Investigating Driver Experience and Augmented Reality Head-Up Displays in Autonomous Vehicles

by

Murat Dikmen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2017

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contribution

The material presented in Section 2.2 was published in the following article:

Dikmen, M., & Burns, C. M. (2016, October). Autonomous Driving in the Real World: Experiences with Tesla Autopilot and Summon. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 225-228).

Author contributions:

Study design – Dikmen, M. (80%), Burns, C. M. (20%)

Data collection and analysis – Dikmen, M. (100%)

Writing – Dikmen, M. (70%), Burns, C. M. (30%)

Abstract

Autonomous driving is on the horizon. Partially automated vehicles recently started to emerge in the market, and companies are dedicated to bringing more automated driving capabilities to the vehicles in the near future. Over the past twenty years, human factors research has increased our understanding of driver behavior and human-vehicle interaction, as well as human-automation interaction considerably. However, as the technological developments accelerate, there is an urgent need to conduct research to understand the challenges of driving a semi-automated vehicle, the role of cognitive and social factors and driver characteristics, and how interactive technology can be used to increase driving safety in this context. This thesis was an attempt to address some of these challenges. In this work, we present two studies on human factors of automated driving. In the first study, we present the results of a survey conducted with Tesla drivers who have been using partially automated driving features of Tesla cars. Our results revealed that current users of this technology are early adopters. Automation failures were common, but drivers were comfortable in dealing with these situations. Additionally, Tesla drivers have high levels of trust in the automated driving capability of their vehicles, and their trust increases as they experience these features more. The results also revealed that drivers don't use owner manuals, and seek out information about their cars by using online sources. The majority of Tesla drivers check multiple information sources when their car software receives an update. Overall these findings show that driver needs are changing as the vehicles become smarter and connected. In the second study, we focused on a future technology, augmented reality head-up displays, and explored how this technology can fit into the smart, connected and autonomous vehicle context. Specifically, we conducted an experiment looking into how these displays can be used to monitor the status of automation in automated driving. Participants watched driving videos enhanced with augmented reality cues. Results showed that drivers adjust their trust in the automated vehicle better when information about the vehicle's sensing capabilities are presented using augmented reality cues, and they have positive attitudes towards these systems. However, there were no major safety-related benefits associated with using these displays. Overall, this work provides several contributions to the knowledge about human-automation interaction in automated driving.

Acknowledgements

First, I would like to express my deepest gratitude to Professor Catherine Burns for providing guidance and support during this process. Her continuous support helped me to overcome the difficulties along the way and become passionate about this work. With her guidance, I always felt confident that I was on the right track. I am very grateful to have a chance to write this thesis under her supervision.

I am also thankful to my friends in Advanced Interface Design Lab, who always inspired me and who have always been a source of joy.

Special thanks to my wife, Neriman Saner. Her companionship provided me the motivation I needed throughout my journey. I would also like to thank my parents. Without their support, this work would not be possible.

Finally, I would like to thank Professor John McPhee and Professor Shi Cao for being my readers. I am thankful for their valuable comments and feedback on this work.

Dedication

I dedicate this thesis to my family who have always supported me throughout my journey.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

Vehicles with advanced automation systems have started to emerge in the market. Currently, more than 30 companies are involved in building advanced driving automation systems and self-driving cars (CB Insights, 2016). As these technologies develop, the nature of driving starts to change fundamentally. With more vehicles becoming smarter and automated, the role of the driver will shift from an active driver to a passive driver and eventually to a passenger. Vehicle automation has been around for some time; however, these technologies were mostly used and tested for research purposes and prototype forms. Over time the technology matured, and recently, commercial systems started appearing in the market (advanced driver assistance systems), ranging from navigational aids to adaptive cruise control with the goal of improving driving experience, and making driving easier and safer. The next step in this evolution is to make the vehicles automated, and eventually replace the driver.

With the advances in automation, human factors research has naturally started examining and evaluating human-automation interaction, the unique challenges automation brings, and opportunities to reduce human error and increase human performance when working with automated systems. Although there has been considerable research on human-automation interaction and driver behavior in automated vehicles, investigation of real world usage of these systems was limited. Additionally, the challenges identified in the past have not been fully addressed yet. Given the rapid advancements in automated driving technology, there is a need to increase these efforts and address human factors challenges of automated driving before these technologies become widely available, to ensure that adoption and use of these systems will be safe and enjoyable.

This thesis attempts to fill this gap and extend our current understanding of driver-automation interaction by presenting two studies we conducted, one survey and one laboratory experiment, to understand how automated driving systems are used in real world and how we can support drivers in this context.

1.1 Motivation

This work was motivated by recent developments in the automotive industry and automated vehicle technology as cars with autonomous driving capabilities started to emerge in the market. While advanced driver assistance systems such as adaptive cruise control, lane keeping assistant, and blind spot monitors have been available for some time (Brookhuis, De Waard, & Janssen, 2001), recently, the combination of these technologies, primarily adaptive cruise control and lane keeping systems allowed the initial step towards autonomous vehicles. Technology is developing rapidly, and manufacturers are adding new features and capabilities to their vehicles. For example, Tesla, in addition to the combination of adaptive cruise control and steering assistance automation that allows hands-free driving, introduced a lane change assistance which allows the car to move to another lane upon the request of the driver. It handles the monitoring task (i.e. whether the lane is available) and speed adjustments along with steering. This and other developments (e.g. automatic overtaking; Milanés et al., 2012) will gradually bring the vehicles closer and closer to become fully automated.

During this transition period, a critical question remains. What will happen to the human driver? How will the human driver deal with the demands of partially automated vehicles and how will they adapt? Although there have been research efforts to understand and deal with problems in automated driving, the use of these technologies beyond laboratories, especially in partially automated vehicles (i.e. level 2 automation, SAE International, 2014**)** just recently started to emerge. Therefore, our aim was to (1) identify challenges of automated driving in real world, (2) explore ways to support drivers through design to adapt to this new situation. To this end, we had several goals in this work:

- Investigate how automated driving is used in the real world
- Identify challenges and opportunities
- Design and test technology to address these challenges

In this research, we first conducted a survey with drivers who are using automated driving features. This work revealed several challenges and opportunities. Next, we conducted an experimental study to test an automation display to address the challenges identified in the survey.

1.2 Structure of the Thesis

This thesis is structured as follows:

In Chapter 1 **-** Introduction, we present an introduction to the thesis, discuss the motivation behind this work, and provide background information about vehicle automation.

In Chapter 2 – Autonomous Driving in the Real World, we present our first study, a survey conducted with Tesla drivers about their experiences with an automated driving system. This chapter is organized in three subsections, with each subsection containing a thematically different analysis. This chapter also features a published paper (section 2.2).

In Chapter 3 – Augmented Reality Displays in Semi-Autonomous Vehicles, we present our second study, a laboratory experiment examining augmented reality head-up displays in automated driving context.

In Chapter 4 - Conclusion, we discuss implications of this work and provide future directions.

Chapters 2 and 3 are self-contained such that relevant background information is presented within the chapters.

1.3 Background on Autonomous Vehicles

1.3.1 The concept of Autonomy

Automation is defined as a system that handles tasks that were previously carried out by humans (Parasuraman & Riley, 1997). Automation is being used in virtually all areas of life, and has many advantages such as handling tasks that are very difficult for humans, and increase safety and efficiency. Vehicle automation, likewise, has potential in increasing road safety, decreasing accidents and overall improve driver conditions (Stanton & Marsden, 1996). Many vehicle automation systems have been developed in the past such as cruise control, adaptive cruise control, lane departure warnings, blind spot monitors, and navigational aids. Stanton and Young (1998) differentiate between two types of vehicle automation: systems that support the driver and systems that replace the driver. Examples of the former type are parking sensors, traffic guidance and blind spot monitors. This type of automation enhances drivers' sensing and decision-making capabilities while not affecting the driving task in significant ways. The latter category includes systems that fundamentally change the driving task. Examples of these systems are cruise control, adaptive cruise control and steering assistance systems. These systems execute some of the primary aspects of driving task such as speed adjustments and steering, and drastically change driver behavior (Young & Stanton, 2007).

Recently, the combination of steering automation and adaptive cruise control allowed vehicles to handle both speed (longitudinal) and steering (lateral) related tasks. Using these systems, the vehicle can stay in the lane, and adjust its speed based on the vehicles in front. This allows hands-free driving under certain circumstances (e.g. highway). However, this is just the beginning of the progress towards fully automated vehicles. Vehicle automation will improve significantly in the near future through advancements such as vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, and developments in sensors and artificial intelligence. These developments will allow the cars to sense the environment more accurately and make better decisions, which are essential for safe driving. Gradually, more driving functions will be automated that were previously handled by human drivers.

The technology that allows automated driving such as connectivity and artificial intelligence will at the same time make the cars smarter. Vehicles of the future will not only feature more autonomous capabilities, but they will become personal companions who understand and support drivers in a number of ways such as communicating with home automation, integrating with personal devices such as smartphones, and providing a smooth and personalized driving experience.

1.3.2 Levels of Autonomy

A key concept when discussing automation is the level of autonomy and the degree of automation. The primary reason behind thinking of automation in terms of levels is that the demands, expectations, and needs for humans and automated systems can drastically differ between different levels of autonomy.

Several taxonomies and levels of automation have been proposed in the past (Endsley, 1999; Parasuraman, Sheridan, & Wickens 2000). The levels usually start with no automation, i.e. human handles all tasks, and end with full automation, i.e. automation handles all tasks without the need for humans. In-between levels allocate functions to humans and automation, with increasingly to automation as the levels increase. For example, Parasuraman et al. (2000), in their 10 levels of automation, describe function allocation in lower levels of automation as: the automation presents action choices (level 2), narrow the set of options (level 3), recommend one action (level4) while in higher levels, the automation executes action and informs the human (level 7), informs the human only upon request (level 8), decides whether or not to inform the human (level 9) and simply ignoring human (level 10). The different levels of automation have varying effects on human performance (Onnasch, Wickens, Li, & Manzey, 2014).

In vehicle automation, the most commonly used taxonomy is developed by Society for Automotive Engineers (SAE) which features six levels of vehicle automation (SAE International, 2014), as shown in Figure 1. These standards have also been adopted by U.S. Department of Transportation (National Highway Traffic Safety Administration, 2016). Given these levels, adaptive cruise control would be considered as level 1, while the combination of adaptive cruise control and lane keeping assistance would be considered as level 2. A key difference between levels 0-2 and levels 3-5 is the agent responsible for monitoring the environment. As the level of automation increases, the monitoring task shifts from human (levels 0,1 and 2) to the system (levels 3,4 and 5). Currently, most advanced vehicles in the market are at level 2, combining multiple functions yet still requiring constant human monitoring. The change from level 2 (partial automation) to level 3 (conditional automation) will require substantial capability from the automation as the sensing systems should be very accurate.

Also, as shown in Figure 1, humans will be responsible for fallback performance until level 4 automation. However, we should note that while level 4 eliminates the need for human control, we cannot assume that drivers will be able to stay in level 4 at all times. For example, while level 4 automation might be suitable for most environments, drivers may still need to switch to lower levels of automation under circumstances where level 4 will not be available. In short, until the vehicle automation reaches level 5 (i.e. human performance is not needed under any circumstance), there will be a need for human driver's capabilities.

SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/ Deceleration	*Monitoring* of Driving Environment	Fallback Performance of *Dynamic Driving Task*	System Capability *(Driving Modes)*
Human driver **monitors the driving environment**						
0	**No Automation**	the full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	**Driver Assistance**	the *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	Human driver and system	Human driver	Human driver	Some driving modes
2	**Partial Automation**	the *driving mode*-specific execution by one or more driver assistance systems of both steering and acceleration/ deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	**System**	Human driver	Human driver	Some driving modes
Automated driving system **("system") monitors the driving environment**						
3	**Conditional Automation**	the *driving mode*-specific performance by an *automated driving system* of all aspects of the dynamic driving task with the expectation that the *human driver* will respond appropriately to a *request to intervene*	System	**System**	Human driver	Some driving modes

4	**High Automation**	the *driving mode*-specific performance by an automated driving system of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*	System	System	**System**	Some driving modes
5	**Full Automation**	the full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	System	**All driving modes**

Figure 1. Six levels of driving automation (SAE International, 2014).

1.3.3 Challenges of Autonomy

One of the main challenges of level 2 automation is that the role of the driver will shift from an active driver to a passive one. Previously, drivers assumed the role of the active driver, handling all driving-related tasks manually. With the introduction of vehicle automation, increasingly more of these tasks will be allocated to the vehicle. The driver, free from the manual driving task, needs to monitor the vehicle and the roadway to make sure that the automation handles these tasks well. If the automation fails, the driver needs to take control timely and revert to manual driving mode. This situation makes the driver “part driver and part passenger” (Casner, Hutchins, & Norman, 2016, p .71). The challenge is whether the drivers will be able to assume this new role properly and timely respond to automation failures. Currently, due to the high failure rates (Dikmen & Burns, 2016; Larsson, 2012), drivers are mostly engaged with the driving task as frequent automation failures are likely keeping the drivers alert and in-the-loop by frequently requesting them to take back the control of the vehicle. The problem starts when the vehicle automation becomes increasingly reliable to the point at which that drivers completely trust and rely on them, leading to automation complacency (Parasuraman & Manzey, 2010). With the comfort of reliable vehicle automation, people will eventually start engaging in a range of activities on a ride such as texting and reading using mobile and wearable technology (De Winter, Happee, Martens, & Stanton, 2014). These activities are detrimental in manual driving and have a direct impact on driving performance (Regan, Lee, & Young, 2008). In automated vehicles, the effects of distractors will be indirect by reducing attention and awareness of the environment, known as the out-of-the-loop situation (Endsley & Kiris, 1995). This may not be an issue so long as the automation is reliable. However, failures will happen. If such situations occur, the driver, who is completely out-of-the-loop, may not be able to handle take-over requests appropriately. A recent fatal Tesla crash provides evidence for these concerns (Golson, 2017). In this accident, the Tesla car was on Autopilot (automated driving mode) and the driver had seven seconds to react to a tractor trailer driving across the highway, yet both the vehicle and the driver failed to react appropriately. This example is considered as the first fatal autonomous car accident, and certainly will not be the last.

Deskilling is another concern in the context of vehicle automation (Stanton, & Marsden, 1997). While little is known about how driving automation will influence driving skills, it is likely that continuous use of automated vehicles may result in degradations in manual driving skills such as reduced reaction speed to hazards. Interestingly, a recent survey found no support for deskilling in driving (Trösterer et al., 2016). The authors concluded that the skilling (i.e. initial training), is more critical than deskilling. We should note that the context of this survey was not specifically automated driving. Regardless, if the initial driving training matters more, this still poses a challenge in automated driving. As the automated driving becomes widely available, novice drivers may rely on these technologies significantly, which in return may hinder proper skill development in manual driving.

To sum up, while vehicle automation has many advantages, it poses some challenges which should be addressed in a timely manner. During the transition from semi-autonomous vehicles to full autonomous vehicles, driver disengagement, loss of awareness, and possibly deskilling are issues that needs to be well understood, and systems should be developed to solve these issues. Although there have been efforts in achieving this goal, given the rapid evolution of technology, researchers, automotive industry, and regulators need to address the unsolved problems and new challenges that emerge in autonomous transportation. We hope this work will contribute to achieving this goal.

Chapter 2
Autonomous Driving in the Real World

2.1 Chapter Introduction

In this chapter, we will present the results of a survey conducted with Tesla drivers about their experiences with an automated driving system (Autopilot) and an automated parking system (Summon). We present the results in three sections to facilitate reader understanding. First, in section 2.2, we will present our findings on the frequency of use, attitudes towards these technologies, and surprises and unexpected situations drivers experienced when using these features. Then, in section 2.3, we will present the results on the use of information sources. Next, we will discuss the findings on trust in Autopilot and Summon in section 2.4. Survey questions we refer to in the following sections can be found in Appendix A.

2.2 Autonomous Driving in the Real World: Experiences with Tesla Autopilot and Summon

2.2.1 Study Overview

As autonomous driving emerges, it is important to understand drivers' experiences with autonomous cars. We report the results of an online survey with Tesla owners using two autonomous driving features, Autopilot and Summon. We found that current users of these features have significant driving experience, high self-rated computer expertise and care about how automation works. Surprisingly, although automation failures are extremely common they were not perceived as risky. The most commonly occurring failures included the failure to detect lanes and uncomfortable speed changes of the vehicle. Additionally, a majority of the drivers emphasized the importance of being alert while driving with autonomous features and aware of the limitations of the current technology. Our main contribution is to provide a picture of attitudes and experiences towards semi-autonomous driving, revealing that some drivers adopting these features may not perceive autonomous driving as risky, even in an environment with regular automation failures.

2.2.2 Study Introduction

Autonomous driving is on the horizon and, in some cases, semi-autonomous features are now available on some models and types of vehicles. As an example of some of the most advanced features currently available, Tesla released its Autopilot and Summon features in October 2015 and January 2016, respectively. Autopilot is a system which provides lateral and longitudinal control and allows hands-free driving, in addition to other functionality such as automatic lane changing. Summon is a parking assistance system which allows drivers to park their cars from outside the vehicle (Tesla Motors, 2016).

The release of these features allow for real world discussions of how people interact with these early autonomous features and how they are influencing driver perceptions and attitudes. Research has raised concerns regarding automated driving such as overreliance (de Waard, van der Hulst, Hoedemaeker, & Brookhuis, 1999), reduced situational awareness (Stanton & Young, 2005; De Winter, Happee, Martens, & Stanton, 2014) and increased engagement with secondary tasks, diverting attention away from the road (Carsten, Lai, Barnard, Jamson, & Merat, 2012; Llaneras, Salinger, & Green, 2013). Given these concerns are largely from laboratory research, it is important to understand whether such concerns are reflected in real world autonomous driving.

2.2.3 Related Work

Many surveys have been conducted in the past to understand people's attitudes towards autonomous cars. Previous work showed that people are attracted to safety and convenience of self-driving cars but were concerned with the lack of control, liability, and cost (Howard & Dai, 2014). The majority of people also have a priori acceptance of autonomous cars (Payre, Cestac, & Delhomme, 2014), yet opinions can be split (Bazilinskyy, Kyriakidis & de Winter, 2015). A recent survey found that majority of people had positive attitudes towards autonomous cars but were concerned with aspects such as security and legal issues (Kyriakidis, Happee, de Winter, 2015). Similarly, another study found that most people had positive opinions about autonomous vehicles while expressing concerns regarding safety (Schoettle & Sivak, 2014). A weakness of these studies, however, is that they were unable to study the attitudes of people who had real world experience with autonomous driving.

In one study of real-life use of autonomous vehicles, Larsson (2012) reported that adaptive cruise control (ACC) users experience frequent limitations of the system and the more they drive with ACC the more they become aware of the system limitations. The same survey also revealed that drivers experience mode errors and concludes that imperfect ACC may be better for driving safety because it keeps the drivers in the loop.

Our research extends these findings by looking at experiences with the next generation of semi-autonomous driving features which combine ACC with steering assistance. We wanted to understand how often drivers use these features, how often do they experience failures, and how does experience with these automation failures influence their attitudes towards the automation.

2.2.4 Method

We conducted an online survey with 162 Tesla Owners. The survey was distributed through online forums and social media during April-May 2016. The survey asked questions about drivers' attitudes towards and experiences with two functionalities built into Tesla Model S cars: Autopilot and Summon. Questions covered frequency of use, satisfaction, ease of learning and knowledge related to Autopilot and Summon. Additionally, we asked participants to report unusual or unexpected behaviors they experienced while using these systems and what they consider a key aspect of safety. The average time to complete the survey was 9.6 minutes.

2.2.5 Results

A total of 121 participants completed the survey fully. The demographics of the sample is summarized in Table 1. The sample was 94.2% male, and had significant driving experience with 89.3% reporting driving experience beyond 10 years. These drivers drive frequently with 79.3% reporting that they drive daily. Participants identified themselves mostly as above average or expert computer users. All means reported in the subsequent analysis correspond to 5-point Likert scales where 5 is high and 1 is low.

Participants reported very high levels of satisfaction with their cars (M = 4.91, SD = .43). Means and standard deviations for self-rated knowledge, ease of learning and importance of knowing how automation makes decisions are shown in Table 2. To summarize, participants reported that it is easy to learn the automated systems, they rated their knowledge level as above average, and importance of knowing how automation makes decisions as above average. In addition, the Autopilot display, the display on the dashboard showing information about the current state of Autopilot such as the detected vehicles on the roadway, was perceived as useful.

Age	**% (N = 121)**
16 - 20	3.3

Age	% (N = 121)
21 - 24	2.5
25 - 34	18.2
35 - 44	25.6
45 - 54	23.1
55 - 64	14.0
65 or older	13.2
Computer Expertise	**% (N = 121)**
Novice	.8
Average	5.0
Above average	38.8
Expert	55.4

Table 1. The demographics of the sample.

90.1% of the participants reported that they actively use Autopilot or have used it in the past. Likewise, 85.2% of the participants reported that they actively use the Summon feature or have used it in the past.

Participants use Autopilot quite frequently with 31.2% saying they use it “always” and 57.8% saying they use it “often”. Participants use Summon less frequently with 49% saying they use it “rarely” and 22% saying “sometimes”.

	Autopilot		Summon	
	Mean	SD	Mean	SD
Knowledge	3.79	.82	3.54	.92
Ease of Learning	4.27	.72	3.97	.84
Importance	3.51	1.08	3.13	1.15

	Autopilot		Summon	
Usefulness of Autopilot Display	4.04	.71	-	-

Table 2. Descriptive statistics for self-rated knowledge, perceived ease of learning, importance of knowing how the system makes decisions, and usefulness of Autopilot display.

2.2.5.1 Automation Limitations and Failures

Of the Autopilot users, 62.4% reported that they have experienced at least one unexpected or unusual behavior from the car while in autonomous driving mode. Further, 13.8% reported that they have experienced at least two unexpected or unusual behaviors from the car. In total, participants reported 91 cases of automation events. Of the Summon users, 21.2% reported that they have experienced at least one unexpected or unusual behavior from the car while using the system. In total, participants reported 27 cases. Perceived risk involved in these events are shown in Figure 2 for Autopilot and Figure 3 for Summon.

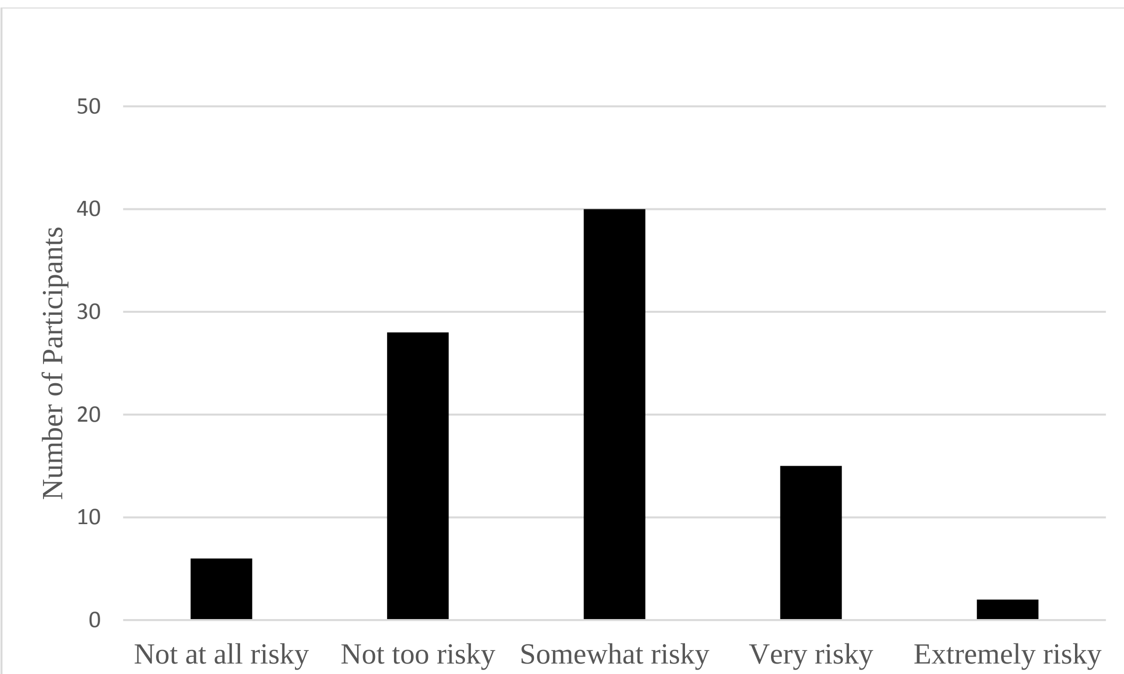

Figure 2. Perceived risk after experiencing an Autopilot failure.

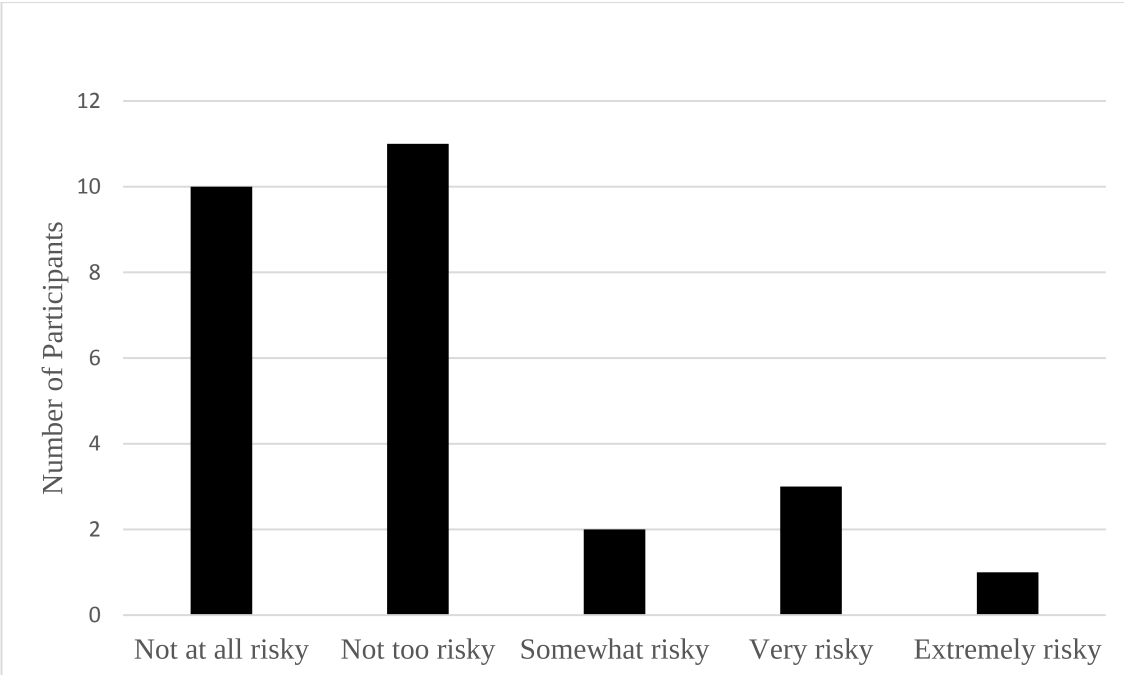

Figure 3. Perceived risk after experiencing a Summon failure.

2.2.5.2 Cases of Unexpected Automation Behaviors

Next, we analyzed the reported cases of unexpected automation behavior. For Autopilot, of the 91 cases analyzed, two major categories of limitations emerged. The first category involved issues with lane detection (74.4% of the cases). These problems included the car trying to take an exit ramp, swerving and veering due to failure to detect the lane, and trying to cross lanes for no apparent reason, sometimes even towards lanes where traffic flowed in the opposite direction. The second category involved problems with speed changes and the adaptive cruise control system. This category includes issues such as sudden braking or uncomfortable acceleration and deceleration (15.6% of the cases). Participants reported that speed related problems mostly occurred in the heavy traffic conditions. Almost all users reported that they took manual control over after the incident and most reported that they re-initiated autonomous driving once the situation that caused automation failure was over. In the majority of the 27 Summon cases, participants reported technical problems such as connection failures between the vehicle and the phone.

2.2.5.3 Statistical Results

There were no differences between age groups in the measured variables, indicated by non-significant ANOVAs. There were also no differences between those who had an Autopilot failure (N=68) and those who did not (N=45) in measured variables, indicated by non-significant t-tests.

Perceived usefulness of Autopilot display was significantly correlated with satisfaction with the car, $r = .22$, $p = .019$, and the ease of learning, $r = .23$, $p = .017$, hinting at the possible contribution of the visual display to the learning process of Autopilot. It was also correlated with importance of knowing how Autopilot makes decisions, $r = .21$, $p = .031$. As expected, the Autopilot display can be used as a means to understand the decision-making process of the car and to obtain situation awareness.

For those who had an Autopilot failure (N=68), perceived risk of the situation was correlated only with importance of knowing how Autopilot makes decisions, $r = .24$, $p = .053$.

2.2.5.4 Safe Driving

Participants emphasized being alert at all times, paying attention to the road environment and keeping hands on the wheel while in autonomous driving mode. They also emphasized the importance of learning the limitations of the technology such as under which conditions the automation can fail. A critical question here is how drivers can learn the specific conditions in which automation is more likely to fail without trial and error? Or should trial and error be part of the learning process, as some participants suggested? We believe addressing this issue requires further research.

2.2.6 Discussion

Based on the results, at first glance, the situation of semi-autonomous driving seems generally positive. Drivers seem to enjoy these technologies, and are aware of the limitations of Autopilot and Summon. In the comments, we observed that drivers were highly motivated to use these technologies safely and have not seen indications of the concerns raised in the past such as engaging with secondary tasks while using Autopilot.

Despite the relatively high frequencies of automation events, these drivers did not consider the automation to be particularly risky. We believe three factors might have contributed to this. First, even though the situations were unexpected, these users were aware that these are new technologies in early release, so they were quite accepting of events with the technology. Second, Tesla owners are unlikely to be representative of general drivers. Tesla drivers are early adopters with high comfort with technology, and are unusually devoted to the development of their vehicles. Third, none of the incidents reported involved a negative outcome, which may also be influencing their perception of risk. Relatively frequent exposure to small events may also be teaching these drivers to stay “in the loop” with the automation.

However, failure rates will decrease eventually and this may trigger different observations of driver performance. In almost all cases covered in the survey, participants reported that they successfully took control and drove manually until in a safe situation again. However, this may not happen always as studies show possible decrements in situational awareness during autonomous driving (Stanton & Young, 1998; 2005). While the argument can be made that imperfect automation will keep the drivers in the loop (Larsson, 2012), it is unreasonable to think that automation will deliberately remain imperfect. Over time, autonomous features will increase in reliability and functionality and this, unfortunately, does present a risk for a lack of situation awareness by drivers who are increasingly “out of the loop”. Further, the drivers in this study were well experienced and very comfortable with technology and may have responded more confidently when experiencing these failures.

Based on the incidents reported, currently, lane keeping is an important issue, especially in situations where lane markings are missing, or the car cannot correctly identify obstacles on the road environment. For the parking system, Summon, the most commonly experienced problem was the operation stopping due to a technical failure such as a connection problem between the phone and the vehicle. An interesting point is that with the rise of semi-autonomous driving, the role of the driver shifts from the active driver to a supervisory role (Banks & Stanton, 2014). This new role can place demands of different nature on the driver. For example, in addition to monitoring the road environment similar to manual driving, the driver also has to monitor whether lane markings are clear or not, or more importantly, whether the car can correctly identify the lane. This and other limitations of the automation might not be always obvious; therefore the communication between automation and the driver becomes crucial in order to maintain situation awareness. The correlations between perceived usefulness of the Autopilot display and ease of learning and importance of knowing how Autopilot makes decisions also indicate the importance of driver-vehicle communication in the autonomous driving context. Further research should address these issues by studying the role of automation displays in obtaining situation awareness in autonomous vehicles.

A limitation of the current study is that our sample is not representative of the general driver population. Considering the computer expertise and knowledge level about autonomous driving functionality, our participants are likely early adopters. Therefore, we must be cautious generalizing our findings. While the focus of this study was on two particular systems, Tesla Autopilot and Summon, we believe the results obtained and issues revealed are applicable to other systems as well.

2.2.7 Conclusion

In this study, we examined the current state of semi-autonomous driving in the real world. Our survey data showed that current users of autonomous driving features of Tesla cars use Autopilot frequently, they are knowledgeable about automation and they find it easy to learn. The frequency of automation failure rate was high; however, most participants did not perceive these incidents as posing a significant risk. Our main contribution is to provide insights into the real world phenomenon of autonomous driving in its early stages, as first generation technology becomes available in the market.

2.3 Use of Information Sources

2.3.1 Overview

Tesla can deliver software updates to the cars over-the-air (Software Updates, 2016) and these updates can have varying degrees of impact on vehicle functionality. They can range from minor small user interface modifications (e.g. changing the color of an object on the in-vehicle display) to major functionality changes, such as enabling automated driving.

In this section, we will present additional findings from the survey we introduced in the previous section (2.2). Specifically, we will present findings on how Tesla owners and non-owners use information sources when they want to learn about the features of their cars, how they access the owner's manual when they need it, and how Tesla drivers learn about the new features of their cars after a software update. We will first discuss why such an analysis is relevant, and then present findings from the survey.

2.3.2 Background

As the vehicles become smarter, connected, and automated, driving experience also evolves significantly. Technologies such as vehicle-to-vehicle communication, vehicle-to-cloud communication and artificial intelligence not only enable autonomous driving capabilities (Koehler, Appel, & Beck, 2016), but also transform the vehicles from static mechanical products to constantly evolving digital products. Companies are already offering connected vehicle services such as streaming services, smartphone connectivity, and smart home integration (Viereckl, Koster, Hirsh, Ahlemann, 2016).

A critical part of this concept is the over-the-air updates: software updates delivered to the vehicle over the internet. These updates not only keep vehicle software secure and up to date, but also allow manufacturers to add new features and functionality to the car. These features can be safety-related (e.g. Tesla Autopilot), or utility and entertainment related (e.g. smartphone-like apps). An important consideration for the success of this upgradeable car concept is to identify user needs, habits, and expectations regarding this new vehicle experience. The following analysis is a first step towards achieving this goal.

While updating software on personal computers, smartphones and consumer devices is a common activity, updating a car is not. There are a few issues associated with upgradeable cars that raise concerns. First, installing new software into the car can result in software malfunctions that can lead to potentially dangerous situations, if these malfunctions occur in safety-critical systems of the vehicle. Second, connectivity raises concerns about security (Greenberg, 2015; Hubaux, Capkun, & Luo, 2004). Third, installing new features and applications can lead to changes in driving behavior as drivers adapt. For example, if the visual layout of the dashboard changes, drivers may need to spend more time when they want to look up information until they are comfortable with the new layout. This can lead to distracted driving (Young, Lee, & Regan, 2008) which is a major concern for driving safety (National Center for Statistics and Analysis, 2016).

When updating software, several factors influence users' decision-making process (Mathur, 2016) such as the type of update (e.g. security vs. new functionality), change logs and trust in the company. Additionally, users go through several stages during an update process such as awareness, deciding, preparation, installation, troubleshooting and post-state (Vaniea, & Rashidi, 2016). Hesitation to apply the updates is common, which results in users researching the features of the update and how their systems will be affected to overcome this hesitation, especially when don't know what the updates will do (Fagan, Khan, & Buck, 2015). Therefore, it is important for users to obtain accurate and useful information to understand the features of the update and form correct mental models, as this will reduce confusion and annoyance regarding the update process (Fagan, Khan, & Nguyen, 2015).

An important issue is how to deliver necessary information to the users regarding vehicle updates. To understand this issue, the following analysis presents information sources used by drivers both in the context of updating car software and the use of information sources in general, including accessing the owner's manual.

2.3.3 Method

We used the same method as described in 2.2.4. In addition to 162 Tesla drivers, the following analysis also presents data from 116 drivers who don't own a Tesla car but participated in the survey. This allowed us to compare Tesla owners with non-owners to better understand how driving an upgradeable vehicle affects drivers' information seeking behavior. In the survey, we asked participants questions about which information sources they use to learn about the features of their cars and how they access owner's manual when they need it. Additionally, we asked Tesla drivers about how they get information about the features of Autopilot and Summon updates.

2.3.4 Results

121 Tesla owners and 101 non-owners completed the survey fully. 96.4% of the participants were male. 49.6% of Tesla owners and 86.1% of non-owners were 34 years or younger. In terms of driving experience, 89.3% of Tesla owners and 42.6% of non-owners reported having more than 10 years of experience. Overall, Tesla owners were older and had more driving experience than non-owners. In the following analysis, for information sources used to learn about the feature of the car and accessing the car manual, we present data from both Tesla owners and non-owners. For information sources regarding the updates, we present data only from Tesla owners.

2.3.4.1 Use of Information Sources to Learn about the Features of the Car

We asked participants about how frequently they consult owner's manual, friends/colleagues, and online resources when they need information about the features of their cars, on a 5-point scale ranging from never to always. Figure 4 shows the mean scores for Tesla owners (N =121) and non-owners (N = 100). The trend was similar for Tesla owners and non-owners. A 2 (ownership) x 3 (source type) repeated-measures ANOVA revealed a main effect of ownership, $F(1, 218) = 20.90$, $p <$ 001, partial $\eta^2 = .09$, a main effect of source type, $F(2, 436) = 318.34$, $p < .001$, partial $\eta^2 = .59$, and a significant interaction between ownership and source type, $F(2, 436) = 4.95$, p = .007, partial $\eta^2 = .02$. As shown in Figure 4, Tesla owners consult information sources less frequently than non-owners. Analysis of simple effects revealed that Tesla owners use online sources significantly more than friends or colleagues as sources ($p < .001$) and the owner's manual ($p < .001$). They also consult their owner's manual more than friends or colleagues as a source, $p < .001$. Non-owners also use online sources more than the owner's manual ($p < .001$) and friends or colleagues as sources ($p < .001$). However, there was no difference between consulting the owner's manual and friends and colleagues, $p = .389$. Additionally, non-owners consult friends / colleagues and online sources more than Tesla owners (both p's $< .001$) but there was no difference in consulting the owner's manual between Tesla drivers and non-owners ($p = .445$).

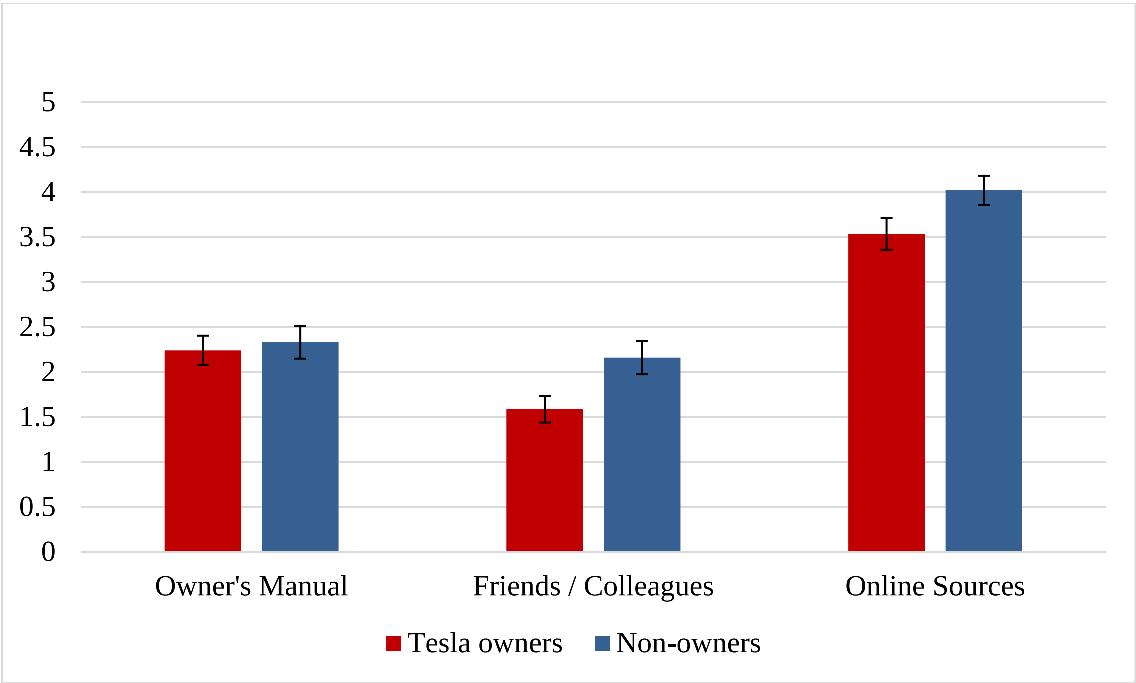

Figure 4. Frequency of using information sources to learn about the features of the car. Error bars represent 95% confidence intervals.

2.3.4.2 Accessing Owner's Manual

Next, we analyzed how drivers access their owner's manual when they need it. Table 3 shows percentages of the various media used to access owner's manuals among Tesla owners and non-owners. Questions for Tesla owners included an additional item, in-vehicle display. Tesla Model S vehicles have a 17" touchscreen display which allows controlling vehicle functions such as A/C but also features multimedia controls, including a web browser. This touchscreen display was one of the iconic features of Tesla cars at the time this study was conducted.

As shown in Table 3, there are considerable differences in accessing owner’s manuals between Tesla drivers and non-owners. First, the rates of using personal computers to access the manual were similar between the two groups. A significant shift can be seen in mobile device use, smartphones and tablets, between Tesla owners and non-owners. Smartphone use is three times higher for non-owners than Tesla owners in accessing the manual. Likewise, tablet use is 1.5 times higher. The primary difference however was the use of in-vehicle display and physical manuals. About 80% of Tesla owners access the manual using the in-vehicle display, and 76.2% of non-owners access the manual in physical form. We should note that some Tesla models don’t come with a physical manual; a digital version is provided to the driver.

Table 3. Media used to access owner's manual for Tesla owners and non-owners.

Media to Access Owner’s Manual	**Tesla Owners** % (N = 121)	**Non-owners** % (N = 101)
Computer	52.1	56.4
Smartphone	14	42.6
Tablet	9.9	14.9
In-Vehicle Display	80.2	N/A
Physical Manual	5.8	76.2
Other	.8	2

When combined (Table 4), we can see that most drivers use two or fewer different media to access owner’s manuals, while the number of media used by non-owners is slightly higher than Tesla owners.

Table 4. Number of media used to access owner's manual for Tesla owners and non-owners**.**

Number of Media Used to Access Manual	Tesla Owners % (N = 121)	Non-owners % (N = 101)
1	48.3	43.6
2	39	29.7
3	11	15.8
4 or more	1.7	9.9

2.3.4.3 Use of Information Sources to Understand Software Updates

Table 5 presents the information sources Tesla drivers used to learn more about the features of the Autopilot and Summon updates. Most participants read the release notes and used online forums to learn about the features that came with the Autopilot update. Only a few people consulted friends, company representatives, and about 30% used websites. We observed a similar pattern for the Summon update. A majority of the participants read release notes, used online forums, and websites. “Other” option included responses such as asking family members or watching videos.

Table 5. Information sources used to learn more about Autopilot and Summon updates.

Information Sources	Autopilot % (N = 109)	Summon % (N = 99)
Reading Release Notes	78.9	81.8
Asking Friends	3.7	7.1
Asking Company Representatives	7.3	5.1
Using Online Forums	74.3	71.7
Using Websites	29.4	28.3
Other	8.3	4

When combined (Table 6), we see that most participants used at least two information sources to learn about the new features of these software updates, with 70.1% for Autopilot and 72.4% for Summon updates.

Table 6. Number of information sources used to learn more about Autopilot and Summon updates.

Number of Information Sources	**Autopilot** % (N = 109)	**Summon** % (N = 99)
1	29.9	27.6
2	42.1	48
3	23.4	22.4
4 or more	4.6	2

2.3.5 Discussion

In this analysis, our goal was to describe how drivers use information sources when they want to learn more about the features of their cars, and in the Tesla case, about updates. We believe these results complement the findings we reported in section 2.2. By comparing Tesla owners to non-owners, we wanted to identify whether driving an upgradable vehicle with autonomous driving capabilities is different than traditional driving. Results from both Tesla owners and non-owners should provide useful guidance for designers and engineers in understanding driver behavior in information seeking activities, especially in the context of connected and upgradeable cars.

The first important finding was that drivers consult online sources significantly more than reading the owner’s manual or asking friends. This might seem a sampling issue, as the surveys were distributed over social media and driving related websites. However, people don’t read manuals, and less so when they are printed (Novick & Ward, 2006). Note that ratings for reading owner’s manuals were 2.24 and 2.33, for Tesla owners and non-owners respectively. These means are just slightly above “rarely”, indicating that the drivers don’t prefer the owner’s manual when they need to access information. An implication of these findings for automotive industry is to reconsider how to provide and deliver useful information to drivers, rather than relying on traditional manuals. For example, interactive voice interfaces show promise in creating a more engaging user help experience (Alvarez et al., 2010; Alvarez, López-de-Ipiña, & Gilbert, 2012).

The way drivers access the owner’s manual when needed showed different patterns between Tesla owners and non-owners. For non-owners, we observed that while half of the drivers access the owner’s manual in a digital format such as personal computers, smartphones or tablets, printed manuals are still the preferred way of accessing information. The use of smartphones is interesting considering the small screen sizes and number of pages found in most vehicle manuals. As mentioned previously, the low preference for printed manuals for Tesla owners was not surprising because most Tesla cars don’t come with a printed manual. More importantly however, Tesla owners don’t prefer accessing the owner’s manual using mobile devices. Considering that majority of these drivers access it using the large in-vehicle display, there is perhaps less need for a mobile device to access the manual inside the car. This also shows how drivers adapt and change their behaviors based on the available technology as the convenience of a large touchscreen display is easier to read than small screen mobile devices. Overall these findings are not surprising, as smartphones and in-vehicle information systems are the top two technologies people want to interact with in self-driving cars, with touchscreens being the preferred method of input (Pfleging, Rang, & Broy, 2016). These findings also indicate that designers of automotive systems should pay attention to in-vehicle information systems as more advanced in-vehicle technology and larger displays become available to use in the cars. A challenge for future human-machine interface (HMI) design in vehicles should be to identify which functions should be allocated to the vehicle HMI, and which functions should be left to individual devices such as smartphones. This is relevant to consumers today who expect some of these functions from their cars. Functionality such as in-vehicle navigation, entertainment, voice control, connectivity and communication systems are indeed the primary problems consumers face today (JD Power, 2017). Our results indicate that if the in-vehicle technologies support a more usable, contextual, and useful way of accessing information (e.g. large in-vehicle display), drivers will prefer and use them.

The results also revealed that most Tesla drivers checked multiple sources when they receive an update. We believe there may be several reasons for this. First, it may be related to the quality of the release notes which are written like user manuals. User manuals are difficult to navigate and usually do not offer proper explanation (Novick & Ward, 2006a; 2006b). It is possible that drivers get an incomplete picture of what the updates really do after reading the release notes. Consequently, they go online, expand their knowledge, or confirm that they understood correctly. The second reason might be to understand how these features will affect them in real life by checking other drivers' opinions and experiences with the features. For example, the release notes state that the Autopilot may fail to work properly due to various reasons (Tesla Motors, 2016). A driver, who has read this statement, might want to see how people experience this limitation in real life. This is also consistent with the finding that when people seek expertise, using documents and people as information sources are both frequent and they are complementary (Herztum, 2014). A third reason might be to simply learn more about the process behind these features. The language used in release notes usually don't describe the technology behind these advanced features in detail, mainly to keep the text simple and understandable. It is possible that drivers want to learn more about how the technology works through interaction with other people online, where drivers and experts share their knowledge about the behind the scenes of this technology. This view is also consistent with earlier findings where people would seek simple and objective information using documents and electronic resources (e.g. Wikipedia) but for complex information such as processes, opinions, and decision-making, they tend to seek other people's knowledge and expertise (Yuan, Rickard, Xia, & Scherer, 2011).

A limitation of this analysis is that while we showed that what people use, we still don't know how they use these sources. For example, which piece of information do drivers obtain from owner's manuals, and how do they integrate this information with the information they gain from online sources and social media? This is an important research question for future research because technology that addresses drivers' needs will likely incorporate information from multiple sources to be relevant for drivers. Likewise, future research should identify factors that influence why people access certain information sources using certain devices. For example, why, when and where do people prefer using their smartphones to look up information about their vehicles as opposed to using a physical manual? Additionally, more research is needed to understand driver's expectations and experiences with upgradability, especially in the context of smartphone-like apps for the car. Answering these questions can provide a better picture of driver information needs and reveal opportunities for future design of successful user assistance systems, in-vehicle information systems and software update processes.

2.3.6 Conclusion

In this analysis, we examined how drivers use information sources to learn about the existing features of their cars and the new features enabled by updates. Overall these findings suggest that drivers are comfortable in using multiple information sources and technology as part of this process. Moreover, drivers look up information about the updates using multiple sources. We expect these trends will be more prominent in the future, when connected ecosystems will become available. The design of future help systems and in-vehicle technologies for upgradable and connected cars should consider how, when and why users demand and access information.

2.4 Trust in Automation

2.4.1 Overview

In this section, we present additional results from the survey we introduced in previous sections (sections 2.2 and 2.3). Specifically, we will present findings on Tesla drivers' trust and confidence in Autopilot and Summon.

Tesla's Autopilot system, along with other advanced driver assistance systems (ADAS) are far from being perfect and failures are common. Given this imperfection, a critical issue is the degree of reliance on Autopilot. If drivers completely rely on Autopilot, negative consequences during automation failures will be inevitable such as the fatal Tesla crash (Golson, 2017). On the other hand, if drivers don't rely on Autopilot at all, the opportunity to save more lives thanks to automation being superior under certain circumstances will be missed. An important concept, trust, can help us in understanding how appropriate reliance can occur. Trust in automation has been a key concept in understanding the use of automated tools and subsequently human-automation team performance. Moreover, trust in technology is an important determinant of user adoption. Understanding how trust is shaped and how it relates to actual experience in the context of autonomous cars is key for safe driving. To this end, we will first discuss relevant literature regarding trust in automation, and then present findings from the survey on Tesla drivers' trust in Autopilot and Summon.

2.4.2 Background

Trust has been a fundamental concept in human-automation interaction (Hoff & Bashir, 2015, Lee & See, 2004; Parasuraman & Riley, 1997). Inappropriate calibration of trust in an automated system can lead to misuse (overreliance) and disuse (underreliance) of automation (Parasuraman & Riley, 1997), and result in decreased performance and less adoption. There has been considerable research on trust in automation (See Hoff & Bashir, 2015, for a review; Schaefer, Chen, Szalma, & Hancock, 2016, for a meta-analysis on factors influencing trust). Lee and See (2004) identified three factors that are critical in trusting an automated agent: performance, process, and purpose. Performance refers to operator's observation of results, process refers to operator's assessment of how the system works, and purpose refers to the intention of the system. These dimensions should match with each other in operator's mind to establish appropriate levels of trust. For example, if observed performance matches the operator's understanding of the system (process), then appropriate levels of trust can be developed.

Trust and reliance on automation increases as perceived reliability of the automation increases (Sanchez, Fisk, & Rogers, 2004; Ross, Szalma, Hancock, Barnett, & Taylor, 2008; Muir, & Moray, 1996). Trust seems to act as a precursor to reliance and mediate the relationship between beliefs and reliance (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003; Wang, Jamieson, & Hollands, 2009). It decreases with automation error (Lee, & Moray, 1992; Bisantz, & Seong, 2001), but providing explanations of why the error occurred (observing the process; Lee & See, 2004) can increase trust and reliance despite the errors (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003). Also, trust is more resilient when an automation error occurs if the operator has the ability to control and compensate for these errors (Muir, & Moray, 1996). In addition, the type of automation error also influences trust and reliance differently (Sanchez, Rogers, Fisk, & Rovira, 2014). For example, increased false alarm rates result in less reliance on automation while alarms that are accurate but not needed by drivers increase trust (Lees, & Lee, 2007). Trust in automation increases over time, especially if there are no major failures (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015), and regardless of prior exposure to automation errors (Hergeth, Lorenz, & Krems, 2016). It can even increase over time without exposure to the automated system (Sauer & Chavaillaz, 2017).

Age can also effect trust in automation. Older people tend to have higher levels of trust in automation (Ho, Wheatley, & Scialfa, 2005; Donmez, Boyle, Lee, & McGehee, 2006; Gold, Körber, Hohenberger, Lechner, & Bengler, 2015). Findings regarding how older people calibrate their trust and reliance are mixed. While some studies showed that they may use different trust calibration strategies (Sanchez, Fisk, & Rogers, 2004; Ezer, Fisk, & Rogers, 2008), others did not (Ho, Wheatley, & Scialfa, 2005).

Taken together, these findings show how important trust is in reliance on automated systems. In the next section, we will present how Tesla drivers trust Autopilot and Summon. Based on the literature, we expected trust to be related to frequency of use, increase over time, negatively affected by experiencing an incident, and increase with age.

2.4.3 Method

We asked participants to rate their trust in Autopilot and Summon on two 5-point Likert scale items measuring trust and confidence in Autopilot and Summon. We averaged these items and created a trust score. Similarly, we asked participants to remember and rate their initial trust and confidence when they first used Autopilot and Summon on a 5-point Likert scale. We averaged these items and created an initial trust score. The items were taken from "Checklist for Trust between People and Automation" scale (Jian, Bisantz, & Drury, 2000) which consists of 12 items to measure trust in automation. The questions are presented in Appendix A.

2.4.4 Results

In the following analysis, we used data from Autopilot users (N = 109) for trust in Autopilot and data from Summon users (N = 99) for trust in Summon. We compared initial and current trust for Autopilot and Summon. We also examined the relationship between trust and other factors discussed in section 2.2.

2.4.4.1 Trust in Autopilot

Overall, participants reported high levels of trust in Autopilot (M = 4.02, SD =.65) and moderate levels of initial trust (M = 2.83, SD = .82). As shown in Table 7, trust in Autopilot was positively correlated with frequency of Autopilot use, self-rated knowledge about Autopilot, ease of learning, and usefulness of Autopilot display. Surprisingly, for those who experienced an Autopilot incident (N = 68), trust was not correlated with how risky they perceived the situation. However, perceived risk was negatively correlated with frequency of use.

Table 7. Correlations between trust and other variables. Correlations between perceived risk and other variables are computed for those who reported an incident.

		Mean	SD	1	2	3	4	5	6	7	8	9
1	Initial Trust	2.83	.82									

2	Current Trust	4.02	.65	.44**							
3	Computer expertise	4.50	.66	.05	.11						
4	Frequency of Use	4.19	.65	.30**	.52**	.14					
5	Knowledge	3.79	.82	.34**	.26**	.28**	.25**				
6	Ease of learning	4.27	.72	.23*	.36**	.18	.29**	.19*			
7	Importance	3.51	1.08	.15	.06	.31**	.08	.38**	-.13		
8	Usefulness	4.06	.70	.20*	.40**	.10	.28**	.12	.22*	.21*	
9	Perceived risk	2.74	.87	-.15	-.09	.06	-.27*	.07	-.13	0.24	-.21

Note: * p < .05, ** p < .01. Knowledge refers to self-rated knowledge about how Autopilot makes decisions. Importance refers to perceived importance of knowing how Autopilot makes decisions. Usefulness refers to perceived usefulness of Autopilot display.

Age was presented as a categorical question in this study, and covered ages from 16 to 65 and older. A one-way ANOVA showed a significant age effect on trust, $F(6, 102) = 2.63$, $p = .02$, partial $\eta^2 = .13$. A trend analysis using polynomial contrasts was also significant, $F(1, 102) = 7.80$, $p = .006$. As shown in Figure 5, trust in Autopilot slightly but significantly decreased with age.

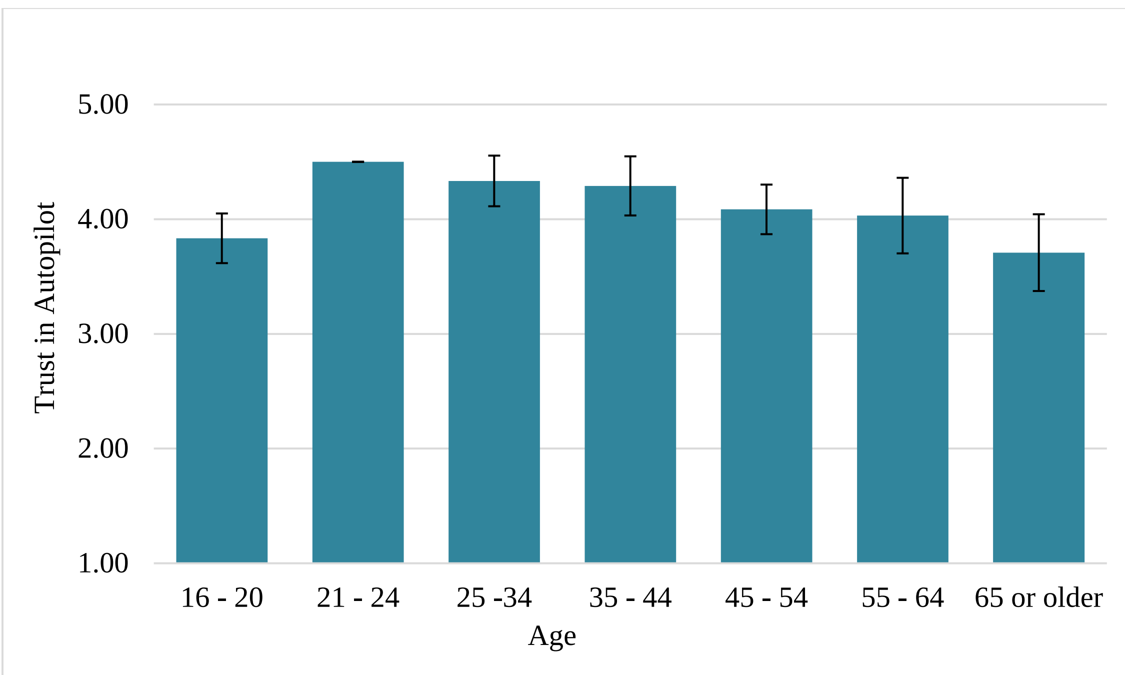

Figure 5. Trust in Autopilot by age. Categories 16-20 had 4 participants, 21-24 had 2 participants, 25-34 had 19 participants, 35-44 had 27 participants, 45-54 had 25 participants, 55-64 had 16 participants and 65 or older had 16 participants. Error bars represent 95% confidence intervals.

Next, we compared Tesla drivers' initial and current trust on Autopilot and how experiencing an incident (Incident group) or not (No Incident group) affects trust. A 2x2 mixed ANOVA with time as a within-subjects factor (Initial Trust, Current Trust) and incident as a between-subjects factor (Incident, No Incident) showed a main effect of trust, $F(1, 107) = 221.05$, $p < .001$, partial $\eta^2 = .67$, and a main effect of incident, $F(1, 107) = 9.59$, $p = .002$, partial $\eta^2 = .08$. The interaction effect was not significant, $p = .086$. As shown in Figure 6, trust in Autopilot was higher than initial trust, and those who experienced an Autopilot incident reported lower levels of trust. Surprisingly, they also reported lower levels of initial trust in Autopilot.

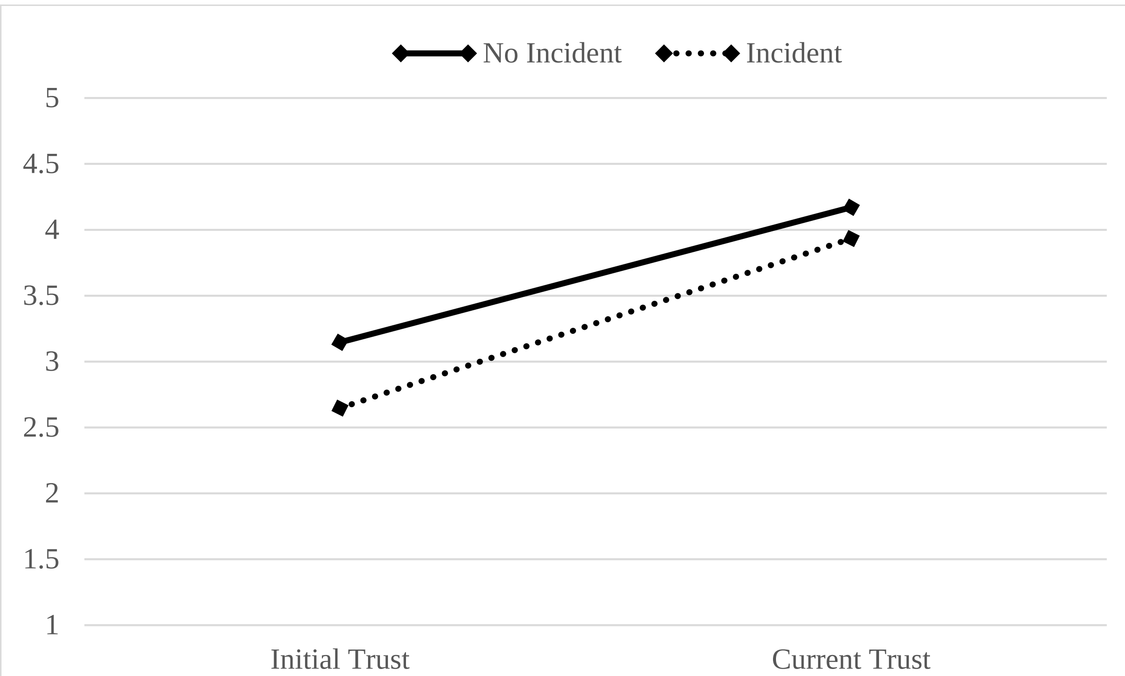

Figure 6. Means of current and initial trust on Autopilot for Incident and No Incident groups.

2.4.4.2 Trust in Summon

Participants ($N = 99$) reported high levels of trust in Summon ($M = 3.80$, $SD = .93$) and moderate levels of initial trust ($M = 3.11$, $SD = 1.01$), similar to Autopilot. As shown in Table 8, trust in Summon was positively correlated with self-rated knowledge about Summon, and ease of learning. Current trust was positively correlated with frequency of use, and initial trust was positively correlated with computer expertise and negatively correlated with perceived. For those who reported a Summon incident ($N = 21$), initial trust but not current trust was negatively associated with perceived risk of the situation. A one-way ANOVA showed no effects of age on current trust in Summon, $F(6, 92) = 1.78$, $p = .108$. Trust in Summon did not differ across age groups ($p > .05$).

A 2x2 mixed ANOVA with time as a within-subjects factor (Initial Trust, Current Trust) and incident as a between-subjects factor (Incident, No Incident) show a main effect of trust, $F(1, 97) = 23.52$, $p < .001$, partial $\eta^2 = .20$. Current trust in Summon was higher than initial trust. The main effect of incident was not significant, $F(1, 97) = 1.05$, $p = .309$; the interaction was not significant as well, $F(1, 97) = 2.74$, $p = .101$. Means are shown in Figure 7.

Table 8. Correlations between trust in Summon and other variables. Correlations between perceived risk and other variables computed for those who reported an incident.

		Mean	SD	1	2	3	4	5	6	7	8
1	Initial Trust	3.11	1.01								
2	Current Trust	3.80	.93	.49**							
3	Computer expertise	4.47	.68	.20*	.03						
4	Frequency of Use	2.67	1.08	.19	.22*	-.14					
5	Knowledge	3.56	.92	.31**	.32**	.18	.18				
6	Ease of Learning	3.99	.84	.35**	.45**	.12	.14	.36**			
7	Importance	3.11	1.17	.10	-.09	.27**	.001	.26**	-.02		
8	Perceived risk	2.14	1.24	-.49*	-.28	-.30	.30	.02	-.19	-.02	

Note: * p < .05, ** p < .01. Knowledge refers to self-rated knowledge about how Summon makes decisions. Importance refers to perceived importance of knowing how Summon makes decisions.

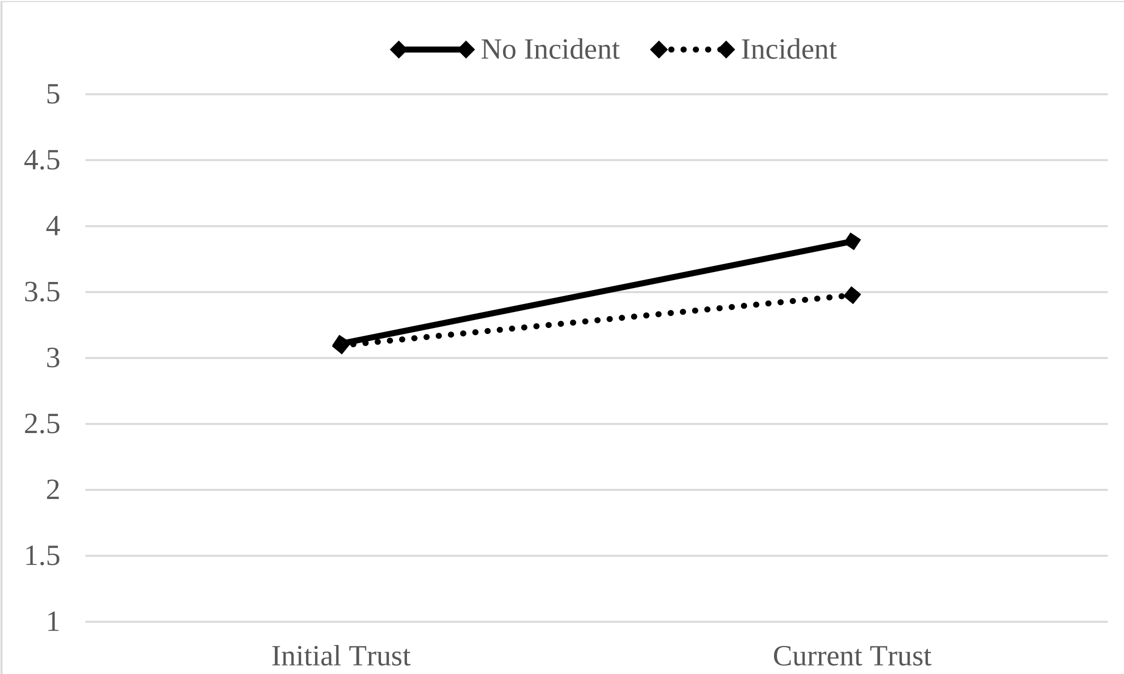

Figure 7. Means of current and initial trust in Summon for Incident and No Incident groups.

2.4.5 Discussion

In this analysis, our goal was to identify how Tesla drivers' trust in Autopilot and Summon relate to attitudes towards these systems, and how experience shapes their trust in these systems. Overall, we observed high levels of trust and moderate levels of initial trust. Trust increased over time regardless of whether participants experienced an incident. Trust in Autopilot but not Summon decreased as the age increased.

High levels of trust reported for both Autopilot and Summon indicate that the drivers are confident in these systems which is in line with previous findings (Dikmen & Burns, 2016). Analysis of correlations revealed interesting patterns. Frequency of use of Autopilot was associated with trust. As expected, those who have higher levels of trust tend to use the system more often (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003). The reverse is also true: The more drivers experience Autopilot and Summon under different circumstances, the more their trust increases, especially if these systems had good performance and reliability in handling different situations. This is also in line with previous findings on the relationship between trust and experience (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015; Hergeth, Lorenz, & Krems, 2016). Ease of learning was also positively correlated with trust for both Autopilot and Summon. The design features of automation such as usability influence trust by altering perceptions of users (Hoff & Bashir, 2015). Likewise, easy to learn characteristics of Autopilot and Summon may have created perceptions of trustworthiness by making the adaptation smooth. The usefulness of Autopilot display was also positively correlated with trust in Autopilot. The main purpose of the Autopilot display is to show the sensing capabilities of the system to the user. At any time during the ride with Autopilot, drivers can glance at this display and see how Autopilot perceives other vehicles on the roadway and whether sensors become active (e.g. ultrasonic sensors). In other words, this display opens the black box of automation and enables the users to observe the process (Lee & See, 2004). If this transparency has a positive effect on trust, it can be an important part of adoption process for autonomous vehicles. However, as Lee and See (2004) notes, having an appropriate level of trust is much more important than just higher levels of trust. While providing transparency can result in better trust calibration (Seong & Bisantz, 2008), further research is needed to identify how these drivers use the Autopilot display. Lastly, self-rated knowledge about how Autopilot or Summon makes decisions was positively correlated with trust in Autopilot and Summon, respectively. In general, knowledge about how these systems work, including their limitations, should result in appropriate trust calibration. However, we don't know the extent to which "self-rated knowledge" matches the real, objective knowledge about how these technologies work. Still, knowledge about how automation makes decisions, especially when it fails, can result in higher levels of trust (Dzindolet et al., 2003). Similarly, awareness of how Autopilot and Summon handles or fail to handle various situations might have resulted in appreciation of these technologies, and subsequently higher levels of trust. However, it is also possible that those who have a priori trust in these systems might be more willing to learn more about how the technology works behind the scenes, and improve their knowledge about the system. Further research is needed to establish how knowledge and mental models, both subjective and objective, relate to trust in autonomous vehicles.

Older people reported slightly lower levels of trust in Autopilot. This finding is contrasts with previous research (e.g. Ho, Wheatley, & Scialfa, 2005 on medication management systems) which showed that older adults have higher levels of trust in automation. One explanation for current findings is that older people tend to have more driving experience than younger drivers and domain expertise has been shown to influence trust and reliance in automated decision aids. For example, farmers (domain experts) rely less on automated aids than non-farmers (domain novices) (Sanchez, Rogers, Fisk, & Rovira, 2014). Another explanation might be the differences in risk perception. Younger drivers tend to perceive situations such as curved roads and rural environments less risky than older drivers (Tränkle, Gelau, & Metker, 1990). It is possible that the perceived risk associated with automated driving might be different across different age groups. Nevertheless, we echo with Schaefer et al. (2016) that there is a need for further research in understanding the relationship between age and trust in automation.

In terms of trust over time, we observed similar results for Autopilot and Summon. Trust increased over time for both Autopilot and Summon. This finding is consistent with previous work (Gold et al, 2015; Hergeth et al., 2016; Sauer & Chavaillaz, 2017). As drivers use these systems more, they likely became more comfortable. Over time, drivers may have adapted to this new environment, whereby they learned how to cooperate with an automated agent. Failures can be a challenge, but they can also provide a learning opportunity.

For Autopilot, those who experienced an incident reported lower levels of both current trust and initial trust. It was surprising to observe the differences between Incident and No Incident groups in initial trust in Autopilot. It is possible that those who experienced an Autopilot incident may have been subject to cognitive biases such as hindsight bias (Roese & Vohs, 2012) and they may have responded based on later negative experiences. However, given other findings, we believe that a more likely reason is that these drivers might indeed have lower levels of trust in Autopilot at first, and this might have led them to be more sensitive to the capabilities of Autopilot, which may have resulted in (a) more likely to consider certain situations as a failure, and (b) motivating drivers to explore the limits and capabilities of Autopilot more to calibrate their trust better. They might, for example, have used Autopilot under circumstances where it is not designed to function. Throughout the comments, we also observed indications of these situations. As one participant pointed out, part of the learning process is testing its limitations. Nevertheless, these findings support the idea that the relationship between trust and automation failures is a complex one, and many factors can influence this process (Hoff & Bashir, 2015). Earlier, we reported that drivers who experienced an incident did not perceive these situations particularly risky (section 2.2). We believe current results on trust support these findings such that experiencing an Autopilot incident does not necessarily cause significant reductions in trust. However, we should note that these ratings don’t necessarily represent drivers’ trust right after experiencing an incident. Trust is a dynamic and evolving process (Lee & See, 2004), and while it decreases after automation faults, gradually it recovers (Lee & Moray, 1992).

Trust in Summon was not influenced by whether participants experienced an incident or not. While there was a trend towards reduced levels of trust for Incident group, current data failed to support this hypothesis, partly due to sample size. Surprisingly, initial trust in Summon was strongly and negatively correlated with perceived risk. This suggests that perhaps failures mostly occurred during initial use of Summon which might have influenced initial trust. Nevertheless, we should note that Autopilot and Summon are qualitatively different automation systems both in terms of the consequences of failures and the level of complexity of the environments where these systems are used. Therefore, trust development process might be affected by different factors for these systems.

This work had several limitations. Unlike laboratory experiments, trust was not assessed immediately after the incidents, and the time interval between the last time the drivers experienced an incident and the survey varies from person to person. A longitudinal study on how trust develops over time with autonomous vehicles would identify both fluctuations in trust and how drivers psychologically deal with automation failures. Also, almost all participants in this study were actively using Autopilot or Summon. We should note that trust in these systems might be different for users to stopped using Tesla cars or these systems due to a major failure or accident. While we observed that trust was associated with multiple factors, identifying exact mechanisms require further research such as how age, knowledge and mental models influence trust. Trust evolves over time, and while trust influences reliance on automation, it is not the only factor (Lee & See, 2004). Future research should examine the affective component of trust in autonomous cars. Our observations throughout this work have been that there is more than meets the eye when it comes to developing a trust relationship with people's own cars, where factors such as their attitudes towards the designer (i.e. the brand or company producing the vehicle), public opinions and social influence might play an important role. At the end of the day, a car is more than just a job-related automation such as automated plants or aircrafts. A car has potential to become part of one's identity, life style, and social world. We believe that some these concepts are reflected in our work as well such as strong tendency to use online forums to connect with other Tesla owners. Therefore, it is critical to develop an understanding of the concept of trust in personal automation such as autonomous vehicles and home automation.

2.4.6 Conclusion

In this analysis, we examined trust in automation in the context of Autopilot and Summon. Overall Tesla drivers who participated in this study have high levels of trust in these technologies. Trust is related to several attitudinal and behavioral factors, and experiences shapes the level of trust in these technologies. While this work was an initial step towards understanding how trust plays a role in real world use of autonomous vehicles, it showed that laboratory findings and concepts developed in the research community are applicable to real world cases as well. We hope these findings will help to understand drivers' trust in autonomous vehicles, as the concept of trust will be fundamental in an automated world.

2.5 Chapter Conclusion

In this chapter, we presented the results of a survey conducted with Tesla drivers about their experiences with two automated systems, Autopilot and Summon. Current users of these technologies are highly comfortable and engaged with these technologies, motivated to learn more about these systems, and use multiple information sources. Automation failures are common but they are not perceived as particularly risky. Users have high levels of trust in Autopilot and Summon, and trust increases over time. These findings are first steps to understand how autonomous vehicles are being used in the real world. We hope this study complements laboratory findings and naturalistic studies on automated driving.

We identified a few key areas which require further research such as understanding nature of trust and how it affects the use of these technologies, how drivers integrate multiple information sources, and given the prevalence of automation failures, how to keep the drivers in the loop and make sure they have a proper understanding of what is going on under the hood. The next chapter describes a study we conducted that addresses some of these questions, particularly the latter one.

Chapter 3
Augmented Reality Head-Up Displays in Automated Driving

3.1 Chapter Introduction

In this chapter, we will present the findings of a study we conducted on augmented reality head-up displays in a simulated environment.

In the previous chapter, we identified several issues that needs to be addressed, such as understanding trust, information access and drivers' motivation to understand better how automation works. The latter point is the basis for this study. Our purpose was to identify ways to increase understanding of how automation works by providing real time information to drivers during automated driving. One of the most convenient ways to achieve this goal is through visual and auditory displays which provide information about the status of the automation, also known as automation displays. The focus of this work was to identify how automation displays influence drivers' attitudes, behavior, and performance in automated driving. For the type of display, we chose to explore the concept of augmented reality head-up displays (AR HUD), head-up displays with augmented reality graphics which are aligned with real world objects, resulting in a conformal display. Our goals and research questions in this study were following:

1. How does presenting varying amounts of information about the vehicle's sensing capabilities via an augmented reality head-up display affects trust, workload, situation awareness, perceived usability, and secondary task engagement?
2. How does representing automation failures affects trust, workload, situation awareness, perceived usability, and secondary task engagement?

In the following sections, we will present an experiment we conducted to achieve these goals. Materials related to this study are presented in Appendix B: Experimental Materials.

3.2 Overview of the Study

The race to build self-driving cars is at full speed. Currently, all major automotive companies are working towards building autonomous vehicles. Some, like Google, aim to produce fully autonomous vehicles that eliminate the driver completely. However, others such as Tesla aim to achieve full autonomy gradually by introducing semi-autonomous driving capabilities into the vehicles and gradually developing them into fully autonomous cars. A safe transition to full autonomy in the next decade requires proper investigation of driver-automation interaction related issues which such as increasing driver situation awareness (Endsley, 2017; Stanton & Young, 2005) and driver distraction (Llaneras, Salinger, & Green, 2013) as the vehicles' automated driving capabilities incrementally increase.

The aim of this research is to identify how automation displays influence driver attitude and behavior. Specifically, this study looks at the effects of using augmented reality head-up displays (AR HUD) on situation awareness, workload, trust, and distraction in a simulated environment. We designed a study in which participants watched driving videos featuring simulated augmented reality visualizations highlighting the objects the automated vehicle identified while engaging a secondary task, and rated their situation awareness, mental workload, and trust.

AR HUD systems promise to provide contextual, meaningful, and timely information to drivers, which will be very important as the vehicles get smarter and more automated, and as the role of the drivers shifts from manual driver to a supervisor, with added cognitive demands resulting from this new role. In this study, we examined how opening the black box of automation using AR HUD impacts driver performance in automated driving.

3.2.1 Related Work

In-vehicle automation displays are a critical part of automated driving. Previously, the driver, who was in charge of the vehicle all the time, had to monitor the environment and control the vehicle accordingly. In automated driving, this monitoring task will be extended to include monitoring of the state of the automation as well. Automation displays play an important role in assisting this new task by providing how the vehicle senses the environment and makes decisions, with the goal of increasing situation awareness (SA; Endsley, 1995). They are the primary method for drivers to understand the state of the system (Banks & Stanton, 2016), and they can provide critical information such as which sensors are activated (e.g. using animation when the blind spot monitor sensor detects an object), how the car senses the objects on the road, whether lane markings are appropriate, and various indicators such as whether or not automated driving is available, and warnings and alert messages (e.g. "please take the control").

An in-vehicle display, whether its related to automated driving functionality or is a conventional in-vehicle display, can be a traditional head-down display, or a head-up display. A more recent technology that started to attract attention is augmented reality head-up displays, a combination of head-up displays and augmented reality visualizations to provide spatial, real-time information about the environment. Augmented reality head-up displays (AR HUD) can provide two ways of presenting information: screen-fixed and world-fixed (Gabbard, Fitch, & Kim, 2014). Screen-fixed displays present information on a fixed location of the screen whereas the world-fixed displays present information in a location that aligns with an object in the real environment to give the perception of "attached graphics". The advantage of world-fixed displays is that they provide contextual information and map it directly onto the real world, minimizing the effort required to attend, perceive, and match the display and real world. Additionally, AR HUD cues can be presented on a head-mounted display, on a dashboard display where the cues are superimposed on real-time camera footage, or on windshield display. The projection onto the windshield was found to be more effective than others in a number of measures including navigation related errors and object detection (Jose, 2015).

AR HUD can effectively convey warnings (Schwarz & Fastenmeier, 2017) and improve the psychological being of drivers by relieving stress and tensions (Hwang, Park, & Kim, 2016)). AR HUD has a major advantage over a traditional head-down display in that drivers can keep their eyes on the road when using AR HUD. This leads to several advantages over traditional in-vehicle displays. Previous work showed that AR HUD results in better navigation performance (Kim & Dey, 2009; Bolton, Burnett, Large, 2015), earlier recognition of turns (Bark, Tran, Fujimura, & Ng-Thow-Hing, 2014), faster responses to road hazards without compromising workload (Kim, Wu, Gabbard, & Polys, 2013), increase awareness of pedestrians (Phan, , Thouvenin, & Frémont, 2016) and smoother breaking when approaching pedestrian crossings (Kim, Miranda Anon, Misu, T., Li, Tawari, & Fujimura, 2016). AR HUD visualization have the power of attracting driver attention, however this can also be detrimental. For example, drivers tend to look at objects longer when highlighted using AR HUD and miss other, possibly important objects compared to not using HUD (McDonald, 2016)

Using augmented reality cues can be an effective way of providing information to the drivers about sensing capabilities of the vehicle. Using such displays, the driver can monitor both the road and the automation's view of the world simultaneously, leading to higher awareness of the state of automation. This can be critical in situations where the vehicle fails to detect an object and thus ignores it, such as failing to notice a parked vehicle. Regarding the use of AR HUD in automated driving, AR cues in the form of highlighting lanes with green (safe) or red (dangerous) resulted in similar reaction times compared to no AR in take over scenarios, however they also resulted in safer maneuvers such as using the brake more in an emergency lane change (Lorenz, Kerschbaum, & Schumann, 2014). Interestingly, these cues also led to checking the side corridors during a lane change less often, suggesting that AR cues can have negative effects as well by attract driver attention such that they may rely on AR cues rather than checking the environment themselves. In a similar study, augmented reality cues in the form of highlighting vehicles on the road, augmented reality cues did not increase response speed to take over requests in automated driving, but resulted in smoother transitions to manual driving and helped the drivers better anticipate the required maneuvers, suggesting an increase in situation awareness (Langlois & Soualmi, 2016). AR HUD can also increase driver engagement with the real world in semi-automated driving by attracting drivers' attention through visualizations. One such concept is presenting a game on AR HUD to keep drivers' attention on the road (Schroeter & Steinberger, 2016).

Despite the efforts to understand and design effective augmented reality head-up displays, there is more research needed to have a complete picture of AR HUD. An important consideration is identifying what should be represented on an AR HUD, especially in the context of automated driving. This experiment therefore sought to identify the effects information type on performance such as situation awareness, workload, trust and secondary task engagement, by providing varying amounts of information about the vehicle's sensing capabilities of the environment. Specifically, we were interested in how providing AR cues related to lead vehicles on the same lane, vehicles in other lanes, and road signs impacts driver attitudes and behavior.

3.2.2 Overview of the Experiment

In this experiment, participants watched several driving videos with simulated AR cues that highlight objects on the road (Figures 8, 9 and 10). These cues could highlight the vehicle on the same lane, other vehicles on the road, and road signs. Additionally, the AR system could be reliable (highlighting objects appropriately) or not reliable (failure to highlight certain objects). Participants, while watching these videos, were also engaged with a secondary task, a word search game on an iPad. The use of a secondary task paradigm is recommended in automated driving studies because they can act as a proxy for reliance on the automation (Gibson et al., 2016). After each video, participants reported their workload, situation awareness, trust in the vehicle and perceived usability of AR cues, in addition to video-specific questions. We chose to simulate AR HUD using videos because we did not have access to a proper AR HUD system.

We expected, based on previous work, that presenting more information about the sensing capabilities of the vehicle (i.e. highlighting both lead vehicles and vehicles in other lanes) would result in higher levels of awareness and increased perceived usefulness. Drivers should be able to obtain information about the state of the automation quicker if such information is presented in a contextual and relevant way. However, a direct consequence of this situation might be increased engagement with the secondary task, if drivers believe they can regain awareness quickly if needed.

3.3 Method

3.3.1 Participants

20 participants took part in the study. The minimum age was 18, and the maximum age was 33, with a mean of 21.8 (*SD* = 3.3). 13 participants were male. Average driving experience was 4.9 years, and on average, participants were driving 170 km per month.

3.3.2 Experimental Design

The experiment was a within-subjects design with 7 levels. Six of them included AR HUD, and the other one was a baseline condition which did not have any AR cues. Six levels of AR HUD were structured as a 3 (Design: Basic, Advanced, Advanced+) x 2 (Reliability: No Failure, Failure) design. These conditions featured AR cues highlighting certain objects on the road. The AR HUD design variations used in the study are shown in Figures 7,8 and 9. Basic design only highlighted the vehicles on the same lane with a yellow line. Advanced design highlighted the vehicles on the same lane with yellow lines as well as vehicles in other lanes with blue. Advanced+ were similar to Advanced design, with the addition of projecting a bigger image of road signs such as exits and service centre signs onto the screen. To manipulate reliability, we removed the AR cues that are supposed to highlight the lead vehicle. Figure 11 shows how a failure scenario was represented in this study. These failures, when they happen, happened only once during a video, and could last between 9.5 and 36.5 seconds (*M* = 19.6 seconds).

Figure 8. Basic Display. In this variation, only the lead vehicles (vehicles on the same lane as the own car) are highlighted.

Figure 9. Advanced Display. This variation highlights lead vehicles and vehicles in other lanes.

Figure 10. Advanced+ Display. This variation highlights lead vehicles, vehicles in other lanes as well as projecting larger images of road signs onto the screen.

Figure 11. AR HUD failures. This is one of the examples where the reliability of AR HUD was manipulated. On top two images, the vehicle on the right moves into the middle lane. On the bottom left, the vehicle identifies it as a lead vehicle and highlights it. On the bottom right, the vehicle “fails” to identify the lead vehicle. Hence, no highlighting occurs.

3.3.3 Videos and Secondary Task

The videos were shot on a nearby highway using a dashcam, the car was driven by a human driver, and the rides were completely safe. Rarely the vehicle changed lanes. The car was either in the middle lane or right lane and was driving within the legal speed limit. Traffic density was similar across videos and but could vary within a video. However, there were no slowdowns due to traffic at any point. All videos started and ended on highway, and the car did not leave the highway. The visualizations are added using a post-processing software. Design of AR cues were inspired by some of the concepts introduced by automotive companies. The colors were chosen somewhat arbitrarily, however we avoided using green and red colors as these have specific meanings in driving. Each video was about three minutes long, and for automation failure conditions, the failure could happen anywhere in the video. We used original 6 videos. For each video, we prepared 6 combinations (Design x Reliability). Baseline video was fixed.

The secondary task was a word search game presented on an iPad. It was a typical word search game where participant had to find words and cross them off.

3.3.4 Procedure

After filling out demographics questionnaire, participants were briefed about the study, and asked to pretend as if they are in the driver seat of this self-driving car, and experienced a short training video (about 1 minute). In the training video, they will be told how to engage automated driving, and how the word search game works. After the training video, participants watched 7 videos (6 AR videos and the baseline). Each video started with participants' hands on the wheel. They could engage automated driving by pressing a button on the steering wheel. We did not force participants to engage or disengage automated driving during the videos. Participants, when engaged in automated driving, could drive hands free and play the word search game on the iPad if they want. At any time if they wanted to take the control back from the car and disengage automated driving, they could do so by putting their hands on the wheel. To engage automated driving again, they pressed the same button. Engaging or disengaging automated driving did not change the flow of the video. However, when they were driving manually (i.e. hands on the wheel) they could not play the iPad game. Instructions were given that the vehicle may fail to detect certain objects and it is the participant's responsibility to make sure the ride is safe. After each video, participants filled out questionnaires measuring subjective situation awareness, subjective workload, trust, usability of the AR display, and ride-related questions. We tried to counterbalance the order of videos as best as possible. However, given the number of possible combinations of 36 videos, a full counterbalance was not possible. After 7 videos, participants filled out a post-experiment questionnaire asking questions about each design. All experimental stimuli were presented on a 27` IPS LED screen with a resolution of 1920x1080 and a refresh rate of 60 hertz. A Logitech G29 steering wheel was used in the experiment.

3.3.5 Measures

3.3.5.1 Primary Measures

After each video, we measured subjective mental workload, subjective situation awareness, trust in the automation, and secondary task performance.

Subjective situation awareness was measured with the Situation Awareness Rating Scale (SART; Taylor, 1990). This scale measures three dimensions of situation awareness, supply, demand, and understanding using 10 questions on 7-point Likert scales. Overall situation awareness score is obtained by using the following formula: SA = Understanding – (Demand-Supply) (Stanton, Salmon, Rafferty, Walker, Baber, & Jenkins, 2013)

Subjective workload was assessed using the National Aeronautics and Space Administration-Task Load Index (NASA-TLX; Hart & Staveland, 1988). This scale measures subjective mental workload by using six sub-scales, namely mental demand, physical demand, temporal demand, performance, effort and frustration. A score from 0 – 100 is obtained using these dimensions.

We measured trust in automation similarly as in the survey we reported in Chapter 2. However, in the current study, we used several additional items from "Checklist for Trust between People and Automation" scale (Jian, Bisantz, & Drury, 2000) which consists of 12 items to measure trust in automation. Some of the items were ambiguous in this context, and we used 6 items. Three of these items measure negative trust (distrust) and others measure positive trust (See Appendix B). These items were asked on a 5-point Likert scale. We then averaged all six items to obtain a trust score.

To measure secondary task engagement, we recorded the number of words participants have crossed off during a video.

Usability of different designs were measured with System Usability Scale (SUS; Brooke, 1996). This scale measures the usability of a system using ten 5-point Likert scale items and provides an overall usability score ranging from 0-100. A score of 70 is considered an acceptable level of usability, based on an analysis of 2,324 surveys (Bangor, Kortum, & Miller, 2008).

3.3.5.2 Secondary Measures

After each video, we asked participants about the number of exits and service centre signs they saw in the video and whether or not there was construction work on the road. While these measures may not be direct predictors of safe driving performance, they could provide insights about the attention participants allocated to monitor the roadway. We also asked about subjective assessments of how much they paid attention to the road and the secondary task, as well as understanding of how automation works. Additionally, at the end of the experiment, we asked participants to rate each interface on several items and open-ended questions about augmented-reality head-up displays.

3.4 Results

For data analysis, we used repeated measure ANOVAs. Only two measures were affected by the reliability of AR HUD: trust and usability scores. For these measures, we used two-way repeated measures ANOVA. For measures that were not affected by AR HUD failures, we collapsed failure and no failure conditions for each display, and ran one-way repeated measures ANOVAs with four levels (No AR, Basic display, Advanced display, Advanced+ display) to facilitate reader understanding. This also allowed easier comparison of AR designs with the baseline (No AR) condition. Greenhouse-Geisser corrections were used when the sphericity assumption was violated.

3.4.1 General Results

Median time to engage automated driving was 4.5 seconds and it did not differ between conditions according to a Friedman's test, $\chi2(3) = 1.62$, $p = .655$. Only a few participants used the take-over feature (putting hands on the wheel and disengaging from the secondary task), and only a few times during the study, therefore there is not enough data for a rich analysis. We believe having no real control over the vehicle played a role here. There were no significant differences in perceived risk of the ride between the videos used in the experiment, $p > .05$. Perceived risk ranged from 2.40 (SD = .99) to 2.75 (SD = .97) on a 5-point Likert scale. Participants perceived the videos low to medium risky.

3.4.2 Situation Awareness

A repeated measures ANOVA with Greenhouse-Geisser corrections showed no differences between SART scores, $F(2.06, 39.12) = .46$, $p = .64$. Participants reported similar levels of subjective situation awareness across conditions.

3.4.3 Mental Workload

A repeated measures ANOVA with Greenhouse-Geisser corrections showed no differences between subjective mental workload between conditions, $F(1.89, 35.97) = .36$, $p = .69$. Participants reported similar levels of mental workload under each condition. There were also no differences in subscales, all p's > .05. Overall workload ranged from 21.13 to 24.42 across conditions.

3.4.4 Secondary Task Performance

A repeated measures ANOVA showed no differences in secondary task performance, $F(3, 57) = .69$, $p = .56$. Participants crossed similar numbers of words in each condition, although the mean scores were slightly higher in AR HUD conditions. Average scores were 5.7, 6.95, 6.9 and 6.8 for No AR, Basic, Advanced and Advanced+ displays, respectively.

3.4.5 Trust

A 3 (Display: Basic, Advanced, Advanced+) x 2 (Reliability: No Failure, Failure) repeated measures ANOVA showed no main effect of display, $F(2, 30) = .79$, $p = .46$, no main effect of reliability, $F(1, 30) = 3.52$, $p = .08$, and no interaction effect, $F(2, 30) = .15$, $p = .44$ on trust in the automated vehicle. Further inspection revealed that two items related to negative trust (being suspicious of the intentions, and being wary of the system) were affected by the reliability of AR cues. Participants were more wary of the system when AR HUD failed ($M = 2.75$, SD = .72) than when the AR HUD was reliable ($M = 2.54$, $SD = .53$), $F(1, 30) = 5.82$, $p = .036$. They also reported higher levels of suspicion about the system's actions when AR HUD failed ($M = 2.42$, $SD = .69$) than when it was reliable ($M = 2.15$, $SD = .54$), but this different was not significant $F(1, 30) = 4.35$, $p = .055$.

To understand how driving with or without AR cues affect trust in automation, we compared reliable AR HUD conditions with No AR condition in trust in the automated vehicle. We chose only the reliable AR HUD conditions as these were similar to No AR where there was indication of failure. A repeated measures ANOVA with four levels (No AR, Basic, Advanced, Advanced+) showed a significant effect of AR HUD, $F(3, 54) = 3.16$, $p = .032$, partial $\eta^2 = .15$. As shown in Figure 12, Advanced+ design led to higher levels of trust than No AR, $p = .01$. Trust was also higher for Basic and Advanced design than the No AR, but these differences did not reach significance, $p = .061$ for Basic vs. No AR, and $p = .07$ for Advanced vs. No AR.

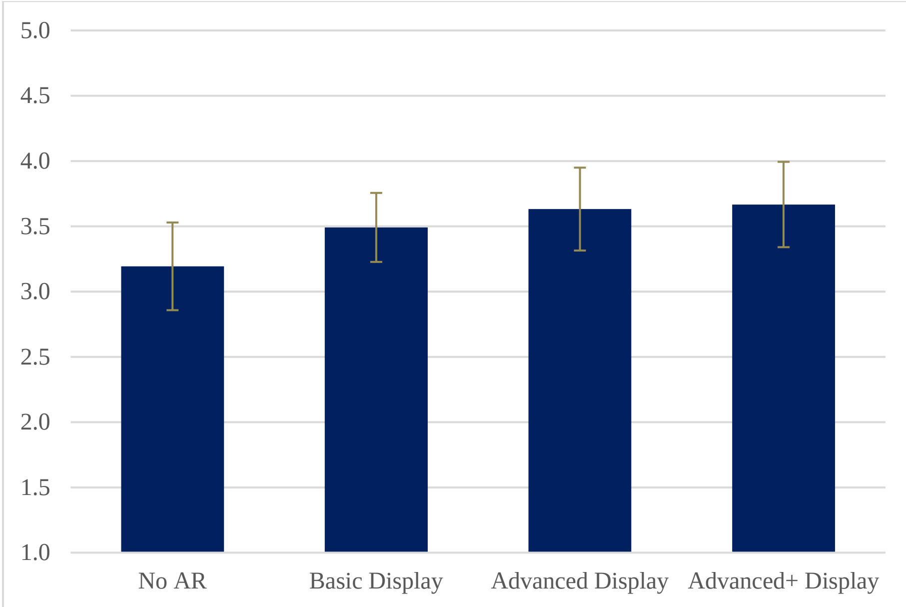

Figure 12. Trust in Automated Vehicle. Error bars represent 95% CI.

3.4.6 Usability

A 3 (Design: Basic, Advanced, Advanced+) x 2 (Reliability: No Failure, Failure) repeated measures ANOVA showed a main effect of reliability on SUS scores, $F(1, 34) = 4.69$, $p = .045$, partial $\eta^2 = .22$. The main effect of design and interaction effect were not significant, $p > .05$. Participants rated the usability of the design less when the system failed (M= 71.02, SD = 12.12) than when the system was reliable (M = 75.05, SD = 10.95). However, all designs were rated over 70. This indicates that AR HUD designs had an acceptable level of usability (Bangor, Kortum, & Miller, 2008).

3.4.7 Attention to the Road and Secondary Task

Subjective assessments of how much participants were able to attend to the road did not differ between four conditions, $F(3, 57) = .26$, $p = .85$. Participants also reported similar levels of attention to the secondary task in each condition, $F(3, 57) = .38$, $p = .77$.

3.4.8 Understanding of How the Automation Works

Subjective assessments of how well participants understood how the automated driving system works was different between conditions, $F(3, 57) = 5.49$, $p = .002$, partial $\eta^2 = .22$. Post-hoc LSD tests showed that participants reported higher levels of understanding when provided with Advanced+ display than Basic display, $p = .028$, and No AR, $p = .012$. They also reported high levels of understanding in Advanced display conditions than Basic display condition, $p = .032$ and No AR, $p = .013$. Overall Advanced and Advanced+ designs led to higher levels of subjective understanding than Basic and No AR conditions.

3.4.9 Design Preference

3.4.9.1 User perception

Participants rated the colors and visual of AR cues as appropriate. On a 5-point Likert scale, participant rated the lead vehicle cues and other vehicle cues as 4.35 ($SD = .59$), and road sign visualizations as 4.1 ($SD = .98$).

In terms of design comparisons, there were no differences between displays in reported awareness of the traffic, $F(2, 38) = 1.33$, $p = .275$. Participants did not think that any of these displays will increase driving safety more than others, Greenhouse-Geisser corrected F (1.23, 23.28) = 1.87, $p = .18$, and rated these designs similarly in terms of potential risk during take-over requests. Designs were perceived as equally capable of providing necessary information, Greenhouse-Geisser corrected F(1.49. 27.90) = 1.68, $p = .208$. There were significant differences between designs in how much distracting they were, Greenhouse-Geisser corrected $F(1.32, 25.22) = 7.75$, $p = .002$, partial $\eta^2 = .29$. Post-hoc LSD tests revealed that Advanced+ display ($M = 2.45$, $SD = 1.15$) was perceived as more distracting than Advanced ($M = 1.95$, SD = .68, $p = .008$) and Basic display ($M = 1.8$, $SD = .70$, $p = .008$. Basic and Advanced displays were perceived equally distracting, $p = .186$. Finally, participants expressed that they would use these systems equally likely, Greenhouse-Geisser corrected $F(1.13, 25.45) = .26$, $p = .69$.

3.4.9.2 Interface Preference

The majority of the participants preferred Advanced+ design (13 participants). 5 participants preferred Advanced design, one participant preferred Basic design and one participant preferred No AR. Open-ended follow-up questions revealed that participants who preferred Advanced and Advanced+ displays found the information provided very useful and appropriate in understanding the surrounding traffic. Some participants who preferred the Advanced display also complained about the distracting nature of highlighting road signs in Advanced+ display.

3.4.9.3 User Needs

Participants provided several recommendations for future design of AR HUD systems. 7 participants requested warnings, alerts, and notifications such as the distance of other cars from own car, collision alerts, possible dangerous situations (e.g. weather conditions). 6 participants wanted to see the cars behind the own car and blind-spot information. 5 participants wanted to see speed related information (both own-car speed and other cars' speed), and 5 participants wanted to see secondary information such as navigation and upcoming traffic.

3.5 Discussion

In this study, we examined augmented reality head-up displays in automated driving context. Overall, participants were positive towards AR HUD visualizations, but AR cues did not seem to provide major safety-related benefits in the context of this experiment. The AR cues led to higher levels of trust in the vehicle and better self-rated understanding of automated driving. In terms of design features, using only safety-related visualizations seem to be superior than adding secondary information such as road signs.

Adding a layer of information in the form of augmented reality cues did not result in a change in situation awareness, mental workload, and secondary task performance. It looks like participants were appropriately engaged with monitoring task regardless of this additional layer of information. One reason for this might be that the task was relatively short, easy, and safe. It is likely that the advantages of AR HUD might be more salient under circumstances where the level of risk is high, when there is not enough information (e.g. nighttime driving; Stanton, & Pinto, 2000), or when the situation is relatively complex (e.g. overtaking). In the current study, the car rarely changed lanes, and there were no road hazards. Still, participants commented that they tried to stay more alert when there were no AR cues. The lack of a difference in subjective mental workload was observed in earlier research as well (Kim, Wu, Gabbard, & Polys, 2013). Overall, participants experienced low mental workload. It seems like AR cues might help in reducing workload if the task requires substantial cognitive demands, such as navigation (Bolton, Burnett, & Large, 2015).

Trust was higher in AR HUD rides, despite the reliability of the automation was the same in each video. It is possible that participants perceive the vehicle as more capable when the sensor information is shown using AR cues. Design features impact trust in the automated system by altering user perception of these systems (Hoff & Bashir, 2015). Making the sensing capabilities more salient might likewise increase perceived trustworthiness of the system. Another explanation might be that the more participants see how automation operates (observing the process, Lee and See, 2004) in a reliable fashion, the higher they trust in the system. Nevertheless, these findings show that using AR HUD to provide information about the state of automation (in this case, the sensing capabilities) can increase trust, and facilitate user adoption, especially for those who are concerned about not knowing why the automation is doing what it is doing.

Trust seem to be affected by automation failures represented in AR HUD. While overall trust was not affected, participants reported higher levels of distrust in the automated vehicle when AR HUD indicated a sensor failure. These results show that participants were sensitive to the failures, and they adjusted their trust accordingly, which supports earlier research on trust calibration when interacting with automation (e.g. Seong & Bisantz, 2008; Wang, Jamieson, & Hollands, 2009). Appropriate trust calibration is critical in automation reliance (Lee & See, 2004), and providing real-time information about automation seem to help in this process. AR cues provide immediate and contextual information about the mapping between real world and automation, and makes automation error more salient, which is especially important in a dynamic driving environment. While providing such salient information about an unreliable system might lead to concerns on the user side, it may play a critical role in developing an appropriate level of trust in autonomous vehicles. Additionally, information about failures would help drivers to align their mental models with mental model of the vehicle, which is critical for achieving safe performance in human-automation interaction (Endsley, 1996; Sarter, Woods, & Billings, 1997)

Usability ratings did not differ between designs, but were affected by reliability, which is consistent with findings on trust. Overall the usability scores were acceptable, and participants liked the visualizations used in the study, as we observed during discussions with participants after the experiment. Taken together with findings on trust, AR HUD can be a positive factor in user adoption of automated vehicles in the future. We should note that a traditional automation display (e.g. dashboard display) can also provide similar benefits. However, based on previous work on AR HUD vs. traditional displays, we believe that the benefits will be more noticeable in AR HUD.

Subjective ratings revealed positive attitudes towards the display as the amount of information presented on the interface increases. However, overall the differences between three displays were negligible. Participants noted that as the amount of information presented increases, so was the potential for distraction. Similar findings were reported by Haeuslschmid, et al. (2015). This can be an important disadvantage, especially during take-over requests or at times when drivers become distracted and need to get back to the monitoring task (situation awareness recovery; Gartenberg, Breslow, McCurry, & Trafton, 2014). While information such as navigation often seen in commercial head-up displays can be perceived positively by drivers, they can become distractions during supervisory control. In terms of design preference, most participants preferred Advanced+ design over other designs, and welcomed the additional information this display provided. An implication of these findings is that the design of AR HUD should consider a possible information - distraction trade-off. Understanding how much information is too much requires more research, but it seems like providing additional information on AR HUD can be preferred by drivers, if such an option is available.

A distinction between this study and earlier studies is that AR HUD systems used in the past were more task oriented, e.g. noticing pedestrians (Kim, Miranda Anon, Misu, T., Li, Tawari, & Fujimura, 2016). In the current study, we tried to simulate a general monitoring task during an automated ride. Another difference was that previous work mostly considered manual driving while we used a highly automated driving context. Overall, the findings indicated several benefits of using AR HUD such as higher levels of trust, and better understanding of vehicle behavior. However, how these benefits translate into driver behavior during the safety-critical phases of automated driving such as handling unexpected situations and take-over requests, requires more research.

3.6 Limitations and Future Research

Current study had several limitations. The videos used in this study were completely safe driving footage. While this might be representative of real-world autonomous driving (i.e. the ride will be safe most of the time), it may have failed to capture the real usefulness of augmented-reality head-up displays, which is to provide the required levels of awareness in safety-critical situations, or when the driver can’t otherwise access the necessary information. It looks like the participants were not reactive to AR HUD cues but rather to the behavior of the car which was safe all the time. Another limitation was that participants had no real control over the car due to the nature of using videos. While trust ratings might serve as a proxy for reliance in automation, being able to respond to take-over requests transition from automated to manual and vice versa could reveal more information about how AR HUD cues will be used.

Future research should examine the usefulness of AR HUD cues in more realistic scenarios, and especially challenging situations where the AR HUD cues might have real safety benefits. It should also consider (1) proactive behavior, i.e. taking the control over in anticipation of an automation failure, and (2) possible negative effects of AR HUD such as complacency and disengagement from monitoring task during longer exposures.

3.7 Conclusion

In this chapter, we presented an experiment that investigated how augmented reality displays can be used in automated driving. Previous research revealed how AR HUD systems perform similarly or different from traditional in-vehicle displays, however the use of AR HUD in automated driving context was relatively understudied.

In the experiment, we examined how AR HUD may influence driver behavior in automated driving in a simulated environment. Performance measures did not show a benefit of AR cues in this context. However, AR HUD resulted in higher levels of trust in automation, increased perceived understanding of the car behavior, and were rated positively. Failures represented in AR HUD visualizations led to lower levels of trust in the automated system, which implies that using such displays can be an effective method to help drivers adjust their trust appropriately. We hope this work contributes to increase our knowledge on augmented reality displays in autonomous vehicles.

Chapter 4
Conclusion

In this thesis, our goal was to identify emerging issues related to automated driving, and address challenges previous research identified. While there has been considerable research in the past, rapid advancements in automotive industry in recent years require even more attention to addressing both existing and recently emerging challenges of human-automation interaction in this context automated driving. This work was an attempt to understand these challenges better and address them appropriately

In the study we presented in Chapter 2, we described the current situation of real-world autonomous driving, how early adopters experience these features, and identified challenges and opportunities related to the real-world usage of autonomous vehicles. Specifically, we identified that the automation failures were common, but drivers try to adapt to this new situation by engaging and learning. We also showed how drivers' trust changes over time, and its relationship to experience and driver characteristics. Also, we identified that drivers demand information in new ways, adjust their behavior based on the technology available, and that they are motivated to seek out knowledge about their vehicles, but not in the traditional manner such as reading manuals.

In the experiment, we presented in Chapter 3, we examined a future interface concept, augmented reality displays, and how it can influence driver behavior during automated driving. Our findings showed that there are possible benefits of using this technology in the automated driving context such as increased levels of trust. Regarding safety-related metrics, AR HUD did not provide a major advantage. There is more research is needed to understand a wider scope of advantages and disadvantages of using such displays. Nevertheless, we showed that AR HUD can play a role in affecting people's trust in the autonomous vehicle, and mapping sensor failures to AR cues can help drivers to become aware of the reliability of the automated vehicle.

4.1 Implications for Research

This work has several implications for future research autonomous driving. First, the survey showed that the driver behavior regarding the use of technology will change as the vehicles become smarter and connected. The concept of upgradeable car will create substantial changes in user needs and demands. Therefore, it is important to study driver needs in this new environment. Previously, driving was mostly considered as a safety-critical only activity, but given the smart and advanced technology equipped in the next generation vehicles, driving will likely evolve into being an integral part of people’s digital ecosystems. Another implication is that our findings highlight the importance of studying driver behavior in the real world. Unlike previous research that raised serious concerns regarding the use of autonomous vehicles, our findings showed a more positive picture of the current situation of automated driving. While simulator studies have been extremely useful in understanding driver behavior, which our findings mostly supported, understanding the actual impact of laboratory findings requires research in the real world. Regarding the role of technology, our findings on augmented reality displays showed that this technology can be useful in addressing some of the challenges of automated driving. It also revealed that there is more research needed to fully understand the impact of AR HUD systems in autonomous vehicles.

4.2 Implications for Design

This work provides several implications for design of future automotive HMIs. Our findings revealed that there is opportunity for automotive HMIs to take over some of the functions of one’s digital life. Designers should seek opportunities in aligning HMIs with what the drivers expect. For example, the design of vehicle manuals should be reconsidered for connected and upgradeable cars. Our experiment on augmented reality head-up displays showed that providing information about the vehicle automation can be important in facilitating user adoption. However, designers should be careful of potential risks associated with such displays. For example, providing non-task related information can be preferred by users but it may be detrimental regarding safety. Design features should be carefully studied to make a positive impact in driver well-being.

4.3 Future Research

Future research should focus on identifying factors that influence critical aspects of safe human-vehicle interaction in the context of automated driving, such as trust and situation awareness. With the partially automated vehicles becoming available in the market, more research should be conducted in the real world to identify new challenges of using such vehicle in real world environments. In-vehicle information displays, likewise should be further studied in the context of connected and autonomous vehicles. Efforts should be made to study and design interactive systems that better fits the new role of the drivers. Finally, given the acceleration of developments in the automotive industry, researchers and designers should start considering the next steps in this evolution. One step ahead of partially automated driving, level 4 and level 5 driving, will completely change the user demands. In a world with no steering wheels, there will be many questions to answer and opportunities to explore, as the vehicles evolve from tools to companions.

4.3.1 Chapter Conclusion

This work was aimed at increasing our understanding of driver behavior in autonomous vehicles by sampling a glimpse of what the future of transportation might look like. Information gained from early adopters, as well as investigation of future technologies that might become part of this evolution, should provide useful directions for researchers and designers. We hope this work facilitates further discussions, provide new perspectives, and ultimately contribute to the effort of creating safer and better technology for the future of driving.

Bibliography

Alvarez, I., Martin, A., Dunbar, J., Taiber, J., Wilson, D. M., & Gilbert, J. E. (2010, November). Voice interfaced vehicle user help. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 42-49). ACM.

Alvarez, I., López-de-Ipiña, M. K., & Gilbert, J. E. (2012, December). The voice user help, a smart vehicle assistant for the elderly. In *International Conference on Ubiquitous Computing and Ambient Intelligence* (pp. 314-321). Springer Berlin Heidelberg.

Bangor, A., Kortum, P., & Miller, J. A. (2008). The system usability scale (SUS): An empirical evaluation. *International Journal of Human-Computer Interaction*, *24*(6), 574-594.

Banks, V. A., & Stanton, N. A. (2014). Hands and feet free driving: ready or not?. *Advances in Human Aspects of Transportation: Part II, 8*, 55-63.

Banks, V. A., & Stanton, N. A. (2016). Keep the driver in control: Automating automobiles of the future. *Applied ergonomics*, *53*, 389-395.

Bark, K., Tran, C., Fujimura, K., & Ng-Thow-Hing, V. (2014, September). Personal Navi: Benefits of an Augmented Reality Navigational Aid Using a See-Thru 3D Volumetric HUD. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1-8). ACM.

Bazilinskyy, P., Kyriakidis, M., & de Winter, J. (2015). An International Crowdsourcing Study into People's Statements on Fully Automated Driving. *Procedia Manufacturing*, *3*, 2534-2542.

Bisantz, A. M., & Seong, Y. (2001). Assessment of operator trust in and utilization of automated decision-aids under different framing conditions. *International Journal of Industrial Ergonomics*, *28*(2), 85-97.

Bolton, A., Burnett, G., & Large, D. R. (2015, September). An investigation of augmented reality presentations of landmark-based navigation using a head-up display. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 56-63). ACM.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Brookhuis, K. A., De Waard, D., & Janssen, W. H. (2001). Behavioural impacts of advanced driver assistance systems–an overview. *EJTIR*, *1*(3), 245-253.

Carsten, O., Lai, F. C., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control task substitution in semiautomated driving: Does it matter what aspects are automated?. *Human factors*, *54*(5), 747-761.

Casner, S. M., Hutchins, E. L., & Norman, D. (2016). The challenges of partially automated driving. Communications of the ACM, 59(5), 70-77.

CB Insights. (2016). 33 Corporations Working On Autonomous Vehicles. Retrieved on March 16, 2017 from https://www.cbinsights.com/blog/autonomous-driverless-vehicles-corporations-list/.

Craig, A. R., Franklin, J. A., & Andrews, G. (1984). A scale to measure locus of control of behaviour. *British Journal of Medical Psychology*, *57*(2), 173-180.

de Waard, D., van der Hulst, M., Hoedemaeker, M., & Brookhuis, K. A. (1999). Driver behavior in an emergency situation in the Automated Highway System. *Transportation human factors*, *1*(1), 67-82.

De Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation research part F: traffic psychology and behaviour*, *27*, 196-217.

Dikmen, M., & Burns, C. M. (2016, October). Autonomous Driving in the Real World: Experiences with Tesla Autopilot and Summon. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 225-228). ACM.

Donmez, B., Boyle, L. N., Lee, J. D., & McGehee, D. V. (2006). Drivers' attitudes toward imperfect distraction mitigation strategies. *Transportation research part F: traffic psychology and behaviour*, *9*(6), 387-398.

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, *58*(6), 697-718.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32-64.

Endsley, M. R. (1996). Automation and situation awareness. *Automation and human performance: Theory and applications*, 163-181.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. Human factors, 37(2), 381-394.

Endsley, M. R. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462-492.

Endsley, M. R. (2017). Autonomous Driving Systems: A Preliminary Naturalistic Study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making*, 1555343417695197.

Eriksson, A., & Stanton, N. (2016). Take-over time in highly automated vehicles: non-critical transitions to and from manual control. *Human Factors*. DOI: 10.1177/0018720816685832

Ezer, N., Fisk, A. D., & Rogers, W. A. (2008). Age-related differences in reliance behavior attributable to costs within a human-decision aid system. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(6), 853-863.

Fagan, M., Khan, M. M. H., & Buck, R. (2015). A study of users' experiences and beliefs about software update messages. *Computers in Human Behavior*, *51*, 504-519.

Fagan, M., Khan, M. M. H., & Nguyen, N. (2015). How does this message make you feel? A study of user perspectives on software update/warning message design. *Human-centric Computing and Information Sciences*, *5*(1), 36.

Gabbard, J. L., Fitch, G. M., & Kim, H. (2014). Behind the Glass: Driver challenges and opportunities for AR automotive applications. *Proceedings of the IEEE*, *102*(2), 124-136.

Gartenberg, D., Breslow, L., McCurry, J. M., & Trafton, J. G. (2014). Situation awareness recovery. *Human factors*, *56*(4), 710-727.

Gibson, M., Lee, J., Venkatraman, V., Price, M., Lewis, J., Montgomery, O., ... & Foley, J. (2016). Situation Awareness, Scenarios, and Secondary Tasks: Measuring Driver Performance and Safety Margins in Highly Automated Vehicles. *SAE International Journal of Passenger Cars-Electronic and Electrical Systems*, *9*(2016-01-0145), 237-242.

Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in automation– Before and after the experience of take-over scenarios in a highly automated vehicle. *Procedia Manufacturing*, *3*, 3025-3032.

Golson, J. (2017, January 19). Driver in fatal Tesla Autopilot crash had seven seconds to take action. Retrieved March 22, 2017, from http://www.theverge.com/2017/1/19/14326604/tesla-autopilot-crash-driver-seven-seconds-inattentive-nhtsa.

Greenberg, A. (2015, July 21). Patch Your Chrysler Now Against a Wireless Hacking Attack. Retrieved March 07, 2017, from https://www.wired.com/2015/07/patch-chrysler-vehicle-now-wireless-hacking-technique/

Haeuslschmid, R., Schnurr, L., Wagner, J., & Butz, A. (2015, September). Contact-analog warnings on windshield displays promote monitoring the road scene. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 64-71). ACM.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, *52*, 139-183.

Hergeth, S., Lorenz, L., & Krems, J. F. (2016). Prior Familiarization With Takeover Requests Affects Drivers' Takeover Performance and Automation Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 0018720816678714.

Hertzum, M. (2014). Expertise seeking: A review. *Information Processing & Management*, *50*(5), 775-795.

Ho, G., Wheatley, D., & Scialfa, C. T. (2005). Age differences in trust and reliance of a medication management system. *Interacting with Computers*, *17*(6), 690-710.

Hoff, K. A., & Bashir, M. (2015). Trust in automation integrating empirical evidence on factors that influence trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(3), 407-434.

Howard, D., & Dai, D. (2014). Public perceptions of self-driving cars: The case of Berkeley, California. In *Transportation Research Board 93rd Annual Meeting* (No. 14-4502).

Hubaux, J. P., Capkun, S., & Luo, J. (2004). The security and privacy of smart vehicles. *IEEE Security & Privacy*, *2*(3), 49-55.

Hwang, Y., Park, B. J., & Kim, K. H. (2016). Effects of Augmented-Reality Head-up Display System Use on Risk Perception and Psychological Changes of Drivers. *ETRI Journal*, 38(4), 757-766.

JD Power. (2017, February 21). 2017 Vehicle Dependability Study. Retrieved March 10, 2017, from http://www.jdpower.com/press-releases/2017-vehicle-dependability-study.

Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53-71.

Jose, R. (2015). A comparative study of using Augmented Reality interfaces for vehicle navigation. Thesis. University of Canterbury. Accessed from https://ir.canterbury.ac.nz/handle/10092/11390.

Kim, H., Miranda Anon, A., Misu, T., Li, N., Tawari, A., & Fujimura, K. (2016, March). Look at Me: Augmented Reality Pedestrian Warning System Using an In-Vehicle Volumetric Head Up Display. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (pp. 294-298). ACM.

Kim, H., Wu, X., Gabbard, J. L., & Polys, N. F. (2013, October). Exploring head-up augmented reality interfaces for crash warning systems. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 224-227). ACM.

Kim, S., & Dey, A. K. (2009, April). Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 133-142). ACM.

Koehler, C., Appel, D., & Beck, H. (2016). Winning Strategies in the Race for Connected Autonomous Cars. *ATZ worldwide*, *118*(11), 20-25.

Kyriakidis, M., Happee, R., & de Winter, J. C. (2015). Public opinion on automated driving: results of an international questionnaire among 5000 respondents. *Transportation research part F: traffic psychology and behaviour*, *32*, 127-140.

Langlois, S., & Soualmi, B. (2016, November). Augmented reality versus classical HUD to take over from automated driving: An aid to smooth reactions and to anticipate maneuvers. In *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on* (pp. 1571-1578). IEEE.

Larsson, A. F. (2012). Driver usage and understanding of adaptive cruise control. *Applied ergonomics*, *43*(3), 501-506.

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243-1270.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50-80.

Lees, M. N., & Lee, J. D. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems. *Ergonomics*, *50*(8), 1264-1286.

Liu, Y. C., & Wen, M. H. (2004). Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. *International Journal of Human-Computer Studies*, *61*(5), 679-697.

Llaneras, R. E., Salinger, J., & Green, C. A. (2013, June). Human factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to the forward roadway. In *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 92-98). Iowa City: Public Policy Center, University of Iowa.

Lorenz, L., Kerschbaum, P., & Schumann, J. (2014, September). Designing take over scenarios for automated driving: How does augmented reality support the driver to get back into the loop?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 58, No. 1, pp. 1681-1685). Sage CA: Los Angeles, CA: SAGE Publications.

Mathur, A. (2016). *A human-centered approach to improving the user experience of software updates* (Doctoral dissertation).

McDonald, Z. (2016). Augmented Reality HUDs: Warning Signs and Drivers' Situation Awareness. Thesis. Rochester Institute of Technology. Accessed from http://scholarworks.rit.edu/theses/9044/

Merat, N., Jamson, A. H., Lai, F. C., & Carsten, O. (2012). Highly automated driving, secondary task performance, and driver state. *Human factors*, *54*(5), 762-771.

Milanés, V., Llorca, D.F., Villagrá, J., Pérez, J., Fernández, C., Parra, I., González, C. and Sotelo, M.A. (2012). Intelligent automatic overtaking system using vision for vehicle detection. *Expert Systems with Applications*, *39*(3), 3362-3373.

Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, *39*(3), 429-460.

National Center for Statistics and Analysis. (2016, April). Distracted driving 2014 (Traffic Safety Facts Research Note. Report No. DOT HS 812 260). Washington, DC: National Highway Traffic Safety Administration.

National Highway Traffic Safety Administration, Department of Transportation. (2016). Federal Automated Vehicles Policy. Retrieved from https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/federal_automated_vehicles_policy.pdf

Novick, D. G., & Ward, K. (2006a, October). Why don't people read the manual?. In *Proceedings of the 24th annual ACM international conference on Design of communication* (pp. 11-18). ACM.

Novick, D. G., & Ward, K. (2006b, October). What users say they want in documentation. In *Proceedings of the 24th annual ACM international conference on Design of communication* (pp. 84-91). ACM.

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors*, *56*(3), 476-488.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230-253.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. Human Factors: The Journal of the Human Factors and Ergonomics Society, 52(3), 381-410.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, *30*(3), 286-297.

Park, H. S., Park, M. W., Won, K. H., Kim, K. H., & Jung, S. K. (2013). In-vehicle AR-HUD system to provide driving-safety information. *ETRI journal*, *35*(6), 1038-1047.

Payre, W., Cestac, J., & Delhomme, P. (2014). Intention to use a fully automated car: Attitudes and a priori acceptability. *Transportation research part F: traffic psychology and behaviour*, *27*, 252-263.

Pfleging, B., Rang, M., & Broy, N. (2016). Investigating user needs for non-driving-related activities during automated driving. In Proceedings of the 15th International Conference on Mobile and

Ubiquitous Multimedia (MUM’ 16). Rovaniemi, Finland, December 12-15, 2016 (pp. 91–99). New York, NY, USA: ACM.

Phan, M. T., Thouvenin, I., & Frémont, V. (2016, November). Enhancing the driver awareness of pedestrian using augmented reality cues. In *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on* (pp. 1298-1304). IEEE.

Regan, M. A., Lee, J. D., & Young, K. (Eds.). (2008). Driver distraction: Theory, effects, and mitigation. CRC Press.

Roese, N. J., & Vohs, K. D. (2012). Hindsight bias. *Perspectives on Psychological Science*, *7*(5), 411-426.

Ross, J. M., Szalma, J. L., Hancock, P. A., Barnett, J. S., & Taylor, G. (2008, September). The effect of automation reliability on user automation trust and reliance in a search-and-rescue scenario. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 52, No. 19, pp. 1340-1344). Sage CA: Los Angeles, CA: SAGE Publications.

Rudin-Brown, C. M., & Parker, H. A. (2004). Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies. Transportation Research Part F: Traffic Psychology and Behaviour, 7(2), 59-76.

SAE International. (2014). Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (J3016). http://standards.sae.org/ j3016_201401/.

Sanchez, J., Fisk, A. D., & Rogers, W. A. (2004, September). Reliability and age-related effects on trust and reliance of a decision support aid. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 48, No. 3, pp. 586-589). Sage CA: Los Angeles, CA: SAGE Publications.

Sanchez, J., Rogers, W. A., Fisk, A. D., & Rovira, E. (2014). Understanding reliance on automation: Effects of error type, error distribution, age and experience. *Theoretical Issues in Ergonomics Science*, *15*(2), 134-160.

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. *Handbook of human factors and ergonomics*, *2*, 1926-1943.

Sauer, J., & Chavaillaz, A. (2017). The use of adaptable automation: Effects of extended skill lay-off and changes in system reliability. *Applied Ergonomics*, *58*, 471-481.

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human factors*, *58*(3), 377-400.

Schoettle, B., & Sivak, M. (2014). Public opinion about self-driving vehicles in China, India, Japan, the U.S., the U.K., and Australia. Retrieved June 8, 2016 from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/109433/103139.pdf.

Schroeter, R., & Steinberger, F. (2016, November). Pokémon DRIVE: towards increased situational awareness in semi-automated driving. In Proceedings of the 28th Australian Conference on Computer-Human Interaction (pp. 25-29). ACM.

Schwarz, F., & Fastenmeier, W. (2017). Augmented reality warnings in vehicles: Effects of modality and specificity on effectiveness. *Accident Analysis & Prevention*, *101*, 55-66.

Seong, Y., & Bisantz, A. M. (2008). The impact of cognitive feedback on judgment performance and trust with decision aids. *International Journal of Industrial Ergonomics*, *38*(7), 608-625.

Software updates. (2016, October 27). Retrieved March 06, 2017, from https://www.tesla.com/support/software-updates.

Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: safety implications of automation in vehicles. *Safety Science*, *24*(1), 35-49.

Stanton, N., & Marsden, P. (1997). Drive-by-wire systems: Some reflections on the trend to automate the driver role. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 211(4), 267-276.

Stanton, N. A., & Pinto, M. (2000). Behavioural compensation by drivers of a simulator when using a vision enhancement system. *Ergonomics*, *43*(9), 1359-1370.

Stanton, N., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2013). *Human factors methods: a practical guide for engineering and design*. Ashgate Publishing, Ltd.

Stanton, N. A., & Young, M. S. (1998). Vehicle automation and driving performance. *Ergonomics*, *41*(7), 1014-1028.

Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, *1*(4), 315-331.

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, *48*(10), 1294-1313.

Taylor, R. M. (1990). Situational Awareness Rating Technique(SART): The development of a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 17 p(SEE N 90-28972 23-53)*.

Tesla Motors. (2016). Model S Owner's Manual. Retrieved May 31, 2016 from https://www.teslamotors.com/sites/default/files/model_s_owners_manual_touchscreen_7.1_das_ap_north_america_r20160112_en_us.pdf.

Tränkle, U., Gelau, C., & Metker, T. (1990). Risk perception and age-specific accidents of young drivers. *Accident Analysis & Prevention*, *22*(2), 119-125.

Trösterer, S., Gärtner, M., Mirnig, A., Meschtscherjakov, A., McCall, R., Louveton, N., Tscheligi, M., & Engel, T. (2016, October). You Never Forget How to Drive: Driver Skilling and Deskilling in the Advent of Autonomous Vehicles. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 209-216). ACM.

Vaniea, K., & Rashidi, Y. (2016, May). Tales of software updates: The process of updating software. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3215-3226). ACM.

Viereckl, R., Koster, A., Hirsh, E., Ahlemann, D. (2016, September 27). Connected car report 2016: Opportunities, risk, and turmoil on the road to autonomous vehicles. Retrieved March 07, 2017, from http://www.strategyand.pwc.com/reports/connected-car-2016-study

Wang, L., Jamieson, G. A., & Hollands, J. G. (2009). Trust and reliance on an automated combat identification system. *Human factors, 51*(3), 281-291.

Young, K., Lee, J. D., & Regan, M. A. (Eds.). (2008). *Driver distraction: Theory, effects, and mitigation*. CRC Press.

Young, M. S., & Stanton, N. A. (2007). Back to the future: Brake reaction times for manual and automated vehicles. *Ergonomics, 50(*1), 46-58.

Yuan, Y. C., Rickard, L. N., Xia, L., & Scherer, C. (2011). The interplay between interpersonal and electronic resources in knowledge seeking among co-located and distributed employees. *Journal of the American Society for Information Science and Technology*, *62*(3), 535-549.

Appendix A
Survey Questions

In this Appendix, we present the questions used in the survey. Duplicate questions were omitted.

Driving Experience

2) What is your age?

[ ] 16 - 20

[ ] 21 - 24

[ ] 25 - 34

[ ] 35 - 44

[ ] 45 - 54

[ ] 55 - 64

[ ] 65 or older

3) What is your gender?

[ ] Female

[ ] Male

[ ] Other:

4) How would you rate yourself as a computer user? (Your general knowledge of using computers)

[ ] Novice user

[ ] Below average user

[ ] Average user

[ ] Above average user

[ ] Expert user

5) How many years of driving experience do you have?

[ ] Less than 1 year

[ ] 1 - 3 years

[ ] 3 - 5 years

[ ] 5 - 10 years

[ ] More than 10 years

6) How often do you drive?

[ ] Every day

[ ] A few days a week

[ ] A few days a month

[ ] A few days a year

7) Do you currently own a Tesla Model S?

[ ] Yes

[ ] No

Your Experience with Model S

8) When did you first own a Tesla Model S?

[ ] 2016

[ ] 2015

[ ] 2014

[ ] 2013

[ ] 2012

[ ] 2011

[ ] 2010

[ ] 2009

[ ] 2008

9) Which year and model are you currently using?

10) Which version of the Tesla software are you currently using on your car?

[ ] I don't know

[ ] 7.1

[ ] 7.0

[ ] 6.2

[ ] 6.1

[ ] 6.0

[ ] 5.14

[ ] 5.12

[ ] 5.11

[ ] 5.9

[ ] 5.8.10

[ ] 5.8.8

[ ] 5.8.7

[ ] 5.8.4

[ ] 5.8

[ ] 5.6

[ ] 5.0

[ ] 5.5

[ ] 4.5

[ ] 4.4

[ ] 4.3

[ ] 4.2

[ ] 4.1

[ ] 4.0

[ ] Other

11) Overall, how would you rate your experience with your car?

[ ] Very dissatisfied

[ ] Somewhat dissatisfied

[ ] Neither satisfied nor dissatisfied

[ ] Somewhat satisfied

[ ] Very satisfied

12) In general, how often do you consult owner's manual when you want to look up information about the features of your car?

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

13) In general, how often do you consult friends/colleagues when you want to look up information about the features of your car?

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

14) In general, how often do you consult online sources (for example, forums and websites) when you want to look up information about the features of your car?

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

15) Are there other sources you use to learn about the features of your car?

(Please describe the sources)

16) How would you rate the usefulness of the owner’s manual in teaching you about the features of the car?

[ ] Not at all useful

[ ] Not very useful

[ ] Somewhat useful

[ ] Very useful

[ ] Extremely useful

17) How do you access the owner’s manual when you need it?

Check all that apply.

[ ] Read on my computer

[ ] Read on my smartphone

[ ] Read on my tablet

[ ] Read on the display in the car

[ ] Read printed manual

[ ] Other:

18) How useful would it be to receive videos explaining new features after an update?

[ ] Not at all useful

[ ] Not very useful

[ ] Somewhat useful

[ ] Very useful

[ ] Extremely useful

19) How useful would it be to have the opportunity to test new features before using them during actual driving? (For example, using a simulator-like system)

[ ] Not at all useful

[ ] Not very useful

[ ] Somewhat useful

[ ] Very useful

[ ] Extremely useful

20) Which of the following applies to you about Autopilot feature?

[ ] I am currently using Autopilot feature or have used it in the past.

[ ] My car does not support Autopilot feature.

[ ] My car supports Autopilot feature but I have not installed the update.

[ ] I don't know what Autopilot is.

Your Experience with Autopilot

21) How often do you use Autopilot?

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

22) When did you update your car after the Autopilot update became available?

[ ] Within a week

[ ] Within a month

[ ] Within two months

[ ] Later than two months

23) How did you learn about the new features of the car after the Autopilot update?

Please check all that apply.

[ ] Read the manual / release notes

[ ] Asked friends/colleagues

[ ] Asked company representatives

[ ] Used online forums

[ ] Used websites

[ ] Other:

24) How confident were you when using Autopilot for the first time?

[ ] Not at all confident

[ ] Not very confident

[ ] Moderately confident

[ ] Very confident

[ ] Extremely confident

25) How much did you trust Autopilot during your initial experience?

[ ] 1 - Didn't trust at all

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Completely trusted

26) How useful is the Autopilot display?

[ ] Not at all useful

[ ] Not very useful

[ ] Somewhat useful

[ ] Very useful

[ ] Extremely useful

27) While driving with Autopilot, have you experienced any unexpected or unusual behavior from the car?

[ ] Yes

[ ] No

28) Can you elaborate on this unexpected situation?

What were the road conditions?

___

What were you doing?

___

What did the car do?

___

Why did you think the car behaved in such a way?

___

What did you do afterwards?

___

29) How would you rate this situation in terms of the risks involved?

[ ] Not at all risky

[ ] Not too risky

[ ] Somewhat risky

[ ] Very risky

[ ] Extremely risky

36) If you would like to introduce Autopilot to a friend/colleague, what would you recommend / emphasize for safe driving?

37) What kind of features and capabilities do you expect from Autopilot in the future?

38) Overall, how would you rate your knowledge about how the car makes decisions when Autopilot is turned on?

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

39) Overall, how would you rate the difficulty of learning how to drive with Autopilot?

[ ] Very difficult

[ ] Difficult

[ ] Moderate

[ ] Easy

[ ] Very Easy

40) For you, how important it is to know how Autopilot makes decisions?

[ ] Not important

[ ] Slightly important

[ ] Moderately important

[ ] Very important

[ ] Extremely important

41) Please indicate the extent to which you agree with the following statements about Autopilot

System here refers to Autopilot

I am suspicious of the system's intent, action, or outputs

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I am wary of the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

The system's actions will have a harmful or injurious outcome

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I am confident in the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

The system is reliable

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I can trust the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I think that I would need the support of a technical person to be able to use this system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I needed to learn a lot of things before I could get going with this system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

Autopilot Update

42) Do you plan to update your software in the near future?

[ ] Yes

[ ] No

43) Can you elaborate on reasons for not updating your car software in the near future?

Summon

47) Which of the following applies to you about the Summon feature?

[ ] I am currently using Summon or have used it in the past

[ ] My car does not support Summon.

[ ] My car supports Summon but I have not installed the update.

[ ] I don't know what Summon is.

Your Experience with Summon

48) How often do you use Summon?

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

49) When did you update your car after the Summon update became available?

[ ] Within a week

[ ] Within a month

[ ] Within two months

[ ] Later than two months

50) How did you learn about the new features of the car after the Summon update?

Please check all that apply.

[ ] Read the manual / release notes

[ ] Asked friends/colleagues

[ ] Asked company representatives

[ ] Used online forums

[ ] Used websites

[ ] Other:

51) How confident were you when using the Summon feature for the first time?

[ ] Not at all confident

[ ] Not very confident

[ ] Moderately confident

[ ] Very confident

[ ] Extremely confident

52) How much did you trust Summon during your initial experience?

[ ] 1 - Didn't trust at all

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Completely trusted

53) While using Summon, have you experienced any unexpected or unusual behavior from the car?

[ ] Yes

[ ] No

54) Can you elaborate on this unexpected situation?

What were the environment conditions?

___

What were you doing?

___

What did the car do?

___

Why did you think the car behaved in such a way?

___

What did you do afterwards?

___

55) How would you rate this situation in terms of the risks involved?

[ ] Not at all risky

[ ] Not too risky

[ ] Somewhat risky

[ ] Very risky

[ ] Extremely risky

62) If you would like to introduce Summon to a friend/colleague, what would you recommend / emphasize for safe driving?

63) What kind of features and capabilities do you expect from Summon in the future?

64) Overall, how would you rate your knowledge about how the car makes decisions when using Summon?

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

65) Overall, how would you rate the difficulty of learning how to park with Summon?

[ ] Very difficult

[ ] Difficult

[ ] Moderate

[ ] Easy

[ ] Very Easy

66) For you, how important it is to know how Summon makes decisions?

[ ] Not important

[ ] Slightly important

[ ] Moderately important

[ ] Very important

[ ] Extremely important

67) Please indicate to the extent that you agree with the following statements about Summon

System here refers to Summon

I am suspicious of the system's intent, action, or outputs

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I am wary of the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

The system's actions will have a harmful or injurious outcome

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I am confident in the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

The system is reliable

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I can trust the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I think that I would need the support of a technical person to be able to use this system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

I needed to learn a lot of things before I could get going with this system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly agree

Summon Update

68) Do you plan to update your software in the near future?

[ ] Yes

[ ] No

69) Can you elaborate on reasons for not updating your car software in the near future?

Non-owners

73) Do you plan to buy a Tesla car in the future?

[ ] Yes

[ ] No

Appendix B
Experimental Material

In this Appendix, we present the questionnaires and scales used in the experiment.

Pre-experiment Questionnaire

This questionnaire was presented at the beginning of the experiment.

Demographics and Driving Experience

What is your age?

What is your gender?

[ ] Male

[ ] Female

[ ] Prefer not to say

[ ] Other:

What is your major?

What is your Driver's License class?

[ ] Class G1

[ ] Class G2

[ ] Full Class G

[ ] Other:

Do you have normal vision or corrected-to-normal vision (e.g. glasses, contact lenses)?

[ ] Yes

[ ] No

How would you rate yourself as a computer user? (Your general knowledge of using computers)

[ ] Novice user

[ ] Below average user

[ ] Average user

[ ] Above average user

[ ] Expert user

For how long do you have your current driver’s license?

For how long have you been driving?

How often do you drive?

[ ] Every day

[ ] Almost every day

[ ] A few days a week

[ ] A few days a month

[ ] A few days a year

On average, how many kilometers do you drive in a month?

Do you have highway driving experience?

[ ] Yes

[ ] No

AR HUD Experience

Do you have experience using an augmented reality display before?

[ ] Yes

[ ] No

[ ] I don't know what an augmented reality display is.

Please rate your knowledge about augmented reality displays

[ ] 1 - Not at all knowledgable

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Extremely knowledgable

Advanced Driver Assistance Systems (ADAS) Experience

Please rate your knowledge regarding the following Driver Assistance Systems

Adaptive Cruise Control

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Automatic Parking

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Collision Avoidance System

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Cruise Control

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Forward Collision Warning

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Blind Spot Monitor

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Lane Departure Warning System

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Lane Change Assistance

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Navigation

[ ] Not at all knowledgable

[ ] Not too knowledgable

[ ] Somewhat knowledgable

[ ] Very knowledgable

[ ] Extremely knowledgable

Do you actively use any of the following Driver Assistance Systems while driving?

Adaptive Cruise Control

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

Automatic Parking

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

Collision Avoidance

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

System Cruise Control

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

Forward Collision Warning

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

Blind Spot Monitor

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

Lane Departure Warning System

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

Lane Change Assistance

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

Navigation

[ ] Never

[ ] Rarely

[ ] Sometimes

[ ] Often

[ ] Always

Locus of Control Scale (Craig, Franklin, & Andrews, 1984)

This scale measures locus of control, beliefs about the cause of the events in one's life. Locus of control can be internal or external. A person with an internal locus of control (a low score in this scale) believes that they can control the flow of the events happening in their lives. A person with an external locus of control (a high score in this scale) believes that the cause of the events is external, and they have little control over the outcomes. Locus of control can be important in understanding trust in automated vehicles (Stanton & Young, 2000). In general, people with an external locus of control should be more comfortable with vehicle automation than people with an internal locus of control.

The following questions are rated on a 5-point Likert scale, ranging from "Strongly Disagree" to "Strongly Agree". After the scores for items 1, 5, 7, 8, 13, 16 are reversed, rating for all items are summed to create a locus of control score. Higher scores indicate an external locus of control, and lower scores indicate an internal locus of control.

1 I can anticipate difficulties and take action to avoid them.

2 A great deal of what happens to me is probably just a matter of chance.

3 Everyone knows that luck or chance determine one’s future.

4 I can control my problem(s) only if I have outside support.

5 When I make plans, I am almost certain that I can make them work.

6 My problem(s) will dominate me all my life.

7 My mistakes and problems are my responsibility to deal with.

8 Becoming a success is a matter of hard work, luck has little or nothing to do with it.

9 My life is controlled by outside actions and events.

10 People are victims of circumstance beyond their control.

11 To continually manage my problems I need professional help.

12 When I am under stress, the tightness in my muscles is due to things outside my control.

13 I believe a person can really be a master of his fate.

14 It is impossible to control my irregular and fast breathing when I am having difficulties.

15 I understand why my problem(s) varies so much form one occasion to the next.

16 I am confident of being able to deal successfully with future problems.

17 In my case maintaining control over my problem(s) is due mostly to luck.

Video Questionnaire

This questionnaire was presented after each video.

National Aeronautics and Space Administration-Task Load Index (NASA-TLX; Hart & Staveland, 1988)

NASA-TLX measures subjective mental workload by using six questions that correspond to six different demands. The items are scored on a 100-point scale divided into 20 equal intervals. Depending on the research question, an overall workload score or scores from relevant subscales can be used.

NASA-TLX features the following items (dimensions):

1. How mentally demanding was the task? (Mental Demand)
2. How physically demanding was the task? (Physical Demand)
3. How hurried or rushed was the pace of the task? (Temporal Demand)
4. How successful were you in accomplishing what you were asked to do? (Performance)
5. How hard did you have to work to accomplish your level of performance? (Effort)
6. How insecure, discouraged, irritated, stressed, and annoyed were you? (Frustration)

Situation Awareness Rating Scale (SART; Taylor, 1990)

SART measures subjective situation awareness by using 10 dimensions of situation awareness (SA). Each dimension is asked on a 7-point Likert scale (1: Low, 7: High). Additionally, the dimensions are grouped into three categories, namely demand, supply, and understanding. The following formula used to calculate situation awareness score:

SA = Understanding – (Demand – Supply).

SART consists of the following items:

DEMAND

Instability of Situation: How changeable is the situation? Is the situation highly unstable and likely to change suddenly (high), or is it very stable and straightforward?

Complexity of Situation: How complicated is the situation? Is it complex with many interrelated components (high) or is it simple and straightforward (low)?

Variability of Situation: How many variables are changing in the situation? Are there a large number of factors varying (high) or are there very few variables changing (low)?

SUPPLY

Arousal: How aroused are you in the situation? Are you alert and ready for activity (high) or do you have a low degree of alertness (low)?

Concentration of Attention: How much are you concentrating on the situation? Are you bringing all your thoughts to bear (high) or is your attention elsewhere (low)?

Division of Attention: How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (high) or focused on only one (low)?

Spare Mental Capacity: How much mental capacity do you have to spare in the situation? Do you have sufficient to attend to many variables (high) or do you have nothing to spare at all (low)?

UNDERSTANDING

Information Quantity: How much information have you gained about the situation? Have you received and understood a great deal of knowledge (high) or very little (low)?

Information Quality: How good is the information you have gained about the situation? Is the knowledge communicated very useful (high) or not useful (low)?

Familiarity with the Situation: How familiar are you with the situation? Do you have a great deal of relevant experience (high) or is it a new situation (low)?

Ride-related Questions

How would you rate this ride in terms of the risks involved?

[ ] 1 - Not at all risky

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Extremely risky

To what extent were you able to focus your attention on monitoring the road and traffic? Mark only one oval.

[ ] 1 - Not at all

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Very much

To what extent were you able to focus your attention on word puzzle?

[ ] 1 - Not at all

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Very much

To what extent did you understand how the system (self--driving car) works?

[ ] 1 - Not at all

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Very much

How many service centre signs did you see during the ride?

Was there construction on the road?

How many exits were there?

Trust (Items taken from "Checklist for Trust between People and Automation", Jian, Bisantz, & Drury, 2000)

Please rate the following statements about the self--driving system you have seen in the video.

I am suspicious of the system's intent, action, or outputs

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly Agree

I am wary of the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly Agree

The system's actions will have a harmful or injurious outcome

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly Agree

I am confident in the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly Agree

The system is reliable I can trust the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither agree nor disagree

[ ] Agree

[ ] Strongly Agree

System Usability Scale (SUS; Brooke, 1996)

This scale measures several aspects of usability such as effectiveness, efficiency, and satisfaction by using 10 items rated on 5-point Likert scales (1: Strongly Disagree, 5: Strongly Agree). Items are scored as follows:

- Reverse the scores for items 2, 4, 6, 8 and 10.
- Subtract 1 from each score.
- Sum the scores, and multiply by 2.5 to obtain an overall usability rating ranging from 0 to 100.

SUS consists of the following items:

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

Post-experiment Questionnaire

This questionnaire was presented at the end of the experiment.

To what extent did the task (experiment) feel similar to real world?

[ ] 1 - Very dissimilar

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Very similar

To what extent have you noticed system failures (e.g. failure to appropriately detect the objects) during the experiment?

[ ] 1 - Never noticed a system

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Always noticed when there was a system failure

AR visualizations

Please rate the following statements about the visuals

The color used for highlighting the lead vehicle was appropriate

[ ] 1 - Strongly disagree

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Strongly agree

The color used for highlighting other vehicles on the road was appropriate

[ ] 1 - Strongly disagree

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Strongly agree

The visuals used for highlighting road signs were appropriate

[ ] 1 - Strongly disagree

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Strongly agree

Attitudes Towards AR HUD Designs

The following set of questions were asked for each AR HUD (Basic, Advanced, Advanced+)

It was easy to understand what the car was doing

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

[ ]

I was aware of what was going on regarding the traffic

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

The interface was distracting

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

It may pose a risk if I have to take the control over and switch to manual driving

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

It will increase driving safety

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

I would actively use this system if I had a semi--autonomous car

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

It is capable of providing necessary information

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

Design Preference

Of the 3 Interfaces, which one would you prefer?

[ ] Interface 1

[ ] Interface 2

[ ] Interface 3

[ ] No Interface

Why?

Questions about Self-Driving Cars in General (Includes items taken from "Checklist for Trust between People and Automation", Jian, Bisantz, & Drury, 2000)

Please rate the following statements about self--driving cars in general (system here refers to self-driving cars)

I am suspicious of the system's intent, action, or outputs

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

I am wary of the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

The system's actions will have a harmful or injurious outcome

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

I am confident in the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

The system is reliable

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

I can trust the system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

I think that I would need the support of a technical person to be able to use this system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

I needed to learn a lot of things before I could get going with this system

[ ] Strongly disagree

[ ] Disagree

[ ] Neither disagree nor agree

[ ] Agree

[ ] Strongly Agree

To what extent do you see yourself using a self--driving car in the near future? (Assuming they will become widely available)

[ ] 1 - Extremely unlikely

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Extremely likely

For you, how important it is to know how a self--driving car makes decisions?

[ ] 1 - Not at all important

[ ] 2

[ ] 3

[ ] 4

[ ] 5 - Extremely important

Open-ended AR HUD Question

In a real world heads-up display for self-driving cars, what would you like to see on the interface (what kind of information/objects/notifications/other features etc.)?
