), drivers had their hands completely off the steering wheel in 119 instances. For the remaining 331 alerts, the drivers had several fingers, one hand, or both hands on the wheel just prior to the onset of the alert. The response types varied based on whether drivers had their hands on or off the wheel prior to the alert. As seen in Table 30, if drivers' hands were on the wheel, the most frequent response type was "No Response." This is surprising – as such alerts typically indicate that the driver needs to takeover steering of the vehicle.

The second most frequent response if drivers' hands were on the wheel was to glance at the instrument cluster, either alone or in combination with another response type. If drivers' hands were off the wheel, the most frequent response type was to reach for the steering wheel, either alone or at the same time as another response.

Table 30. Response types to alerts in the L2 MFA NDS.

Response Type	% Hands on Wheel	% Hands off Wheel
Engage brakes	0.6%	0%
Instrument cluster glance	35.1%	11.8%
Reach for steering wheel	7.9%	43.7%
Other	2.1%	0.8%
Adjust grip on steering wheel	1.8%	0.8%
Instrument cluster glance + Reach for steering wheel	7.0%	26.1%
Instrument cluster glance + Adjust grip	7.0%	0%
No Response	38.4%	16.8%

Research Question 13: How long does it take drivers to respond to ADAS alerts?

Response times to the various alerts were assessed from the time the alert began to the first indication of a response from the driver. In the VCC L2 NDS, drivers responded significantly faster to FCW alerts than to immediate takeover alerts (t = -5.85, p < 0.0001), with an average response time of 0.15 seconds for FCW alerts compared to 0.51 seconds for immediate takeover alerts. Reflecting the lack of urgency associated with HOW prompts in the VCC L2 NDS, the drivers' average response time was 2.23 seconds. Interestingly, the response times to alerts in the L2 MFA NDS did not differ significantly based on the position of drivers' hands (p = 0.36). The average response time when drivers' hands were off the wheel was 0.99 seconds; when their hands were on the wheel, the average response time was 0.89 seconds.

Research Question 14: Is driver drowsiness more prevalent when drivers receive ADAS alerts?

PERCLOS 1 was performed on all alerts and prompts from both data sets that had valid data (i.e., one minute of data available with at least 80% of video frames usable for analysis). Results were then compared to the corresponding ADAS status baseline PERCLOS 1 results from Research Question 10.

PERCLOS 1 on Alerts from the VCC L2 NDS

For the FCW alerts, which do not require any specific automation to be active (i.e., such alerts can occur during manual driving), there were 55 events with valid PERCLOS 1 data. However, none of these events reached the 12% threshold to be defined as drowsy driving events.

For the immediate takeover alerts, which occur under conditions when L2 (i.e., ACC and LKA) is active, there were 68 events with valid PERCLOS 1 data. Six of these events reached the 12% threshold to be defined as drowsy driving events. Therefore, the prevalence of drowsy driving during the occurrence of immediate takeover alerts was 8.8%. Compared to the baseline prevalence (0.6%) of drowsy driving when L2 was active (see Table 26), it appears the prevalence of drowsy driving was higher when drivers received immediate takeover alerts than under normal driving conditions with both systems active.

HOW prompts only occur under conditions when L2 is active. In the VCC L2 NDS, there were 340 events with valid PERCLOS 1 data, 12 of which reached the pre-defined drowsy driving threshold. Therefore, the prevalence of drowsy driving during the occurrence of HOW prompts was 3.5%. This is higher than the prevalence of drowsy driving (i.e., 0.6%) under normal driving conditions with both systems active.

PERCLOS 1 on Alerts from the L2 MFA NDS

For the L2 MFA NDS alerts, which occur under L2 active conditions, there were 318 hands-on-wheel events with valid PERCLOS data, 12 of which reached the pre-defined drowsy driving threshold. In terms of hands-off-wheel events, there were 117 events with valid PERCLOS data, three of which reached the drowsy driving threshold. Thus, the prevalence of drowsy driving was 3.8% during alerts when drivers had their hands on the steering wheel and 2.6% when drivers had their hands off the steering wheel. These results are lower when compared to the prevalence of drowsy driving under normal driving conditions with both systems active, which was 5.4% in the L2 MFA NDS (see Table 26).

Discussion and Conclusions

Decades of prior research in other domains such nuclear power and commercial flight have highlighted the potential consequences of partial automation. Such consequences include operator confusion over control authority, reduced workload resulting in boredom that could lead to drowsiness or fatigue or an over-reliance on a system that results in reduced attention to the primary task.

One important difference between those industries and the current study is the level of training and practice received. Commercial aviators or nuclear power plant operators, for example, receive thousands of hours of training and practice prior to certification as operators. Conversely, users of automotive technology only receive a minimal set of written and verbal instructions prior to use on public roadways in live traffic.

Saad et al. (2004) proposed a two-phase model involved in the development of behavioral adaptation – the learning phase and the integration phase. Building on this model to incorporate the results of the current study, we propose a three-phase model of ADAS operation:

1. The "novelty" phase of system use. Since ADAS are generally new to users and training and practice are minimal, it may be hypothesized — as was borne out, at least anecdotally, in Russell et al. (2018) — that drivers are learning the system while using it in real time. This leads to some level of "testing" system use and limitations in live traffic. If drivers are not cautious in this testing and learning approach, it seems reasonable that risky behaviors and safety-critical circumstances could be the result. Moreover, drivers who do not have experience with the systems may not trust the systems completely.
2. The post-novelty operational phase. This phase may be equivocated in other domains to a fully trained, but new, operator. Risks here may include cases wherein the driver has a mental model of system operation, including cases of transition and control authority. This is where overreliance on system capabilities potentially develops, and one could expect to encounter SCEs in such circumstances.
3. The experienced user phase. This is a phase in which the driver has used the system repeatedly, understands the system well (e.g., has "tested" the system and perhaps

even over-relies on its limited capabilities), and uses the system frequently. This phase is where the potential for work underload could manifest itself. An increasing level of overreliance could also occur with increased experience, particularly in cases where the novelty and post-novelty phases did not result in many (or any) SCEs that led to caution on the part of the user.

There were substantial differences between the results from the VCC L2 NDS and the L2 MFA studies. Many of these differences could be due to dissimilarities in participants' exposure to the operation phases described above. As detailed in the Methods section, several aspects of data collection were different across the studies. These differences, which may inform important behaviors in using ADAS features, included:

1. Length of time in the study. It can be inferred from the NDS data used in the current study that a novelty phase existed wherein drivers "tested" the systems. Although the exact duration of a typical novelty phase for a user is unclear, it is reasonable to assume that it lasts at least several weeks. In essence, it may be assumed that participants in the L2 MFA NDS were always in the novelty phase, since they had equipped cars for one month at a time. By contrast, the VCC L2 NDS data perhaps provided a better snapshot into long-term adaptation and overreliance effects.
2. Level of training received. The L2 MFA NDS subjects were provided detailed, comprehensive training that was precisely what the manufacturers recommended. Although it is possible the vehicle owners in the VCC L2 NDS received some orientation from the various dealers, recent studies (e.g., Abraham et al., 2017) suggest that automotive dealers do not adequately "train" drivers on ADAS use. Moreover, drivers rarely consult the owner's manual in full (Leonard, 2001), and such manuals may not be adequate in delivering instructions relative to driving automation (Boelhouwer et al., 2019).
3. Provided versus owned vehicle. There are known differences as to how drivers behave — and the associated risks — when operating familiar (owned) or unfamiliar (e.g., rental or leased) vehicles (Perel, 1983). Such differences potentially led to several of the result discrepancies between the two studies. However, the results of the current study were contrary to prior research. Drivers in L2 MFA NDS had

lower SCE rates than VCC L2 NDS, despite the fact the study vehicles were provided to the drivers for a relatively short period of time, similar to a rental or leased vehicle.

In general, the results of the VCC L2 NDS showed a greater impact on a number of the driver behavioral measures when both lateral and longitudinal automation features (i.e., L2: ACC and LKA) were active compared to when these systems were available. Eye-glance metrics indicated drivers had their eyes off-road and on non-driving-related tasks more frequently and for longer durations when L2 was active compared to any other ADAS status (i.e., active or available) (see Table 18 and Table 22). In addition, when L2 was active compared to available in the VCC L2 NDS, drivers had approximately 1.5 times the odds of engaging in a secondary task and 1.8 times the odds if the secondary task was primarily visual, manual, or visual-manual in nature. When these measures are interpreted as surrogates for driver distraction, it can be concluded that the use of both lateral and longitudinal automation features (i.e., L2 automation) results in greater driver distraction. However, it should be acknowledged that this increase in STE with L2 active does not necessarily mean ADAS are unsafe. If a driver is going to engage in a secondary task while driving, the optimal time to do so would be when L2 is active and momentary diversions of the driver's attention would be less risky than when the driver is in full control of the vehicle. It is also possible that drivers specifically engage automation features when they plan to engage in a secondary task; thus, while the VCC L2 NDS results indicate the use of L2 automation culminated in an increased occurrence of distracted driving behaviors, perhaps this is intentional on the part of the driver.

The VCC L2 NDS results may also suggest that drivers trust the ADAS features, at least with some level of extended use. However, were the systems to fail, especially when the driver's attention is focused elsewhere, it is questionable whether the driver would have the ability to recover control of the vehicle in time to prevent a crash. While the SCE rates for ADAS Active compared to ADAS Available/Inactive were not significantly different in the VCC L2 NDS, this may be due to the systems operating as designed or drivers trusting the systems in situations where the surrounding roadway and traffic scenarios were such that it was relatively safe for the driver to trust the systems. In other words, as long as the automation features work as they should, and in operational scenarios where they work well, they may help the driver avoid an SCE. Yet, caution should always be taken not to over-trust and over-rely on the systems to maintain driver safety (Inagaki & Itoh, 2013).

Results from the L2 MFA NDS suggest that drivers who had little to no experience with ADAS features did not place the same level of trust in the systems as the drivers in the VCC L2 NDS. L2 MFA NDS drivers with no systems active (i.e., under manual driving conditions) were more likely to engage in a secondary task than with L1 active (OR = 1.23) or L2 active (OR = 1.38). These drivers had nearly 1.3 times the odds of engaging in a secondary task when no systems were active compared to one system active if the task was primarily visual, manual, or visual-manual and approximately 1.5 times the odds of STE when no systems were active compared to L2 active if the task was cognitive in nature. Eye-glance metrics also indicated manual driving was more likely to involve longer and more frequent eyes-off-road glances and on non-driving-related tasks than when ADAS were in use (see Table 18 and Table 22). These results were similar to those demonstrated by Russell et al. (2018), who, using this same data set, found distracting behaviors were just as prevalent during periods of manual driving or with L1 active as they were when L2 was active. Thus, these results suggest that, depending on the driver, his/her experience with the ADAS features, and the trust he/she has in these systems, manual driving can be just as risky, if not riskier, than using ADAS, at least in some circumstances.

The L2 MFA NDS results provide support for the existence of a novelty phase of ADAS operation, wherein drivers are still testing and learning how to use the ADAS. Additional support can be found by examining a small number of vehicles in the VCC L2 NDS that were owned by drivers for less than three months upon enrollment in the study. Here, five vehicles and drivers contributed one-quarter of the SCEs to the data set, 14 of which occurred with ADAS active and 23 that occurred when ADAS were available or inactive. Despite this seemingly high number of SCEs, these five "novel" drivers in the VCC L2 NDS did not have higher SCE rates than the study sample as a whole due to their higher number of driving minutes. Thus, while their SCE count seemed high, these five drivers also accounted for more minutes of driving, resulting in a per-driver SCE rate of 0.20 per 1,000 minutes. Interestingly, when these five "novel" drivers were excluded from the VCC L2 NDS data set, the remaining "experienced" drivers had twice the odds of engaging in a visual, manual, or visual-manual secondary task when L2 was active. This may be an indication of experienced drivers' over-trust in the systems.

To investigate the potential impact of technological differences in lane centering versus lane keeping assist (LKA) on distracted driving behavior, STE and eye-glance metrics were selected for the subset of Teslas within the VCC L2 NDS (i.e., the vehicle make and model that was known to be equipped with true lane-centering technology). Drivers of those vehicles equipped with lane centering did not exhibit any particular differences in STE when ADAS were active versus when ADAS were available. However, when L2 was active, the proportion of long glances (i.e., > 2 seconds) on non-driving-related tasks was higher than the study sample as a whole (9.3% versus 6.7%, respectively). This indicates that drivers may feel more comfortable looking away from the road at non-driving-related tasks when their vehicle is equipped with a proactive lane-keeping system, such as lane centering, rather than a reactive system, such as LKA. This is a difference that should be investigated further, with additional studies targeting vehicles equipped with lane-centering systems. It also further highlights how many different factors play a role in driver behavior when it comes to advanced automation features.

The low baseline observation of drowsy driving in the VCC L2 NDS was contrary to what was expected, especially under scenarios when L2 was active. Previous research into automation and underload (Young & Stanton, 2004; 2006) and the relationship between underload, monotony, and passive task-related fatigue (Matthews et al., 2009) would seem to indicate that, when both lateral and longitudinal systems are active (i.e., L2 active), drivers would be more likely to experience driver drowsiness. However, this was not the case in the current study, which may be an indication that the use of these ADAS features does not negatively impact driver alertness. Again, although preliminary, these results suggest there is not substantial workload underload that leads to bouts of unusual drowsiness when operating with L2 automation relative to manual driving. It is worth reiterating, though, that the overall observation of drowsy driving events, regardless of ADAS status, was low in the VCC L2 NDS (0.6%), which may imply there is something different about these drivers and their trips. For example, if the majority of the trips were taken during the day and for short duration, then the observation of driver drowsiness would be low. Drowsy driving prevalence was higher in the L2 MFA NDS, particularly when both systems were active (5.4%), which may suggest drivers with little to no experience using the ADAS are more impacted by drowsiness than experienced drivers.

Similar to previous research using the SHRP 2 NDS (Dingus et al., 2016), the observation of performance errors in both the VCC L2 NDS and L2 MFA NDS data sets was low (1.6% and 1.7%, respectively). Judgment errors, however, were much higher in the current study than in the SHRP 2 NDS. This difference was due to the higher occurrence of speeding (i.e., ≥ 10 mph above the posted speed limit) in the current study. In the VCC L2 NDS, the occurrence of speeding decreased from 17.5% to 11% when drivers went from L2 Available to L2 Active. Thus, the use of ADAS decreased the occurrence of speeding. Drivers in the L2 MFA NDS showed the opposite trend, with speeding being more prevalent when L2 was active compared to when no systems were active (19% versus 16%, respectively). It is possible that the more inexperienced drivers (i.e., inexperienced with automated features) in the L2 MFA NDS misunderstood how the systems worked and assumed the system would adjust the speed accordingly when entering a new speed zone. Such drivers may not have realized they had to reset/adjust their speed when the speed zones changed, resulting in a greater propensity to exceed the speed limit when L2 was active. Further investigation is merited.

The results from the eye-glance and STE analyses of the VCC L2 NDS provide concerning safety implications for the use of lateral and longitudinal automation features combined (i.e., L2). When activated simultaneously, these systems resulted in drivers engaging more frequently in distracting secondary tasks (58%) than when both systems were available (47%). This resulted in greater EORT than when systems were available but not active (i.e., 2.02 seconds L2 Active versus 1.29 seconds L2 Available). This was not the case for drivers in the L2 MFA NDS, who appeared to still be in the novelty phase of operation. Drivers in the L2 MFA NDS engaged more frequently in secondary tasks with no systems active (i.e., under manual driving conditions) than with L2 active (69% versus 60%). They also spent more time with their eye off-road with no systems active compared to when L2 was active (2.70 seconds versus 2.51 seconds). Here, behavioral adaptation did not seem to occur within the first month of driving a vehicle equipped with ADAS. As outlined in Sullivan et al. (2016), behavioral adaptation likely occurs after the driver has developed an idea of how the ADAS operate and has incorporated that insight into his/her own driving. In other words, drivers first acquire a vehicle equipped with advanced safety features (e.g., ACC and LKA), then they learn and test these features to determine the circumstances under which the systems work and when they do not work (or do not work as reliably). Once the driver has determined situations wherein he/she feels comfortable trusting the ADAS to relieve him/her of vehicular control, potentially negative behavioral adaptation begins to occur. Adding to this danger is the possibility the driver may not fully understand the limitations of the ADAS and may overestimate its ability to prevent crashes (McDonald et al., 2018). As such, drivers need to be aware of potential pitfalls that exist even after they have learned when and how to use the systems. Unintentional dangers exist for novel drivers, as well as experienced drivers. Becoming complacent and over-reliant on the systems are very real dangers, especially for drivers who have not experienced any repercussions of such behavior during the novelty or post-novelty phase. Negative behavioral adaptation could thus have real and tragic consequences for drivers of ADAS-equipped vehicles. The possible safety benefits of these systems are undeniable, but it is imperative that potential lack of driver awareness of unintentional dangers be evaluated and addressed.

Limitations

The following limitations should be considered when interpreting the results of this study. First, there is a potential interpretive caveat that should be raised relating to the comparison of periods of time when ADAS were active and when ADAS were available but inactive. Given that the drivers had complete control over when to use, or not use, the various ADAS features, there may be instances when the driver elected not to use the ADAS simply because he/she would prefer to drive themselves. On the contrary, there may be instances when the driver specifically chose to use L2 automation because he/she intended to engage in a secondary task that required them to take their eyes off the forward roadway for a long duration of time. The choice to use the ADAS was dependent on driver-based reasoning in every instance, and this reasoning was not available to the researcher; thus, there may be alternative interpretations to the differences found between ADAS active and ADAS available.

Second, there were cases where small sample sizes prevented strong statistical analyses. For instance, the prevalence of driver behaviors was relatively low, as was the number of SCEs. This presented a problem when attempting to break the numbers down into specific

ADAS statuses, especially for the VCC L2 NDS that comprised six ADAS groupings (e.g., ACC, LKA, L2 active and ACC, LKA, L2 available). As a result, epochs were grouped according to system active or system available/inactive, thereby still allowing for the comparison of interest to be made.

Third, it should be noted that, relative to STE, it is inherently difficult to assess cognitive distraction in a naturalistic driving environment. NDSs preclude the use of intrusive instrumentation typically used to measure and assess brain activity. As such, only observable instances of cognitive distraction (e.g., interacting with passengers, talking/listening on a hands-free cell phone) were used in the current study; instances of mind-wandering, or being "lost in thought," were not included.

Additionally, processing of future advanced-vehicle NDS data could benefit from refinement of the machine-learning process. The process requires the same amount of training and testing for one individual vehicle make and model as for multiple vehicles of the same make and model. Thus, the VCC L2 NDS study population, which originally comprised 20 additional vehicles excluded from the current study, was too diverse to complete machine-learning on all 50 vehicles. The vehicles that were excluded typically did not have easily distinguishable dash icons to represent ADAS status. It would, therefore, be beneficial for future studies to pay closer attention to, and be more selective of, the vehicles chosen to be included in the study design. Another issue that impacted the machine-learning process was glare on the dash, which interfered with the view of the ADAS icons. When glare obstructed the view of a small portion of the icon, ADAS status (i.e., active, available, or inactive) became unknown. Such portion of a trip was automatically excluded from the baseline epoch pool, meaning the available data pool became smaller. Unfortunately, this was not an easy problem to solve as the majority of driving was completed during daylight hours and the dashes of some of the vehicles (e.g., Tesla) seemed particularly susceptible to glare.

Finally, NDSs involve the participation of volunteer drivers; as such, these studies may involve self-selection bias.

Acknowledgments

This research program was sponsored by the AAAFTS. The authors would like to thank William Horrey from the AAAFTS for his support and guidance throughout the course of this study and to two anonymous reviewers for their helpful comments and suggestions. The authors wish to acknowledge the data reduction team at VTTI for its contributions in reducing the VCC L2 NDS and the L2 MFA NDS data. Special thanks is also given to Mindy Buchanan-King for technical writing and review of the document and to Zeb Bowden for his assistance with data mining and troubleshooting. The findings and conclusions of this paper are those of the authors and do not necessarily represent the views of the AAAFTS.

References

Abraham, H., McAnulty, H., Mehler, B., & Reimer, B. (2017). Case study of today's automotive dealerships: Introduction and delivery of advanced driver assistance systems. *Transportation Research Record*, *2660*, 7-14.

Benson, A. J., Tefft, B. C., Svancara, A. M., & Horrey, W. J. (2018). *Potential Reductions in Crashes, Injuries, and Deaths from Large-Scale Deployment of Advanced Driver Assistance Systems* (Research Brief). Washington, D.C.: AAA Foundation for Traffic Safety.

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., ... & Morgan, J. F. (2015). *Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts* (DOT HS 812 182). Washington, D.C.: National Highway Traffic Safety Administration.

Blanco, M., Atwood, J., Russell, S., Trimble, T., McClafferty, J., & Perez, M. (2016). *Automated Vehicle Crash Rate Comparison Using Naturalistic Data.* Blacksburg, VA: Virginia Tech Transportation Institute.

Boelhouwer, A., van den Beukel, A. P., van der Voort, M. C., & Martens, M. H. (2019). Should I take over? Does system knowledge help drivers in making take-over decisions while driving a partially automated car? *Transportation Research Part F*, *60*, 669-684.

Calkins, L. N. & Zlatoper, T. J. (2001). The effects of mandatory seat belt laws on motor vehicle fatalities in the United States. *Social Science Quarterly, 82*(4), 716-732.

Cicchino, J. B. (2017a). Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates. *Accident Analysis and Prevention, 99*, 142-152.

Cicchino, J. B. (2017b). *Effects of Lane Departure Warning on Police-Reported Crash Rates.* Arlington, VA: Insurance Institute for Highway Safety.

Cicchino, J. B. (2017c). *Effects of Blind Spot Monitoring Systems on Police-Reported Lane-Change Crashes.* Arlington, VA: Insurance Institute for Highway Safety.

de Winter, J. C. F., 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, 27,* 196-217.

Dingus, T., Klauer, S., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., . . . Knipling, R. (2006). *The 100-Car Naturalistic Driving Study, Phase II - Results of the 100-Car Field Experiment* (DOT HS 810 593). Washington, D.C.: National Highway Traffic Safety Administration.

Dingus, T. A., Hankey, J. M., Antin, J. F., Lee, S. E., Eichelberger, L., Stulce, K. E., . . . Stowe, L. (2015). *Naturalistic Driving Study: Technical Coordination and Quality Control* (SHRP 2 Report S2-S06-RW-1). Washington, D.C.: Transportation Research Board.

Dingus, T., Guo, F., Lee, S., Antin, J., 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 Science, 113*(10), 2636-2641.

Dingus, T. A., Owens, J. M., Guo, F., Fang, Y., Perez, M., McClafferty, J., Buchanan-King, M., & Fitch, G. M. (2019). The prevalence of and crash risk associated with primarily secondary tasks. *Safety Science, 119*, 98-105. https://doi.org/10.1016/j.ssci.2019.01.005

Fitch, G. M., Blanco, M., Camden, M., Olson, R., McClafferty, J., Morgan, J. F., . . . Hanowski, R. J. (2012). *Field Demonstration of Heavy Vehicle Camera/Video Imaging Systems: Final Report* (DTNH22-05-D-01019). Washington, D.C.: U.S. Department of Transportation.

Fitch, G. M., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., Olson, R. L., . . . Dingus, T. A. (2013). *The Impact of Hand-held and Hands-free Cell Phone Use on Driving Performance and Safety-Critical Event Risk* (DOT HS 811 757). Washington, D.C.: U.S. Department of Transportation.

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. *IEEE Access, 7*. Available from: https://www.researchgate.net/publication/321180547_MIT_Advanced_Vehicle_Technolog y_Study_Large-Scale_Naturalistic_Driving_Study_of_Driver_Behavior_and_Interaction_with_Automation

Fridman, L., Brown, D. E., Kindelsberger, J., Angell, L., Mehler, B., & Reimer, B. (2019). Human Side of Tesla Autopilot: Exploration of Functional Vigilance in Real-World Human-Machine Collaboration. Available from: https://www.researchgate.net/publication/332462137_Human_Side_of_Tesla_Autopilot_Exploration_of_Functional_Vigilance_in_Real-World_Human-Machine_Collaboration

Gimeno, P. T., Cerezuela, G. P., & Montanes, M. C. (2006). On the concept and measurement of driver drowsiness, fatigue and inattention: Implications for countermeasures. *International Journal of Vehicle Design, 42*, 67-86.

Guo, F. (2019). Statistical methods for naturalistic driving studies. *Annual Review of Statistics and Its Application, 6*(1), 309-328.

Guo, F. & Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. *Accident Analysis & Prevention*, *61*, 3-9.

Hanowski, R. J., Blanco, M., Nakata, A., Hickman, J. S., Schaudt, W. A., Fumero, M. C., . . . Madison, P. (2008). *The Drowsy Driver Warning System Field Operational Test: Data Collection Methods Final Report* (DOT HS 810 035). Washington, D.C.: U.S. Department of Transportation.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 770-778. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7780459&tag=1

Inagaki, T. & Itoh, M. (2013). Human's overtrust in and overreliance on advanced driver assistance systems: A theoretical framework. *International Journal of Vehicular Technology, 2013,* 1-8.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data* (DOT-HS-810-594). Washington, D.C.: National Highway Traffic Safety Administration.

Klauer, S. G., Guo, F., Sudweeks, J., & Dingus, T. A. (2010). *An Analysis of Driver Inattention Using a Case-Crossover Approach on 100-Car Data: Final Report* (DOT HS 811 334). Washington, D.C.: National Highway Traffic Safety Administration.

Leonard, S. D. (2001). Relation of Owner's Manuals to Safety. *Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 125-130). Iowa City, IA: Public Policy Center.

Lin, R., Ma, L., & Zhang, W. (2018). An interview study exploring Tesla drivers' behavioral adaptation. *Applied Ergonomics, 72,* 37-47.

Llaneras, R. E., Salinger, J., & Green, C. A. (2013). Human factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to the forward roadway. *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.* Iowa City, IA: University of Iowa. https://ir.uiowa.edu/drivingassessment/2013/papers/15/

Matthews, G., Saxby, D., Warm, J. S., & Hitchcock, E. (2009). Driving fatigue as an impairment in effort-regulation. *Proceedings of 2009 International Conference on Fatigue Management in Transportation Operations.* Boston, MA.

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

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual: Driver behavior when resuming control from a highly automated vehicle. *Transportation Research Part F, 27*, 274-282.

Motor1.com (2019). *Driver Assistance Systems: The Ultimate Guide to Car Safety Tech.* https://www.motor1.com/features/346112/driver-assistance-systems-guide/

Nachreiner, F. (1995). Standards for ergonomics principles relating to the design of work systems and to mental workload. *Applied Ergonomics, 26*, 259-263.

OECD (1990). *Behavioural Adaptations to Changes in the Road Transport System*. Paris, France: Organisation for Economic Co-operation and Development, Road Transport Research.

Perel, M. (1983). *Vehicle Familiarity and Safety* (DOT HS-S06 509). Washington, D.C.: National Highway Traffic Safety Administration.

Peterson, S.P. & Hoffer, G.E. (1994). The impact of airbag adoption on relative personal injury and absolute collision insurance claims. *Journal of Consumer Research, 20,* 657-662.

Rudin-Brown, C. M., & Parker, H. A. (2004). Behavioral adaptation to adaptive cruise control (ACC): Implications for preventive strategies. *Transportation Research Part F, 7*, 59-76.

Russell, S. M., Blanco, M., Atwood, J., Schaudt, W. A., Fitchett, V. L., & Tidwell, S. (2018). *Naturalistic Study of Level 2 Driving Automation Functions* (DOT HS 812 642). Washington, D.C.: National Highway Traffic Safety Administration.

Saad, F., Hjälmdahl, M., Cañas, J., Alonso, M., Garayo, P., Macchi, L., Nathan, F. et al., (2004). *Literature Review of Behavioural Effects* (AIDE report IST-1-507674-IP). Gothenburg, Sweden: Adaptive Integrated Driver-Vehicle Interface.

SAE International (2018). *Surface Vehicle Recommended Practice: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles* (J3016). Warrendale, PA: Society of Automotive Engineers.

Schoettle, B., & Sivak, M. (2015). *A Preliminary Analysis of Real-World Crashes Involving - Self-Driving Vehicles*. Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Simons-Morton, B. G., Ouimet, M. C., Zhang, Z., Klauer, S. G., & Lee, S. E. (2011). The effect of passengers and risk-taking friends on risky driving and crashes/near crashes among novice teenagers. *Journal of Adolescent Health*, *49*(6), 587-593.

Smithson, M. & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, *11*(1), 54-71.

Strand, N., Nilsson, J., Karlsson, I. C. M., & Nilsson, L. (2014). Semi-automated versus highly automated driving in critical situations caused by automation failures. *Transportation Research Part F, 27,* 218-228.

Sullivan, J. M., Flannagan, M. J., Pradhan, A. K., & Bao, S. (2016). *Literature Review of Behavioral Adaptation to Advanced Driver Assistance Systems*. Washington, D.C.: AAA Foundation for Traffic Safety.

Vaa, T. (2013). Psychology of behavioral adaptation. In C. Rudin-Brown & S. L. Jamson (Eds.) *Behavioral Adaptation and Road Safety: Theory, Evidence and Action*. Boca Raton, FL: CRC Press.

VTTI (2015). *Researcher Dictionary for Safety Critical Event Video Reduction Data* (v4.1). Blacksburg, VA: Virginia Tech Transportation Institute. http://hdl.handle.net/10919/56719

Young, M. S. & Stanton, N. A. (2002). Attention and automation: New perspectives on mental underload and performance. *Theoretical Issues in Ergonomics Science, 3*, 178-194.

Young, M. S. & Stanton, N. A. (2004). Taking the load off: Investigations of how adaptive cruise control affects mental workload. *Ergonomics, 47*, 1014-1035.

Young, M. S. & Stanton, N. A. (2006). In loco intellegentia: Human factors for the future European train driver. *International Journal of Industrial and Systems Engineering, 1*, 485-501.