Psychology Theses & Dissertations   Psychology

Spring 2019

Whose Drive Is It Anyway? Using Multiple Sequential Drives to Establish Patterns of Learned Trust, Error Cost, and Non-Active Trust Repair While Considering Daytime and Nighttime Differences as a Proxy for Difficulty

Scott Mishler
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/psychology_etds

Part of the Cognitive Psychology Commons, and the Transportation Commons

WHOSE DRIVE IS IT ANYWAY? USING MULTIPLE SEQUENTIAL DRIVES TO ESTABLISH PATTERNS OF LEARNED TRUST, ERROR COST, AND NON-ACTIVE TRUST REPAIR WHILE CONSIDERING DAYTIME AND NIGHTTIME DIFFERENCES AS A PROXY FOR DIFFICULTY

by

Scott Mishler
B.S. December 2016, Purdue University

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY
May 2019

Approved by:

Jing Chen (Director)

James P. Bliss (Member)

Kristin Heron (Member)

ABSTRACT

WHOSE DRIVE IS IT ANYWAY? USING MULTIPLE SEQUENTIAL DRIVES TO
ESTABLISH PATTERNS OF LEARNED TRUST, ERROR COST, AND TRUST REPAIR
WHILE CONSIDERING DAYTIME AND NIGHTTIME DIFFERENCES AS A PROXY FOR
DIFFICULTY

Scott Mishler
Old Dominion University, 2019
Director: Dr. Jing Chen

Semi-autonomous driving is a complex task domain with a broad range of problems to consider. The human operator's role in semi-autonomous driving is crucial because safety and performance depends on how the operator interacts with the system. Drive difficulty has not been extensively studied in automated driving systems and thus is not well understood. Additionally, few studies have studied trust development, decline, or repair over multiple drives for automated driving systems. The goal of this study was to test the effect of perceived driving difficulty on human trust in the automation and how trust is dynamically learned, reduced due to automation errors, and repaired over a seven-drive series. The experiment used 2 task difficulty conditions (easy vs. difficult) x 3 error type conditions (no error, takeover request or TOR, failure) x 7 drives mixed design. Lighting condition was used as a proxy for driving difficulty because decreased visibility for potential hazards could make monitoring the road difficult. During the experiment, 122 undergraduate participants drove an automated vehicle seven times in either a daytime (i.e., "easy") or nighttime (i.e., "difficult") condition. Participants experienced a critical hazard event in the fourth drive, in which the automation perfectly avoided the hazard ("no error" condition), issued a takeover request ("TOR" condition), or failed to notice and respond to the hazard ("failure" condition). Participants completed trust ratings after

each drive to establish trust development. Results showed that trust improved through the first three drives, demonstrating proper trust calibration. The TOR and automation failure conditions saw significant decreases in trust after the critical hazard in drive four, whereas trust was unaffected for the no error condition. Trust naturally repaired in the TOR and failure conditions after the critical event but did not recover to previous levels before the critical event. There was no evidence of perceived difficulty differences between the daytime and nighttime conditions. Thus, a consistent lack of trust differences was found between lighting conditions. This study demonstrated how trust develops and responds to errors in automated driving systems, informing future research for trust repair interventions and design of automated driving systems.

This thesis is dedicated to my incredible parents, Ty and Christy Mishler for their endless

love and support.

ACKNOWLEDGMENTS

NOMENCLATURE

ADOS Automated Driving Opinion Survey

DDT Dynamic Driving Task

RT Reaction Time

TOR Takeover Request

TTC Time To Collision

UAS Usefulness of Automated System

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Autonomous vehicles are quickly becoming a prominent factor in society. Many companies new to automobiles like Tesla, Uber, and Google are trying to develop autonomous vehicles and the already established automotive companies like GM, Toyota, and BMW are implementing more autonomous features. Features like lane-assist and automatic braking are already present in many new cars, and fully autonomous vehicles are currently being tested. However, this technology is not perfect and sometimes errors occur due to the current limitations of the automation. These automation errors create doubt for potential consumers. The level of public trust in autonomous vehicles could determine how they will be utilized in the future. If humans do not trust the automation to do its job, they will disuse the system, losing any benefits it can provide. Conversely, if humans rely on the automation too much, it will lead to misuse of the system. Misuse will cause any errors from the system to go unnoticed by the human, potentially resulting in fatal consequences (Parasuraman & Riley, 1997).

Considerable research has examined human trust in automation and the various factors that influence it (Hoff & Bashir, 2015; Lee & Moray, 1994; Lee & See, 2004; Schaefer, Chen, Szalma, & Hancock, 2015). Additionally, much research has focused on the best ways to design automated vehicles to ensure trust is at the proper level of calibration (Fredrick, Mikael, & Jana, 2017; Helldin, Falkman, Riveiro, & Davidsson, 2013; Verberne, Ham, & Midden, 2012). Among different types of trust, situational trust considers factors that are dependent on context and therefore affect the development of trust (Hoff & Bashir, 2015). Many external variables in automated driving are beyond the driver's control, such as the design and features of a system, setting and environment, and task difficulty. Factors like expertise, self-confidence, and previous

experiences are a part of the individual that can vary depending on the task. All these factors are important to consider because they directly influence trust development and trust level (Hoff & Bashir, 2015).

Even though these various situational factors can greatly influence trust, a search of the literature yielded few studies examining how a difference in driving setting, environment, or context can have differential effects on trust. Something as common as driving to work during the day when the sun is out and then driving home when the sun has already set demonstrates a huge change in driving context. These two situations are different for manual and automated driving and could change how someone interacts with the automation. Nighttime driving could be perceived as more difficult than daytime driving due to lower visibility and decreased object details (Konstantopoulos, Chapman, & Crundall, 2010; Leibowitz & Owens, 1997; Plainis & Murray, 2002). Some passengers in a manual vehicle might be less trusting of a human driver during nighttime because of the added difficulty. However, automation could reduce concern for nighttime driving compared to a human-driven vehicle. Automated vehicles have more navigation and vision tools to navigate at night than human drivers, so the automation is likely safer than a human driver at night. However, depending on an individual's perception of the situation, their trust could vary for these two lighting conditions.

Research examining how trust in automated systems changes with automation errors in easy and difficult tasks has been done in non-driving scenarios. Madhavan, Wiegmann, and Lacson (2006) found in basic signal detection experiments that if an automated aid fails on a task that is easy for the operator to perform, this has a stronger negative effect on trust than for a difficult-to-perform task. Similarly, human drivers may suffer greater decreases in trust when an automated vehicle fails to avoid a hazard in an easy driving condition than in a difficult driving

condition. Driving conditions are constantly changing and this could affect human behavior. Thus, the current study focused on daytime and nighttime driving conditions, examining how automation errors in each context affected trust in automation. This finding would inform autonomous vehicle designers whether the automation system might need to adapt depending on the external conditions to account for the human factor. Additionally, the study demonstrated a trajectory of trust development over seven drives, encapsulating learned trust, decline of trust due to errors, and rebuilding of trust to add to the theoretical literature of trust development in an automated driving scenario.

1.1 Autonomous Vehicles

Autonomous vehicles were once only a dream of science fiction but are now currently being studied and developed for consumer use. Autonomous vehicles will afford humans safer travel to and from any destination. In 2015, 35,092 people died in motor vehicle accidents and 2,443,000 were injured (National Center for Statistics and Analysis, 2016). Because approximately 94% of these incidents were caused by human error (Singh, 2018), autonomous vehicles could help to limit the mistakes of human drivers (Anderson et al., 2016; Fagnant & Kockelman, 2015; Kalra & Paddock, 2016). Automation has significant benefits over human drivers. It is never subject to adverse emotional or physiological states, has better lines of sight and vision of the road and traffic, has faster processing speed for multiple concurrent sources of data, and better technical driving skill than most human operators (Kalra & Paddock, 2016).

The safety of the driver relies on proper assessment of the capabilities of the automation in relation to the operational environment, what the automation is going to do, and if he or she needs to intervene (Katrakazas, Quddus, Chen, & Deka, 2015; Molloy & Parasuraman, 1996; Parasuraman, Sheridan, & Wickens, 2008). Automated driving systems may not be able to

perfectly identify the hazards in the road every single time. Automation is not 100% perfect, and is still susceptible to errors or certain conditions that it is not currently able to operate in. As a result, the human driver may need to take over control of the car or assist the automation (Leonard et al., 2009). Identification and execution of safe maneuvers that are obvious to the human driver may be more difficult for the automated system (Brown & Laurier, 2017; Endsley, 2017; Parasuraman, Hancock, & Olofinboba, 1997). Therefore, the human driver might need to compensate for the limitations of automation. However, not all autonomous vehicles have equivalent capabilities. These differences are an important consideration for research and design of autonomous vehicles.

Autonomous vehicles are defined on a continuum based on how much human involvement is required (SAE, 2016). Level 5 refers to a fully autonomous vehicle that requires no human monitoring or interaction, while Level 0 vehicles have no autonomous function and require the human to operate all aspects of the driving task. The intermediate levels have a mix of different amounts of human and automation driving. Currently, the highest level of commercially available cars is level 2 automation. Level 2 automation cannot operate effectively on all types of roads or driving conditions, limited mostly to highway driving. The vehicle is capable of lateral and longitudinal motion, but the driver must continuously monitor the environment and take over immediately if the system's limitations are exceeded. Even though the current highest level of automation is Level 2, the driver may often see it as something like Level 3 or even Level 4, with which they can, and do, engage in secondary tasks, neglecting their responsibility to monitor the operational driving environment and be ready to take over (Carsten, Lai, Barnard, Jamson, & Merat, 2012; De Winter, Happee, Martens, & Stanton, 2014).

Automation is always improving and is very close to level 3 with some companies like Audi claiming that the A8 is already capable of level 3 automation. At level 3, the car is capable of monitoring the driving environment and the driver must only be prepared to respond to a "request to intervene" if the automation determines it needs the human driver to take over. Therefore, it is necessary to consider both the current level of automation, and what is coming in the near future to be prepared when technology catches up. The current study exists between Level 2 and Level 3, with a system that can perform sustained lateral and longitudinal vehicle motion control, recognize and avoid objects in the driving environment, but still relies on the human to monitor the system in case of error and act as a fallback through requests to intervene.

1.2 Takeover Requests

Takeover requests (TOR) are requests to intervene issued by an autonomous vehicle to the driver telling the driver to resume control of the vehicle, because the system is no longer capable of safe driving (Dogan et al., 2017; Endsley & Kiris, 1995; Koo et al., 2015; Sirkin, Martelaro, Johns, & Ju, 2017). TORs can arise from a change in road conditions or because of a lack of necessary information for effective performance. An example is encountering unexpected objects in the road (Leonard et al., 2009; Walch, Lange, Baumann, & Weber, 2015).

The automation could issue warnings about a hazard it has noticed and needs the driver's assistance to take over control of the vehicle to maintain safe driving behavior. Proper TOR behavior, as explained by Walch et al. (2015), is done by simultaneously interrupting all secondary behavior (e.g. music or visualizations on the infotainment display), alerting the driver about the situation, and issuing a request for the driver to take over. If the driver realizes the situation is a problem, he or she can then resume driving manually.

The automation can issue a TOR if a situation is considered outside of its operating capabilities. However, the automation may not always judge this quickly enough. Additionally, the automation might not detect a hazard in the driving lane. In this case, the human driver must be ready to take over control of the vehicle manually, without any warning or announcement of a TOR. This is a critical situation and constitutes a complete failure of the automation because the human was required to intervene without warning. In contrast, a TOR is not technically considered a failure of the automation because it is issuing a warning based on its limitations (SAE, 2016). Detection failure situations should become rare as the technology improves and advances toward higher levels of automation, although these situations still exist.

1.3 Trust in Automation

For automated driving to be effective, the human driver will need to appropriately calibrate his or her level of trust in the automation to safely monitor the system for any errors, but still allow the system to perform its job and gain the benefits of the automated driving. Lee and See (2004) have defined human-automation trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (p. 54). For the purposes of this study, the same definition will be used, emphasizing human trust in an automated system with uncertainties in the automation's actions and vulnerabilities of the system to fail. The operator should trust that a system will accurately perform the job it is programmed to do in a real-world situation where the automation can fail with unpredictable consequences.

Human-automation systems like semi-autonomous vehicles require both parties to work together for the system to function optimally (Parasuraman & Riley, 1997). Trust calibration allows human operators to balance his or her trust in the system with the actual capabilities of the

system (Lee & Moray, 1992). If a human driver distrusts the automation, he or she may be less likely to allow the automation to perform its intended function. In this case, the human may try to take over control from the automation early, or neglect to engage the automation entirely. Equally, a human driver might overtrust the automation, resulting in the human not being available to help or takeover when the automation requires so. Understanding the dimensions of trust in automation will allow designers to create systems that maximize human-automation collaboration, thus increasing safety.

1.3.1 Dispositional Trust and Learned Trust. Everyone has certain innate factors that influence a person's likelihood to trust an automated system (Hoff & Bashir, 2015). A person's age, culture, gender, and personality traits can all affect the tendency to trust automation. *Dispositional trust* is this inborn propensity to trust that is relatively stable over time (Hoff & Bashir, 2015). In contrast, *learned trust* is trust that people have accrued over time through experience (Hoff & Bashir, 2015). *Initial learned trust* is the level of trust an individual has in a specific automated system, even if he or she has never used it before. Nearly all individuals experienced many types of automation, either through first-hand interaction with the automation, or through second-hand knowledge from media or friends. These experiences can shape an individual's opinions about and future interactions with automation. Before measuring how a new automation system affects trust, it is important to understand and control for an individual's dispositional and initial learned trust.

1.3.2 Dynamically Learned Trust. Operators of automated machines build trust through experience with the system (Hoff & Bashir, 2015). *Dynamically learned trust* is affected by design features, system performance, and situational factors outside of the system (Hoff & Bashir, 2015). In contrast to dispositional trust and initial learned trust, dynamically learned trust

is built and adapted through interaction with an automation system. Over time, operators can understand the capabilities and limitations of the system and properly calibrate their trust accordingly. With a reliable automation system, the level of trust should increase to match an equivalent level of reliability (Parasuraman & Riley, 1997; Chen et al., 2018). Unfortunately, studies measuring the development of trust over multiple events are rare, with most simply administering pre- and post-experiment trust measures (but see Bliss et al., accepted).

1.3.3 Trust decline due to error. When an automation error occurs, it is likely to reduce an operator's trust in the system. This automation error is seen to decrease the reliability of the automation, thus making a user less likely to trust the automation to perform the task correctly (Hancock et al., 2011; Hoff & Bashir, 2015; Parasuraman & Riley, 1997; Schaefer, Chen, Szalma, & Hancock, 2015). Automation errors are one of the most influential factors affecting trust, depending on the severity of the error. For minor errors, trust decrement might be small and inconsequential. However, major errors like a vehicle crash would be much more potent (de Visser, Pak, & Shaw, 2018). The decline of trust could hurt the performance of the human-automation team if the level of trust does not rebound from the decrease.

Both overreliance and disuse can result in suboptimal use of the system. Two potential results are errors of commission, in which a human operator can make action that is incorrect while following the suggestion of the automation, or omission, in which a human operator fails to act due to missing a signal by the automation (Parasuraman & Riley 1997). Errors of commission commonly occur in a reliance-oriented system in which the automation notifies the human of a need for action and the human must decide if the warning is legitimate. Errors of omission commonly occur in a compliance-oriented system in which the automation shows when the system is working normally, and the human must notice when a fault has occurred (Lee &

See, 2004). Automated driving systems currently occupy a space between reliance- and compliance-oriented systems because the human must pay attention to warnings and requests to takeover control, but they must also monitor the system for a failure of the automation recognize to recognize a hazard. The human driver's trust level helps maintain the delicate balance between avoiding commission errors when reacting to potential warnings and avoiding omission errors when the automation fails to detect and avoid hazards. The amount of trust decline could depend on the type of system error and what was required from the human operator.

     1.3.4 Trust repair. de Visser et al. (2018) review different strategies for repairing trust after a system error. They argue that in the same way that the best human-human teams try to intentionally build rapport fix trust if any problems occur (Duhigg, 2016), human-automation teams should act similarly to promote the best performance of the system. There is limited work in the area of trust repair for human-automation systems; however, de Visser et al. appraised studies about trust repair and provide a list of several active trust repair strategies (e.g. apologize, deny, empathize, explain). These strategies seem to be promising, but many are new or require more testing. The authors provide a theoretical trajectory model showing how different strategies could repair trust over time and include a theoretical baseline recovery pattern. The trajectory shows fast declines directly after an error, and a slowly rising, stepped line for the baseline recovery (de Visser et al., 2018). The baseline recovery rate demonstrates how trust can recover over time without any intentional repair strategy. The baseline recovery pattern can provide theoretical evidence for how trust repair occurs without a repair strategy. This pattern can then be used to compare the effectiveness of alternative repair interventions. As of now, there has been little evidence of such repair trajectories and even de Visser et al. call on researchers to consider

trust repair for automation studies. As a result, evidence of trust repair is needed to help build the theoretical literature.

1.4 Trust for Easy and Difficult Tasks

As mentioned before, many design aspects of automated systems can help ensure appropriate trust and reliance in automated systems. However, there are also external environmental factors that influence the system that can affect the human's trust in the automation. Each driving experience can be very different. The road, traffic, setting, and dozens of other conditions can change for each drive and humans can perceive these differences and may change their behavior and actions to match (Fredrick et al., 2017). The human driver may deem some situations easy and believe that the automation should have no trouble. But in other situations, the human driver might understand that the situation is difficult for them and the automation might also have difficulty with the task. This difference in task difficulty could affect an individual's level of trust in the automation and how they perceive and react to errors of the automation in these contexts.

As the difficulty of a particular task changes, so can a human's trust in automation. If an automated system is performing a difficult task, a human operator will account for the strengths and weaknesses of the system and respond accordingly (Madhavan et al., 2006). Madhavan et al. (2006) performed a study to measure the effects of automation failures on easy tasks. Their participants performed a target detection task in which they were told to identify target letters within an array of distractor letters. Some of the trials had small numbers of distractors making it easy, and other trials had many distractors to make it more difficult. They found that when automation made errors in a task that the operator perceives is easy, their trust was lower than the group with errors in a difficult task, they relied less on the automation, and their confidence

about their own ability increased (Madhavan et al., 2006). This finding falls in line with prior research (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003). The authors also stated that this effect of easy errors was likely due to two factors: conspicuity of automation errors in an easy task and the tendency of humans to have inflated ideas about their own ability when the task appears easy (Madhavan et al., 2006; Dunning, Johnson, Ehrlinger, & Kruger, 2003). A key point of note is that Madhavan et al. (2006) measured if the participant perceived the task as easy. If someone does not perceive a task as more easy or difficult than another, the finding is not likely to be shown.

The task that Madhavan et al. (2006) used was a basic target detection task that they reasoned was similar to baggage screening. However, the basic nature of letter discrimination cannot always be readily applied to the real-world and might not translate well to more complex tasks. Dzindolet et al. (2003) had a conceptually similar task that involved spotting a camouflaged soldier in a photograph of a nature scene. These studies both demonstrated humans and automation can have different performance for the same task and the difficulty of the task can influence the human's perception of the capabilities of the automated system. Because task difficulty has been shown to affect trust, it is important to expand this finding and test it in an applied setting. This way, it can help inform real-world design because automation might need to behave differently according to the current context.

The way Madhavan et al. (2006) tested trust differences was through direct manipulation of the complexity of the task. The task incorporated more or less targets, thus making the targets easier or harder to find and increasing required search time. They were able to quantifiably change the difficulty of the task by having either 5 targets out of 1000 stimuli or 90 targets out of

1000. They can attach a number to how many targets they changed, but it might not be a linear relationship in task-difficulty change. Thus, quantifying relative difficulty is challenging.

A task can also be made more difficult by changing the scene, setting, or scenario. The target detection task could be made more difficult by dimming the screen to obscure the readability of the letters or obscuring the periphery of vision so that the participant could only see what was in the fovea. A lighting manipulation for difficulty could be quantified similar to critical target decreases by decreasing lighting levels by a specific percentage or restricting the field of view by a specific amount. The key difference between these two manipulations is that Madhavan et al. (2006) changed the aspect of task itself to make it more difficult, whereas an alternative approach is to change the scene, setting, or environment to change the difficulty of the task. To simplify and clarify the idea for the rest of the document, the term "lighting condition" will be used to refer to daytime and nighttime that could affect the difficulty of a task. The change in lighting condition is a new concept from that of previous research and will contribute to previous theory by expanding the idea of what could make a task easy or difficult.

1.5 Current Study

Current literature has examined the development of trust in automation and the factors that influence it. However, little research has been done to investigate the effect of driving condition, particularly in applied scenarios like semi-autonomous driving where trust is developed over time. Prior studies have examined trust with takeover scenarios, but only investigated how trust changed from before one drive with a TOR to after that experience (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015; Hergeth, Lorenz, Krems, & Toenert, 2015). Other research has found that when an automated aid makes errors in an easy task it leads to more distrust than when an automated aid makes errors in a difficult task (Madhavan et al.,

2006). Additionally, Madhavan et al. found that automation reliance is more likely with highly reliable automation for a difficult task.

The purpose of this study was to investigate how different lighting conditions could affect users' trust levels in the automation in response to realistic automation failures. The study examined how trust is built over time and the effect that automation failure has on the human driver's trust. The study expanded trust development over time by getting trust ratings after each of the seven drives. This way, the building, decline, and repair of trust was observed within the experiment, in contrast to typical pre/post trust differences which do not show growth or decline patterns. The study also developed the theory behind how trust is affected when automation fails in daytime and nighttime conditions as a proxy for easy versus difficult driving conditions. Participants were asked to rate drive difficulty in an attempt to quantify differences between the daytime and nighttime conditions. By examining this phenomenon in an applied context, it further demonstrated how this critical difference can cause unique changes in the development of trust, leading to design recommendations for autonomous vehicle developers.

Three separate types of automation were used to properly assess trust in a realistic scenario. A no error (perfect baseline) condition, a TOR condition (one take-over request) and an automation failure condition (no warning/takeover request) were tested to examine if different automation errors could affect trust, especially when lighting conditions were different. Participants engaged in seven simulated drives with trust ratings collected after each drive to assess their development of trust over time. The no error condition showed how trust was normally developed over time in the scenario. This perfect baseline was necessary for understanding how errors in the other two conditions affect trust. The TOR and automation failure showed how trust declined for an unreliable system. Additionally, testing two different

automation errors was helpful because both problems represent realistic situations that occur in real-world driving. However, these errors focus on different causes of the automation failure and require different actions from the human driver. For a TOR, the automation understands its limitations and requests a takeover by the human driver (not a miss), whereas in the automation failure condition, the human driver must take over control because they have noticed a problem that the automation has missed. The six experimental groups in the study are shown in Table 1.

Table 1. The six experimental groups used in the study.

Group	Lighting Condition	Automation Error Type	Number of Participants
Group 1	Day	No Error	21
Group 2	Day	TOR	21
Group 3	Day	Failure	21
Group 4	Night	No Error	17
Group 5	Night	TOR	21
Group 6	Night	Failure	21

Note. TOR = Takeover Request.

Separate drives allowed the participant to experience the capabilities of the system over various scenario and unique experiences. A problem with some lab-based studies is the lack of real-world validity due to a short timeframe for experience with the automation in single-session studies (Bailey & Scerbo, 2007; Molloy & Parasuraman, 1996). In this study, instead of using one single drive with only a pre- and post-drive trust measure, multiple drives allowed for measuring trust development after each drive over time. Even though this study was still completed in a single session, the multiple drives with trust measures still showed development

over time. Thus, the design allowed investigation into the dynamics of building trust, reducing trust after an incident, and then possible trust recovery.

Previous studies have implemented a multiple-drive framework to test take-over request related questions (Naujoks, Mai, & Neukum, 2014; Van Der Heiden, Iqbal, & Janssen, 2017; Walch et al., 2015; Zeeb, Buchner, & Schrauf, 2015, 2016). Additionally, most of these studies took around 30-45 minutes to complete the driving portion. For the current study, the total driving time was about 40 minutes, fitting well within the range of the other studies. All the drives had common features that could be seen in normal, everyday drives, such as traffic lights and merging cars. These common features, listed in Table 2, are all events that could potentially be hazardous and cause a crash if the situation is not handled properly. Presenting several hazards forced participants to make sure the driving automation system safely avoids each one. However, for this study, only the construction hazard on the three-minute-long fourth drive) was the critical to maintain consistency.

Table 2. Each of the six assorted hazards present in the seven drives.

Six Hazards
Traffic light
Merging car
Pedestrian crossing
Stop sign
Construction
Slowing Car

Trust repair can be improved in various ways (de Visser et al., 2018); however this study did not use a specific active repair strategy. Trust repair through only perfect trials demonstrates a theoretical baseline without an active repair method. When a trust repair baseline is established, later improvements and approaches can be compared. For now, most automation systems do not have advanced trust repair strategies, so testing the basic method of repair is currently most useful.

Trust and usefulness are likely highly correlated constructs. However, usefulness can demonstrate a more direct application of how likely participants would be to use of the system outside of the experimental setting. After the participant is done with all the drives, a Usefulness of Automation Survey (UAS) based on Gold et al. (2015) captured how likely they are to want to use this system. Their level of trust in the system is important, but the UAS further clarified the participant's perceived safety gain by using the automation and how much they would intend to use the system afterwards (Gold et al., 2015).

1.6 Hypotheses

Hypothesis 1: Ratings of task difficulty should be lower for the daytime condition than for the nighttime condition. Nighttime driving is objectively more difficult than daytime driving due to visual problems associated with low luminance, which can lead to increased reaction times and slower object recognition (Konstantopoulos et al., 2010; Leibowitz & Owens, 1977; Plainis & Murray, 2002).

Hypothesis 2a: Trust in automation in the daytime and nighttime no error conditions was expected to be equivalent. Previous research by Madhavan et al. (2006) and Dzindolet et al. (2003) showed that highly reliable automation showed no trust differences, regardless of

difficulty. Participant's trust diverged only when the automation made errors in easy and difficult condition in previous studies. Therefore, with no error, there should be no differences in trust.

Hypothesis 2b: The daytime trust was expected to be lower than nighttime trust for the TOR and Automation Failure conditions because the daytime task is perceived as easier. The perceived error on drive 4 should reduce participant trust. If Hypothesis 1 is not supported, this hypothesis would not be expected because the differences in trust in Madhavan et al. (2006), Dzindolet et al. (2003), and Schwark, Dolgov, Graves, and Hor (2012) are only found when participants perceive a task to be difficult.

Hypothesis 2c: An interaction was expected, such that the daytime condition has a larger decrease in trust than nighttime, and that difference is even greater in the failure condition than the TOR condition. This prediction was based on results in Madhavan et al. (2006) and Dzindolet et al. (2003). Failures in an easy task show greater decreases of trust than failures in a difficult task. TORs are a less extreme type of error, so they should have less of an effect. Similar to Hypothesis 2b, this interaction would not be expected if Hypothesis 1 was not supported.

Hypothesis 3: Level of trust was expected to increase as the number of drives increases, but then drop after the takeover event in the TOR condition and the Automation Failure condition, with a larger drop in trust for the Automation Failure. Trust was also expected to recover to higher levels after subsequent drives after the takeover event (de Visser et al., 2018).

Hypothesis 4: Subjective assessment of usefulness via the UAS should yield a result pattern similar to trust in automation, such that, a significant main effect of error type was expected for UAS. This hypothesis was based on Hoff and Bashir (2015) because an errorless system should be more useful and trustworthy than one that produces errors.

Hypothesis 5a: The main effects of lighting condition and automation error type were expected to be significant. RT was expected to be faster for the TOR condition than for the failure condition, and RT in the daytime condition was expected to be faster than that in the nighttime condition. This would not be a new finding because warnings have been well established to improve RT (Porter, Irani, & Mondor, 2008; Ruscio, Ciceri, & Biassoni, 2015; Xiang, Yan, Weng, & Li, 2016). However, it helps to demonstrate the magnitude of the differences between groups, which is important for comparison with other research to help inform designers.

Hypothesis 5b: Accuracy was expected to follow the same pattern as described in Hypothesis 5a. Accuracy would be better for the TOR condition than for the Automation Failure condition. Because the participant would be warned about the upcoming critical event in the TOR condition, they would have more time to react to the scene and make a correct avoidance maneuver.

CHAPTER 2

METHOD

2.1 Design

The study employed a 2 (Lighting Condition: Daytime, Nighttime) x 3 (Automation Error Type: No Error, TOR, and Failure) x 7 (Drive number: Drive 1-7) mixed design. Lighting condition and automation error type were manipulated between subjects and drive number was manipulated within subjects.

2.1.1 Independent Variables. There were three independent variables in this study. The first was lighting condition. The daytime and nighttime conditions represented the difficulty of the task, easy and difficult, respectively. The daytime setting was used as the easy condition because all potential hazards could be clearly seen from the instant they appeared on the horizon. The nighttime setting was more difficult because it obscured the hazards, not allowing the driver to fully see the entire scene until the objects were closer (Konstantopoulos, Chapman, & Crundall, 2010).

The second independent variable was the automation error type. There were three automation error types: no error, TOR, and failure. For all three conditions, a critical construction hazard appeared in the driver's lane during the fourth drive. The hazard required moving to the left-hand lane to avoid a collision. See Figure 3 for images of the critical construction hazard. The automation responded to the critical hazard in three different ways, depending on the condition. First, in the no error condition, the automation had no errors and thus this condition served as the baseline. The second condition had a standard TOR during which the automation issued a three-beep warning telling the participant to take over control. The TOR was issued as soon as the hazard could be seen, about 480 feet from the hazard. In the

last condition, automation failure, the automation failed to notice and avoid the critical road hazard, and thus the participants were required to notice it and manually take over control. Both the TOR and automation failure conditions demonstrate realistic scenarios that happen during current autonomous car operation (Brown & Laurier, 2017; Endsley, 2017). Trust should be affected differently for these two groups because even though they both fail to perform, the TOR still provides a warning (Hoff & Bashir, 2015; Parasuraman et al., 1997; Parasuraman & Riley, 1997; Walch, Lange, Baumann, & Weber, 2015).

Both the lighting condition and error type condition were manipulated between subjects to avoid participants being exposed to an error more than once. There were six groups with about 20 participants tested in each group (see Table 1). Participants were randomly assigned to a daytime or a nighttime condition and randomly assigned to either no error, TOR, or failure conditions. Random assignment of participants was carried out in blocks of 12 so that two participants were in each of the six conditions for every 12 experimental sessions. Random assignment helped to ensure participants were evenly assigned to groups and that each condition was filled consistently throughout the data collection period. The random assignment was determined to be successful after conducting separate 2 (Lighting condition: Daytime, Nighttime) x 3 (Error type: No Error, TOR, Failure) Analyses of Variances between all groups for age, gender, driving experience, pre-drive trust, or ADOS ratings, respectively. No significant differences were found between groups using a $p = .05$ cutoff.

No Error. The no error condition featured automation that committed no errors and did not require the participant to take over control. The participant was required to have his or her hands on the wheel and foot on the pedal to be ready to respond at any point. As a result, although participants did not have to intervene, they were still required to attend to the drive.

TOR. For the TOR condition, the automation noticed the critical hazard and issued a TOR. The TOR consisted of a 1000-hz sinusoidal earcon (as used by Borojeni, Chuang, Heuten, & Boll, 2016) to signal the request. The participant needed to take over manual control of the vehicle after receiving this request to avoid crashing. The auditory warning was presented about six seconds before the critical event to give the driver a short warning before the event. Previous research has shown that a two-second transition time is not sufficient for a driver to resume control, but five seconds should be enough (Mok et al., 2015). Other studies have confirmed that 5-7 seconds is sufficient (De Winter et al., 2014; Gold, Damböck, Lorenz, & Bengler, 2013). A review of driver takeover times in automated vehicles has shown that one of the most commonly used TOR lead times has been 3 seconds with a mean takeover reaction time of 1.14 seconds (Eriksson & Stanton, 2017). However, the present experiment also included an automation failure that required more time for participants to realize they needed to take over because the automation was not working. Pilot testing of the experiment showed that 6 seconds was enough time in both conditions, similar to the forewarning provided by De Winter et al. (2014) and Gold et al. (2013). The 6-second time to contact (TTC) gave participants enough time to handle the critical event while also leaving a potential for crashes if responses were not fast enough.

Automation Failure. In the automation failure condition, the automation did not make an appropriate avoidance behavior for the critical event. Moreover, it did not signal the driver about the issue as in the TOR condition but continued as if there was no hazard, resulting in a crash if the driver did not adjust. It was solely up to the driver to notice the problem and correct for the issue.

Drives and Critical Event. Participants encountered seven separate drives with all but one of the drives lasting about five to six minutes each. The fourth drive was three minutes long

because a critical hazard event happened at the end of that drive, approximating the midpoint of the other 6 drives. The critical hazard was a construction zone in the driver's lane, requiring the vehicle to change lanes to avoid crashing. For the TOR condition, the automation issued a TOR and the automation failure condition gave no warning but continued toward the hazard with an approximate 6-second TTC. In the no error condition the car successfully navigated around the critical hazard. The drive ended directly after the critical hazard event. The critical hazard event occurred at three minutes to provide participants enough time to acclimate to the drive but prevented them from expecting something to happen at the end of the drive. The fourth drive was chosen because it was the middle drive and would allow data collection showing if trust improved before the critical event and if trust recovered afterward. The first three drives should show increases in trust due to experience with the system. The fourth drive should lower trust because of the need for takeover. Then the final three drives could show either continued low trust or repair of trust on the subsequent errorless drives.

Each drive was designed to be unique, with different locations, building, objects, and potential hazards. Although the drives were unique, they all contained a similar number of about five potential hazard events and an attempt was made to equate them all for interest and difficulty. The goal was to make participants feel like each drive was like a normal trip they would take to different locations.

2.1.2 Dependent Variables. The objective dependent measures were accuracy (whether the participant safely avoided crashing in the TOR or automation failure conditions), RT (First response of at least 1 degree turn of steering wheel from event onset/Braking action). Subjective dependent measures were the subjective Human-Computer Trust questionnaire (Appendix D) based on Madsen and Gregor (2000), the Automated Driving Opinion Survey (ADOS), including

driving experience questions (Appendix C), based on Kyriakidis, Happee, and De Winter

(2015), the Usefulness of Automation Survey (UAS; Appendix E) based on Gold et al. (2015).

Demographic information was collected via a demographic questionnaire (Appendix B).

Participants completed the trust questionnaire eight times, once after the practice demonstration

and each of the drives. All measures are further described in Section 2.3 Measures.

2.2 Participants

I conducted an a priori power analysis with Power Analysis and Sample Size (PASS)

software (https://www.ncss.com/software/pass/) using the mixed models simulation to estimate a

desired sample size for the current study (Appendix A). Using a 2 (A: Lighting Condition) x 3

(B: Error Type) x 7 (C: Drive) mixed design with the A and B terms as between-groups and C as

within-groups, the software performed 100 simulations to estimate the power of an N of 90, 120,

and 150. Results showed that an N of about 120 would provide over 80% power for the A, C,

AxC, BxC, and AxBxC factors and interactions. However, the B factor and AxB interactions

were estimated lower at about 50% and 60% power, respectively. Although the power was

expected to be lower for B and AxB, the power only rose to about 65% with an N of 150.

Because of diminishing returns for increasing the N, and balancing the time and resources of the

lab, I decided that an N of 120 would be sufficient to find most of the desired effects. However,

the potential for underpowered factors and Type II error should be considered with any non-

significant results found in the study.

As a result, 133 participants were recruited for the study. However, 11 of those

participants were excluded from the study due to issues with data collection caused by simulator

crashes or data collection errors. The remaining 122 participants (84 female and 38 male) yielded

about 20 participants in each of the six groups (see Table 1). The participants were required to

have a valid driver's license and reported that their average driving experience was 3.73 (SD = 1.23) years and had a mean age of 19.57 (SD = 3.00). The participants' ethnicities were Caucasian (41), African American (63), Asian (6), and more than one race (12). All participants had normal or corrected-to-normal hearing and vision. Participants were undergraduate psychology students recruited from Old Dominion University's SONA (odupsychology.sona-systems.com) pool. They were given class credits for their time.

2.3 Apparatus and Stimuli

The participants faced a table mounted computer display and steering wheel (Logitech G920 driving Force). Three pedals were at their feet (clutch, brake, accelerator, from left to right respectively). The clutch was not used because the vehicle had automatic transition (see Figure 1). A Dell P2717H 27-inch monitor with a 1920x1080 resolution was used to display the driving environment. The participants sat approximately 32 inches from the screen resulting in a horizontal visual angle of about 45.75° (see Figure 2).

Figure 1. The Logitech G920 racing wheel and pedals used in the experiment (https://www.logitechg.com/en-us/product/g920-driving-force)
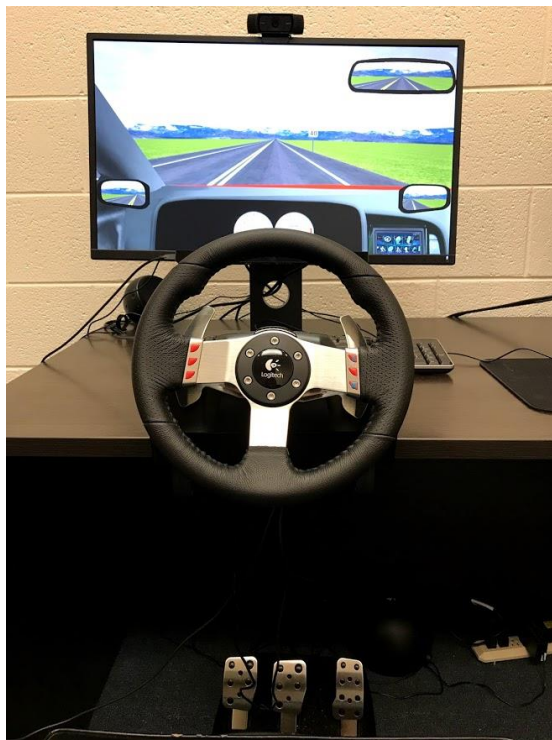
Figure 2. The setup of the driving simulator from the view of the participant.

2.3.1 STISIM. The driving simulation software used was STISIM (stisimdrive.com) driving simulator software, version 3.14.02. Daytime driving occurred with lighting parameters set to replicate daylight conditions (100% ambient, 50% diffuse, 25% specular). From the STISIM manual, "Ambient specifies how intense or bright the light will be, diffuse specifies how much shadowing there will be on non-lit surfaces, and specular specifies how shiny an object appears". The daytime values were set to default. Nighttime driving occurred with lighting parameters set to replicate nighttime conditions (25% ambient, 0% diffuse, 0% specular). These values are those specified in the manual for a nighttime driving scenario. For nighttime driving, the car's simulated headlights were turned on. No specific information on headlight lighting levels was given in the manual, but the headlights are meant to replicate realistic driving. The headlights shined on the road directly in front of the vehicle, allowing for adequate vision for

objects within a short span of the front of the vehicle, but more obscured vision further out. See Figure 3 for an example of the lighting conditions in the daytime and headlights in a nighttime drive.

Figure 3. Screenshots of the construction hazard for both lighting conditions. Daytime is on the left and Nighttime is on the right. The construction hazard first visible for the top images and fully visible in the bottom images.

2.4 Measures

    2.4.1 Demographics. A basic demographics sheet (see Appendix B) was used to gather demographic and handedness information about the participants to report distributions of gender, age, race, and handedness.

    2.4.2 Automated Driving Opinion Survey (ADOS). This survey was used to gather manual driving experience, automated driving experience, and opinions about automated

vehicles from each of the participants. Automated driving experience was not expected to show much variability but was included to help control for the potential that some participants might be experienced with automated driving. The survey was adapted from Kyriakidis et al.'s (2015) survey but was modified to use only questions about driver behavior and attitudes about automated vehicles. The authors used the survey to investigate public opinions about automated driving across many cultures. Therefore, the original survey contained many unrelated questions, including income, disability, personality, willingness to pay for automated driving systems, data privacy concerns, country of origin, and computer usage. Therefore, the current study included 23 of the original 41 questions related to manual and automated driving behaviors and opinions. This scale was administered to participants at the beginning of the experiment so that experience with the system did not influence responses (see Appendix C).

   2.4.3 Trust in Automation Scale. Trust is often measured via the subjective Human-Computer Trust questionnaire developed by Madsen and Gregor (2000). The measure asks various questions to assess five key aspects of trust: personal attachment, reliability, understandability, technical competence, and faith. Madsen and Gregor refined and validated the scales, finding that the understandability, reliability, faith, and personal attachment scales demonstrated good scale reliability, Cronbach's alpha = .84, .85, .88, and .90, respectively. The perceived technical competence scale fared the worst with Cronbach's alpha equal to .74, although it is still considered acceptable. The authors stated that the wording for the technical competence questions may have been hard for participants to clearly understand, so some phrasing changes could be necessary. A principal components analysis for the single factor

human-computer trust conducted by Madsen and Gregor demonstrated that all questions were highly correlated with human-computer trust.

Participants in the present study were asked to fill out a Human-Computer Trust questionnaire adapted from the scale designed by Madsen and Gregor (2000) after each of the seven drives. The scale included 24 questions that assess human-automation trust based on five separate criteria to encompass the multi-dimensional aspects of trust. Madsen and Gregor determined that the question, "The system correctly uses the information I enter", would improve the scale reliability if it was removed. Therefore, the current study replaced that question with an attention check requiring the participant to mark a specific answer specified by the experimenter. Minor phrasing modifications were made to the questions to make them relevant to an automated driving task. Additionally, participants were told that, "This survey sometimes uses language like, 'actions', 'decision', and 'problem'. Please interpret these in the context of the driving automation system. 'Problems' are potential hazards on the road, 'actions' are movements that the vehicle makes, and 'decisions' are actions taken in response to problems where multiple different actions could be made". The researcher was always present to make sure participants clearly understood what the questionnaire was asking. Questions were rated on a Likert scale from 1 ("strongly disagree") to 7 ("strongly agree"). Examples of questions are, "I can rely on the system to function properly" and "I believe actions from the automated driving system are safe even when I don't know for certain that it is correct." By administering the scale after each drive, it allowed a dynamic understanding of trust, showing initial trust, how each drive changes trust, and then a final overall rating of trust for the automated driving system (see Appendix D).

2.4.4 Perceived Difficulty Rating. Participants rated the drive difficulty after each drive. Participants were asked, "How would you rate the ease or difficulty of performing this drive?".

They responded using a Likert scale from 1-7 (extremely easy-extremely difficult). Perceived drive difficulty ratings helped to demonstrate if there were distinctions between different drives, lighting conditions, and automation error types.

2.4.5 Usefulness of Automated System (UAS). Based on the survey administered by Gold et al. (2015), participants completed the Usefulness of Automation System survey after all seven drives had been completed. The survey asked about perceived discharge of the driver due to automation (how much the driver believed they could let the automation do the task for them), safety gain from the automation, hazards to safety caused by the automation, perceived control of conduct, and intention of use. The five individual sections will be collectively referred to as Usefulness if the Automated System (UAS) for simplicity. All these categories were analyzed together to see if there were differences among the independent variables. The UAS helped verify the practical application and potential use for each of the systems (see Appendix E).

2.5 Procedure

Participants were greeted and welcomed into the lab at their scheduled times. Each participant sat in a non-rolling chair in front of the driving simulator. He or she was encouraged to adjust the position of the chair to comfortably reach the wheel and pedals. Each participant read and signed an informed consent form. Next, participants provided demographic information about themselves (Appendix B). Participants then completed the ADOS to get information about their driving experiences and opinions about automated driving (Appendix C). Participants were randomly assigned to one of the six experimental groups. Participants were then given instructions for the following tasks and then completed each task. See Figure 4 for a graphical representation of the procedure.
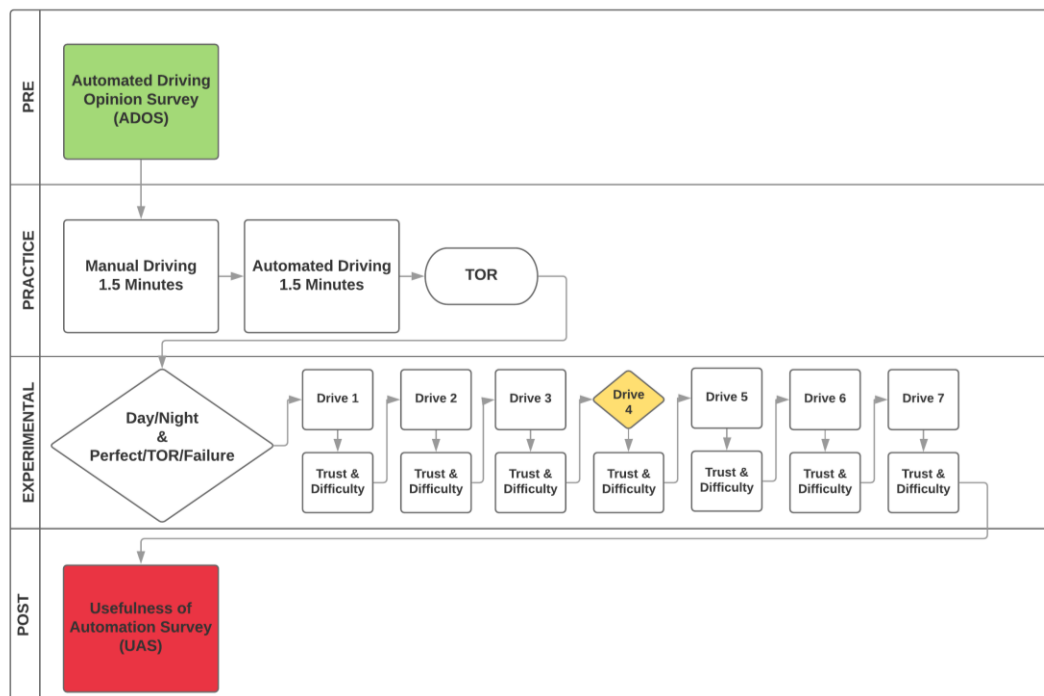
Figure 4. The experiment procedure. The experiment started with the green rectangle (ADOS) and followed the arrows to end at the red rectangle (UAS).

2.5.1 Practice Task. The practice session had two phases. The first was about 1.5 minutes long and required manual driving (automatic gearbox, but not autonomous vehicle) to familiarize the participant with the simulator. The next phase required about 1.5 minutes of automated driving to familiarize the participant with the behaviors of the automated driving. This second phase was a demonstration of the automation's capabilities, so participants could become familiar with how the system works. Participants were required to have their hands on the wheel and their feet on the pedals during the experiment. This reflected the requirements of Level 2 driving automation systems that require the driver to be engaged and ready to take control at any

time (SAE, 2016). It also helped capture consistent reaction times (Greenlee, DeLucia, &

Newton, 2018).

The automated practice section ended with a forced takeover so that the participants

could become familiar with the takeover process. The takeover process used an ambiguous

earcon (1000hz sinusoidal tone, as used in Borojeni, Chuang, Heuten, & Boll, 2016; Naujoks,

Mai, & Neukum, 2014) to alert the driver that he or she would be taking control of the vehicle.

There was no critical event associated with the practice takeover, because it could prime the

participant for a hazard later in the experiment. Participants were explicitly told that this is a

training scenario to show them what a TOR could look like and to give them experience with

automation. The experimenter pointed out exactly when the TOR would occur and told

participants that they were now in full manual control of the vehicle again and could switch lanes

or maneuver however they wanted until the drive ended in about 25 seconds. Participants could

voluntarily repeat the practice section if they did not get a proper understanding of the simulator

or the situation. However, it was not necessary for anyone to repeat the practice.

2.5.2 Experimental Task. For the experimental task, the participant was told to maintain

concentration on the driving environment and be prepared for the small chance that the vehicle

could issue a Takeover Request. Participants were not told a specific percentage chance for

likelihood of a TOR. Instead, participants were able to see how the automation functioned in the

demonstration and could develop an estimate of their own. The initial learned trust was measured

to help account for individual differences. Automotive companies do not tell consumers a true

reliability percentage for any autonomous functions of their vehicle. The users typically must

find this out for themselves or other sources like friends and media. Reliability data are hard to

accurately collect for current commercial vehicles. It requires numerous user reports, because

companies could not feasibly test drive the required number of miles to demonstrate reliability rates (Kalra & Paddock, 2016). To simulate more realistic and accurate trust calibration, participants saw the automation demonstration and were assessed on their preconceived trust.

The participants completed all seven drives as explained in the design for the drive and critical event phases. After each drive they completed the Human-Computer Trust Questionnaire (Appendix D), with the appended perceived drive difficulty question. Participants were told to rate the automated driving system as a whole, incorporating all of the drives together, not individually. After completing the drives, participants filled out the Usefulness of Automation Scale (Appendix E). Once they completed the questionnaires and drives, they were thanked for their time, granted credit, and escorted out of the lab.

CHAPTER 3

RESULTS

3.1 Subjective Drive Difficulty

Participants estimated subjective drive difficulty on a scale on 1-7 (Extremely easy – Extremely difficult) after each drive. Difficulty was analyzed using a 2x3x7 mixed factor ANOVA with lighting condition and error type as the between-subjects IVs, and drive number as the within-subjects IV. The difficulty ratings were not significantly different between daytime and nighttime conditions $F(1, 116) = 1.33$, $p = .251$, $\eta_p^2 = .01$, which contrasts the expected differences in perceived difficulty stated in Hypothesis 1. The failure to support Hypothesis 1 indicates that Hypotheses 2a, 2b, and 2c would no longer be expected. This expectation is demonstrated in Section 3.2 (Subjective Trust).

Difficulty ratings were significantly different for error conditions $F(1, 116) = 6.26$, $p = .003$, $\eta_p^2 = .10$. Specifically, the failure condition had significantly higher ratings of difficulty than both the TOR, $t(116) = 2.46$, $p = .04$, and no error conditions, $t(116) = 3.42$, $p = .002$. The difference between TOR and no error was not significant, $t(116) = 1.02$, $p = .565$. The main effect of drive was significant, $F(1, 116) = 7.45$, $p < .001$, $\eta_p^2 = .06$, and planned contrast showed that drive 4 had significantly higher ratings of drive difficulty than the other drives $F(1, 116) = 26.12$, $p < .001$, $\eta_p^2 = .18$. Results seem to indicate that participants rated the difficulty of the drive based on how the automation performed, rather than considering the lighting conditions of the drive. See Figure 5 for graph showing difficulty for each drive separated by error type.
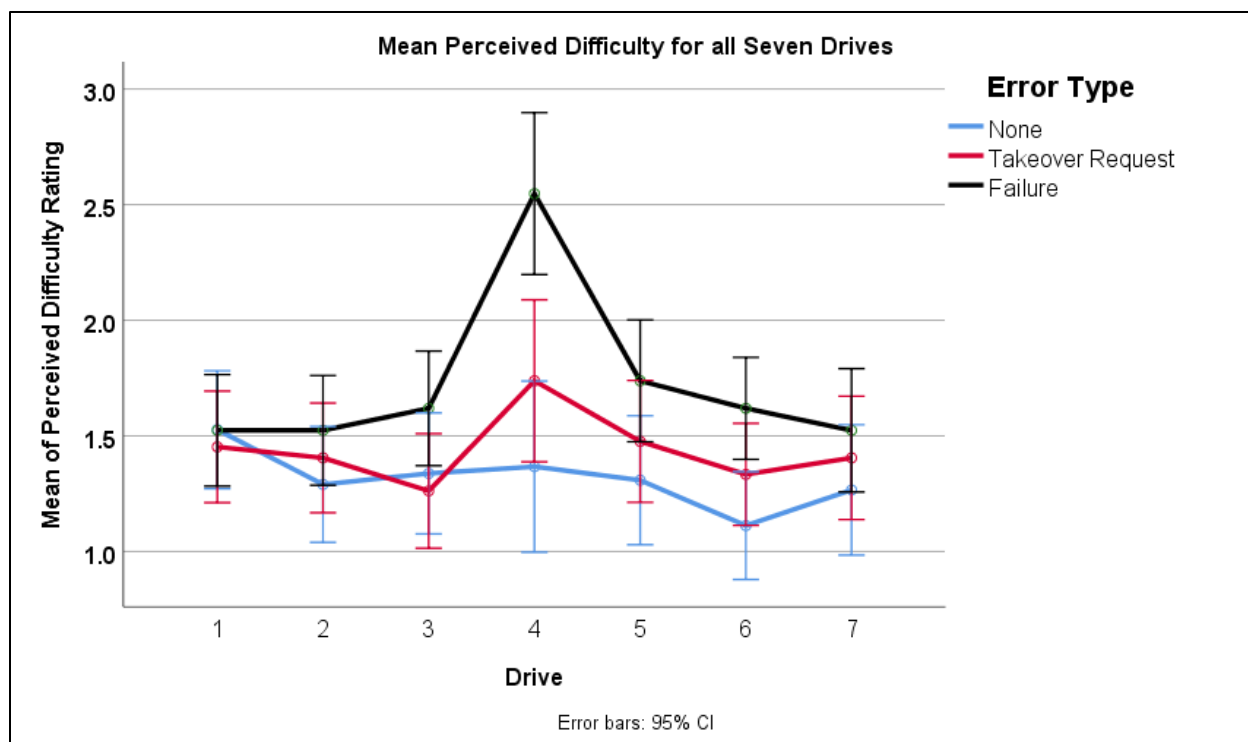
Figure 5. Subjective difficulty for each drive. Each line represents a different error type condition. Error bars are 95% confidence intervals.

3.2 Subjective Trust

Trust was analyzed first by demonstrating no difference between all conditions for pre-drive trust. This was done to ensure no groups started at a different level of dispositional and initial learned trust. Then, trust was analyzed across drives 1-7 to document trust growth or decline patterns. After the discussion of overall trust findings below, specific findings for dynamically learned trust, decline of trust, and trust repair were addressed. Dynamically learned trust was discussed separately from the other sections because, as expected, no differences were found between groups, so they were collapsed and analyzed together.

3.2.1 Pre-drive Trust. The pre-drive trust measure, taken after the participants participated in practice and saw the demonstration of the automation's capabilities, was intended

to measure their dispositional trust, specifically for the automated vehicle. Because participants may vary concerning their comprehension of automation, the demonstration was meant to clarify how participants would understand what automation was in this experiment. We also wanted to ensure that participants' later ratings of trust related to their baseline initial learned trust ratings.

Pre-drive trust was analyzed using a 2x3 Analysis of Variance (ANOVA) with lighting condition (daytime, nighttime) and error type (no error, TOR, failure) as independent variables. Neither the main effects nor the interaction was significant, Fs <1. Equivalency testing using two one-sided t-tests (TOST) was used to verify equivalence among the six groups. All compared groups' 95% confidence interval for equivalence was within the equivalence interval of [-.5, .5], ps < .05, meaning a claim of equivalence is supported. Table 3 shows the results of the t tests with the greater of the two p values listed for each group comparison. Daytime and nighttime groups was also equivalent, $t(58) = 3.32$, $p = .001$.

Table 3. Results of the TOST showing the greater of the two p values for each comparison.

	None	TOR	Failure
None	-	$t(41) = 2.01$, $p = .024$	$t(41) = 3.17$, $p = .001$
TOR	$t(41) = 2.01$, $p = .024$	-	$t(41) = 1.76$, $p = .041$
Failure	$t(41) = 3.17$, $p = .001$	$t(41) = 1.76$, $p = .041$	-

The Automated Driving Opinion Survey (ADOS) was compared to pre-drive trust to see if preconceived opinions about automated driving matched with trust in the automation after the demonstration. TOST was used to check equivalency and the 95% confidence interval for equivalence was within the equivalence interval of [-.5, .5], ps < .05, supporting a claim of

equivalence. This finding demonstrates that the automation demonstration did not significantly change trust levels from previously held opinions of automated vehicles.

3.2.2 Overall Trust – Drives 1-7. To test Hypothesis 3 which expected overall trust across the seven drives to rise through the first 3 drives, decrease for the critical drive 4, and to repair for the final 3 drives, a 2x3x7 mixed factor ANOVA was used with lighting condition and error type as the between-subjects IVs, and Drive Number as the within-subjects IV (see Figure 6). Results showed that participants' average trust ratings varied across the drives, $F(6, 115) = 6.11$, $p < .001$, $\eta_p^2 = .50$. A trend analysis was used to analyze the overall pattern of trust using within-groups polynomial contrasts. The analysis revealed a significant cubic trend, demonstrating an upward trend through the first three drives, a downward trend for drive 4, and then an upward trend for the final 3 drives, $F(1, 116) = 13.33$, $p < .001$, $\eta_p^2 = .10$. There was also a significant main effect of error type, $F(2, 116) = 4.56$, $p = .012$, $\eta_p^2 = .07$. Planned contrast showed that participants in the no error condition were significantly more trusting than those in the TOR condition, $t(116) = 2.77$, $p = .006$, and in the failure condition, $t(116) = 2.44$, $p = .016$. The difference between TOR and failure was not significant, $t(116) = .34$, $p = .734$. There was also a significant interaction between drive and error type, $F(12, 696) = 3.96$, $p < .001$, $\eta_p^2 = .06$. This interaction resulted from the decline of trust in drive 4, and the slow repair in subsequent drives.

The results support Hypothesis 3, demonstrating dynamically learned trust, decline of trust due to errors, and trust repair over time. The development of trust was one of the major foci of the study, so the analyses are split into the Dynamically Learned Trust, Trust in drive 4, and Trust Repair sections. The main effect of lighting condition was not significant, $F(1, 116) = .54$, $p = .466$, $\eta_p^2 = .01$. The interaction between drive and lighting condition was not significant, $F(6,$

696) = 1.90, $p = .079$, $\eta_p^2 = .02$, nor was the 3-way interaction among drive, lighting condition, and error type, $F(12, 696) = 1.58$, $p = .093$, $\eta_p^2 = .03$. A Pearson correlation between overall trust and overall difficulty showed a significant negative correlation, $r(122) = -.20$, $p = .027$.
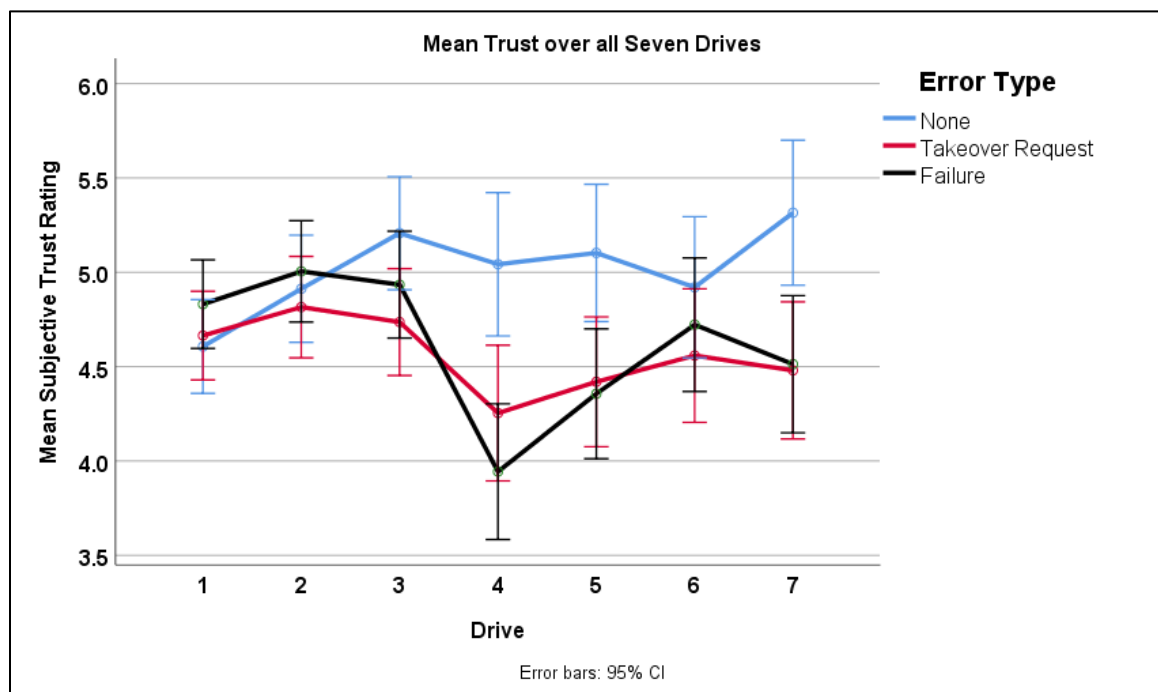
Figure 6. Trust for all seven drives. Each line represents a different error type condition. Error bars are 95% confidence intervals.

3.2.3 Dynamically Learned Trust. Pre-drive trust and trust for drives 1-3 were analyzed using a 2x3x4 mixed-factors ANOVA with lighting condition and error type as between-subjects conditions and drive as the within-subjects factor. This separate analysis was conducted to examine dynamically learned trust before the critical event in drive 4. Trust was not significantly different across error types, $F(2, 116) = 1.53$, $p = .221$, $\eta_p^2 = .03$, or between the two lighting conditions, $F(1, 116) = .04$, $p = .835$, $\eta_p^2 < .01$. As per Hypotheses 2a and 3, this result was expected for error type because at this point, no error intervention had occurred. Lighting

condition was also not expected to show differences until drive 4, where the error could affect

their trust due to perceived difficulty differences. Because between-groups comparisons were not

significantly different, they were merged to examine dynamically learned trust. After merging, a

repeated-measures ANOVA for drive showed that trust was significantly different as a function

of drive number $F(3, 363) = 7.45$, $p < .001$, $\eta_p^2 = .06$. A trend analysis showed a significant

positive linear trend with trust increasing from pre-drive to drive 3, $F(1, 116) = 15.09$, $p < .001$,

$\eta_p^2 = .12$. This result supports Hypothesis 3, which expected a positive trend for trust as the

number of drives increases before the critical error event in drive 4. See Figure 7 for graph of

results.

Figure 7. Initial learned trust and trust at drives 1-3. All conditions were collapsed together.
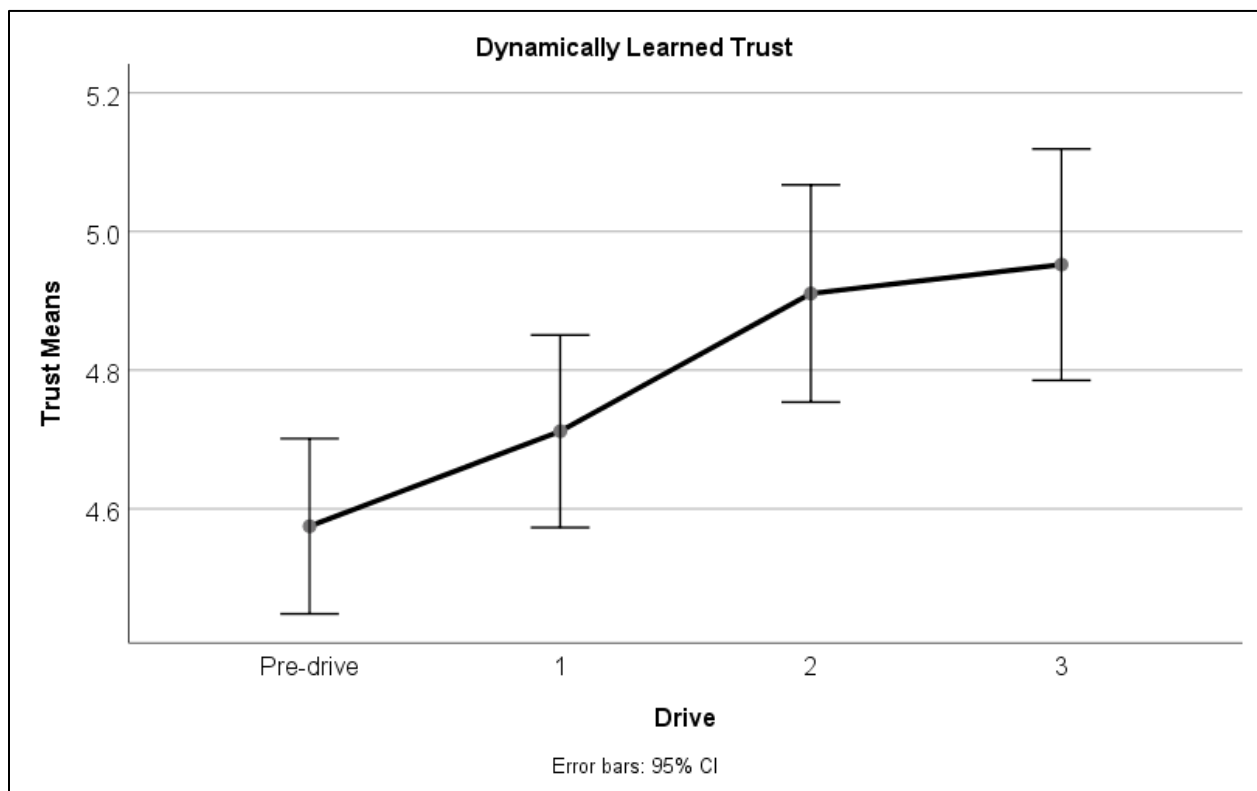
Error bars are 95% confidence intervals.

3.2.4 Trust in Drive 4. There was a significant effect of error type for trust after the critical fourth drive, $F(2, 116) = 9.09$, $p < .001$, $\eta_p^2 = .14$. The effect of lighting condition was not significant, $F(1, 116) = .07$, $p = .795$, $\eta_p^2 = .01$. Pairwise comparisons showed that trust was significantly higher for no error than TOR, $t(116) = 2.99$, $p = .003$, and Failure, $t(116) = 4.16$, $p < .001$. However, TOR and Failure were not significantly different, $t(116) = 1.21$, $p = .229$. These results partially supported Hypothesis 3. Trust was expected to stay constant for no error and decrease for TOR and failure in drive 4. However, because TOR is a warning, it was expected to decrease trust less than a complete automation failure.

3.2.5 Trust Repair. Further analysis using a 2x3x4 mixed ANOVA for between-subjects lighting condition, error type and within-subjects drives 4-7 was conducted to demonstrate how trust could be repaired over time after additional errorless drives. There was a significant main effect of error type $F(2, 116) = 6.94$, $p < .001$, $\eta_p^2 = .11$. Pairwise comparisons showed that trust was higher through all drives for the no error condition than the TOR condition, $t(116) = 3.15$, $p = .002$, and the Failure condition, $t(116) = 3.36$, $p = .001$. This finding supported Hypothesis 3 because there was no decrease of trust during drive 4 in the no error condition and thus maintained the same level for no error, and it was still low in the TOR and failure condition. There was no significant difference between TOR and failure, $t(116) = .21$, $p = .830$. Because TOR and failure were not statistically different, they were collapsed together to evaluate the baseline repair rate after a critical error. Using a repeated measures ANOVA for drives 4-7, it showed a main effect of drive, $F(3, 249) = 6.25$, $p < .001$, $\eta_p^2 = .07$. A trend analysis showed a positive linear trend for trust increasing from drive 4 to drive 7, $F(1, 116) = 13.44$, $p < .001$, $\eta_p^2 = .14$. This result partially supports Hypothesis 3, demonstrating trust repair directly after an automation error when the automation performs perfectly. However, the magnitude of trust

rating increase after drive 5 is not very large, indicating that the repair is slow. See Figure 8 for graph of trust repair after a critical event for the combined TOR and failure conditions.
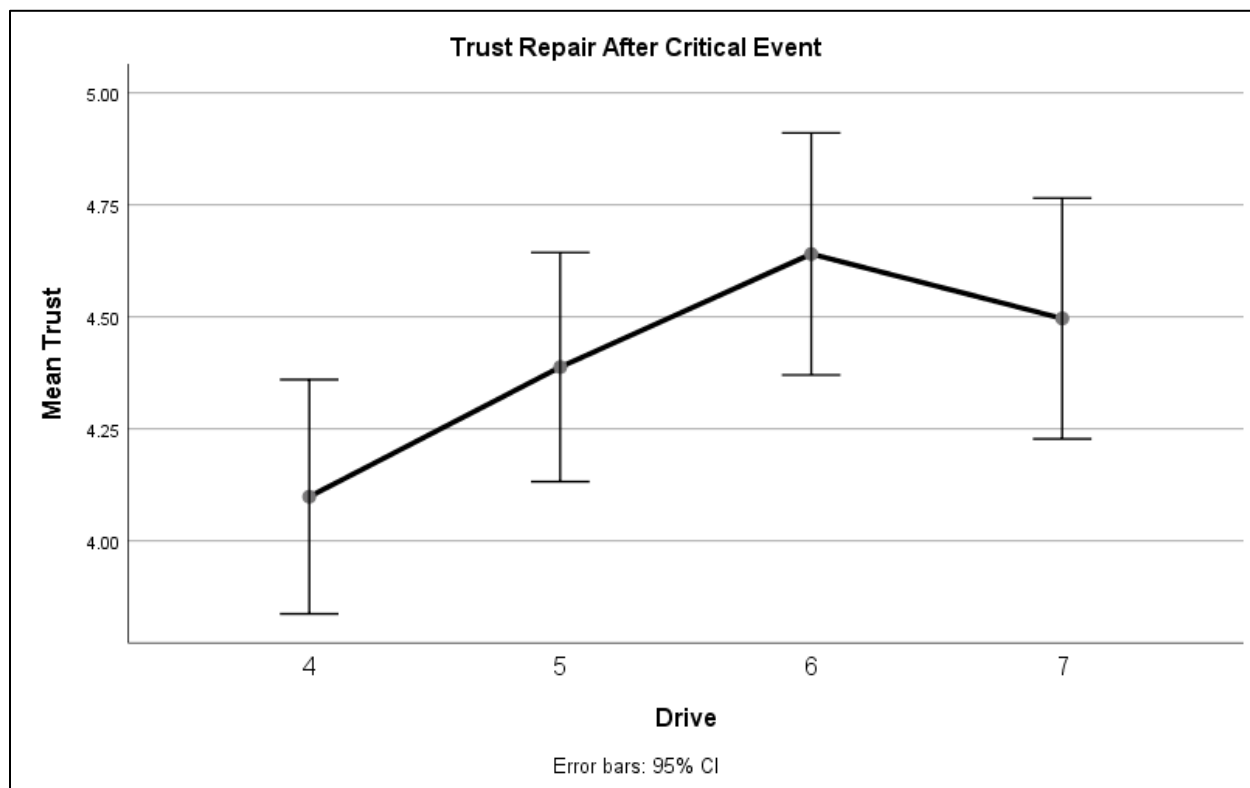
Figure 8. Trust repair after a critical event in drive 4 with TOR and failure condition merged. Error bars are 95% confidence intervals.

3.3 Performance Measures - Drive 4

Performance measures included accuracy and RT in response to the critical construction hazard in Drive 4. Accuracy was defined by whether the participant safely avoided crashing in the TOR and automation failure conditions, and RT was recorded from event onset to first response of 1-degree wheel turn or braking action. Accuracy and RT were not analyzed for the no error condition because there was no crash in that condition, thus no data were available to

analyze. Recall that Hypothesis 5 predicted the TOR condition to be faster and more accurate than the failure condition, and daytime to be faster and more accurate than nighttime. To test this hypothesis, a 2x2 between-subjects ANOVA was conducted with lighting condition (daytime and nighttime), and Automation Error Type (TOR, and Automation Failure).

3.3.1 Accuracy. Accuracy was analyzed using a binary logistic regression to ascertain the effects of daytime or nighttime lighting conditions, TOR, and automation failure on the likelihood that a participant would crash. The logistic regression model was statistically significant, $\chi^2(2) = 41.053$, $p < .001$. The model explained 54.1% (Nagelkerke R^2) of the variance in accuracy and correctly classified 79.8% of cases. Holding lighting condition constant, participants in the failure condition were 69.74 times more likely to crash than participants in the TOR condition, Wald $\chi^2 = 15.82$, $p < .001$. However, lighting condition was not significant, Wald $\chi^2 = .82$, $p = .365$, odds ratio = 1.74. There were lower accuracy rates for the failure condition ($M = 38\%$) compared to the TOR condition ($M = 98\%$). Numerically, participants in the failure condition performed worse in the nighttime (29% or 6 out of 21 participants avoided crashing) compared to the daytime (48% or 10 out of 21 participants avoided crashing). These results partially supported Hypothesis 5b, demonstrating that automation failure had more crashes than TOR. However, it did not confirm the expected main effect of lighting condition. See Figure 9 for graph of results.
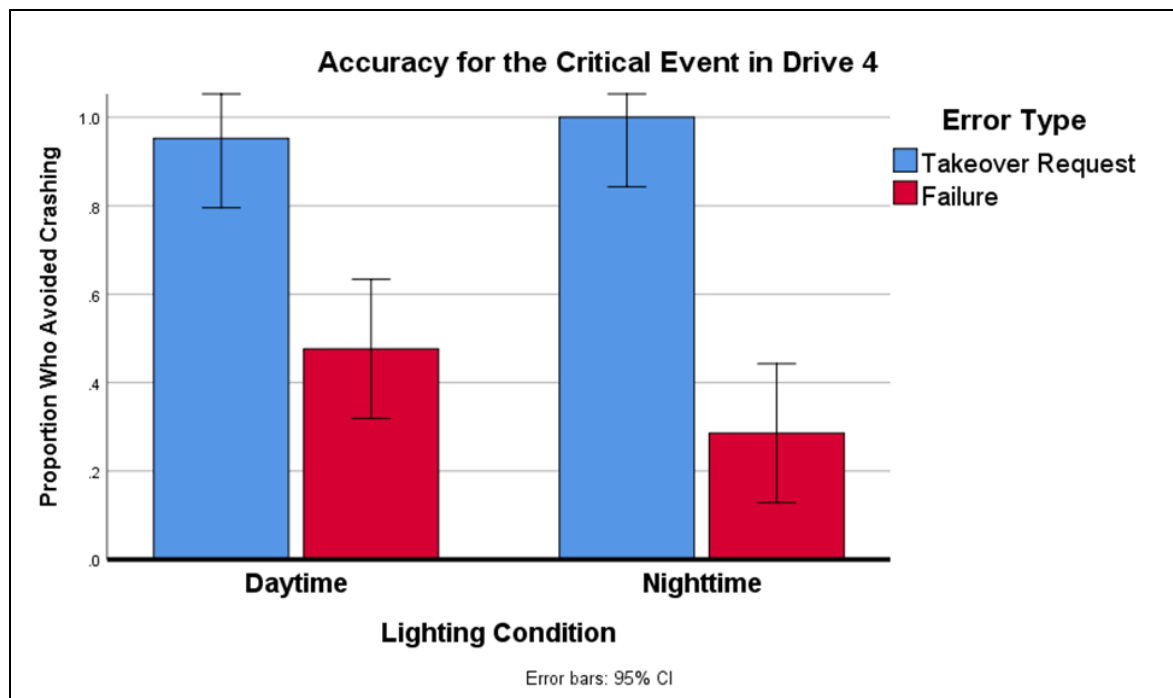
Figure 9. Accuracy as a function of error type and lighting condition.

3.3.2 Reaction time. RT was operationalized as the time interval between the moment the hazard was visible, which is the same moment the TOR was issued, and the first reaction (1-degree wheel turn). A total of 27 (32%) participants who crashed were not included in this analysis.

A significant main effect was found for the error type, such that average RTs in the TOR condition were significantly faster than those in the automation failure condition (Ms = 3.20s and 4.05s, respectively), $F(3, 56) = 6.62$, $p = .013$, $\eta_p^2 = .11$. No significant effect of lighting condition was found, $F(3, 56) = 1.34$, $p = .253$, $\eta_p^2 = .03$. Because crash trials were not used, it might not reflect the true performance because more people crashed in the nighttime condition than the daytime. However, even when considering participants who crashed as the maximum RT (5.82 s) instead of eliminating them from analysis, the conclusions were still the same for the

error condition (Ms = 3.26s for TOR and 5.15s Failure), $F(3, 56) = 54.35$, $p < .001$, $\eta_p^2 = .41$, and

for the lighting condition (Ms = 4.02s for Daytime and 4.39s for Nighttime), $F(3, 56) = 2.02$, $p =$

.159, $\eta_p^2 = .03$. Hypothesis 5a was also partially confirmed because RT was faster for

participants in the TOR as compared to the failure condition. However, the expected difference

in lighting condition did not occur (see Figure 10).
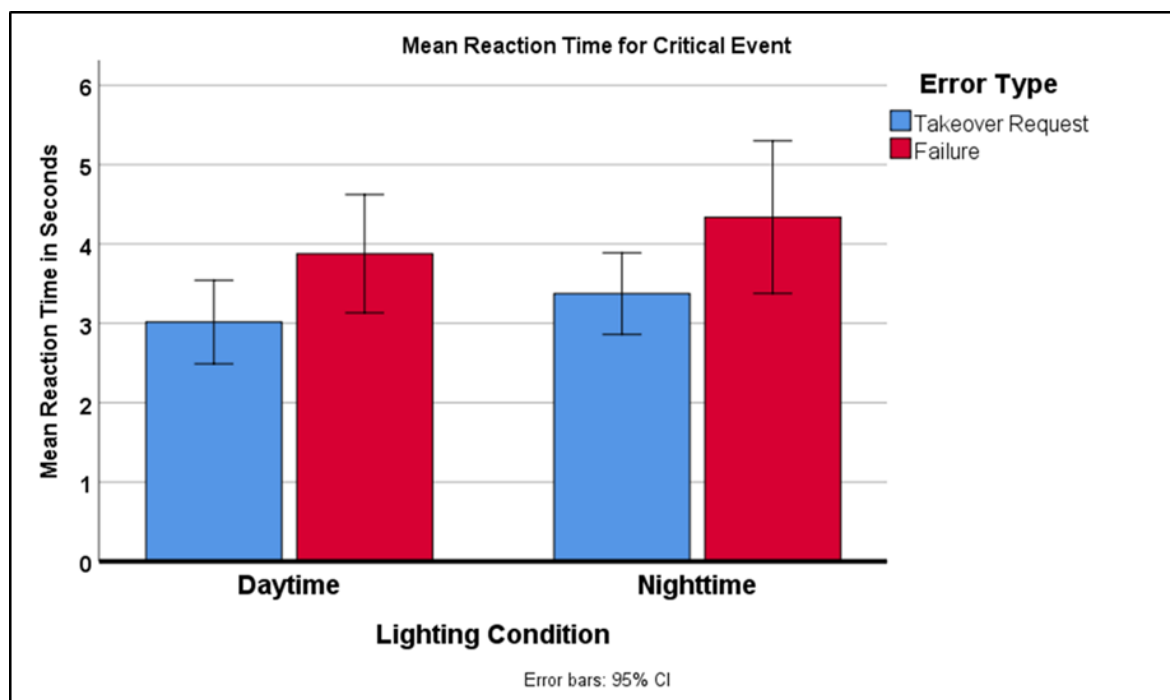
Figure 10. RT for error type and lighting condition with crash trials removed. Error bars are 95%

confidence intervals.

3.4 Usefulness of Automation (UAS)

The usefulness of the automation was rated at the end of the experiment. UAS data were

missing for six participants due to technical problems with collection and those participants were

excluded from the UAS analysis. To test Hypothesis 4 that UAS should show that errorless

automation was found most useful, followed by TOR, and failure to be least useful, a 2x3 ANOVA with lighting condition and error type was conducted to evaluate differences between groups. Results show that the main effect of error type was significant $F(2, 108) = 3.11$, $p = .048$, $\eta_p^2 = .06$, such that errorless automation was rated as significantly more useful than failure automation, $t(108) = 2.43$, $p = .017$. The difference between TOR and no error was marginally significant, $t(108) = 1.95$, $p = .054$, showing that TOR was found to be not as useful as the no error automation. TOR was not significantly different than failure, $t(108) = 0.45$, $p = .651$, indicating that the two had similar usefulness. The main effect of lighting condition was not significant $F(1, 108) = .083$, $p = .774$, $\eta_p^2 = .01$, nor was the interaction between lighting condition and error type, $F(2, 108) = 1.86$, $p = .160$, $\eta_p^2 = .03$. These results support Hypothesis 4, which expected UAS results to match with those of trust. The automated driving system in the no error condition was trusted more and found to be more useful than both the automated driving system in the TOR and Failure conditions. See Figure 11 for graph of results.
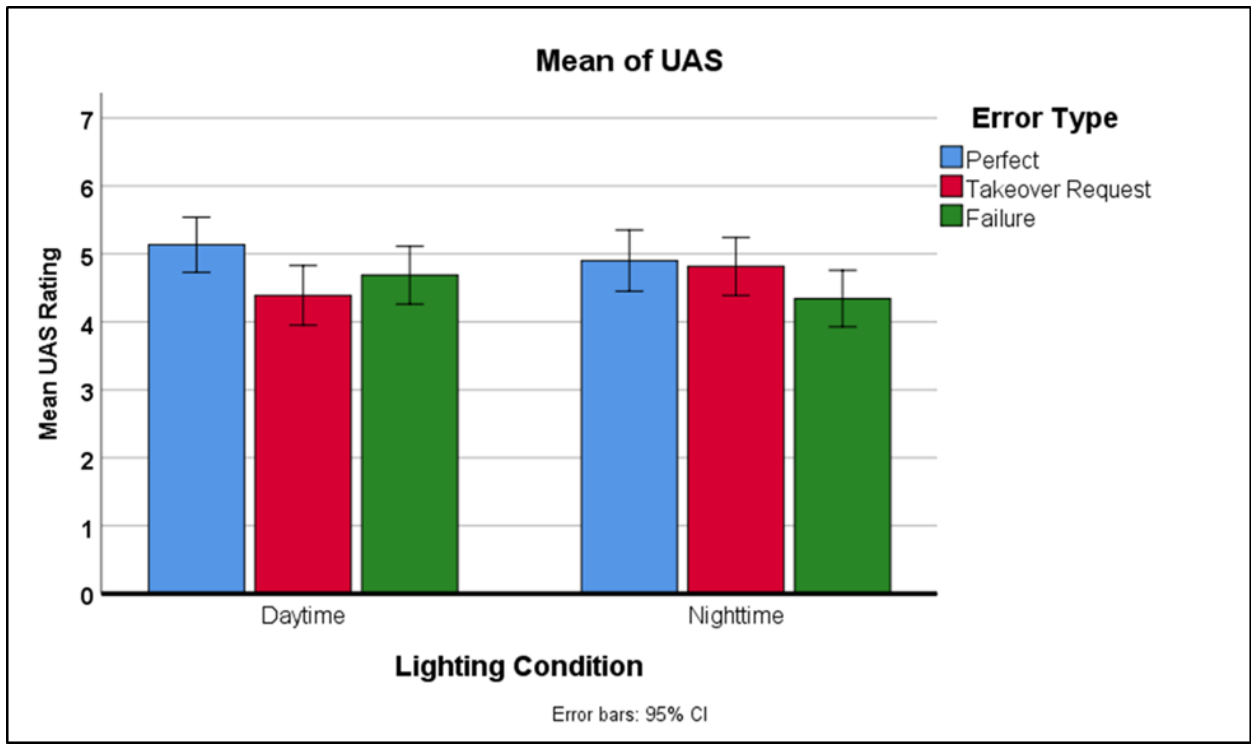
Figure 11. UAS for error type and lighting condition. Error bars are 95% confidence intervals.

CHAPTER 4

DISCUSSION

This study investigated the patterns of trust displayed by drivers of autonomous vehicles across seven drives in daytime and nighttime environments for three different types of automation responses (no error, TOR, and failure) to critical events. Trust was assessed after each drive to investigate learned trust, trust decline, and trust repair for drivers of autonomous vehicles. Daytime and nighttime conditions were expected to be a proxy for drive difficulty, although results did not confirm this. Ratings of drive difficulty did not differ between daytime and nighttime conditions. However, ratings of drive difficulty were different between error conditions, suggesting that participants rated difficulty based on performance of the automation as opposed to the lighting conditions. Results showed differences in trust for the error conditions, meaning that the performance of the car can influence participants' level of trust. Additionally, results provided evidence for dynamically learned trust through experience with the system, trust declines due to system errors, and baseline trust repair after an error.

4.1 Subjective Trust

    4.1.1 Pre-drive Trust. The demonstration of what the automation could do helped participants establish a reasonable baseline that was founded on actual experience and not on preconceived notions of what automated driving could or should be. As the results showed, participants had similar trust levels for each of the three error type conditions or between the daytime and nighttime conditions. An interesting finding was that the Automated Driving Opinion Survey (ADOS) was shown to be equivalent to the pre-drive trust. The ADOS gathered participants' opinions about automated driving to discern what trust levels were like before experiencing the automation. This result shows that participants' previously held opinions for

automated vehicles matched trust in the automated vehicle in the demonstration. This result was expected because most individuals have neutral, somewhat positive, or positive opinions about using automated driving (Kyriakidis et al., 2015) and the demonstration showed them positive aspects about automated driving.

4.1.2 Overall Trust. As predicted, participants in the no error automation condition trusted the automation more than those in the TOR and failure conditions. This was expected because the no error condition did not demonstrate any problems, so there was no reason for participants to lose trust in the automation. In contrast, automation errors can degrade trust (Hoff & Bashir, 2015; Lee & See, 2004) as can TOR (Hergeth, Lorenz, Krems, & Toenert, 2015). However, the finding that TORs can decrease trust was in contrast to findings from previous works that showed no differences between trust before and after a TOR (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015; Körber, Prasch, & Bengler, 2018). These authors argue that TORs are not perceived as failures, so they should not decrease trust in the automation. Nevertheless, the driver might not understand why the vehicle is issuing a TOR and may perceive it to be a result of an automation failure, especially if they are not well versed in the limitations of automation. Automation transparency is the idea that an automated system provides the user with information about what it is doing and why it is doing it. Previous research has shown that added transparency can help improve operator trust and performance (Adams & Webb, 2003; Lyons et al., 2016). Explaining the specific reasons for a TOR to a human driver could help the driver better understand and utilize the system.

In the present experiment, the TOR was issued without explanation to avoid an unexpected construction hazard. Gold et al., (2015) and Körber et al., (2018) measured trust pre- and post-drive but allowed participants to continue driving after the TOR. Körber et al., (2018)

manipulated if an explanation was given for why the TOR occurred, which showed no trust difference from the control. However they manipulated TOR behavior within groups, meaning participants saw a TOR three separate times. Gold et al., (2015) did not offer explanations for TORs, but still had participants experience three TORs before the post-drive trust measure. The present experiment measured trust directly after the first and only critical hazard event and corresponding vehicle error. Perhaps in the other studies, trust was repaired back to normal levels through either explanations or multiple TORs before measurement. This explanation is supported by Hergeth et al. (2015) who found small decreases in trust directly after TOR that repaired over time, which mirrors the findings of the present study. Even though the present findings did not show trust repair back to pre-error levels, it did reach the level of initial learned trust and had a continuing positive trend. Differences between this study and Gold et al. and Korber et. al, like multiple TORs or explanations for TORs, could be the additional difference needed to repair trust back to normal levels. However, neither of those studies measure trust directly after the TOR, so it is difficult to compare their findings to those from the present study. The findings of this study add to previous understandings of how trust is affected by TORs by showing that the timing of trust measurement could provide differential resutls. It demonstrates that TORs can decrease trust in automated driving sitiatons, particularly if individuals are unaware of why a TOR was issued.

Because the construction in the driving lane was rapidly approaching, participants in the TOR condition might have felt that they did not have sufficient time to take over control from the automation. Even though the critical construction hazard scenario required participants to takeover at a relatively normal transition time of about 6 seconds, it still required fast action (Eriksson & Stanton, 2017). Large and colleagues (2019) found that acceptance of warnings was

lowest when the TTC was shortest. However, their short TTC was two seconds as compared to six seconds in the present study. Further information from participants about their thought processes is needed before any conclusion could be made.

4.1.3 Dynamically Learned Trust. Pre-drive and drives 1-3 showed an overall increase in trust. This demonstrates that there was a building of trust as participants gained more experience with the system. This demonstrates that as participants became more familiar with an errorless driving automation, their trust in the automation improved. Initial levels of trust in the system were moderate, which makes sense because they had little to no experience with the system. The automated driving system was errorless through the first three drives and trust should increase because an errorless automation system should have high levels of trust. Thus, the increase over the first three drives demonstrates proper trust calibration (Hancock et al., 2011; Lee & See, 2004; Parasuraman & Riley, 1997). This initial development of trust is a promising display of how trust might grow in an initial experience with an automated vehicle. However, the participants in the no error automation condition did not continue the positive trend through all seven drives. Even when alarms and warnings are highly reliable, human performance and trust in automation may not always match the reliability level (Bliss & Acton, 2003; Chen et al., 2018). Participants in the present study had to learn the reliability of the system over time because they were not explicitly told a reliability percentage. Previous research shows that there can be differences if people are explicitly told a systems reliability versus if they need to infer the reliability through experience (i.e., the description-experience gap; Hertwig, Barron, Weber, & Erev, 2004; Chen et al., 2018). Human-human trust and human-automation trust can take a long time to learn and develop (Adams & Webb, 2003). Therefore, it could take

more time and experience with an automated driving system than participants had in this study to develop high levels of trust.

In the current study, the automation was able to demonstrate its capabilities several times during each drive by avoiding multiple hazards. However, participants might not have perceived all the potential hazards present in each drive. Real-world automated vehicles must constantly monitor the environment and surrounding objects to continue driving forward and predict and avoid any potential upcoming hazards. Research on vigilance has shown that participants' attention and ability to monitor decreases with time spent on task (Cabrall, Happee, & de Winter, 2016; Mackworth, 1948). As the participants spent more time with the automated driving system which required no input from them, they may have become less vigilant and prone to mind-wandering. The easiness of the task in the no error condition probably contributed to a lack of vigilance and inattention to the automation's performance. Vigilance and attentional measures would be helpful to see how participants are engaging in the driving task over time.

Various studies have focused on how to increase user acceptance and use of autonomous vehicles through matching personal driving style with automation performance (De Gelder et al., 2016; Kuderer, Gulati, & Burgard, 2015; Li, Li, Cheng, & Green, 2017). De Visser et al. (2018) give a hypothetical situation where a human driver is getting anxious because his automated vehicle is switching lanes rapidly and the following distance is uncomfortably close. The automation used in the current experiment had many instances of varied braking behaviors at signs and lights or passing and avoidance behaviors for vehicles and objects. Therefore, it is reasonable that some participants did not feel entirely comfortable with the driving strategy of the automated vehicle. Thus, this could be a valid explanation for why trust in the no error condition stagnated. People do not yet have much experience with automated vehicles, and thus

it could take longer to adapt to their driving style or become comfortable with automated driving. It is important for designers and researchers to consider that it might take a while for individuals to adapt to an automated driving style. Trust and adaptation to automated driving might take longer to develop than previously expected.

4.1.4 Drive 4 Trust. Looking specifically at drive 4 where error type was manipulated, trust was significantly different for the no error condition, but there was no significant difference between TOR and failure. As discussed previously in the overall trust section, the TOR was expected to decrease trust less because it was a warning as opposed to an automation failure. Reasons for this finding echo the reasons discussed for the lack of overall trust differences. However, the decrease in trust directly after the critical hazard seems to be at least numerically greater for failure than for TOR. Follow-up equivalence test determined that equivalence of trust between TOR and failure for drive 4 could not be concluded. Therefore, this should not be taken as conclusive evidence that the two are different. The future research section discusses plans to further test the trust decline between TOR and failure.

4.1.5 Trust Repair. For the three drives after drive 4, trust was slowly repaired for the TOR and Failure drives. However, the increase did not continue upward, but remained at the same level for the last two drives. The findings of the study indicated that although trust can rebound after a critical event, it still did not reach trust levels where no problem had occurred. TORs and automation failures can have a lasting effect on trust, which can potentially hurt later human-automation collaboration. Participants likely did not know the exact cause of the problem and could have worried that another error could happen again, with no way to predict it.

This experiment did not use any specific trust repair intervention strategy but relied only on subsequent errorless trials to help establish a theoretical baseline for trust repair for automated

driving (de Visser et al., 2018). The baseline repair in this case shows that a small increase did occur, but recovery did not continue. The automated driving system in the current study technically 'ignored' the failure, but this would be typical of situations in automated driving, in which the system did not notice a hazard, thus it wouldn't know if a repair strategy was even necessary. Hence, the system is not intentionally ignoring the failure, it just was not capable of responding to the failure. It is not clear where TORs fit into the repair strategy framework list proposed by de Visser et al., (2018) because they offer more information that a failure, but still do not invoke a repair strategy. The warning that a hazard is upcoming is a correct notification, but the automation has failed to do the job of driving autonomously. Perhaps it is a less explicit form of 'Recognize' in that the automation admits it cannot properly function and requests help. No supporting evidence or research has been provided for the recognize repair strategy by de Visser et al., (2018). If the failure to maintain automated driving is thought of as the error, then the TOR could be conceptualized as a 'Request Help' repair strategy to alleviate a loss of trust and to prevent a crash. If automated vehicle systems are to be human-automation collaboration teams, requesting help when one member of the team is unable to manage the problem is appropriate. Further research with automated driving system repair strategies can help identify which strategies work best and how TORs relate to trust repair.

4.2 Drive Difficulty

Difficulty rating was expected to be higher for nighttime driving. However, the observed difference was not significant. There are objective differences between daytime and nighttime driving that the subjective difficulty rating must not have captured. Nighttime driving has lower levels of overall luminance leading to hampered focal vision which could cause difficulty noticing and recognizing low luminance objects (Leibowitz & Owens, 1982). Additionally,

Plainis and Murray, (2002) showed that low luminance conditions lead to slower reaction times and can be a reason for the increase in rate and severity of crashes during the nighttime (Clarke et al., 2006). A study by Konstantopoulos, Chapman, and Crundall (2010) showed that drivers had longer fixations on objects during nighttime driving, resulting in longer processing times. These three studies demonstrate the objective differences between daytime and nighttime driving, indicating that driving at night is more difficult. One self-report study showed that night driving is perceived as more difficult and demanding compared to daytime driving. Participants could not compare lighting levels because they only saw either the daytime or nighttime condition, so they could not compare them might not have considered light levels in their rating.

Further evidence that difficulty reflected a different construct than lighting was the drive and error condition interaction. Participants in the failure condition rated drive 4 as significantly more difficult than those in TOR and no error. The drives were identical other than the automation performance, so participants must have considered automation performance when estimating difficulty. In addition, the significant difference in accuracy performance between TOR and no error conditions indicate that the participants' own performance was affecting their difficulty rating. More participants crashed in the failure condition than the TOR condition, so they may have rated it more difficult.

Drive difficulty for the drives after the critical event (drives 5-7) was higher for the failure condition than the no error condition. Even though the drives were identical, participants who saw the failure in drive 4 rated subsequent drives as more difficult. This finding suggests that the failure led participants to believe the drives were more difficult due to the automation's lower performance. The failure could be acting as a cognitive anchor (Epley & Gilovich, 2006; Tversky & Kahneman, 1973, 1992). The failure event caused participants to rate drive 4 at high

difficulty, anchoring participants ratings at a higher reference level. Then for subsequent drives, ratings for difficulty are higher due to the high anchor.

Findings from previous research show that there are no significant trust differences if the participant does not perceive the task as difficult (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003; Madhavan, Wiegmann, & Lacson, 2006; Schwark, Dolgov, Graves, & Hor, 2012). These previous studies mirror the findings of the current study because participants did not rate the two lighting conditions as different for difficulty. Thus, no trust differences were found between the lighting conditions due to lack of perceived difficulty.

4.3 Driver Performance

The performance results help inform how much participants gained or lost in terms of safety by having either an errorless aid, a warning (TOR), or no automation help (failure). Participants in the no error condition did not need to respond because the automation avoided the critical event for them. Therefore, no performance data were obtained in the no error condition.

4.3.1 Accuracy. Participants in the nighttime condition were expected to have more crashes in the failure condition due to it being more difficult to see the upcoming hazard. There were numerically more people that crashed in the nighttime than the daytime (36% versus 29%), but it was not statistically significant. More than 50% of participants in both the daytime and nighttime failure conditions crashed, indicating the task was likely difficult regardless of lighting conditions. This could have been because drivers had only six seconds to realize that the automation had failed and to react to the potential collision. In contrast, very few participants in the TOR condition crashed (1%). This indicates that the taking-over task itself was not too difficult, but taking-over without a warning made the task more difficult. Current data could not inform whether participants were simply trusting that the automation would avoid the hazard or

if attention had waned. Future researchers should investigate vigilance and use eye tracking to determine attention.

4.3.2 RT. As expected, participants who received the TOR were significantly faster than those in the failure condition. The warning helped participants recognize the need to act because the automated system was not going to avoid the hazard. Participants in the failure condition were likely waiting for the automation to act, and it took longer to realize that action was required. Often, participants realized too late, leading to the crashes described earlier.

These performance measures are not new findings, but rather a confirmation of past research that warnings improve RT and accuracy (Porter, Irani, & Mondor, 2008; Ruscio, Ciceri, & Biassoni, 2015; Xiang, Yan, Weng, & Li, 2016). However, this confirmation served to demonstrate that the experimental construct was valid and any discussion about trust between these conditions could be considered appropriate. In addition, the RT measures helped demonstrate how these scenarios compare. Showing that the errorless automation was safer than the TOR condition demonstrates how important it is to limit the amount of times TOR are required. Additionally, these results show baseline performance for TORs and automation failures and can be used in comparison with different warning types to test their effectiveness.

4.4 Limitations and Future Research

The current study used a seven-drive framework to look at trust development over time. Although the seven-drive timespan is longer than typical automated driving studies, it is still a relatively short period. Future researchers should investigate effectiveness of trust repair strategies using longitudinal designs. More experience and interactions with an automated driving system could change the way a user trusts and interacts with the system (Adams & Webb, 2003; Hoff & Bashir, 2015). As automated driving systems become more popular in the

commercial market, it will be important to measure how people are using, trusting, and learning about their systems.

No active trust repair strategy was used for this study. However, this was done intentionally to examine how trust could be repaired naturally without any intervention so that later studies could be compared to the baseline. The current design could be modified to integrate actual trust repair strategies such as apologize, explain, and emotionally regulate, as discussed in de Visser et al. (2018). Some of these strategies incorporate a more transparent system that tells the operator about why the automation is performing a certain way. The explain strategy is an explicit and active form of automation transparency where the automation intentionally tries to repair trust. Automation transparency can provide varying levels of information regarding its actions or present them less explicitly or only upon request. Comparing the baseline repair of trust found in this study to active trust repair strategies should demonstrate effectiveness of the repair strategy and inform designers of the best strategies for trust repair.

Other methods of TOR could be implemented to further investigate how 'Request Help' could be achieved. For example, semantic TORs could be issued that explain exactly why the automation needs the human driver to take over. Instead of uninformative beeps as is traditional in most TORs, specific information regarding the systems limitations could help assuage some of the misunderstandings associated with TORs. Special focus towards getting the human driver and the automated driving system to work as a collaborative partnership, as in human-human relationships could be helpful for performance and utility.

Anthropomorphized agents could be used to foster a relationship closer to that of a human-human team. For example, an automated driving system could have a personalized voice, much like that of a GPS or smart home assistants like the Google Assistant or Amazon Alexa

(Forster, Naujoks, & Neukum, 2017). Anthropomorphism has been shown to increase trust and trust resilience in automated systems (Hoff & Bashir, 2015; de Visser et al., 2016), and some research has already been applied the idea to automated driving vehicles (Forster, Naujoks, & Neukum, 2017; Häuslschmid et al., 2017, Waytz, Heafner, & Epley, 2014). Future researchers should implement an anthropomorphized agent to personalize the automated driving system to investigate if trust resilience can be improved.

Another problem of the current study is that trust was self-reported on a Likert scale. Self-report trust scales tend to have moderate results overall and people tend to underestimate the reliability of automated support systems (Sanchez, Fisk, & Rogers, 2004), which could be a reason trust did not continue to grow. Additionally, the trust measure was given to participants eight separate times, which could have caused survey fatigue. In future studies it would be beneficial to have a shorter questionnaire so that participants do not spend so much time filling out surveys. Additionally, objective measures of trust such as eye tracking and physiological measures could be implemented as either a complimentary measure or a sole measure so that the drive is not interrupted.

This study utilized a convenience sample of undergraduate participants, which is not necessarily externally valid and generalizable to all populations. Replicating the experiment with different populations would help designers know how to design automated driving systems to fit different markets. Additionally, testing professional truck drivers or ride-sharing drivers like Uber or Lyft could yield potentially interesting differences from the general population.

Only two types of errors were tested in this experiment. In real-world situations, many different things could go wrong in an automated driving vehicle. Researchers should also consider investigating development of trust for different types of errors such as automation lane departure

and incorrect hazard warnings. The timing and number of errors can also be adjusted to see if trust decline and repair still follows similar trajectories. Different driving styles for the automated driving system could be tested to confirm if preferred driving style can impact trust. Perhaps driving style can be adjusted for each participant, the rates of trust can increase faster and higher.

Drive difficulty should still be investigated, but there are likely more effective ways of manipulating difficulty. Lighting condition could benefit from a within-groups design to allow drivers to compare the drive between conditions. Different types of difficulty such as high-traffic versus low-traffic situations or adverse weather conditions could also be implemented as a form of difficulty. As automation improves, more drivers will become engaged with secondary, non-driving-related tasks, which could result in different trust development over time. Because the automation is improved, users might have higher levels of initial learned trust resulting in a lower magnitude of dynamically learned trust development. Additionally, lower engagement with the system could result in the human-out-of-the-loop, leading to more severe consequences due to longer takeover times and larger impacts on trust (Eriksson & Stanton, 2017). Consequently, secondary tasks can also be intermixed to determine if different patterns of trust occur.

4.5 Conclusion

The results of the current study provide a theoretical baseline of how trust in automated driving systems develops over time and how automation errors shape this development of trust. The study did not demonstrate differences for trust between daytime and nighttime driving, which was due to a lack of perceived difficulty difference between the lighting conditions. Trust in automated vehicles was slow to develop and never reached high levels. TORs and automation failures have similar effects, decreasing trust directly after a critical event. Trust repair without

an active repair strategy is observed, although it does not reach previous levels of trust before the

critical event. Designers should pay careful attention to the amount of time drivers have spent

with automation because trust could take longer to develop than previously expected.

Additionally, TORs and automation failures can both be costly, and trust is slow to repair.

Therefore, limiting errors and implementing active trust repair systems could prove useful for

maintaining efficient human-automation collaboration.

REFERENCES

Adams, B. D., & Webb, R. D. G. (2003). Model of trust development in small teams. *Report to Department of National Defence. DRDC No. CR-2003*, *16*.

Anderson, J., Kalra, N., Stanley, K., Sorensen, P., Samaras, C., & Oluwatola, O. (2016). *Autonomous vehicle technology: A guide for policymakers*. Santa Monica: Rand Corporation, doi:10.7249/RR443-2

Bailey, N. R., & Scerbo, M. W. (2007). Automation-induced complacency for monitoring highly reliable systems: The role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science*, *8*, 321–348. doi:10.1080/14639220500535301

Bliss, J. P., & Acton, S. A. (2003). Alarm mistrust in automobiles: How collision alarm reliability affects driving. *Applied Ergonomics*, *34*, 499-509. doi:10.1016/j.apergo.2003.07.003

Bliss, J. P., Hunt, G., Rice, S., & Geels, K. (accepted). Signal trust and situational experience: Theoretical and experimental evidence. In P. Hancock & M. Mouloua, *Automation and Human Performance: Theory and Applications (2nd ed.)*. CRC Press.

Borojeni, S. S., Chuang, L., Heuten, W., & Boll, S. (2016). Assisting drivers with ambient take-over requests in highly automated driving. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Ann Arbor, MI, USA, 237–244. doi:10.1145/3003715.3005409

Brown, B., & Laurier, E. (2017). The trouble with autopilots: Assisted and autonomous driving on the social road. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA, 416–429. doi:10.1145/3025453.3025462

Cabrall, C. D. D., Happee, R., & De Winter, J. C. F. (2016). From Mackworth's clock to the

open road: A literature review on driver vigilance task operationalization. *Transportation Research Part F: Traffic Psychology and Behaviour*, *40*, 169-189.

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

Chancey, E. T., Bliss, J. P., Yamani, Y., & Handley, H. A. H. (2017). Trust and the compliance–reliance paradigm: The effects of risk, error bias, and reliability on trust and dependence. *Human Factors*, *59*, 333–345. doi:10.1177/0018720816682648

Chen, J., Mishler, S., Hu, B., Li, N., & Proctor, R. W. (2018). The description-experience gap in the effect of warning reliability on user trust and performance in a phishing-detection context. *International Journal of Human-Computer Studies*, *119*, 35-47. doi:10.1016/j.ijhcs.2018.05.010

Clarke, D. D., Ward, P., Bartle, C., & Truman, W. (2006). Young driver accidents in the UK: The influence of age, experience, and time of day. *Accident Analysis & Prevention*, *38*, 871-878. doi:10.1016/j.aap.2006.02.013

De Gelder, E., Cara, I., Uittenbogaard, J., Kroon, L., van Iersel, S., & Hogema, J. (2016). Towards personalised automated driving: Prediction of preferred ACC behaviour based on manual driving. *2016 IEEE Intelligent Vehicles Symposium (IV),* Gothenburg, Sweden, 1211-1216. doi:10.1109/IVS.2016.7535544

De Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, *22*, 331. doi:10.1037/xap0000092

De Visser, E. J., Pak, R., & Shaw, T. H. (2018). From 'automation' to 'autonomy': The importance of trust repair in human–machine interaction. *Ergonomics*, *139*, 1–19. doi:10.1080/00140139.2018.1457725

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

Dogan, E., Rahal, M. C., Deborne, R., Delhomme, P., Kemeny, A., & Perrin, J. (2017). Transition of control in a partially automated vehicle: Effects of anticipation and non-driving-related task involvement. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 205–215. doi:10.1016/j.trf.2017.01.012

Duhigg, C. (2016). *Smarter Faster Better: The Transformative Power Of Real Productivity*. New York:Random House.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*, 83–87. doi:10.1111/1467-8721.01235

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*, 697–718. doi:10.1016/S1071-5819(03)00038-7

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

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of

control in automation. *Human Factors*, *37*, 381–394. doi:10.1518/001872095779064555

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science*, *17*, 311-318.

Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: Noncritical transitions to and from manual control. *Human Factors*, *59*, 689–705. doi:10.1177/0018720816685832

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, *77*, 167–181. doi:10.1016/j.tra.2015.04.003

Forster, Y., Naujoks, F., & Neukum, A. (2017). Increasing anthropomorphism and trust in automated driving functions by adding speech output. *2017 IEEE Intelligent Vehicles Symposium (IV),* Redondo Beach, CA, USA, 365-372. doi:10.1109/IVS.2017.7995746

Fredrick, E., Mikael, J., & Jana, S. (2017). Creating appropriate trust for autonomous vehicle systems: a framework for human-machine interaction design. *95th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA, *48*, 1–7. doi:10.1111/ijlh.12426

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). Take over! How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting*, San Diego, CA, USA, 1938–1942. doi:10.1177/1541931213571433

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*, 3025–3032. doi:10.1016/j.promfg.2015.07.847

Greenlee, E. T., DeLucia, P. R., & Newton, D. C. (2018). Driver vigilance in automated

vehicles: Hazard detection failures are a matter of time. *Human Factors*, *60*, 465-476. doi:10.1177/0018720818761711

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*, 517-527. doi:10.1177/0018720811417254

Häuslschmid, R., von Buelow, M., Pfleging, B., & Butz, A. (2017). Supporting trust in autonomous driving. *Proceedings of the 22nd International Conference on Intelligent User Interfaces,* Limassol, Cyprus, 319-329. doi:10.1145/3025171.3025198

Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013). Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Eindhoven, Netherlands, 210–217. doi:10.1145/2516540.2516554

Hergeth, S., Lorenz, L., Krems, J. F., & Toenert, L. (2015). Effects of take-over requests and cultural background on automation trust in highly automated driving. *2015 Driving Assessment Conference,* Salt Lake City, Utah, USA, 331–337. doi:10.17077/drivingassessment.1591

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534–539. doi:10.1111/j.0956-7976.2004.00715.x

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust . *Human Factors, 57*, 407–434. doi:10.1177/0018720814547570

Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes*, *51,* 296-312.

Kalra, N., & Paddock, S. M. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice*, *94*, 182–193 doi:10.1016/j.tra.2016.09.010

Katrakazas, C., Quddus, M., Chen, W. H., & Deka, L. (2015). Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions. *Transportation Research Part C: Emerging Technologies, 60*, 416–442. doi:10.1016/j.trc.2015.09.011

Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis and Prevention*, *42*, 827–834. doi:10.1016/j.aap.2009.09.022

Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., & Nass, C. (2015). Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing*, *9*, 269–275. doi:10.1007/s12008-014-0227-2

Körber, M., Prasch, L., & Bengler, K. (2018). Why do I have to drive now? *Post hoc* explanations of takeover requests. *Human Factors*, *6*, 305-323.

Kuderer, M., Gulati, S., & Burgard, W. (2015). Learning driving styles for autonomous vehicles from demonstration. *2015 IEEE International Conference on Robotics and Automation,* Seattle, Washington, USA, 2641-2646. doi:10.1109/ICRA.2015.7139555

Kyriakidis, M., Happee, R., & de Winter, J. C. F. (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. doi:10.1016/j.trf.2015.04.014

Large, D. R., Kim, H., Merenda, C., Leong, S., Harvey, C., Burnett, G., & Gabbard, J. (2019). Investigating the effect of urgency and modality of pedestrian alert warnings on driver acceptance and performance. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 11-24. doi:10.1016/j.trf.2018.09.028

Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, *40*, 153–184. doi:10.1006/ijhc.1994.1007

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*, 50–80. doi:10.1518/hfes.46.1.50_30392

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*, 1243–1270. doi:10.1080/00140139208967392

Leibowitz, H. W., & Owens, D. A. (1977). Nighttime driving accidents and selective visual degradation. *Science*, *197*, 422-423.

Leonard, J., How, J., Teller, S., Berger, M., Campbell, S., Fiore, G., … Williams, J. (2009). A perception-driven autonomous urban vehicle. *Journal of Field Robotics, 25, 727-774.* doi:10.1002/rob.20262

Li, G., Li, S. E., Cheng, B., & Green, P. (2017). Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities. *Transportation Research Part C: Emerging Technologies*, *74*, 113-125.

Lyons, J. B., Koltai, K. S., Ho, N. T., Johnson, W. B., Smith, D. E., & Shively, R. J. (2016). Engineering trust in complex automated systems. *Ergonomics in Design*, *24*, 13-17. doi:10.1177%2F1064804615611272

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual research.

Quarterly Journal of Experimental Psychology, 1, 6-21. doi:10.1080/17470214808416738

Madhavan, P., Wiegmann, D. A., & Lacson, F. C. (2006). Automation failures on tasks easily performed by operators undermine trust in automated aids. *Human Factors*, *48*, 241–256. doi:10.1518/001872006777724408

Madsen, M., & Gregor, S. (2000). Measuring human-computer trust. *11th Australasian Conference on Information Systems,* Brisbane, Australia, *53*, 6-8

Mok, B., Johns, M., Lee, K. J., Miller, D., Sirkin, D., Ive, P., & Ju, W. (2015). Emergency, automation off: Unstructured transition timing for distracted drivers of automated vehicles. *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, Gran Canaria, Spain, 2458-2464. doi:10.1109/ITSC.2015.396

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors*, *38*, 311–322. doi:10.1177/001872089606380211

National Center for Statistics and Analysis. (2016). 2015 motor vehicle crashes: Overview. (Traffic Safety Facts Research Note. Report No. DOT HS 812 318). Washington, DC: National Highway Traffic Safety Administration.

Naujoks, F., Mai, C., & Neukum, A. (2014). The effect of urgency of take-over requests during highly automated driving under distraction conditions. *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics*, Krakow, Poland, 2099–2106.

Parasuraman, R., Hancock, P. A., & Olofinboba, O. (1997). Alarm effectiveness in driver-centered collision warning systems. *Ergonomics*, *40*, 390–399.

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

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, *2*, 140–160. doi:10.1518/155534308X284417

Plainis, S., & Murray, I. J. (2002). Reaction times as an index of visual conspicuity when driving at night. *Ophthalmic and Physiological Optics*, *22*, 409-415.

Porter, M. M., Irani, P., & Mondor, T. A. (2008). effect of auditory road safety alerts on brake response times of younger and older male drivers: A simulator study. *Transportation Research Record*, 2069, 41–47. doi:10.3141/2069-06

Ruscio, D., Ciceri, M. R., & Biassoni, F. (2015). How does a collision warning system shape driver's brake response time? The influence of expectancy and automation complacency on real-life emergency braking. *Accident Analysis & Prevention*, 77, 72-81.

SAE. (2016). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. *SAE International*. doi:10.4271/J3016_201609

Sanchez, J., Fisk, A. D., & Rogers, W. A. (2004). Reliability and age-related effects on trust and reliance of a decision support aid. *Proceedings of the Human Factors and Ergonomics Society 48^h annual meeting*, New Orleans, LA, USA, *48*, 586–589. doi:10.1177/154193120404800366

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2015). A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. *Human Factors*, *58*, 377–400. doi:10.1177/0018720816634228

Schwark, J. D., Dolgov, I., Hor, D., & Graves, W. (2013). Gender and personality trait measures impact degree of affect change in a hedonic computing paradigm. *International Journal of*

Human-Computer Interaction, *29*, 327-337.

Singh, S. (2018). *Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey*. (Traffic Safety Facts Crash Stats. Report No. DOT HS 812 506). Washington, DC: National Highway Traffic Safety Administration.

Sirkin, D., Martelaro, N., Johns, M., & Ju, W. (2017). Toward measurement of situation awareness in autonomous vehicles. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, CO, USA, 405–415. doi:10.1145/3025453.3025822

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.

Van Der Heiden, R. M. A., Iqbal, S. T., & Janssen, C. P. (2017). Priming drivers before handover in semi-autonomous cars. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, CO, USA, 392-404. doi:10.1145/3025453.3025507

Verberne, F. M. F., Ham, J., & Midden, C. J. H. (2012). Trust in smart systems: Sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars. *Human Factors*, *54*, 799–810. doi:10.1177/0018720812443825

Walch, M., Lange, K., Baumann, M., & Weber, M. (2015). Autonomous driving: Investigating the feasibility of car-driver handover assistance. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Nottingham, UK, 11–18. doi:10.1145/2799250.2799268

Xiang, W., Yan, X., Weng, J., & Li, X. (2016). Effect of auditory in-vehicle warning information on drivers' brake response time to red-light running vehicles during collision avoidance. *Transportation research part F: Traffic Psychology and Behaviour, 40,* 56-67. doi:10.1177/1541931218621440

Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, *52*, 113–117. doi:10.1016/j.jesp.2014.01.005

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accident Analysis and Prevention*, *78*, 212–221. doi:10.1016/j.aap.2015.02.023

Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis and Prevention*, *92*, 230–239. doi:10.1016/j.aap.2016.04.002

APPENDIX A

A PRIORI POWER ANALYSIS

Using a 2 (A: Lighting Condition) x 3 (B: Error Type) x 7 (C: Drive) mixed design with all interaction terms and the A and B terms as between-groups and C as within-groups, the power analysis software performed 100 simulations to estimate the power of an N of 90, 120, and 150.
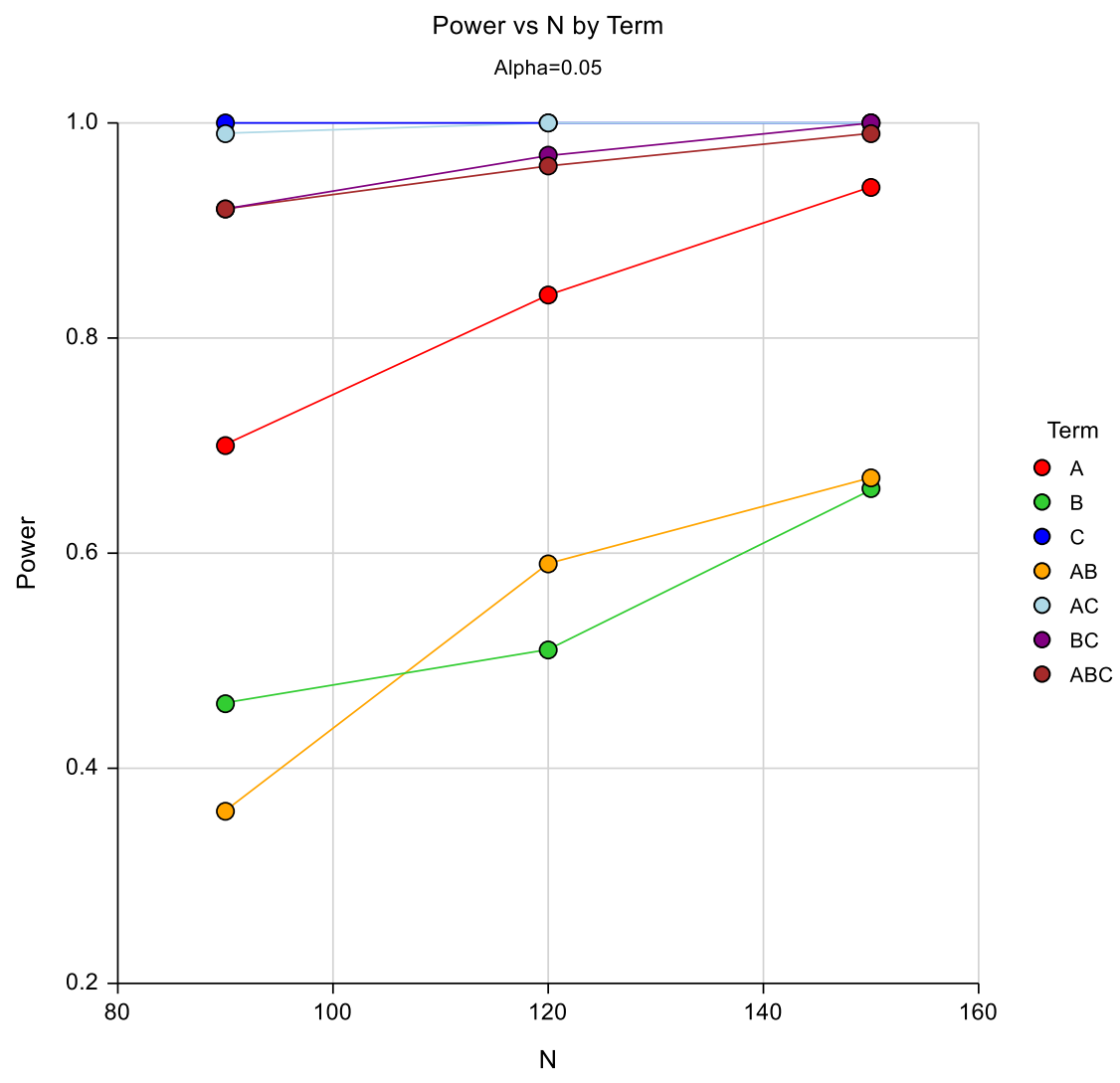
APPENDIX B

DEMOGRAPHIC INFORMATION SURVEY

Date	RA	Expt	Subject No.	Age	Gender		American Indian/ Alaska Native	Asian	African/ African American	Caucasian	Native Hawaiian/ Pacific Islander	More than one race	Other/ Unknown	Dominant Hand L/R	Vision (Y/N)	Hearing (Y/N)
					M	F										

Gender | *Ethnicity (please select one)* | *Dominant Hand* | *Normal or Corrected to Normal*

APPENDIX C

AUTOMATED DRIVING OPINION SURVEY

The purpose of this survey is to gather your experience with manual driving and your opinions on automated driving. The following is a description of the levels of driving automation from the 2016 SAE Taxonomy and Definitions for Driving Automation systems that is referenced in the following survey.

- **Manual driving:** The human driver executes the driving task him/herself using the steering wheel and pedals.

- **Partially Driving Automation:** The automated driving system takes over both speed and steering control on some roads. However, the system cannot handle all possible situations. Therefore, the driver shall permanently monitor the road and be prepared to take over control at any time.

- **Conditional Driving Automation:** The automated driving system takes over both speed and steering control on most roads. The driver is not required to permanently monitor the road. If automation cannot handle a situation it provides a take-over request, and the driver must take-over control with a time buffer of 7 s.

- **Highly Driving Automation:** The automated driving system takes over both speed and steering control on all roads. The driver is not required to permanently monitor the road. There is no expectation of the user to respond to a request to intervene.

- **Full Driving Automation:** The system takes over speed and steering control completely and permanently, on all roads and in all situations. The driver sets a destination via a touchscreen. The driver cannot drive manually, because the vehicle does not have a steering wheel.

APPENDIX C (Continued)

Question	Unit/Coding
1. Have you read and understood the above instructions?	1 = Yes
2. The definitions given in the instructions are clear to me.	1= Disagree Strongly, 7= Agree Strongly
4. What is your primary mode of transportation?	1= Private Vehicle, 2= Public Transportation, 3= Motorcycle, 4= Walking, 5= Other
5. At what age did you obtain your first driver's license?	Year
6. On average, how often did you drive a vehicle in the last 12 months?	1=Never, 6=Every Day
7. About how many miles did you drive in the last 12 months?	1=0, 2=1-5000, …, 11= more than 50,000
8. Have you ever heard of the Google Driverless Car (Waymo) or other driverless cars?	2=No, 1=Yes
9. The idea of fully automated driving is fascinating.	1= Disagree Strongly, 7= Agree Strongly
10. Manual driving is enjoyable.	1= Disagree Strongly, 7= Agree Strongly
11. Partially Driving Automation will be enjoyable.	1= Disagree Strongly, 7= Agree Strongly
12. Conditional Driving Automation will be enjoyable.	1= Disagree Strongly, 7= Agree Strongly
13. Highly Driving Automation will be enjoyable.	1= Disagree Strongly, 7= Agree Strongly
14. Fully Driving Automation will be enjoyable.	1= Disagree Strongly, 7= Agree Strongly
15. Partially Driving Automation will be easier than manual driving.	1= Disagree Strongly, 7= Agree Strongly
16. Conditional Driving Automation will be easier than manual driving.	1= Disagree Strongly, 7= Agree Strongly
17. Highly Driving Automation will be easier than manual driving.	1= Disagree Strongly, 7= Agree Strongly
18. Fully Driving Automation will be easier than manual driving.	1= Disagree Strongly, 7= Agree Strongly
19. Some modern cars are equipped with Adaptive Cruise Control (a system that can automatically follow another car). How often did you use Adaptive Cruise Control (ACC) when driving in the last 12 months?	1=Never, 6=Every Day, (-1 = I do not have ACC, -2= I do not know what ACC is)
20. I would be comfortable driving in a fully automated driving vehicle without a steering wheel.	1= Disagree Strongly, 7= Agree Strongly
21. The idea of fully automated driving is silly. Scientists should focus on other, more important, research topics.	1= Disagree Strongly, 7= Agree Strongly
22. I believe that within 30 years, automated driving systems will be so advanced that it will be irresponsible to drive manually.	1= Disagree Strongly, 7= Agree Strongly

APPENDIX D

HUMAN-COMPUTER TRUST QUESTIONNAIRE

All questions were rated on a 1-7 (*Strongly Disagree-Strongly Agree*) Likert scale.

Instructions - "This survey sometimes uses language like, "actions", "decision", and "problem".

Please interpret these in the context of the driving automation system. 'Problems' are potential

hazards on the road, 'Actions' are movements that the vehicle makes, and 'Decisions' are actions

taken in response to problems where multiple different actions could be made."

1. The automated driving system always acts in a way that I would agree with
2. The automated driving system performs reliably
3. The automated driving system responds the same way under the same conditions at different times
4. I can rely on the system to function properly
5. The automated driving system analyzes problems consistently
6. The automated driving system uses appropriate methods to reach decisions
7. The automated driving system has sound knowledge about navigating the driving environment built into it
8. The actions the automated driving system produces is as good as that which a highly competent person could produce
9. Please select the "Correct answer" option – Response 2 was changed to "Correct answer"
10. The automated driving system makes use of all the knowledge and information available to it to produce its solution to the problem
11. I know what will happen the next time I use the automated driving system because I understand how it behaves
12. I understand how the automated driving system will make decisions
13. Although I may not know exactly how the automated driving system works, I know how to use it
14. It is easy to follow what the automated driving system does
15. I recognize what I need to do when the automated driving system is taking action
16. I believe actions from the automated driving system are safe even when I don't know for certain that it is correct
17. When I am uncertain about an action, I believe the automated driving system rather than myself
18. If I am not sure about an action, I have faith that the automated driving system will provide the best solution
19. When the automated driving system takes unusual actions, I am confident that the action is correct
20. Even if I have no reason to expect the automated driving system will be able to solve a difficult problem, I still feel certain that it will
21. I would feel a sense of loss if the automated driving system was unavailable and I could no longer use it
22. I feel a sense of attachment to using the automated driving system
23. I find the automated driving system suitable to my style of driving
24. I like using the automated driving system for driving
25. I have a personal preference for driving with the automated driving system

APPENDIX E

USEFULNESS OF AUTOMATED SYSTEM SURVEY

Questions marked with an asterisk were reverse coded. Questions marked with a

superscript x were not included in the calculation of usefulness due to a monetary requirement or

implication.

Variable	Question	Scale
Discharge of the driver due to automation	Highly Automated Driving decreases my problems while driving.	1 = Strongly Disagree, 7= Strongly Agree
	Highly Automated Driving enables me to manage useful activities while driving.	1 = Strongly Disagree, 7= Strongly Agree
	The system saves time that I would have lost driving manually.	1 = Strongly Disagree, 7= Strongly Agree
Safety Gain	The system increases road safety	1 = Strongly Disagree, 7= Strongly Agree
	The system prevents traffic violations	1 = Strongly Disagree, 7= Strongly Agree
	The system supports the driver to detect hazards in time	1 = Strongly Disagree, 7= Strongly Agree
	The system contributes to reduce crash risk	1 = Strongly Disagree, 7= Strongly Agree
Safety Loss	The system distracts from detecting hazards in time*	1 = Strongly Disagree, 7= Strongly Agree
	I drive safer than the vehicle in HAD mode*	1 = Strongly Disagree, 7= Strongly Agree
	Highly Automated Driving is vulnerable for new hazards like hacker attack and issues with data safety*	1 = Strongly Disagree, 7= Strongly Agree
	To me, new risks that emerge from Highly Automated Driving appear to be more serious*	1 = Strongly Disagree, 7= Strongly Agree
Perceived control of conduct	It is likely that I can use the system	1 = Strongly Disagree, 7= Strongly Agree
	There is no reason why I should not be able to use it	1 = Strongly Disagree, 7= Strongly Agree
	Whether I can use Highly Automated Driving is dependent upon me	1 = Strongly Disagree, 7= Strongly Agree
	I probably could not operate Highly Automated Driving*	1 = Strongly Disagree, 7= Strongly Agree
	I could not afford a Highly Automated Driving system[x]	1 = Strongly Disagree, 7= Strongly Agree
	I don't have any money for additional functions in my car[x]	1 = Strongly Disagree, 7= Strongly Agree
	For me, additional comfort functions are of high value[x]	1 = Strongly Disagree, 7= Strongly Agree
	Highly Automated Driving is not available for my vehicle type[x]	1 = Strongly Disagree, 7= Strongly Agree
Intention of use	I would like to have this system in my car	1 = Strongly Disagree, 7= Strongly Agree
	I would consider the use of the system	1 = Strongly Disagree, 7= Strongly Agree
	I would not use the system in any case	1 = Strongly Disagree, 7= Strongly Agree
	I would purchase the system together with my next car[x]	1 = Strongly Disagree, 7= Strongly Agree

VITA

Department of Psychology

Old Dominion University

Norfolk, VA 23529

Email: smishler17@gmail.com

Website: https://sites.google.com/view/scottmishler

EDUCATION

2017 - May 2019	**M.S., Psychology**, Old Dominion University, VA
	Thesis: *Whose drive is it anyway? Using multiple sequential drives to establish patterns of learned trust, error cost, and trust repair while considering daytime and nighttime differences as a proxy for difficulty*
2016 - 2017	**Master's Student, Engineering Psychology**, New Mexico State University, NM
	First Year Project: *Driver's type of response to auditory car warnings in a semi-autonomous vehicle mediates safety*
2012 - 2016	**B.S., Psychology** - Purdue University, IN
	Focus: Human Factors, Cognitive Psychology
	Senior Research Project: *Can the working memory representation of a stimuli mediate the Simon effect through variation in color category*

SELECT PUBLICATIONS AND CONFERENCE PROCEEDINGS

Chen, J., **Mishler, S.**, Hu, B., Li, N., Proctor, R. W. (2018). The Description-Experience Gap in the Effect of Warning Reliability on User Trust and Performance in a Phishing Detection Context. *International Journal of Human-Computer Studies,* 119, 35-47.

Yamaguchi, M., Chen, J., **Mishler, S**., & Proctor, R. W. (2017). Flowers and Spiders in Spatial Stimulus-Response Compatibility: Does Affective Valence Influence Selection of Task-Sets or Selection of Responses?. *Cognition and Emotion*, 1-15.

Chen, J., **Mishler, S.**, Hu, B. (2018). Conveying automation reliability and automation error type: An empirical study in the cyber domain. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* Austin, TX, USA, *62*, 172–173.

Mishler, S., Chen, J. (2018). Effect of response method on driver responses to auditory warnings in simulated semi-autonomous driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Austin, TX, USA, 62,* 1934–1938.

Mishler, S., Chen, J., Sabic, E., Hu, B., Li, N., & Proctor, R. W. (2017). Description-experience gap: The role of feedback and description in human trust in automation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* Los Angeles, CA, USA, 313.