Driver Behavioral Adaptation to In-Vehicle Technologies: Influence of Demands and Exposures

Yiyun Peng

A dissertation

Submitted in partial fulfillment of the

Requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Linda Ng Boyle, Chair

Christina M. Mastrangelo

John D. Lee

Zelda B. Zabinsky

Program Authorized to Offer Degree:

Industrial and Systems Engineering

University of Washington

Abstract

Driver Behavioral Adaptation to In-Vehicle Technologies: Influence of Demands and Exposures

Yiyun Peng

Chair of the Supervisory Committee:

Associate Professor Linda Ng Boyle

Industrial & Systems Engineering, Civil & Environment Engineering

In-Vehicle Information Systems (IVIS) can assist drivers by increasing both safety and efficiency, but may also divert drivers' attention away from the road and cause distractions. The goal of this dissertation is to examine the effect of entering and reading text while driving, and how the drivers adapt their behavior over time under different traffic conditions. Two experiments have been conducted for this purpose. The main objective of the first experiment is to understand drivers' behavior adaptation with varying task demands. The findings showed that drivers had longer eyes-off-road (EOR) time for tasks with higher demands. However, the increase of glance duration had a tendency to flatten out with the increasing task demands, which suggests a risk compensation behavior. The objective of the second experiment is to examine the drivers' behavior adaptation on using IVIS under different traffic conditions and over time. The results showed that EOR was significantly longer over time and shorter when there was traffic on the road. However, drivers were able to improve lateral control over time. This suggests that, with practice, drivers were able to have better control of the vehicle, but also tend to look longer off the road toward the IVIS tasks as they gain confidence, which may impose greater risks on the driver under complex driving situations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DEDICATION

To my family, for their unconditional love and support.

ACKNOWLEDGEMENTS

I'd like to express my deepest appreciation to my advisor Dr. Linda Boyle, for taking me on as a student six years ago, which has largely changed my life, and opened up great new possibilities for my future. I'd like to thank her for always being supportive, and guide me with patient and trust. The skills that I've learned and the experience I've had during working with Dr. Boyle will benefit me lifelong wherever I go. I'd also like to express my special thanks to my committee members, Dr. Christina Mastrangelo, Dr. John Lee, Dr. Zelda Zabinsky, Dr. Ted Klastorin, and Dr. Anne Goodchild, for their insightful comments and suggestions on my research and dissertation work.

I'd like to thank my colleagues and friends from the Human Factors and Statistical Lab, especially Dr. David Neyens, Dr. Huimin Xiong, Yuqing Wu, Erika Miller, and Chun-Cheng Chang for their great helps and suggestions on my study and research. Great thanks to Sean Xuan, Daniel Ullom, and Kayla Fukuda for their hard work on the data collection and reduction.

Finally, I'd like to recognize Federal Motor Carrier Safety Administration (FMCSA) and National Highway Traffic Safety Administration (NHTSA) for funding my research assistantship during my PhD study. Portion of this dissertation was based on the Text Reading and Text Input Assessment project funded by NHTSA (Contract # DTNH22-11-D-00237/0006), and I'd also like to express my thanks to the National Advanced Driving Simulator (NADS) for providing technical support on the experimental devices.

INTRODUCTION

Human beings tend to adapt continuously to the changing environment. For drivers, behavioral adaptation is defined as the change of behavior that occurs following the changes of the road/vehicle/user system (OECD, 1990). Behavioral adaptation may be both negative and positive. Counterproductive behavioral adaptation refers to the phenomenon where drivers become riskier when being supported by safety-enhancing systems, whereas positive behavioral adaptation refers to the phenomenon where drivers become safer. Behavior adaptation has been examined in a number of studies on its effect on in-vehicle driving assistant systems such as adaptive cruise control, and crash avoidance systems such as forward collision warning systems and lane departure warning systems. However, only a few studies examined the adaptation related to in-vehicle information system (IVIS) use.

IVIS can provide non-driving-related services such as radios and communications (e.g., phone calls, emails, and text messages). They can also provide drivers with a wealth of driving-related real-time information, such as roadway conditions, time delays, and alternative routes (Blanco, Biever, Gallagher, & Dingus, 2006; J. D. Lee, 1997; Vashitz, Shinar, & Blum, 2008). Properly designed IVISs should assist drivers in improving safety and efficiency. However, they may also distract drivers from the primary driving task if too demanding. The purpose of this dissertation is to understand driver behavioral adaptation under different IVIS task demands, driving demands, and over time.

Study Aims

Aim 1: *Examine behavior adaptation on IVIS tasks with varied task demands.* A driving simulator study with 28 drivers across all age range was conducted for this purpose. IVIS task demands were manipulated by varying the task types (text entry and text reading), text lengths (short, medium, and long), and the task irrelevant text (presence and absence). Drivers' glance behavior and vehicle control performance were examined in order to understand how drivers adapt their behavior under these different task demands.

Aim 2: *Examine behavior adaptation on IVIS tasks with varied driving demands, over time, and for drivers with different risk levels.* A second driving simulator study with 28 young drivers under 30 years old was conducted with the same IVIS tasks, but in two different driving scenarios (with and without traffic), and over 3 sessions in 7 consecutive days. Drivers were clustered into different groups based on their risk levels. Their glance behavior and vehicle control performance were examined and compared for different driving scenario and over time.

Aim 3: *Examine the relationship between the change in drivers' perceptions and actual driving performance over time.* Using the results from the second experiment, a correlation study was proposed to examine the relationship between drivers' perceptions on their adaptation and their actual adaption over time. This was aimed to understand whether drivers could correctly realize their behavior change, and what aspects of driving were more relevant when assessing their own performance.

Chapter 1 contains literature reviews of relevant previous studies on IVIS, driver distraction, and driver behavioral adaptations. Chapter 2 describes the statistical challenges for longitudinal experiments and the main statistical modeling technique adopted in this dissertation. Chapter 3 describes the first experiment and the results that answer Aim 1. Chapter 4 describes the second experiment and the results that answer Aim 2. Chapter 5 discusses the data analysis and results for Aim 3. Chapter 6 summarizes the overall findings in this dissertation and discusses the contributions and possible future research topics.

CHAPTER 1: BACKGROUND

This chapter introduces in-vehicle information systems (IVISs) and the related research of their influences on driver distraction. The impact of IVIS task types and characteristics on driving performance and glance behavior are discussed. Drivers tend to adapt according to specific driving environment and their experience levels when interacting with in-car technologies and secondary tasks. Previous findings on driver adaptation related to driving demand and long-term use are also discussed in this chapter. The objectives of the dissertation are introduced toward the end of the chapter.

In-Vehicle Information Systems and Driver Distraction

In-vehicle information systems (IVISs) can provide non-driving-related services such as radios and communications (e.g., phone calls, emails, and text messages). They can also provide drivers with a wealth of driving-related real-time information, such as roadway conditions, time delays, and alternative routes (Blanco et al., 2006; J. D. Lee, 1997; Vashitz et al., 2008). Properly designed IVISs should assist drivers in improving safety and efficiency. However, they may also distract drivers from the primary driving task if not designed properly.

A number of studies have shown the impact of IVIS on driving performance. Texting and interacting with navigation systems have been shown to increase drivers' reaction time and impact their ability to maintain proper lane position (Drews, Yazdani, Godfrey, Cooper, & Strayer, 2009; Hosking, Young, & Regan, 2006). Anttila & Luoma (2005) showed that both visual and cognitive IVIS tasks tended to increase improper yielding of the right-of-way to other vehicles and driving speed at urban intersections. Tsimhoni and Green (2001) found that map reading tasks increased standard deviation of lane position (SDLP) by almost 80% compared to driving only. Drivers, especially older drivers over 60 years old, significantly reduced their driving speeding while using entertainment systems compared to being engaged in phone conversations or not being distracted at all (Horberry, Anderson, Regan, Triggs, & Brownb, 2006). Lee, et al. (2001) showed that using a speech-based in-vehicle email system could increase reaction time to sudden events by 30%.

Drivers also spend more time looking at in-vehicle devices when performing IVIS tasks while driving (Chiang, Brooks, & Weir, 2004a), which reduces the chance to see unexpected events. When visual scanning is required by the IVIS tasks, drivers' visual attentions usually shift back and forth between the roadways and the in-vehicle display. Too much time looking away from the road can lead to larger lane deviations and slower responses to other vehicles on the road (Dingus, Antin, Hulse, & Wierwille, 1989; Donmez, Boyle, & Lee, 2007). Total off-road glances exceeding two seconds in a six second period during a safety-critical event doubles the risk of crashes/near-crashes (Klauer, Dingus, Neale, Sudweeks, & Ramsay, 2006). Early study by Zwahlen, Adams, and Debald (1987) also suggested that drivers' lane keeping performance became unacceptable when looking at an in-vehicle display for more than 2 seconds. In general, individual glances that exceed two seconds have been considered as critical for secondary tasks such as IVIS tasks. Therefore, designing in-vehicle displays that reduce drivers' off-road glances is essential for car manufactures to improve driver safety.

Driver Behavioral Adaptation and Risk Compensation

It is human nature to adapt continuously to the changing environment in order to survive. In the transportation domain, this means that drivers tend to adapt their behavior given the car they are driving, the technologies inside and outside the vehicle, and roadways continuously over time, in order to reach their destinations without being involved in crashes. OECD (1990) defined driver behavioral adaptation as the behaviors that occurred following the changes of the road/vehicle/user system but was unintended by the system engineers and designers. This behavioral adaptation may negatively or positively affect driver safety. Counterproductive behavioral adaptation, which affects safety in a negative way, refers to the phenomenon where drivers become riskier when being supported by safety-enhancing systems. For example, a study of Sagberg, Fosser, and Setermo (1997) revealed that although the purpose of antilock brake systems (ABS) was to increase driver safety, taxi drivers with such systems in their vehicles actually maintained significantly shorter time headway than drivers without ABS. Therefore, the safety effect of antilock brake systems was cancelled out by the drivers' negative behavioral adaptation. Risk compensation, which on the other hand positively affects driver safety, refers to drivers' responses to the changed perceived risk. Vaa (2007) further argued that risk compensation is a special type of behavioral adaptation, which occurs without involving consciousness.

There are a number of studies that have discussed driver behavioral adaptations to in-vehicle technologies such as driving assistant systems, intelligent transportation systems, and navigation systems. The following section discusses behavioral adaptation and risk compensation due to IVIS task demands, driving demands, and long-term use of the system.

Influence of IVIS Task Demands

IVIS tasks with different types, difficulties, and characteristics pose different demands on drivers and require varied attentions. Different types of IVIS tasks can affect vehicle control performance and glance behavior differently. Sending and receiving text messages using an embedded IVIS was shown to cause less performance degradation than using a handheld mobile phone (Owens, McLaughlin, & Sudweeks, 2011). However, speech and manual controlled IVIS tasks that include audio, phone number selection, and address entry tasks still decreased lateral control performance and increased reaction time (Maciej & Vollrath, 2009).

Ranney, Baldwin, Parmer, Martin, and Mazzae (2012) conducted a study to compare driving performance during several text entry related in-vehicle tasks. The tasks tested in their study include dialing a contact, destination entry, 10-digit dialing, and texting, using radio tuning as benchmark. The results showed that all entry tasks resulted in significantly more lane departure events and larger standard deviation of headway compared to the radio tuning task. Specifically, text messaging tasks resulted in the greatest performance degradation, followed by destination entry and phone dialing tasks. However, the study also found a strong positive correlation between task duration and lane departure frequency, and when the task duration was taken into account, only text messaging still had significantly larger number of lane departure events compared to radio tuning. Another study by Ranney, Baldwin, Parmer, Martin, and Mazzae (2011) also showed that text messaging resulted a larger SDLP and longer response time to task detection, compared to radio tuning. Destination entry task also took considerably longer time to complete compared to other tasks, which can expose drivers to additional risks.

The difficulty of the IVIS tasks may also influence the magnitude of the distraction. Based on literature reviews performed by previous studies (Metz, 2009; Metz, Schomig, & Kruger, 2011), the average number of off-road glances may range between 2.2 for simple radio tasks and 13.8 for complex infotainment tasks, while the mean single glance durations were similar and around 1.3 seconds. Horrey and Wickens (2007) illustrated that, even though the mean glance duration was similar for simple and complex IVIS tasks, the complex IVIS tasks significantly increased the proportion of long glances (i.e. more than 1.6 seconds), which were accounting for the majority of collision events. They thus further argued that the influence of complex IVIS should be examined using the tails of the glance distributions but not the means as the former measurement was more closely related to crash risks.

Studies have also shown that drivers cannot fully adjust their priorities in driving and recover their driving performance with an increased IVIS task demand. Merat, Lai, and Jamson (2011) showed that high (e.g., route selection) and medium difficulty (e.g., destination selection) navigation tasks resulted in larger speed variation and SDLP than low difficulty tasks (e.g., read visual information). An increase on IVIS task demand was also found to associate with shorter time-to-collision and reduced anticipation of sudden events (Jamson & Merat, 2005). Long duration map reading task was also found to increase mean glance duration by about 0.5 seconds compared to the short ones when performed on straight roadways (Tsimhoni & Green, 2001).

The design configuration can impact drivers' ability to use IVISs while driving. Design specifications such as control types (i.e., visual-manual vs. speech), information formats (e.g., paragraph, graph, table), and text characteristics (e.g., number of characters, fonts, color) can all influence driver performance. Tsimhoni, Smith, and Green (2004) compared speech and keyboard based address entry tasks while driving. The study showed that entering an address using the touchscreen keyboard resulted in worse lateral control performance, but there was no significant difference between driving only and speech recognition conditions. The following distance almost doubled while using the keyboard compared to using word-based speech recognition. Another study by Maciej and Vollrath (2009) found that drivers spent 30-40% of the time looking at the IVIS screen during address entry by keyboard and this was almost reduced by half when using speech control. However, although speech-controlled IVISs reduced distractions, it still resulted in worse SDLP and reaction time. Text presented in paragraph format degrades performance significantly more than information presented in graph or table format

(Blanco et al., 2006). An IVIS display would most likely require some text even with speech control and graphics were used. Thus, examining the effects of text characteristics on driving performance is still needed.

A number of text characteristics associated with IVIS has been shown to influence driving performance. In the transportation domain, changeable messages signs have been extensively examined with respect to reading text while driving (Dudek, 1992, 1997). Guidelines developed for traveler information systems also provide insights that are applicable to IVIS. A report by Campbell et al. (1998) suggests that in-vehicle displays should use clear and simple fonts (e.g. sans serif), with 7:1 preferred symbol contrast level, and no more than four information units displayed while a vehicle is in motion. In this context, information units refer to “the amount of information presented in terms of key nouns and adjectives contained within a message”. Labiale (1996) showed that the number of glances increased with the length of message, while the mean duration of each glance remained relatively stable. However, drivers’ performance in memory recall decreased proportionally to the text length displayed on in-vehicle systems, and their speed and vehicle control significantly declined for long messages (10 to 18 units of information). The Japan Automobile Manufacturers Association (JAMA) provide guidelines that suggest for dynamic information (e.g., traffic information from outside of the vehicle) presented in text, the number of letters displayed at a time should not exceed 31, so that drivers can read the text without rushing (JAMA, 2004). This guideline of 31 characters can correspond to 108 to 134 alphabet characters in the English language (Boyle et al., 2013).

In a study of Schieber, Holtz, B., and McCall (2008), drivers’ glance behavior for text reading tasks with varied lengths were compared. The results suggested that older drivers spend significantly longer time to complete the tasks as the text lengths increased. Older drivers also on average required an additional 2 more glances at the IVIS display when reading 6 lines of messages, compared to younger drivers. The length of the messages also had a significant influence on mean glance duration as it increased from 1.06 seconds for reading a one-line message to 1.25 seconds for a 6-line message. In addition, the SDLP was larger for longer message lengths.

In addition to the characteristics of the text elements, another important aspect for the design of in-vehicle displays is the impact of task-irrelevant texts around the targeted text that drivers are searching while driving. Although some irrelevant text is oftentimes necessary (text on dials, time display, etc.), having too much or poorly arranged irrelevant text may unnecessarily clutter the display. For example, Wolfe & Pashler (1988) showed that visual clutter could affect the speed and accuracy of searching for targeted text. Random locations of irrelevant text may also draw the drivers' attention, which can pose a greater distraction. However, few studies have examined the effects of task-irrelevant text on driving performance.

Influence of Driving Demands

Interacting with an IVIS while driving on an open rural road with no other vehicles or objects around may not cause any safety consequences. However, performing the same task while driving on a busy freeway could be safety critical. Lee et al. (2008) suggest that IVIS tasks lead to distraction and safety consequences when both the task and roadway demand are high and exceed the driver's capacity to respond to critical roadway events. Thus, to fully understand the influence of IVIS, it is important to examine drivers' interactions with such systems under various driving situations, especially the ones that pose high driving demand. Drivers may also realize this difference in driving demand and to some extent compensate the increased risk by adjusting their driving behavior while interacting with IVISs.

Lee et al. (2001) showed that drivers had longer reaction time to braking events while driving in a more complex driving environment, but there was no interaction between the driving environment complexity and IVIS tasks. A study by Jamson and Merat (2005) did find such an interaction between the task demand and driving and suggested that when the demand of IVIS was too high, drivers were likely to give up on the IVIS task and focus on the driving task, which resulted in a relatively less degradation on their reaction time. Schomig, Metz, and Kruger (2011) found that drivers lowered their driving speeds in highly demanding driving situations in general and further reduced the speed while performing IVIS tasks. However, they did not reduce their driving speeding while performing IVIS tasks in low demanding situations.

It is also essential to understand whether drivers realize the competing demands and modify their glance behavior to attend more toward the roadways instead of the IVIS. Metz et al. (2011) compared drivers' glance behavior while using IVIS on non-critical (e.g., rural straight road) and critical driving scenario (e.g., pedestrian crossing), and the results showed that drivers on average spent 4% more of their time on the road when driving situation is critical. Chiang, Brooks, and Weir (2004b) conducted an on-road study to examine the influence of destination entry tasks on glance behavior under different driving environment. The results suggested that drivers spent slightly larger proportion of time looking at the road on the freeway. In addition, drivers spent similar total amount of time to complete the entry tasks while driving on the freeway and city road, but the task completion time was more varied among drivers while on the freeway.

The Horrey and Wickens (2007) study compared the influence of IVIS complexity and driving demand (introduced by high or low wind turbulence) on the distribution of drivers' glance durations. Their study results showed that complex IVIS increased the proportion of long glances over 1.6 seconds, however, this proportion of long glances was not reduced for the driving condition with higher demand. This suggests that drivers do not adjust their glance behavior to have shorter glances when the driving demand is high, and this yields an even higher risk of having crashes when performing complex IVIS tasks.

However, other studies did find different results than Horrey and Wickens (2007) in terms of the effect of driving demand on glance durations. For example, when using road curvatures to represent the visual demand of driving, the mean glance duration was found to decrease from 1.8 seconds on straight road to 1.2 seconds on 194m radius curves for map reading tasks (Tsimhoni & Green, 2001). The mean time between glances also significantly increased from 1.1 seconds to 1.8 seconds when the curve became sharper. However, the total glance duration were found unchanged as the road curvature became sharper (Tsimhoni et al., 2004).

Influence of Long-Term Use

Drivers may also change their behavior as they continue to engage in IVIS tasks while driving. That is, their initial interaction with an in-vehicle device may not represent how they would actually engage with the device with continual use. This behavioral adaptation over time can be both positive and negative. Positive adaptation over time suggests that drivers improve their driving safety after long-term use of the system (e.g., improved turn signal use after using a lane departure assistant system). Negative adaptation over time, however, suggests that after prolonged use of the system, drivers become more risky with less safer driving behavior (e.g., shorter following distance, longer glances off the road, etc.).

A number of studies have examined the behavioral adaptation effects over time. Donmez, Boyle, and Lee (2008) compared drivers' driving performance and engagement in IVIS tasks under retrospective feedback, combined retrospective and concurrent feedback, and no feedback conditions over a sequence of four drives. The results showed that drivers had longer glance durations to the IVIS in the third and forth drives, compared to the first two drives, which suggested that drivers were more comfortable performing the IVIS tasks after more exposure to the system. However, they also had slower braking responses in the last two drives, which indicated a worse driving performance associated with the longer off-road glances. Drivers who were provided with feedback alerting them for being too distracted were found to have relatively shorter glances and faster braking responses during the last two drives, compared to those who were not exposed to any feedback.

The behavioral adaptation relates to the time effect can also be considered as a practice effect on the IVIS tasks over time. With practice, drivers may become more familiar with the system and tasks, and thus change their behavior or be able to improve their driving performance during dual-task situations over time. Shinar, Tractinsky, and Compton (2005) conducted a study to examine the practice effect of phone tasks on driving using five driving sessions. The authors found that drivers actually improved lane keeping and speed control while talking on a phone over time. This improvement is further affected by driving demand (i.e., speed limit), as the practice effect on speed control increased when driving with higher speed limit. A study of Chisholm, Caird, and Lockhart (2008) examined the whether practice of iPod tasks while driving could reduce the influence of distraction. Nineteen drivers were recruited to complete driving sessions in six consecutive weeks, in which they were asked to use an iPod while driving and encountered a number of critical events such as pedestrian entering roadway, vehicle pullout, and lead vehicle braking. An improvement on response time to hazard events was found after using the iPod over time. However, the driving performance with complex iPod tasks was still worse than in the driving-only condition without iPod tasks, which suggested that drivers were unable to improve dual-task performance to a safe level even with practice and long term use.

Rouzikhah, King, and Rakotonirainy (2013) examined the practice effects of several in-vehicle tasks, including eco-driving system, CD changing, and navigation, using two driving sessions over two consecutive days. The study showed that drivers' mental workload for navigation tasks were significantly reduced on day 2 compared to day 1, with an improved performance on driving tasks. However, the effect of practice was different on different tasks and it was more promising for complex tasks compared to simpler ones.

Dingus et al. (1997) compared drivers' glance behavior and driving performance while using a navigation system with and without experience with the system. The study found that drivers glanced at a navigation display less with shorter durations after six weeks of use. However, their lateral and longitudinal driving performance remained.

Strayer, Jason, and Drews (2011) discussed the interference of cellphone on driving and argued even though practice may reduce the impairments due to multitasking, the distraction effect caused by cellphone is unlikely to disappear over time. This is because driving involves reacting to sudden and unexpected events, which is unlikely to be automated with practice. A simulator study by the authors in fact showed support of this hypothesis. When drivers were asked to drive with cellphone tasks in the same scenario for four consecutive days, the number of collisions occurring in the pedestrian crossing event reduced from 41 on day one to 18 on day four. However, there were still twice as many as collisions under dual task condition compared to driving-only condition even on day four. In addition, when given a new driving scenario on day four, the number of collisions under dual task condition significantly increased again and the collision rate did not differ from that on day one.

In summary, given time, drivers appear to be able to improve their driving performance in some aspects under multitasking conditions such as when interacting with IVIS while driving. However, this does not mean that they become safer and less distracted. A mixed finding was found on their off-road glance durations, and their ability to detect and quickly react to new and sudden events may not be able to improve even with practice.

Individual Differences on Driver Distraction and Behavioral Adaptation

Individuals are largely different in terms of risk taking and multitasking abilities, and thus, the magnitudes of driver distraction that posed by IVIS tasks may also differ among drivers. The average performance, although to some degree represent the influence of IVIS tasks, may underestimate their impacts on risk takers or inexperienced drivers, who are more likely to have crashes than the average population.

Studies on individual differences are often time compared on sub-population level to examine the distraction effects among drivers with different age, gender, and driving expereinces. Younger drivers have been shown to be affected more by secondary tasks and have higher likeihood of invoving in crashes as they tend to direct their attention to the roadways less effectively, neglect hazards, and more willing to engage in distracting activities and take risks (Deery, 1999; Ferguson, 2003; Fisher et al., 2002; Williams, 2003). In addition, young drivers also have different scan patterns compared to older and more experienced drivers while driving. Wikman, Nieminen, and Summala (1998) found that young and inexperienced drivers tend to look away from the roadway with longer glances compared to others. In fact, the study showed that 29% of young drivers had glances longer three seconds. A study of Lansdown (2009) showed that novice drivers deviated out of lane more often compared to experienced drivers, and male drivers drove signifcantly faster with greater speed variation compared to females. In addition, novice drivers also had more off-road glances with longer durations compared to experienced drivers when interacting with an in-vehicle system.

Even within the same demographic population, drivers still show different glance behavior and driving performance due to different personality traits, driving habits, experiences, and willingness to take risks. Forbes (2009) found that drivers that were more skilled with computers had better lateral control ability when entering destinations while driving. Donmez, Boyle, and Lee (2010) conducted a study with 53 drivers between 18 and 21 years old and examined their glance behavior during in-vehicle tasks while driving. A cluster analysis with the number and duration of glances to the IVIS was conducted and three clusters of drivers with different risk levels were identified. The high risk group of drivers were shown to have longest glances off the road, moderately large number of off-road glances, and shortest minimum time to collision. About 72% of the low risk drivers were female drivers, whereas only 27% of the moderate risk drivers were female. However, the high risk cluster had approximately equal proportions of males and females.

Study has shown that heavy multitaskers tend to be more distracted and affected by irrelevent tasks, compared to people who don't usually multitask (Ophir, Nass, & Wagner, 2009). However, Strayer et al. (2011) has shown that although secondary tasks impact the driving performance for majority of drivers, a small percentage of drivers are actually "supertaskers" whose driving performance is not affected. The study further examined the frontal cortext of the "supertaskers" and compared with other drivers using fMRI and found differences that can explain their ability to multitask.

Given drivers' different responses to distracting tasks, it is possible that they also exhibit different adaptive behavior under changing situations, such as with different task and driving demands, being exposed to new techonologies, and after long-term use of IVIS. For example, Donmez et al. (2010) found that high risk drivers were more likely to adapt and improve their driving behavior while interacting with in-vehicle device after using a real-time feedback system that warned of unsafe behavior. Similarly, McGehee, Raby, Carney, Lee, and Reyes (2007) conducted a field study to examine the effect of concurrent and cumulative feedback on teenage drivers. The study results revealed one group of drivers had significantly high frequency of safety-critical events than another group and the feedback significantly reduced the event rate of the high risk group and the reduction rate was as high as 72% after the first nine weeks of feedback intervention. No significant improvements were found for the less risky group of drivers. However, few studies discussed the behavioral adaptation under different IVIS task and driving demands for drivers with different risk levels.

Objectives and Hypotheses

The main objective of this dissertation is to understand drivers' behavioral adaptation on IVIS tasks mediated by varied task demands, driving demands, long-term use, and driver risk levels. The overall hypothesis is that, although drivers tend to compensate for increased risk, the actual risk they are exposed to may still increase with IVIS task demands as it is out of their capacities to maintain the same risk level. However, there is a limit with respect to the maximum risk they are willing to take even when the task demands continue to increase (Aim 1). Further, the amount of risk that drivers are willing to take while performing IVIS tasks is affected by the driving demands, driver risk seeking levels, and their familiarity with the system (Aim 2). Off-road glance duration and frequency to in-vehicle tasks have been shown to be highly correlated with crash likelihood (Wierwille & Tijerina, 1998). Thus, the risk associated with IVIS tasks was assessed by drivers' glance behavior. Aim 1, which studies the behavioral adaptation with IVIS task demands, is examined using Experiment 1 and discussed in Chapter 3. Aim 2, which studies other factors for behavioral adaptation on IVIS tasks are examined using Experiment 2, and discussed in Chapter 4 (Figure 1).

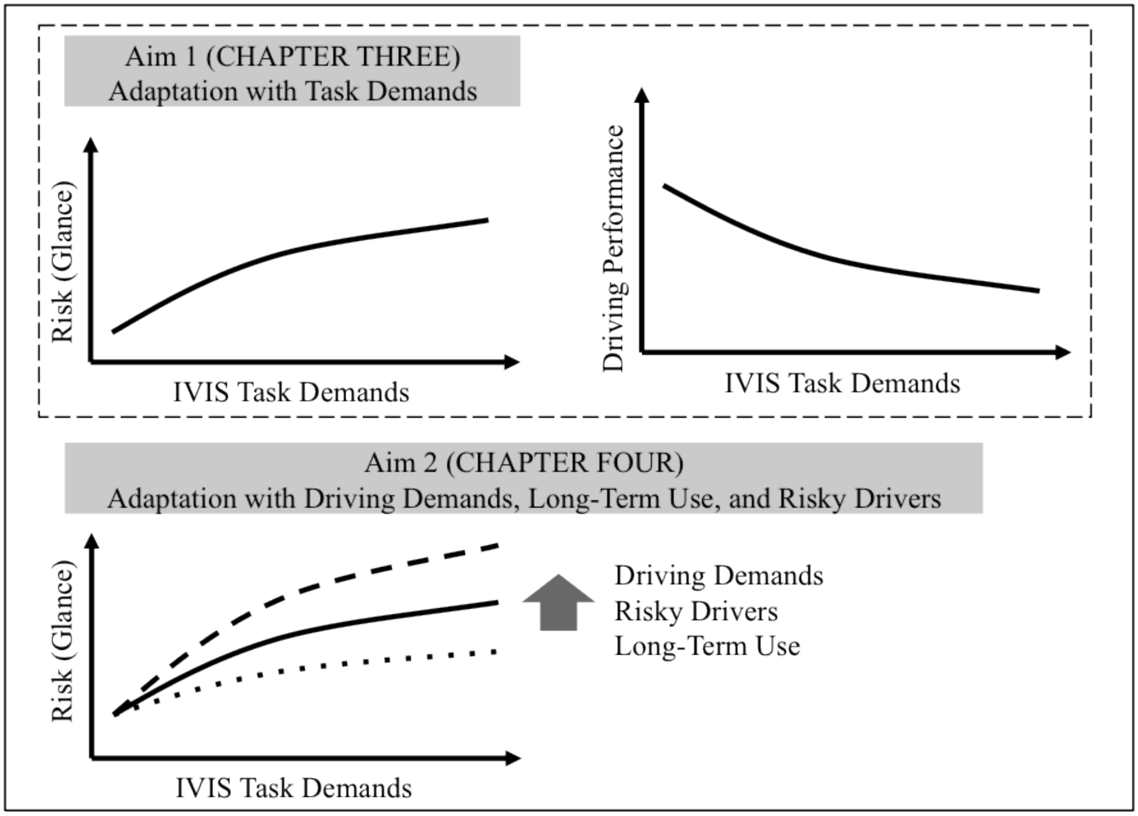

Figure 1. Expected nonlinear relationship between task demands, eye glances, and driving performance as measured in Aims 1 and 2 (Chapter 3 & 4)

Some previous studies have concluded that drivers' interactions with IVIS improved with practice based on the improved vehicle control performance, while others argue that the distraction effects of IVIS tasks cannot be practiced away as they did not see fewer collisions associated with IVIS tasks even with practice. This seemingly contrary result may explain the changing relationship between driving performance and glance behavior over time. Thus, Aim 3 of this dissertation is to test this association. It is hypothesized that with practice, drivers adapt to the IVIS tasks with improved driving performance (i.e., vehicle control ability), and this improved driving performance may give drivers confidence on their ability to multitask and thus, resulting in less safe glance behavior which may increase their actual risk on the road (Figure 4). Aim 3 is examined in Chapter 5, using the data from Experiment 2.

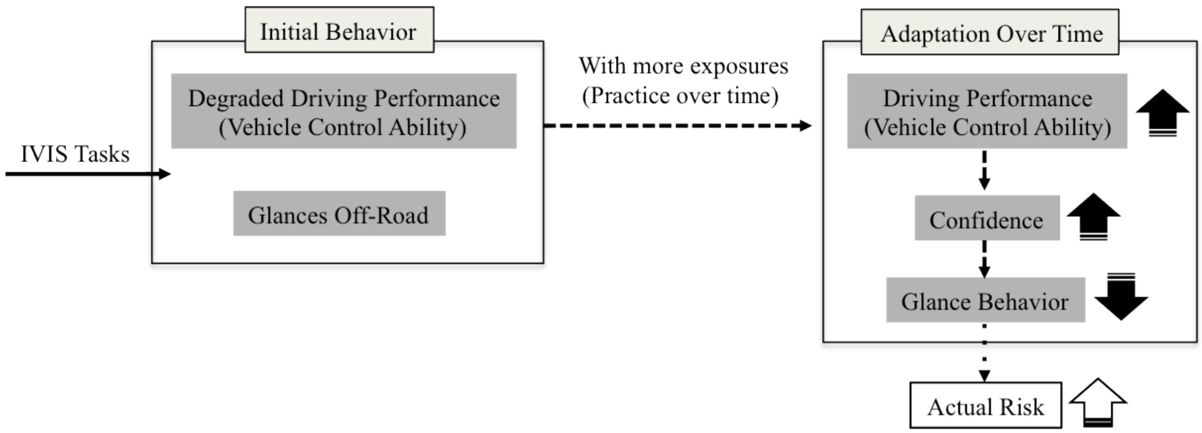

Figure 2. Hypothesized relationship between driving performance and glance behavior associated with behavioral adaptation over time (Aim 3, Chapter 5)

Chapter Summary

This chapter discusses previous findings on IVIS and its associated driver distractions, as well as drivers' behavioral adaptation relate to in-vehicle technologies. Texting on a mobile device is clearly non-driving related and can negatively affect driver safety. However, some IVISs are specifically designed for use while driving (e.g., route and traffic information, lane assist) and minimizing the distracting effects in these systems is critical. IVIS will continue to evolve and drivers will also adapt their behavior to "better" multitask while driving in order to use the system efficiently. Thus, it is important to recognize this behavioral adaptation in order to understand potential safety consequences relate to IVIS use under various situations, which is the goal of this dissertation. The overall hypotheses and study aims are introduced at the end of this chapter. These hypotheses are elaborated in more details in Chapter 3, 4, and 5 respectively. The next chapter discusses the essential statistical model that has been used to examine the hypotheses.

CHAPTER 2: RANDOM COEFFICIENT MODELING

This chapter discusses the main statistical modeling method adopted in this dissertation. In order to test the driver behavioral adaptations over time, a longitudinal study design, where each participant was asked to drive in several sessions over days, was used for this dissertation. Studies that have repeated-measures or observations over time could involve large amount of data, with complex correlation structures. Thus, more advanced regression techniques that can take the within-subject correlations into account are needed. Random coefficient regression modeling is one method that has been commonly used for this purpose. This chapter introduces the fundamentals of the random coefficient modeling techniques, and the specific models used for this study.

Data Analysis Challenges and Techniques for Longitudinal Data

When studies are longitudinal in nature where outcomes are measured repeatedly for each subject, the observations from the same subject are likely to be correlated. Thus, the assumption for ordinal least squares (OLS) regressions is violated and OLS cannot be used for analyzing longitudinal data. Special statistical methods that can take into account the correlations between observations within subjects need to be adopted.

A number of methods have been developed to analyze repeated measures data, including paired t-test and repeated measures ANOVA, which have been widely used for analyzing correlated data with continuous outcomes. However, longitudinal studies often times have complex study designs with multiple independent variables, in which case requires more sophisticated methods than paired t-test. In addition, the outcome variables of interest may not be continuous, and the researchers may also be interested in obtaining the effect magnitudes of the independent variables, and performing predictions for the outcome variables. Thus, repeated-measures ANOVAs cannot meet the data analysis requirements in these cases either and sophisticated regression techniques are warranted.

Two commonly adopted regression methods for longitudinal methods are generalize estimating equations (GEE) and random coefficient modeling. Both methods can be used to account for the within-subject correlations, and analyze continuous or discrete outcome variables. GEE model uses a working variance-covariance matrix to estimate the within-subject correlations, and answer population-level modeling questions (Wakefield, 2013). That is, GEE estimate the marginal model $E[Y_i] = x_i\beta$, with $W_i = var(Y_i)$ being the working variance-covariance matrix, which assumes that $cov(Y_i, Y_{i'}) = 0$ for $i \neq i'$ (i.e., observations from different subjects are independent). GEE model is also considered as the "population-average" model and is most appropriate when the average performance of the population is of interest (Twisk, 2004).

Random coefficient model, on the other hand, is considered as "subject specific" (Twisk, 2004). Different from GEE, which estimate the average performance, random coefficient model allows each subject to have his/her own regression line (i.e., different coefficient estimates for different individuals), and assumes that the individual coefficient estimates follow certain distributions. A detailed description of the random coefficient modeling technique is discussed in the following section.

Random Coefficient Modeling

Random coefficient modeling was first introduced by Ware and Laird (1982) in a biometrics context. The random coefficient model has also been referred as linear mixed-effects model, or multilevel models in different research domains (Verbeke, Molenberghs, & Rizopoulos, 2010). However, the fundamental ideas are the same: the model consists both fixed effects and random effects, where the fixed effects are same across individuals and the random effects can vary for different individuals.

Model Equation and Example

Let $\boldsymbol{Y_i} = [Y_{i1}, \dots, Y_{in_i}]^T, i = 1, \dots m$ be the outcomes for *i*-th individual. Let $\boldsymbol{\beta}$ be a $(k+1) \times 1$ vector of fixed effects, and $\boldsymbol{b_i}$ be a $(q+1) \times 1$ vector of random effects. Let $\boldsymbol{x_i} = [\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{in_i}]$ be the design matrix for the fixed effects, and $\boldsymbol{z_i} = [\boldsymbol{z}_{i1}, \dots, \boldsymbol{z}_{in_i}]$ be the design matrix for the random effects. Then the random coefficient model can be expressed as (Wakefield, 2013; Ware & Laird, 1982):

$$\boldsymbol{y_i} = \boldsymbol{x_i\beta} + \boldsymbol{z_i b_i} + \boldsymbol{\epsilon_i},$$

where $\boldsymbol{\epsilon_i}$ is a $n_i \times 1$ vector of error terms for $i = 1, \dots m$, and $\boldsymbol{\epsilon_i} \sim N(\boldsymbol{0}, \boldsymbol{R_i})$ with $\boldsymbol{R_i}$ being the covariance matrix for the error terms. For simplicity, we can assume that $\boldsymbol{R_i} = \sigma^2 \boldsymbol{I_{ni}}$, which implies that the responses of individual *i* are independent given $\boldsymbol{b_i}$ and $\boldsymbol{\beta}$ (i.e., conditional-independent model). Additionally, $\boldsymbol{b_i} \sim N(\boldsymbol{0}, \boldsymbol{D})$ where $\boldsymbol{D}$ is the covariance matrix for the random effects, and $cov(\boldsymbol{b_i}, \boldsymbol{b_{i'}}) = 0$ for $i \neq i'$. The marginal model for $\boldsymbol{Y_i}$ can be written as $E[\boldsymbol{Y_i}] = \boldsymbol{x_i\beta}, var(\boldsymbol{Y_i}) = \boldsymbol{R_i} + \boldsymbol{z_i D z_i^T}, cov(\boldsymbol{Y_i}, \boldsymbol{Y_{i'}}) = 0$ if $i \neq i'$, for $i = 1, \dots m$.

For a longitudinal study, the random effects usually contain a random intercept and random slope (for the time variable). In other words, the baseline value is varied for each individual, so does the development of certain variables over time (Twisk, 2004). Consider a simple example where the outcome *y* was measured *T* times for each subject *i*. The model estimates can be shown in Figure 3, where the bold line represents the population intercept and slope on *y* over time *t* estimated by the fixed effects, and the thin lines represents the subject-specific intercept and slope obtained through the random effects.

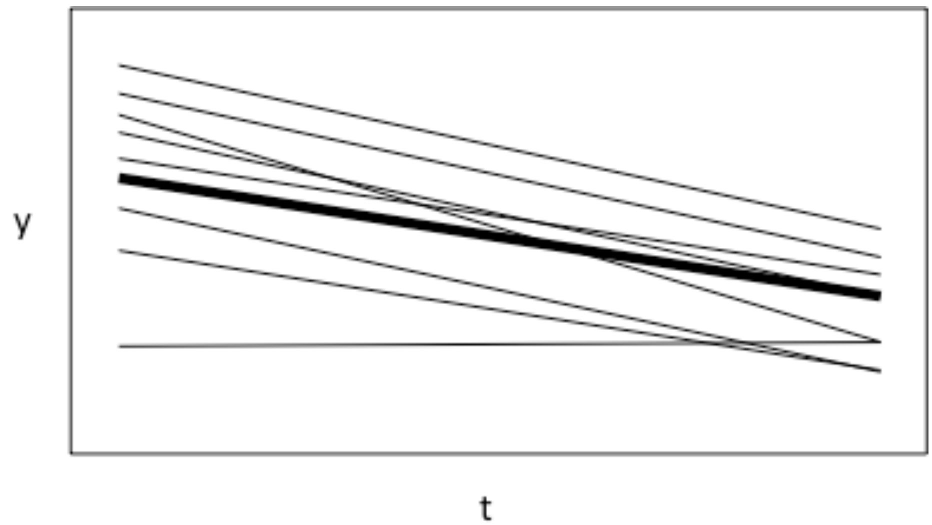

Figure 3. Hypothetical example of longitudinal continuous data and model estimates of random coefficient model

Such model in the above simple hypothetical example can be written as:

$$y_{it} = \beta_0 + b_{0i} + (\beta_1 + b_{1i})t + \epsilon_{it},$$

where y_{it} is the outcome for subject *i* at *time t, i=1,2,…m and t=1,2,..T.* The equation can also be written in the matrix format as:

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \\ \vdots \\ y_{mT} \end{bmatrix} = [1 \quad t] \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} b_{01} & b_{11} \\ b_{02} & b_{12} \\ \vdots & \vdots \\ b_{0m} & b_{1m} \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{iT} \\ \vdots \\ \epsilon_{mT} \end{bmatrix},$$

where $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma_\epsilon^2 \boldsymbol{I}), \boldsymbol{b}_{\mathbf{0,1}} \sim N(\mathbf{0}, \boldsymbol{D}), \boldsymbol{D} = \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix}$. That is, the intercept has standard deviation σ_0, the slope of time has standard deviation of σ_1. In addition, the intercept and slope are correlated with covariance σ_{01}.

Likelihood Inferences for Random Coefficient Models

Inferences of the fixed effects can be made using maximum likelihood estimates. A detailed explanation of the inferences methods is described in Wakefield (2013) and briefly summarized here. Consider the example showed in Figure 3, where there are both random intercept and slope for time t. The likelihood can be written in the following form:

$$p(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{D},\sigma_\epsilon^2) = \int_b p(\boldsymbol{y}|\boldsymbol{b},\boldsymbol{\beta},\boldsymbol{D},\sigma_\epsilon^2)\times p(\boldsymbol{b}|\boldsymbol{\beta},\boldsymbol{D},\sigma_\epsilon^2)d\boldsymbol{b} =$$

$$\prod \int_{bi} p(\boldsymbol{y_i}|\boldsymbol{b}_i,\boldsymbol{\beta},\sigma_\epsilon^2)\times p(\boldsymbol{b_i}|\boldsymbol{D})d\boldsymbol{b_i}.$$

Given a convolution of normal is still normal, there is

$$\boldsymbol{y_i}|\boldsymbol{\beta},\boldsymbol{D},\sigma_\epsilon^2 \sim N(\boldsymbol{\mu}_i,\boldsymbol{V}_i),$$

where $\boldsymbol{\mu_i} = \boldsymbol{x_i\beta}, \boldsymbol{V_i} = \sigma_\epsilon^2\boldsymbol{I} + \boldsymbol{z_i D z_i^T}$.

Thus, the log-likelihood can be written as

$$l(\boldsymbol{\beta},\boldsymbol{D},\sigma_\epsilon^2) = -\frac{1}{2}\sum \log|\boldsymbol{V}_i| - \frac{1}{2}\sum(\boldsymbol{y_i} - \boldsymbol{x_i\beta})^T \boldsymbol{V_i}^{-1}(\boldsymbol{y_i} - \boldsymbol{x_i\beta})$$

The MLE for $\boldsymbol{\beta}$ can then be obtained via maximization of the log-likelihood by solving the score equations for $\boldsymbol{\beta}$

$$\frac{\partial \boldsymbol{l}}{\partial \boldsymbol{\beta}} = \sum \boldsymbol{x}_i^T \boldsymbol{V_i}^{-1}(\boldsymbol{y}_i - \boldsymbol{x}_i\boldsymbol{\beta})$$

The solved MLE estimate for $\boldsymbol{\beta}$ yields $\widehat{\boldsymbol{\beta}} = (\sum \boldsymbol{x_i^T V_i}^{-1}\boldsymbol{x_i})^{-\mathbf{1}} \sum(\boldsymbol{x_i^T V_i}^{-1}\boldsymbol{y_i})$. Note that the $\widehat{\boldsymbol{\beta}}$ becomes the ordinary least squares estimator when $\boldsymbol{D} = 0$.

Given that there is no closed-form solution of the MLE estimate of the variance components $\boldsymbol{D}$ and σ_ϵ^2, the expectation-maximization or Newton-Raphson algorithm is usually applied to obtain $\widehat{\boldsymbol{D}}$ and $\widehat{\sigma_\epsilon^2}$. The random effect $\widehat{\boldsymbol{b_i}}$ can be estimated as $\widehat{\boldsymbol{b_i}} = \boldsymbol{D z_i^T V_i}^{-1}(\boldsymbol{y_i} - \boldsymbol{x_i\beta})$. Normal approximation confidence intervals can also be calculated for both the fixed effects and the variance-covariance parameter.

Comparing Random Coefficient Models and GEE

Both GEE and random coefficient model can obtain valid results for longitudinal data, and specific variance-covariance structures can be used for both methods. There is no clear answer to the question on which of the two methods is better (Twisk, 2004). However, the two methods are different in terms of fundamental concepts and purpose. Given that GEE estimates the "population averaged" effects, and random coefficient model estimates the "subject specific" effects, it is generally agreed that it is more appropriate to use random coefficient models when one is interested in obtaining the individual-level inference for a subject either observed or unobserved in the sample, or the within- and between- subject variability (Twisk, 2004; Wakefield, 2013). In other words, the random intercept and slope from random coefficient model allows one to predict the responses for a new subject who is not in the study sample but from the same subject population, given the coefficient estimates on a "similar" subject. It can also provide information on the variability of the responses among subjects over time. Therefore, random coefficient model is more useful than GEE model when predicting individual level performance is of more interest, especially if the individual difference is large.

For this dissertation, it is expected that there is a somewhat large difference among individual drivers in behavioral adaptations and glance behavior. Thus, a subject specific model can provide more information than a population-averaged model, and therefore, the random coefficient modeling technique is adopted.

Chapter Summary

This chapter introduced the data analysis problem for longitudinal data in which observations are correlated within-subjects, and the commonly adopted modeling techniques. Two regression methods, GEE and random coefficient models, are further discussed in details.

The reason to choose random coefficient model over GEE for this dissertation was also discussed at the end of the chapter. The specific models used in the study are presented in the following chapters.

CHAPTER 3: EXPERIMENT ONE: BEHAVORAL ADAPTATION AND IVIS TASK DEMANDS

This chapter presents Experiment 1, which examined the relationship between IVIS task demands and behavioral adaptation. The data analyses in this experiment were based on a driving simulator study with 28 drivers, where IVIS task demands were manipulated by task type, text length, and task irrelevant text.

Objective and Hypotheses

The main objective of Experiment 1 is to understand how drivers behave and perform differently with varying task demands. There are two hypotheses associated with this experiment.

Hypothesis 1.1: Drivers increase the duration of off-road glances when the IVIS tasks have higher demands, even though they may compensate for the increasing risk. Specifically, secondary tasks may increase individual eyes-off-road (EOR) duration, proportion of EOR time, and total EOR time during a task. However, this increase may not be linear due to risk compensation.

Hypothesis 1.2: IVIS tasks with higher demands will negatively impact vehicle control performance. Specifically, IVIS tasks may increase the number of lane departure events, and SDLP. Higher demands of IVIS tasks may also be associated with longer and more varied following distances.

Methodology

Participants

Twenty-eight participants (15 males and 13 females) with valid drivers' license were recruited from the Seattle, Washington area. They were from four age groups (18-24, 25-39, 40-54, and 55-75 years old) with equal number of participants, including three to four females per group. The selection of sample size, age group, and screening protocol were based on the recommendations described in the US DOT-NHTSA Visual-Manual Distraction Guidelines (NHTSA, 2012). All participants were native English speakers, reported being in good general health, and driving at least 7000 miles per year when they were recruited. Participants were also asked if they felt comfortable using computers, touchscreens, and communicating via text messages to ensure that they had some interactions with computing technology and did not feel uneasy about it. Participants were compensated $20 per hour for the study plus parking validation.

All drivers had at least some college education. Two drivers in the youngest group reported driving less than once a week, and the remainder reported driving at least once a week. On average, participants obtained their driver's license while still a teenager (Table 1). Seven drivers had one or more moving violations in the past three years, and two drivers had not-at-fault crashes in the past three years.

Table 1. Driver demographics and driving history in Experiment 1

Age Group	Age			Age received permit	Age received license	Moving violations in past 3 years		Crashes in past 3 years	
	Mean	sd	Range			0	1+	0	1+
Age 18-24	21.3	0.5	21-22	15.1	16.5	5	2	6	0
Age 25-39	32.4	5.4	26-39	17.3	18.3	4	2	6	0
Age 40-54	46.9	4.6	40-54	15.7	16.7	5	1	5	1
Age 55-75	59.7	3.0	56-64	15.3	16.0	5	2	6	1

Experimental Design

The study used within-subject design with 12 IVIS task conditions and 1 driving-only condition. The 12 IVIS task conditions were formed through manipulation of task type (2 levels), text length (3 levels), and irrelevant text (2 levels) (Figure 4). The 12 task conditions were compared to a driving-only condition with no IVIS tasks for vehicle control performance measurements and with each other for eye glance measurements.

The two levels of task type were text entry and text reading. The text entry tasks involved participants entering a word using a touchscreen keyboard and the text reading tasks included reading non-scrolling phrases on the same touchscreen. All text entry words and text reading phrases were unique and randomly selected without replacement.

The three levels of task length were categorized as short, medium, and long. For the variable text entry, the short, medium, and long were associated with 4, 6, and 12 characters, respectively. The texts used for text entry were based on street names found in road database files. The three text reading phrases had lengths of 20 to 40, 60 to 80, and 120 to 140 characters, for the short, medium, and long conditions respectively. Text phrases were designed to include character strings of traffic and routing information as typically observed on changeable message signs.

There were two levels for the variable irrelevant text factor, text present or text absent (Figure 5). The row location of the target text was purposely randomized across the task relevant and task irrelevant condition to remove any potential confounding effects. The task irrelevant text in this experiment was defined as the text that appeared outside the box. The information presented in the irrelevant text was still driving related, but different than the target text within the box, and thus was used to add visual clutter to the display. The font used for the tasks was Arial narrow, with 1/8 inch height, displayed at a distance of about 2.3 feet away from driver's eye point, resulting in a visual angle of 0.26 degrees.

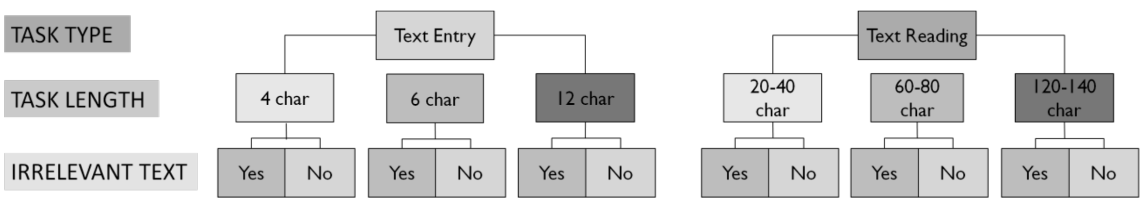

Figure 4. In-vehicle task conditions

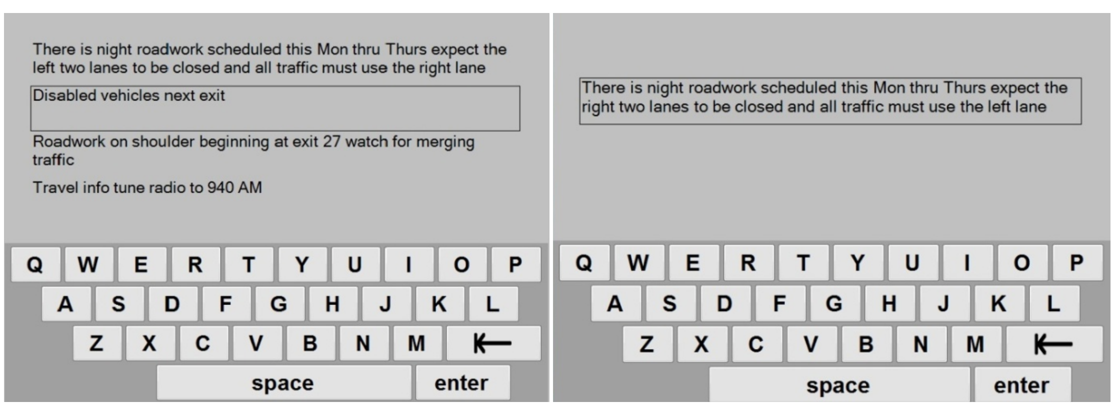

Figure 5. Example of the top part of the touch screens depicting with (left) and without (right) irrelevant text

Apparatus

A low-cost, low-fidelity driving simulator (NADS MiniSim) and an eye tracker (Seeing Machines faceLAB) was used for the experiment. The NADS MiniSim included three screens (3.0' [wide] by 1.7' [tall] each) that were placed about 4.5' away from driver's eye point. A 7" touchscreen display with QWERTY keyboard included only keys needed for the tasks. The touchscreen display was attached approximately 10 inches to the right side of the simulator's steering wheel, with 16 to 20 degrees of visual angle at distance 2.1' to 2.7'. All data were collected at 60Hz. There was no predictive typing feature built into the keyboard. The simulated road was a four-lane undivided rural road with a speed limit of 55 mph and no horizontal or vertical curve. All tasks were performed while driving in the left lane with no lane changes. Drivers were asked to drive safely, and maintain speeds at 50 mph (≈ 80.5km/h) and a headway distance of 2-second as best they could while following a lead vehicle travelling between 47.5 (≈ 76.4 km/h) and 52.5 mph (≈ 85.5km/h). The speed variation was based on the sum of three sinusoids with amplitudes of 0.3, 0.5, and 0.2 mph, frequency of 0.2, 0.067, and 0.033 Hz, and phase offset of 1, 0, and 14 sec. No lead vehicle braking events were included in the scenario.

Procedures

Participants received training on the tasks and the simulator after being provided an informed consent. For the text entry task, the computer provided an auditory cue consisting of a word that the participant would then enter using the keyboard on the touchscreen. Participants could revise the sequence of keystrokes by pressing the backspace key and re-entering the character. Once they were satisfied with the sequence of keystrokes, they needed to press the "enter" key to proceed to the next task. For the reading task, the computer provided a beep that directed the participant to read a phrase displayed within a box on the touchscreen. The participant then pressed the “enter” key to indicate they have read and comprehended the phrase. The computer then verbally stated a sentence related to the phrase and the participants had to select either true or false to indicate whether the statement heard matched the phrase just read (Table 2). The T/F statements were of similar length, regardless of the length of the reading phrases. For both entry and reading tasks, participants were instructed to provide their best guess if they did not know the answer. The task procedures were summarized in Figure 6.

Table 2. Example of the True/False statements for short, medium, and long reading phrases

Reading Phrases	T/F Statements
Traction tires or chains recommended	No need for traction tires or chains
Caution icy roads ahead traction tires or chains recommended	Icy roads ahead are expected
Use caution icy road conditions ahead traction tires or chains required for vehicles towing or over the weight of 10000 lbs	Icy road conditions ahead

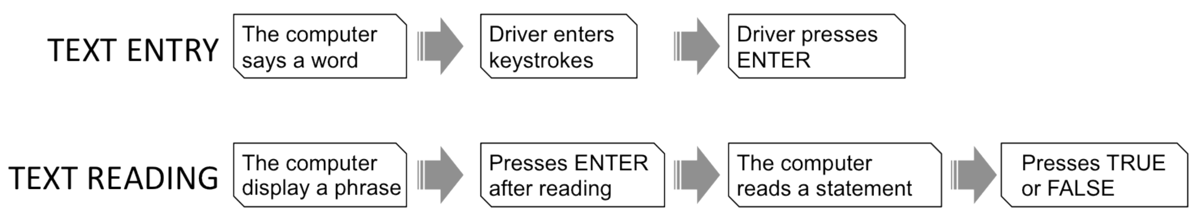

Figure 6. Text entry and text reading task procedures

A practice drive was completed before the main study to familiarize participants with driving in the simulator and experimental tasks. This main study included two simulator trials and two static trials. In each simulator trial, participants performed 18 tasks (9 readings and 9 entries) in a random order (i.e., each participant encountered each text reading or entry condition three times, for a total of 36 tasks). Each task was triggered 15 seconds after the participant completed the previous one (i.e., either pressing “enter” for text entry or “true/false” for text reading). Each trial lasted approximately 15 minutes including a 5-minute baseline drive toward the end of each trial, when all IVIS tasks were completed. Participants completed another set of 36 tasks in two static trials (18 tasks in each) without driving. The static trials lasted about 8 to 10 minutes in duration. A 5-minute break separated each trial. To minimize order effects, 50% of participants performed simulator trials first, and the other 50% performed the static trials first.

Dependent Variables

IVIS task performance. Descriptive statistics for IVIS task performance were analyzed to show the engagement and demands of the tasks. Task performance was measured by keypunch correction rate, task error rate, and task completion time. The keypunch correction rate was examined for text entry tasks only. For each task condition, it was defined as the total number of tasks in which the "backspace" key was pressed once or more, divided by the total number of tasks performed by all drivers combined (i.e. $n=28\times3=84$).

The task error rate was defined as the total number of tasks in which the final answers given by the drivers was incorrect, divided by the total number of tasks performed by all drivers combined. Based on experimenters' observations and verbally expressed by the participants, a number of entry errors were made in the text entry tasks due to participants' confusion on similar sounding words (e.g., Dale vs. Gale). It is likely that there were a substantial number of errors made due to these hearing errors. However, given that it was difficult to discern the differences between hearing and spelling errors, the error rates for text entry tasks were not presented in this dissertation to avoid misleading results.

The task completion time was also examined and calculated from the moment the task started to the moment the participant pressed the "enter" key. That is, the time spend on the comprehension question for reading tasks was excluded when calculating task completion time.

Eye glance behavior. Several eye glance measurements were examined and used for the hypothesis 1.1 of Experiment 1. These included EOR time, total EOR time, and proportion of EOR time. The maximum EOR time was defined as the longest duration of all individual glances in a task. The total EOR time was defined as the total duration of individual glances during a task. The proportion of EOR time was defined as the ratio of total EOR time to the task completion time.

Vehicle control performance. Vehicle control performance was measured in terms of lateral and longitudinal control. SDLP measured drivers' lateral deviation from the centerline. The number of lane departures indicated how frequently the outer side of the driver's car crossed one side of the lane marks. Time headway was defined as the time distance from the front bumper of the driver's vehicle to the rear bumper of the lead vehicle. The mean and standard deviation of the time headway were used to measure drivers' longitudinal control performance.

For the 12 IVIS task conditions, the dependent variables were calculated over the duration of each task (i.e., task completion time). The mean duration for each of the 12 task conditions was computed for each driver, and 12 corresponding driving-only segments with the same durations were extracted from the driving-only period at the end of each simulator trial (6 random segments from each trial). For the 6 driving-only segments extracted from each simulator trial, each segment was 15 seconds apart (same duration as the time between each IVIS task). The dependent variables were then calculated over the duration of each driving-only segment.

Data Analysis

The 28 participants completed three replications of 12 different task conditions for a total of 84 data points per condition. There were some constraints in the first road segments such that the first task occurred too early in the simulator drives. The lead vehicle, as well the majority of participants (n=19), did not reach the desired driving speed (i.e., 50mph) in the very first task of each simulator trial. Additionally, the following distance during this first task was shorter than the average following distances during other tasks. Therefore, this first task was removed (or two tasks per participant) in order to eliminate any possible confounding due to low speed and short following distances. Two additional tasks were removed for one participant because a lane change occurred during the task even though the participant was instructed not to change lanes during training. After removing these observations, at least one instance of each task condition was still available for each participant, and at least 19 participants had three instances of each condition.

Random coefficient models (using R2.12.1, package 'nlme') were used to examine the effects of text entry and reading on drivers' eye glance behavior, and their ability to maintain lateral and longitudinal control while driving (i.e., hypotheses 1.1 and 1.2). Dependent variables were log transformed to meet the linear regression assumptions when needed. Forward selection method was used for selecting main and interaction effects and the final models were chosen based on the lowest AIC values. Statistical significance was assessed at α=0.05.

The driving-only condition revealed that the duration of driving affected the dependent variables (e.g., SDLP increased as the time spent driving increased). Therefore, duration of driving was included in the vehicle control performance models as a covariate. Task completion time (or duration of driving) was log transformed for two of the models (SDLP and SD of time headway) to meet the residual assumptions regarding heterogeneity. Main effect of driver age groups and gender were also tested in the models.

In the eye glance behavior models, the main and interaction effects of task type, text length, and task irrelevant texts were tests. In the vehicle control performance models, each task condition was coded as a binary variable to compare with the driving-only condition (baseline), and with each other using pairwise comparisons. There was no significant difference observed for the irrelevant text condition, and thus the two levels of irrelevant text were combined for each task type and length in the final models for a smaller AIC and model simplicity (i.e., 6 task conditions, long entry, medium entry, short entry, long reading, medium reading, and short reading, were each compared with driving-only).

Results

Descriptive Statistics on In-Vehicle Task Performance

Descriptive statistics were summarized for the task error rates (for text reading tasks only), keypunch correction rates (for text entry tasks only), and task completion times. These data were compared between the static (in-vehicle task only) and simulator (in-vehicle task with driving) trials.

In the simulator trials, drivers did not necessarily make more errors when reading longer phrases. The error rates for medium and long text reading tasks were actually lower than that for short text reading tasks in the static trials. Error rates for reading tasks were generally higher in simulator trials than in static trials (Figure 7). However, drivers did comprehend most of the reading phrases (at least 84%), which suggests that they were engaged in the IVIS tasks.

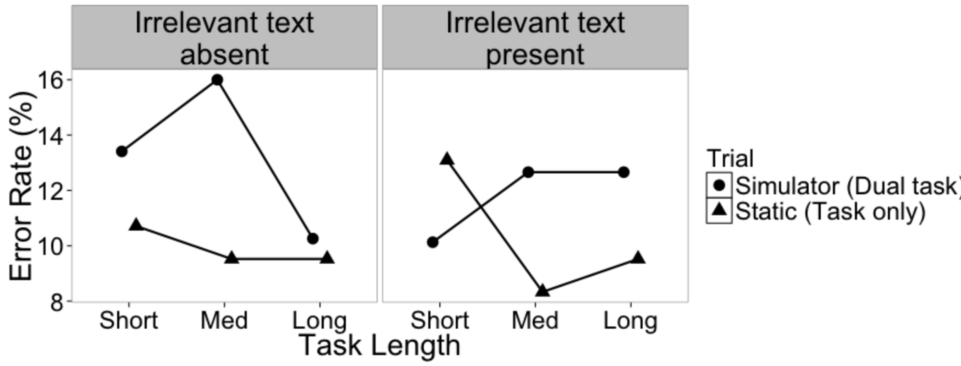

Figure 7. Text reading task error rates

The keypunch correction rate for entry tasks appears to increase linearly as the text length increases. There was no apparent difference in correction rate for short text entry tasks between simulator and static trials (Simulator: 13.3%, Static: 13.1%). However, the difference in correction rates was larger for longer task lengths: 30.4% (static) and 40.4% (simulator) (Figure 8). This suggests that drivers made more keypunch errors while driving and entering texts at the same time compared to entering texts alone without driving, especially when there were more characters to enter.

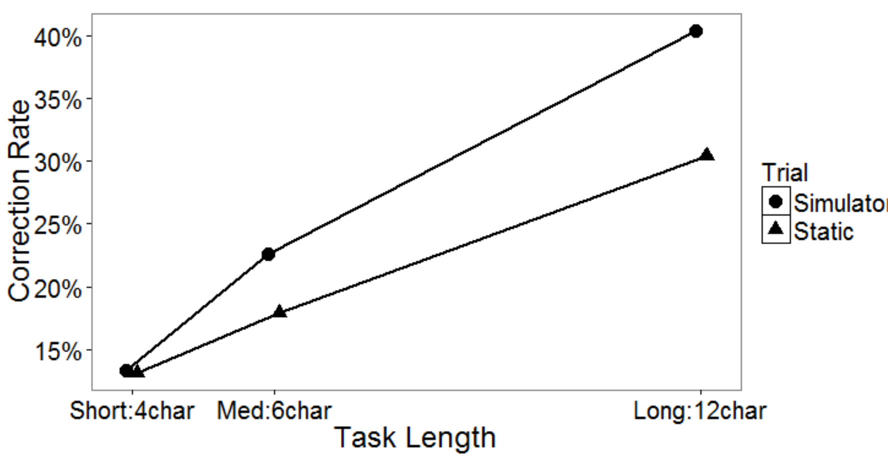

Figure 8. Correction rates for text entry tasks

As expected, text entry and text reading with longer words/phrases took longer for participants to complete while driving, which suggest an increasing task demand for drivers. The task completion time increased linearly with the number of characters to enter and read (Figure 9). Entering text with words as short as 4 characters took approximately the same time as reading long phrases with ~130 characters, which suggests that in the current study, text entry tasks were in general more demanding than text reading tasks. Entering a 12-letter word on average took drivers 22.5 seconds to complete, with large between-subject standard deviations (irrelevant text present: SD=7.1s, irrelevant text absent: SD=7.8s). The task completion time for text entry was very similar with and without the task irrelevant text. However, the display of irrelevant text increased the task completion time of text reading tasks by 2.5s, 1.5s, and 1.7s for short, medium, and long length conditions, respectively.

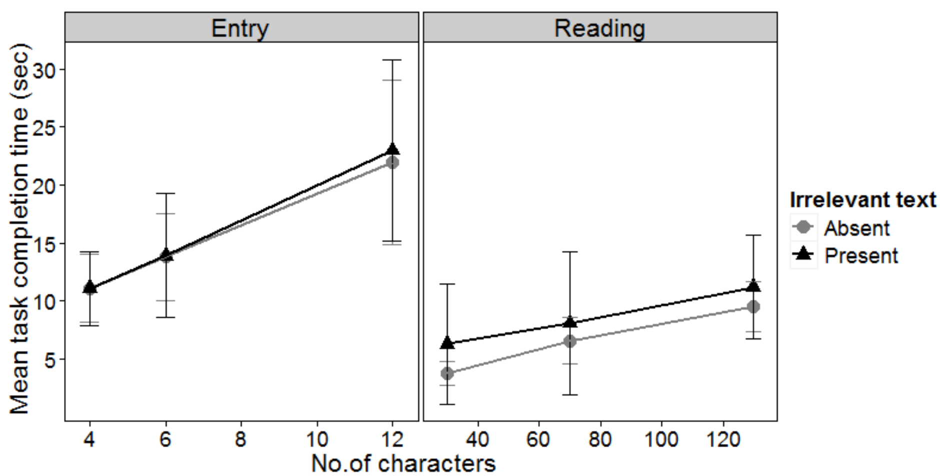

Figure 9. Mean task completion time

Hypothesis 1.1: Glance behavior

The hypothesis 1 states that drivers increase the duration of their off-road glances when the IVIS has higher demands, even though they may compensate for the increasing risk. The descriptive statistics of the maximum EOR time, total EOR time, and proportion of EOR time are summarized in this section, followed by statistical models that prove the hypothesis.

Descriptive Statistics

Table 3 shows the mean and between-subject standard deviation for maximum, total, and proportion of glance. The maximum glance (i.e., max EOR time) ranged from 2.5 seconds to 3.3 seconds for text entry tasks, and ranged from 1.4 seconds to 2.2 seconds for text reading tasks. There was also a larger variation on the maximum EOR time among drivers for text entry tasks than for text reading tasks. Additionally, the maximum EOR time increased as the text lengths increased, but were similar for both irrelevant text present and absent conditions.

The total EOR time ranged from 5.7 seconds to 15 seconds for text entry but only ranged from 2.3 seconds to 7.5 seconds for text reading tasks. The long text reading task had total EOR time ranged between the short and medium text reading tasks. The total EOR time was half a second to one second longer when task irrelevant text present compared to when irrelevant text absent.

The proportion of EOR time ranged from 52.5% to 66% for text entry tasks, and 58.3% to 68.9% for text reading tasks. It increased as the text lengths increased, and the presence of task irrelevant text appeared to increase the proportion of EOR time for some tasks as well.

Table 3. Descriptive statistics for max, sum, and proportion of EOR time (Experiment 1)

Task Type	Task Length	Irrelevent Text	Max EOR Time (s)		Sum EOR Time (s)		Proportion of EOR Time (%)	
			mean	between subj sd	mean	between subj sd	mean	between subj sd
Text Entry	**Short**	**Absent**	2.56	0.91	5.68	1.53	52.48%	8%
		Present	2.51	0.95	6.18	1.74	56.27%	8%
	Medium	**Absent**	2.82	1.09	7.90	2.03	57.76%	7%
		Present	2.82	0.84	8.38	1.94	61.03%	7%
	Long	**Absent**	3.26	1.48	14.23	4.37	64.48%	5%
		Present	3.23	1.11	15.17	5.45	65.98%	6%
Text Reading	**Short**	**Absent**	1.40	0.26	2.26	0.38	61.75%	9%
		Present	1.51	0.37	3.24	2.41	58.33%	10%
	Medium	**Absent**	1.96	0.62	4.21	1.17	63.33%	10%
		Present	2.05	0.57	5.24	2.97	66.52%	8%
	Long	**Absent**	2.10	0.60	6.47	1.44	68.14%	8%
		Present	2.23	0.62	7.53	2.37	68.86%	9%

Maximum EOR time

The maximum EOR time was log transformed before entering the regression model. The model results showed significant main and interaction effects for task type and text length (Table 4). Multiple comparisons revealed that the medium and long text reading tasks did not have significantly different maximum EOR time ($p = 0.36$). Additionally, comparing the text entry to the text reading tasks, the shortest text entry task had significantly longer maximum EOR time than longest text reading task ($p = 0.036$).

Further, the model estimated that the average maximum EOR time for text entry tasks with 4, 6, and 12 characters were 2.4 (95%CI = [2.2s, 2.6s]), 2.7 (95%CI = [2.4s, 2.9s]), and 3.0 (95%CI = [2.8s, 3.4s]) seconds, respectively. The estimated maximum EOR time for text reading tasks with 20 ~ 40, 60 ~ 80, and 120 ~ 140 characters were 1.4 (95%CI = [1.3s, 1.5s]), 1.9 (95%CI = [1.7s, 2.1s]), and 2.1 (95%CI = [1.9s, 2.3s]) seconds, respectively. It can be seen from Figure 10 that the increase on maximum EOR time was not linear as the text length increased for text entry and reading tasks. Rather, the increase on maximum EOR time slowed down when the text length was longer, which can be considered as consequence of drivers' risk compensation (i.e., positive adaptation) to the increasing task demands.

Table 4. Model coefficient estimates for log maximum EOR time (Experiment 1)

	Coef. Est.	Std.Error	DF	t-value	p-value
(Intercept)	0.35	0.051	303	6.93	<0.0001
Text Entry (vs. Text Reading)	0.52	0.060	303	8.71	<0.0001
Length Medium (vs. Short)	0.30	0.039	303	7.67	<0.0001
Length Long (vs. Short)	0.39	0.039	303	9.88	<0.0001
Text Entry * Length Medium	-0.19	0.050	303	-3.84	0.0001
Text Entry * Length Long	-0.15	0.050	303	-2.98	0.0032
Model fit	Log Likelihood		AIC	L. Ratio	p-value
At convergence (df=12)	17.71		-11.42	347.079	<0.0001
Null (df=2)	-155.83		315.7		

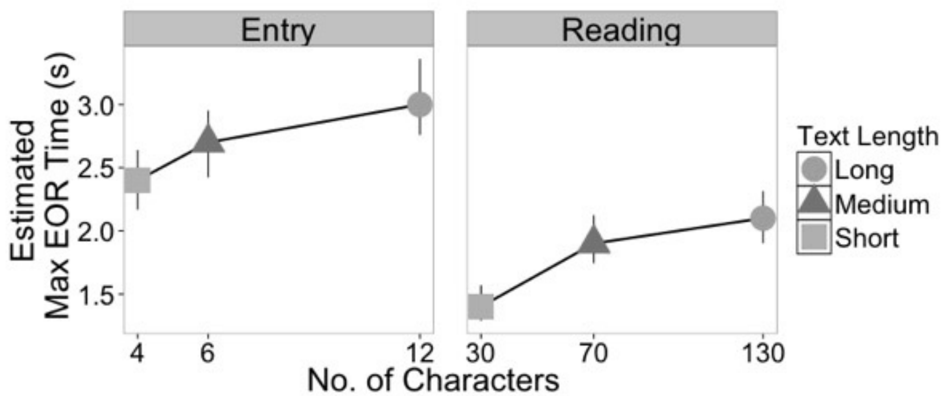

Figure 10. Estimated maximum EOR time (Experiment 1)

One other finding from the model is that, there was a large individual difference among drivers in terms of the length of individual glances. Figure 11 shows the estimated maximum EOR time for each driver and it can be seen that, although most drivers had an increase on their maximum EOR time with the text length, the increase rate was varied among drivers (i.e., different slopes). Although most middle age and older drivers had reasonable glance durations even for long text entry tasks, a few younger drivers had maximum EOR time as long as 4 to 5 seconds even for short text entry tasks. This individual difference is much smaller for text reading than text entry.

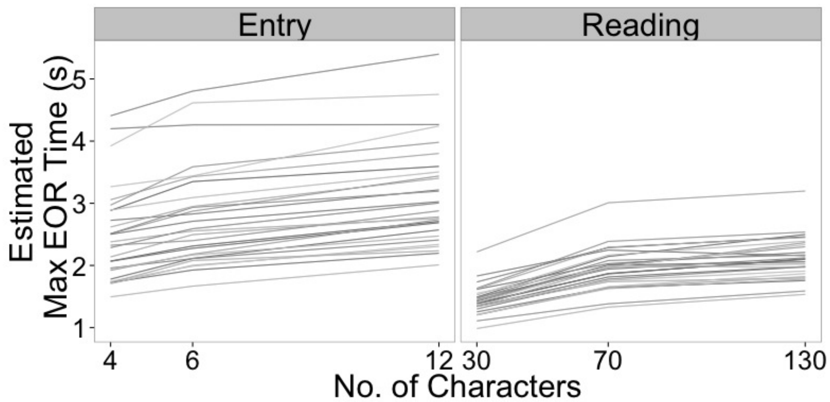

Figure 11. Estimated maximum EOR time for individual drivers (Experiment 1)

Percentage of EOR time

The percentage of EOR time was significantly higher for text reading tasks than text entry tasks (Table 5). For both task types, the model estimated that there was a 9.5% increase on the percentage of EOR time from short length to long length text, and a 4.8% increase from short length of medium length. In fact, the short text reading task had similar percentage of EOR time compared to medium text entry task (p = 0.96), and that of the medium text reading task was similar compared to long text entry task (p = 0.99). Similar to the maximum EOR time, the increase on percentage of EOR time with text length was not linear and had a slow growth with the number of text characters to enter or read (Figure 12). Additionally, the present of task irrelevant text also significantly increased the percentage of EOR time by 1.9% for both text reading and entry tasks. No significant interaction terms or gender and age effects were found in the model.

Table 5. Model coefficient estimates for percentage of EOR time (Experiment 1)

	Coef. Est.	Std.Error	DF	t-value	p-value
(Intercept)	0.593	0.013	304	46.33	<0.0001
Text Entry (vs. Text Reading)	-0.054	0.013	304	-4.26	<0.0001
Length Medium (vs. Short)	0.048	0.008	304	6.06	<0.0001
Length Long (vs. Short)	0.095	0.008	304	11.98	<0.0001
Irrelevant Text Present (vs. Absent)	0.019	0.006	304	2.93	0.0037
Model fit	Log Likelihood		AIC	L. Ratio	p-value
At convergence (df=13)	426.57		-827.1	217.77	<0.0001
Null (df=2)	317.68		-631.4		

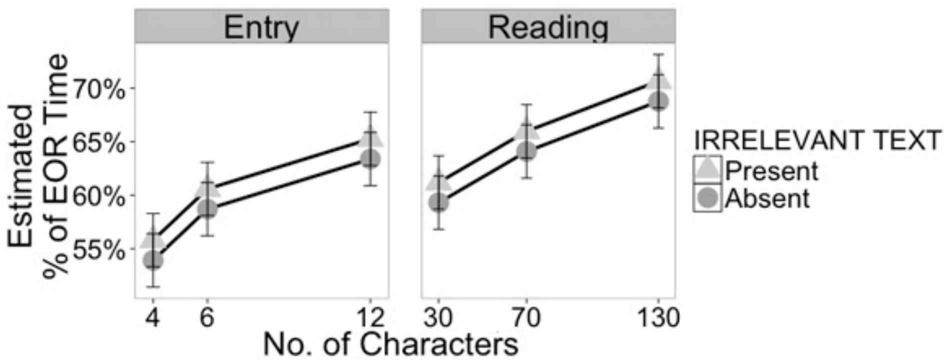

Figure 12. Estimated % of EOR time (Experiment 1)

Total EOR time

Significant main effects were found for task type, text length, and task irrelevant text. The interaction term between task type and text length, task type and task irrelevant text were also significant (Table 6). Specifically, the total EOR time for long text entry tasks was 1.8 times of that for medium text entry tasks, and 2.4 times of that for short text entry tasks. The total EOR time for long text reading tasks was 1.5 times of that for medium text reading tasks, and 2.7 times of that for short text reading tasks. The presence of task irrelevant text slightly increased the total EOR time, and this effect was larger for text reading tasks than task entry tasks. When irrelevant text was presented, the total EOR time was 1 second and 1.2 seconds longer for text entry and text reading tasks, respectively. Unlike the maximum EOR time and the percentage of

EOR time, the increase of total EOR time was in general linear as the text length increased (Figure 13).

Table 6. Model coefficient estimates for total EOR time (Experiment 1)

	Coef. Est.	Std.Error	DF	t-value	p-value
(Intercept)	0.83	0.051	301	16.36	<0.0001
Text Entry (vs. Text Reading)	0.89	0.059	301	14.93	<0.0001
Length Medium (vs. Short)	0.57	0.039	301	14.64	<0.0001
Length Long (vs. Short)	1.00	0.039	301	25.63	<0.0001
Irrelevant Text Present (vs. Absent)	0.17	0.044	301	3.88	0.0001
Text Entry * Length Medium	-0.25	0.055	301	-4.50	0.0000
Text Entry * Length Long	-0.10	0.055	301	-1.85	0.0648
Text Entry * Irrelevant Text Present	-0.10	0.045	301	-2.29	0.0230
Model fit	Log Likelihood		AIC	L. Ratio	p-value
At convergence (df=16)	-5.63		43.25	607.927	<0.0001
Null (df=2)	-309.59		623.2		

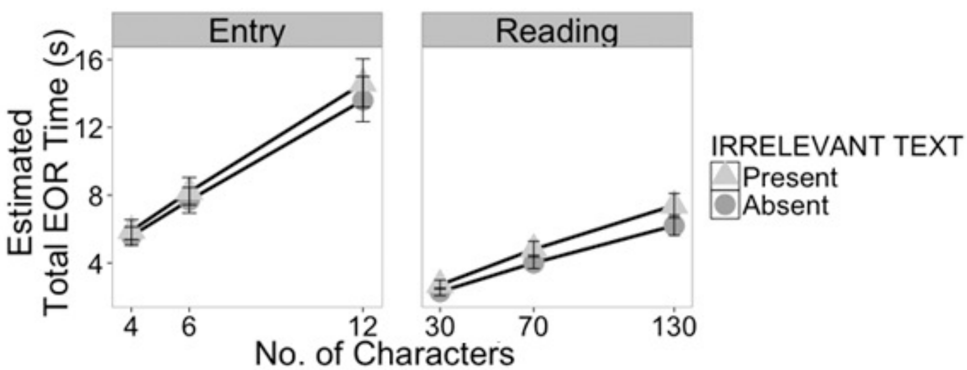

Figure 13. Estimated total EOR time (Experiment 1)

Hypothesis 1.2: Vehicle control performance

The hypothesis 2 states that IVIS tasks with higher demands will negatively impact vehicle control performance. This hypothesis is assessed by drivers' lateral control ability (i.e., SDLP and lane departure events), as well as longitudinal control ability (i.e., mean and standard deviation of following distances). Drivers' vehicle control performance measurements under varied task demands were compared with driving only condition and with each other. The following section summarizes the descriptive statistics and the results from random coefficient models.

Lateral control

There were a total of 46 lane departure events in the 12 task conditions, and 2 in the driving-only conditions. Among the 46 lane departure events, 21 occurred during long text entry tasks, and 11 occurred in medium text entry tasks. No lane departures occurred during short text reading tasks (Table 7). On average, the SDLP was very similar with or without irrelevant text. The average SDLP of text entry tasks was larger than that of text reading tasks, and the SDLP increased as the task length increased.

Table 7. Summary of lateral control performance (Experiment 1)

Type	Length	Irrelevant	SDLP (ft)		No. of lane departures	Total No. of tasks
			Mean	Inter-subj s.d.		
Entry	Short	No	0.52	0.19	3	80
		Yes	0.49	0.17	2	78
	Medium	No	0.60	0.16	7	80
		Yes	0.61	0.20	4	79
	Long	No	0.73	0.16	12	81
		Yes	0.73	0.22	9	80
Reading	Short	No	0.22	0.11	0	78
		Yes	0.28	0.12	0	79
	Medium	No	0.38	0.13	2	75
		Yes	0.42	0.18	1	79
	Long	No	0.46	0.14	4	82
		Yes	0.46	0.14	2	79

The task duration (log transformed) had a significant effect on SDLP as expected, that is, the SDLP was significantly larger when the task duration was longer. When controlling for task completion time, all text entry and reading conditions except short text reading had significantly larger SDLP compared to the driving-only condition (Table 9). The SDLP was about 30% larger while performing text entry tasks, and it was 23% larger while performing medium and long text reading tasks, compared to the same durations driving-only. Specifically, the SDLP increased 0.10, 0.15, and 0.16 feet for short, medium, and long text entry tasks compared to driving only, and increased 0.07 and 0.08 feet for medium and long text reading tasks compared to driving only (Figure 14). However, short reading tasks were not significantly different compared to driving-only condition in terms of SDLP.

Pairwise comparisons showed that SDLP did not differ across the three text entry lengths, once adjusted for task completion time. This suggests that the impact of text length on SDLP may be due to the mediating effect of task completion time. When controlled for task completion time, medium and long reading were not significantly different than entry tasks either. The medium and long length reading both had significantly larger SDLP than short reading tasks.

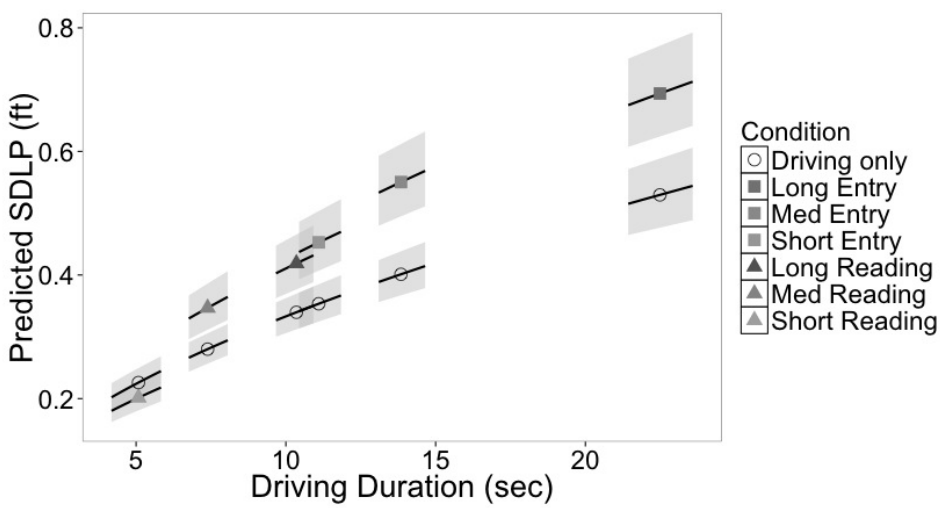

Figure 14. Predicted SDLP by test conditions (Experiment 1)

Longitudinal control

Most drivers kept a larger distance with the lead vehicle than the instructed two-second time headway. The mean time headway for driving-only segments, in spite of driving length, was 4.1 seconds (median = 3.1 s, inter-subject SD = 2.3 s). The mean time headway for each of the 12 task conditions was between 4.44 to 5.14 seconds, with inter-subject SD between 1.8 to 2.7 seconds (Table 8). There was no apparent increase in the mean time headway from short to long text entry or reading tasks.

Table 8. Descriptive Statistics for time headway (Experiment 1)

			Mean time headway (s)		SD of time headway (s)	
Type	**Length**	**Irrelevant**	**Mean**	**Inter-subj s.d.**	**Mean**	**Inter-subj s.d.**
Entry	Short	No	4.78	2.51	0.18	0.11
		Yes	4.46	1.78	0.16	0.08
	Medium	No	4.64	2.58	0.21	0.15
		Yes	4.63	1.93	0.21	0.16
	Long	No	5.14	2.58	0.32	0.23
		Yes	4.85	2.27	0.32	0.23
Reading	Short	No	4.44	2.28	0.06	0.05
		Yes	4.90	2.69	0.10	0.12
	Medium	No	4.53	2.44	0.10	0.07
		Yes	4.85	2.42	0.10	0.07
	Long	No	4.85	2.30	0.15	0.09
		Yes	4.48	1.84	0.14	0.09

Task completion time was included in the model, which had a significant effect on mean time headway (Table 9). However, the coefficient estimate was very small for task completion time. All lengths of text entry and reading resulted in significantly larger mean time headway than driving-only of comparable durations. This is consistent with the previous studies that showed drivers tend to keep a larger headway while performing secondary tasks in order to increase the safety margin. The model predicts that performing text entry and reading tasks of any length increases the mean time headway from 16% to 21%, compared to driving only (Figure 15). No significance differences were observed among different task lengths for text entry and readings.

Table 9. Random coefficient models predicting lateral and longitudinal control (Experiment 1)

			Lateral	Longitudinal	
			Model 1 log (SDLP)	**Model 2** log(mean time headway)	**Model 3** log(SD of time headway)
Variable		**DF**	**Estimated coefficients with standard error**		
(Intercept)		1251	-2.42 (0.10)***	1.17 (0.08)***	-4.27 (0.14)***
Long Entry	**vs. Driving only**	1251	0.27 (0.06)***	0.18 (0.04)***	0.40 (0.07)***
Long Reading		1251	0.21 (0.05)***	0.19 (0.03)***	0.32 (0.06)***
Medium Entry		1251	0.32 (0.06)***	0.15 (0.03)***	0.36 (0.07)***
Medium Reading		1251	0.21 (0.06)***	0.19 (0.03)***	0.23 (0.07)***
Short Entry		1251	0.25 (0.05)***	0.17 (0.03)***	0.39 (0.06)***
Short Reading		1251	-0.11 (0.06)	0.18 (0.04)***	0.25 (0.08)***
log(Task Duration)		1251	0.57 (0.04)***	NA	0.84 (0.05)***
Task Duration		1251	NA	0.005 (0.002)**	NA
Gender Male		26	NA	NA	-0.34 (0.12)**

Model Fit	AIC	df	L. Ratio	AIC	df	L. Ratio	AIC	df	L. Ratio
At convergence	2250	12	622***	985	12	1165***	2671	13	852***
Null	2853	2		2130	2		3502	2	

Note: * p<0.05, ** p<0.01, *** p<0.001

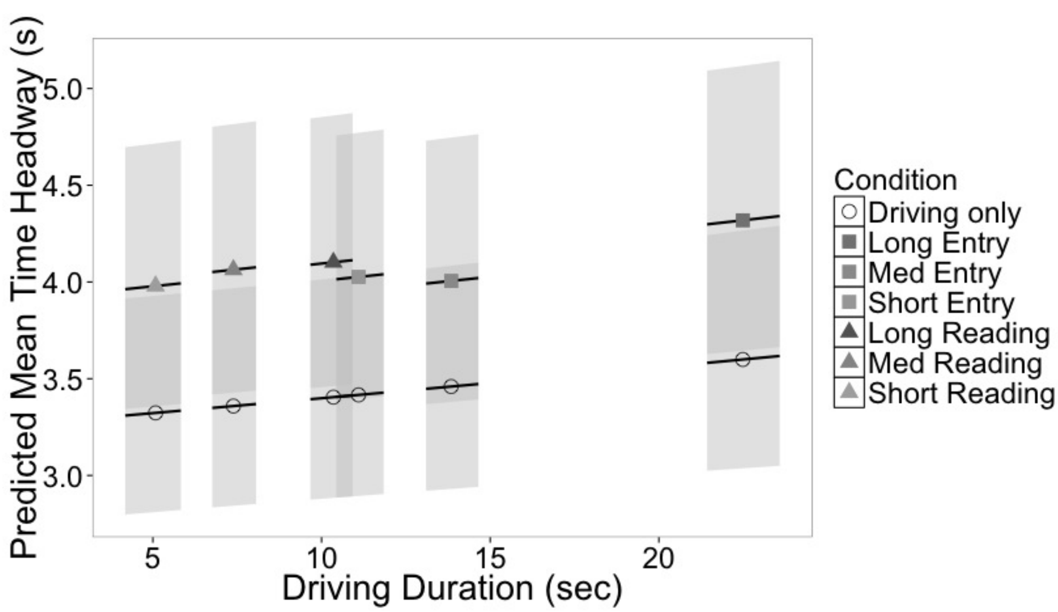

Figure 15. Predicted mean time headway by test conditions (Experiment 1)

The average standard deviation of time headway for driving-only, regardless of driving length, was 0.11s (inter-subject SD = 0.04s). The time headway variation was larger for all text entry tasks than text reading tasks, and increased with the text length (Table 8). The task-irrelevant text did not increase the time headway variation. The regression model showed that both text entry and text reading, regardless of length, resulted in significantly larger standard deviation of time headway while controlling for the log transformed task completion time (Table 9). Male drivers had significantly smaller variation in time headway than female drivers. No significant interaction was observed between gender and task conditions. The model predicted that performing text entry tasks of any length increases the time headway variation by about 45%, compared to the same duration of driving-only. Performing text-reading tasks of any length increases the time headway variation by about 30%, compared to the same duration of driving-only. The time headway for female drivers was 40.5% more variable than for male drivers under the same condition (Figure 16). No significance differences were found between different task lengths for text entry and reading.

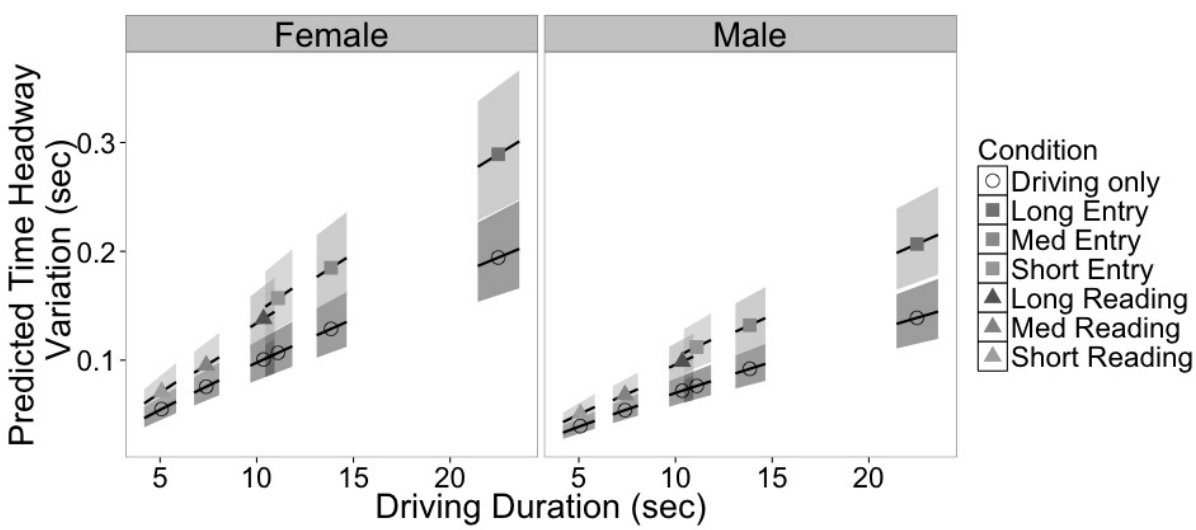

Figure 16. Predicted time headway variation by text condition (Experiment 1)

Discussions

This chapter described and analyzed the data from Experiment 1 of this dissertation, which assessed the hypotheses with regard to drivers' behavioral adaptation with IVIS task demands. A driving simulator study with 28 drivers across four age groups was used to examine drivers' glance behavior and vehicle control performance changes under a set of text entry and reading tasks.

As hypothesized, the findings from this first experiment suggest that drivers changed their glance behavior as the IVIS task demand increased, with longer individual glances toward the IVIS. However, the slow growth on maximum and percentage of EOR time indicates that drivers do not always increase their off-road glance durations with the increasing IVIS demand. Rather, they try to compensate it as possible so the glances off the road was not too long, and when the task demands exceed a certain threshold, the individual off-road glance duration may stop to increase. In addition, the text entry task had larger demand than text reading tasks, but was associated with lower proportion of EOR time. This may suggest a risk compensation strategy of the drivers as well: they try to look longer on the road to compensate for the long off-road glances they had to take in order to finish the text entry tasks. However, different from the other two eye glance measures, the total EOR time increase linearly with the task demand. This shows that the total EOR time measures task demands but not drivers' glance strategies. That is, under certain circumstances, drivers have to spend certain amount of glances on the IVIS in order to finish the task, and they cannot shorten it in order to compensate for driving safety.

The long text entry tasks were associated with approximately half of the lane departure events, the largest mean SDLP, and the largest mean and standard deviation of time headway. It needs to be noted that when calculating SDLP and standard deviation of time headways for tasks with different lengths, the artifact of sampling needs to be taken into account. A longer sampling window may simply increase the likelihood of including items from the tail of a distribution and result in larger variations in the outcomes. This can be seen from the increase of SDLP and standard deviation of time headway for driving-only segments with longer durations. When the task completion time was included in the model, short text reading tasks did not significantly affect SDLP compared to driving only. All other task conditions still significantly increased SDLP, as well as the mean and standard deviation of time headway, when compared to driving only without any secondary task engagement.

The Ranney et al. (2012) study did not reveal any significant differences in terms number of lane departures among IVIS tasks once controlled for task duration. Similar to the Ranney study, this dissertation also showed that long text entry tasks were not significantly different from medium and short text entry, or text reading, when task duration was controlled in the model. That is, when drivers used the same amount of time on the task, text entry does not necessarily result in worse vehicle control performance than text reading, and longer text does not necessarily result in worse vehicle control performance than shorter text, even with the longer off-road glance durations. However, the difference in SDLP and standard deviation of time headway between driving-only and task conditions became larger as the task completion time increased. That is, performance degradation is more obvious with long and more demanding tasks.

There are several reasons why long tasks may increase drivers' risks. One possible reason why drivers may have larger risk when performing long tasks is that it increases the exposure to distractions and therefore the risk increases simply because drivers are at risk for a longer period of time. Lee et al.(2012) however, suggested that when tasks are longer, drivers tend to neglect the driving task and induce long glances away from the roadways. This is somewhat consistent with the observations from this dissertation. When typing or entering texts, drivers tend to chunk the tasks and look back to the road after finishing each chunk. When the text lengths increased, it might become harder for them to chunk the task into short segments, and thus had to neglect the driving task and increase the individual off-road glance durations in order to finish the task.

In summary, drivers adapt their glance behavior with different IVIS task demands. When the IVIS task is more time demanding, they spend longer time looking at the IVIS in each glance, with larger proportion of time during the task. However, they may also realize there is an increased risk with looking away for too long, and adjust their glances accordingly. There appears to be a ceiling effect on the increase of glance durations. On the other hand, their vehicle control performance degraded with the increased IVIS task demands. Similarly with the Ranney et al. (2012) study, the current experiment also found that the task duration plays a key factor in drivers' driving performance assessment. Once task duration was accounted for in the analytical model, there was no significant difference in drivers' lateral and longitudinal control ability among different tasks.

There are several limitations in this study. Due to the design of the tasks and driving scenario, the driving-only block occurred only after the participants completed the IVIS tasks in each simulator trial. This non-randomized portion of the design confounds any learning effect with the findings associated with enhanced vehicle control performance in the driving-only condition. The study design is not capable of determining whether any improvement in vehicle control performance is due to the omission of an IVIS tasks or due to increased familiarization in the simulator. However, given that each participant was given a practice drive at the beginning of the experiment (with and without IVIS tasks for five minutes each), it is expected that any additional learning effects were minimal. In addition, the number of participants in each gender and age group also prohibited any interaction terms from being examined in the regression models.

Experiment 1 only examined drivers' adaptation to IVIS under different task demands on open roadways with no traffic, and on their first exposure to the system. Drivers may also change their behavior when they drive under a more complex driving environment and after they have been exposed to the system more often. In addition, it was also revealed from Experiment 1 that drivers behave very differently even when performing the same tasks. Thus, the next chapter will discuss Experiment 2, which examines other factors affect driver behavioral adaptation to IVIS, such as traffic, long term use, and individual differences.

CHAPTER 4: EXPERIMENT TWO - BEHAVIORAL ADAPTATION WITH DRIVING DEMANDS, LONG-TERM USE, AND DRIVER RISK LEVELS

This chapter discusses the second experiment for this dissertation. A similar set of IVIS tasks was examined under two different traffic conditions over three sessions in one week.

Objective and Hypotheses

The objective of Experiment 2 is to understand how drivers' behavioral adaptation is affected by driving demands, long-term use, and driver risk levels. The two hypotheses associated with this experiment are shown as below.

Hypothesis 2.1: Drivers may have worse glance behavior when interacting with IVIS over time (negative adaptation), but better glance behavior when driving with ambient traffic on the road compared to situations with no traffic (risk compensation). In addition, driver risk levels may also have an impact on their glance behavior.

Hypothesis 2.2: Drivers may have better vehicle control performance over time, and the traffic condition and driver risk levels may also have impacts on their vehicle control performance.

Methodology

Participants

Twenty-eight drivers (14 males and 14 females) were recruited from the Seattle, Washington area. They were between 18 to 30 years old (16 were under 25 and 12 were from 25 to 30), native English speakers, and driving at least 7000 miles per year. Other recruitment inclusion criteria include being in good general health conditions, comfortable using computers, touchscreens, and communicating via text messages. Participants were compensated for their time in the study, and provided with parking validation if needed.

All drivers had at least a high school diploma, and on average obtained their driver's license at 17 to 18 years old (Table 10). Five drivers reported that they drove less than once a week, and the rest of the drivers reported driving at least once weekly. For the participants between 18 to 25 years old, 50% of them had at least one moving violation, and 37.5% of them had at least one crash in the past three years. For participants above 25 years old, 16.7% of them had moving violations or crashes in the past three years. However, none of the crashes that the participants reported was their fault.

Table 10. Driver demographics and driving history (Experiment 2)

Age Group	Gender	Age		Age received license	Moving violations in past 3 years		Crashes in past 3 years	
		Mean	sd		0	1+	0	1+
Age 18-24	Male	21.1	1.81	17.1	4	4	4	4
	Female	21.6	1.30	17.0	4	4	6	2
Age 25-30	Male	27.8	1.47	18.0	5	1	5	1
	Female	27.3	2.34	17.3	5	1	5	1

Experimental Design

The study used a within-subject design with 6 IVIS task conditions of interest (2 task types x 3 text lengths). All task conditions were conducted under two traffic conditions and repeated over three driving sessions in 7 consecutive days (Figure 17).

The 6 IVIS task conditions include short (4 char), medium (6 char), and long (12 char) text entry, and short (20 – 40 char), medium (60 – 80 char), and long (120 – 140 char) text reading. The IVIS tasks in Experiment 2 are very similar to those in Experiment 1, and were tested using the same facility.

The two traffic conditions were with and without ongoing traffic from both directions. For the "without traffic" condition, there was only one lead vehicle on the simulated roadway and no other vehicles. For the "with traffic" condition, in addition to the lead vehicle, there were other vehicles passing from the opposite direction approximately every 7 seconds, one vehicle drove behind the driver, and two other vehicles drove right to the driver (Figure 18). There was no lead vehicle braking events in any of the driving scenarios.

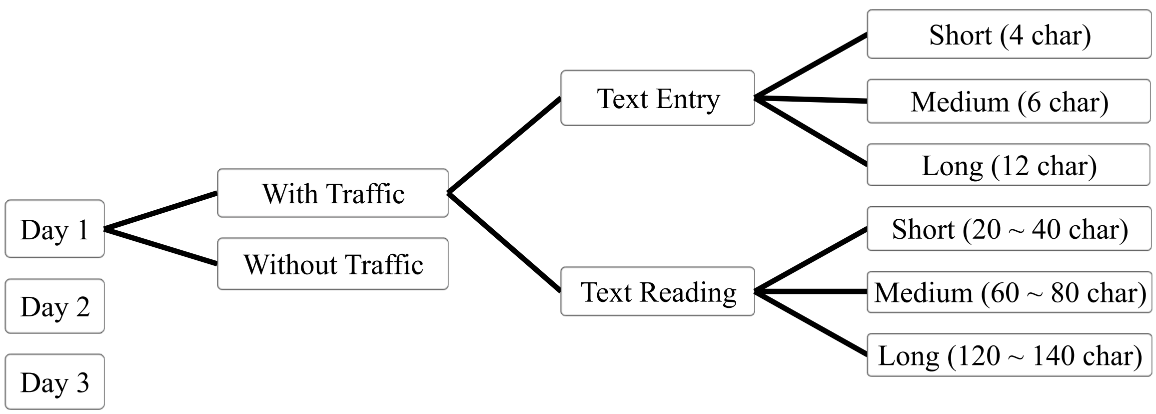

Figure 17. Experimental design of Experiment 2

Figure 18. Driving scenario with traffic

In order to examine the hypotheses, several driving habit and opinion questions were asked in a pre-study questionnaire (Appendix 1) to obtain drivers' risk levels. Drivers were asked to rate how frequently they conduct the behavior in the statements, or how safe they considered them (Table 11). Both driving habit and opinion questions were rated on 5-point Likert scales.

The driving habit questions were rated as “never”, “rarely”, “sometimes”, “mostly” or “always”; the opinion questions were rated as “extremely dangerous”, “somewhat dangerous”, “neither”, “somewhat safe”, or “extremely safe”. The questionnaire questions can be categorized into four categories related to speeding, aggressive driving, traffic rules, and driver distraction.

Table 11. Driver habit and opinion questions

Speeding
Drive at least 10 mph over the speed limit
Driving the speed limit in adverse weather (e.g., snow, heavy rain)
Disregard the speed limit late at night or early in the morning
Aggressive Driving
Tailgate the car in front of you
Cutting in front of other cars
Turn without signaling
Switching back and forth between lanes to drive through traffic
Use angry or insulting gestures toward other drivers
Use your horn
Traffic Rules
Drive through the intersection when you see a yellow light
Not come to a full stop at a stop sign
Run a red light if nobody is around
Pass a school bus while it’s red lights are flashing
Driver Distraction
Using cell phone while driving
Typing address on a GPS device while driving

Procedures

The experiment involved 3 sessions in 7 consecutive days. The study protocol for Experiment 2 was similar to Experiment 1, except that for each session, the participants were asked to drive 2 trials without ambient traffic, and 2 trials with ambient traffic. Same as in Experiment 1, drivers were asked to drive 5 minutes after they finished all tasks in each trial as the driving only condition. The first session (Day 1) contained one practice trial and 4 main study trials. Half of the participants were randomly selected to perform the traffic trials first to minimize the order effects. During the second and the third session (Day 2 and 3), participants only performed the 4 main study trials.

Data Analysis

Drivers' risk levels were determined by conducting cluster analysis using their responses on the driving habit and opinion questions. Cluster analysis is a commonly used statistical method for identifying homogeneous groups of subjects based on given attributes (Aldenderfer & Blashfield, 1984). Specifically in this study, cluster analysis evaluates distance or similarity between drivers based on their driving habits and opinions to create sub-groups of drivers so that different sub-groups are significantly different in terms of risk.

The habit and opinion questions were coded as 1 to 5 where 1 was for "never" (for habit questions) and "extremely dangerous" (for opinion questions), and 5 for "always" and "extremely safe". Drivers' ratings on the questions from each category (i.e., speeding, aggressive driving, traffic rules, and driver distraction) were averaged. Drivers were then clustered based on their averaged ratings for these four categories.

The four variables were not highly correlated ($r < .90$), so strong collinearity was not a concern for the analysis (Mooi & Sarstedt, 2011). In addition, the sample size was considered sufficient with five clustering variables ($n \geq 2^m$,where m = number of clustering variables, n = sample size) (Formann, 1984). A hierarchical clustering method with Euclidean distance measure and Ward's method was used to minimizes the increase of variance when any two clusters are formed (Romesburg, 2004). The clustering solution was validated using several internal (i.e., measures the compactness, connectivity, and separation) measures using *R* package *clValid* (Brock, Pihur, Datta, & Datta, 2011). The resulting clusters were used as independent variables in the regression models to examine the hypotheses.

The 28 participants completed 6 replications of 6 different task conditions for a total of 168 data points per session, traffic, and IVIS task condition. The dependent variables, maximum EOR time and proportion of EOR time, SDLP and standard deviation of time headways, were averaged over the 6 replications for each participant for statistical analysis. Some partial data from 6 drivers were not used due to data missing problem; however, there were still data points from at least 27 drivers left for each condition for analysis.

Random coefficient models were used to examine the vehicle control performance and eye glance behavior (using R 2.12.1, package "nlme"). The six IVIS tasks conditions (i.e., long text entry, medium text entry, short text entry, long text reading, medium text reading, and short text reading) were examined in the models with short text reading condition as the baseline level when modeling glance behavior variables, and with the driving only condition as the baseline level when modeling vehicle control performance variables. The method to extract segments for the driving only condition was the same as in Experiment 1. The traffic condition was coded as a binary variable with the "without traffic" scenario as the baseline level. The driver risk levels obtained from the cluster analysis were also coded as categorical variable and tested in the models. The time variable was coded as a continuous variable ranging from 1 (Day1) to 3 (Day3). Task durations were also included in the vehicle control performance models as covariate since it was shown to have significant effects in the first experiment. Age groups (i.e., 18-24 years old vs. 25 ~ 30 years old) and genders were also tested in the models. The dependent variable maximum EOR time, SDLP, and standard deviation of time headways, were log transformed to meet the regression assumptions on residuals. Forward selection method was used for variable selection. The inclusion and exclusion of random intercepts and slopes, fixed main and interaction effects, and the final model were determined using AIC values and likelihood ratio tests.

Results

Driver risk levels

The driver risk levels were determined using cluster analysis. The dendrogram appeared to suggest two clusters (Figure 19), which were named high and low risk groups based on drivers' responses to the driving habits and opinion questions. The internal validation results also supported the two-cluster solution as it performed the best in terms of connectivity, closeness, and separation. There were 14 drivers in the high-risk group and 14 drivers in the low-risk group, with equal number of drivers from the 18-24 age range and 25-30 age range for each group. In addition, the high-risk group had 9 male drivers while the low-risk group had 6 male drivers. For both the low- and high-risk drivers, there was the same proportion of drivers involved in at least one crash in the past three years (28.6%). For moving violations, 50% of the high-risk group drivers had at least one moving violations in the past three years, while only 21.4% drivers in the low-risk group had moving violations in the past three years. However, Fisher's exact test did not show significant difference in terms of number of moving violations between the two groups (p-value = 0.23), but this could be more associated with the sample size.

There also seems to be an outlier (#17 in Figure 19), which was caused by the participant's extremely high ratings on the driver distraction questions (i.e., considered using cell phones and typing address during driving was "extremely safe"). The cluster analysis was conducted also without this outlier and the result suggested same clustering for all other participants as to include the outlier. Therefore, the clustering solution with the outlier was used as the final results.

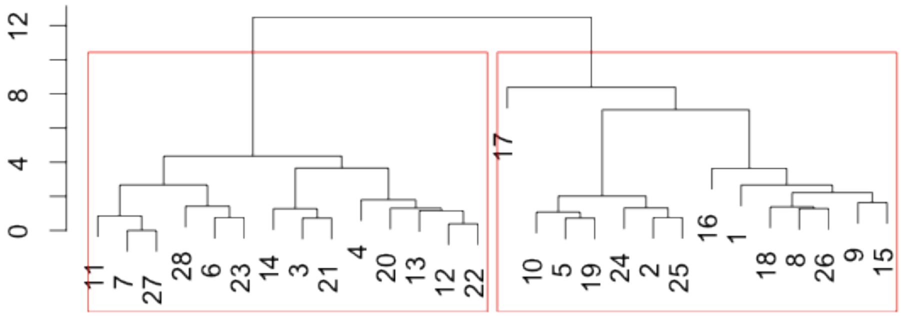

Figure 19. Dendrogram for cluster analysis

Simple linear regressions were used to test the differences between the two clusters of drivers on their driving habits and opinions. Drivers in the high-risk group had significantly higher ratings on the questions related to speeding (p<0.0001), aggressive driving (p=0.024), and traffic rules (p=0.0001). In other words, the high-risk group drivers thought driving fast, aggressively, and disobey traffic lights under certain situations were not as dangerous as the low risk group drivers thought, and they also slightly more frequently did so. However, the two groups of drivers had similar opinions (p=0.285) on driver distraction and on average all agreed that it was somewhat dangerous to enter address or use cell phones while driving (Figure 20). The driver clusters were then taken into the models for the maximum EOR time and proportion of EOR time in the following analyses.

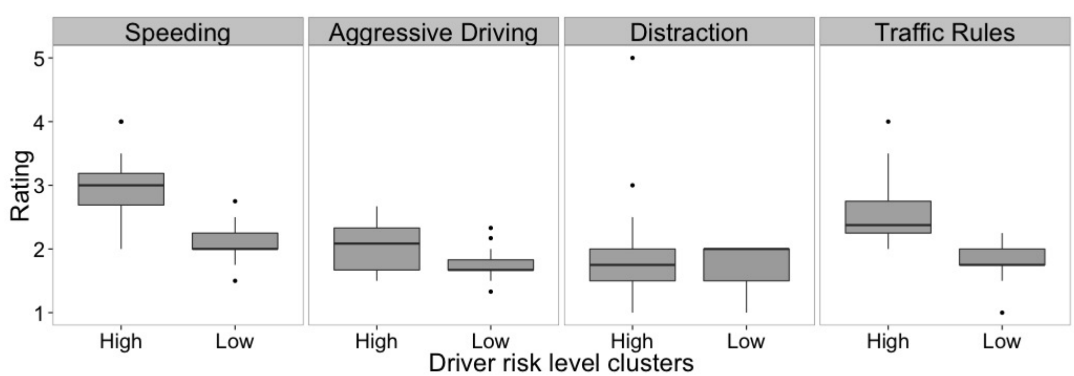

Figure 20. Driving habits and safety opinions for drivers with high and low risk levels

Hypothesis 2.1:Maximum EOR Time

Averaging over all 6 IVIS task conditions, the mean maximum EOR time on day 1 was 2.6 and 2.3 seconds under “without traffic” and “with traffic” conditions, respectively. The maximum EOR time increased approximately 0.2 seconds each day for both scenarios. On day 3, the average maximum EOR time across all task conditions were 3.0 and 2.8 seconds for “without traffic” and “with traffic” conditions, respectively.

Similar to the results from Experiment 1, drivers had different glance behavior for reading and entry tasks. The average maximum EOR time for all text reading tasks was approximately 2.0 seconds, and remained similar over the three experimental sessions. The average maximum EOR time for all text entry tasks were 3.0, 3.3, and 3.7 seconds for Day1, 2, and 3, respectively. In addition, this increase on maximum EOR time for text entry tasks was larger for high risk drivers compared to the low risk ones. Figure 21 shows the distributions of maximum EOR time for high and low risk drivers over time. It can be seen that for both risk level groups, the maximum EOR time densities were similar from day1 to day3 for text reading tasks. However, it shifted toward the right over time (i.e., longer maximum EOR time) for the high risk drivers for text entry tasks.

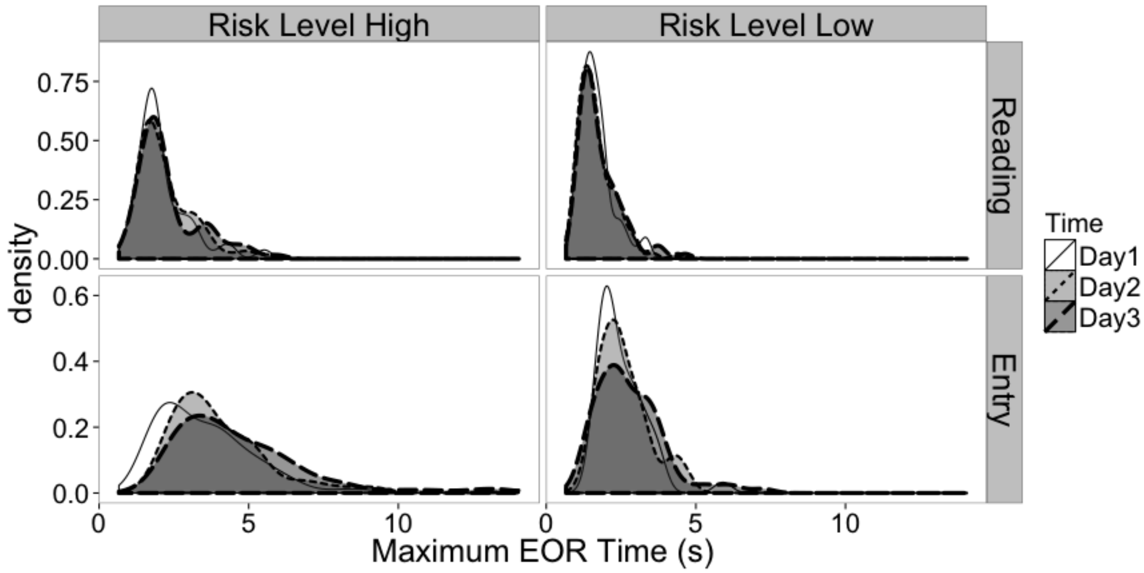

Figure 21. Distributions for max EOR time by driver risk group, time, and task types (Experiment 2)

The maximum EOR time was predicted using the random coefficient model. The final model equation is:

$$\log(Max\ EOR\ Time)_{ijkt} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t + (\beta_2 + b_{2i})Task\ Demands_j + (\beta_3 + b_{3i})Traffic_k + \beta_4 Gender + \beta_5 Driver\ Risk\ Levels + \beta_6 t \times Task\ Demands_j + \beta_7 Driver\ Risk\ Level \times Task\ Demands_j + \beta_8 Driver\ Risk\ Level \times t + \beta_9 Driver\ Risk\ Level \times Task\ Demands_j \times t + \varepsilon_{ijkt}.$$

The dependent variable in the equation is the log transformed maximum EOR time for driver *i*, under task demand *j*, traffic condition *k*, at time *t*. In addition, β_0 is the fixed effect for the model intercept, and β_1 to β_9 are the fixed effects for variable time, task demands, traffic, gender, and driver risk levels, and their interactions. b_{0i} is the random effect of the model intercept, and b_{1i}, b_{2i}, b_{3i} are the random effects of the variable time, task demands, and traffic conditions, respectively. The random effects follow the following distributions:

$$b_{0i} \sim N(0, \sigma_0^2), b_{1i} \sim N(0, \sigma_1^2), b_{2i} \sim N(0, \sigma_2^2), b_{3i} \sim N(0, \sigma_3^2),$$

$$cov(b_{0i}, b_{1i}) = \sigma_{01}, \varepsilon_{ijkt} \sim N(0, \sigma_e^2)$$

The model results showed that there was a significant effect on task demands, traffic condition, gender, and driver risk levels (Table 12). There was also a significant interaction effect between task demands and time, as well as among time, task demands and driver risk levels. The maximum EOR time was about 8.2% longer when driving without traffic, and 25.6% longer for male drivers compared to female drivers. Consistent with the results from Experiment 1, the maximum EOR time was significantly longer for text entry tasks than text reading tasks, and increased (but nonlinearly) with text length.

The high risk drivers had similar maximum EOR time as low risk drivers when performing short and medium length text reading tasks, but had significantly longer maximum EOR time when performing long text reading tasks as well as text entry tasks with any length (Figure 22). Specifically, high risk drivers had approximately 1.4 times longer maximum EOR time than low risk drivers when performing medium and long length text entry tasks, and had 1.2 times longer maximum EOR time when performing short text entry tasks and long text reading tasks.

Table 12. Random coefficient model results for maximum EOR time (Experiment 2)

FIX EFFECTS		Value	Std.Error	DF	t-value	p-value
(Intercept)		0.255	0.101	944	2.53	0.012
Time		0.015	0.034	944	0.45	0.656
Entry.Medium		**0.678**	**0.074**	**944**	**9.12**	**<0.0001**
Entry.Short		**0.515**	**0.074**	**944**	**6.92**	**<0.0001**
Reading.Long	**(vs. Reading.Short)**	**0.545**	**0.074**	**944**	**7.32**	**<0.0001**
Reading.Medium		**0.301**	**0.074**	**944**	**4.04**	**0.0001**
Entry.Long		**0.764**	**0.075**	**944**	**10.22**	**<0.0001**
With Traffic (vs. Without Traffic)		**-0.079**	**0.023**	**944**	**-3.47**	**0.0005**
Male (vs. Female)		**0.229**	**0.088**	**25**	**2.60**	**0.015**
Driver Risk Low (vs. Driver Risk High)		-0.044	0.120	25	-0.36	0.720
Entry.Medium		**0.109**	**0.028**	**944**	**3.83**	**0.0001**
Entry.Short		**0.117**	**0.028**	**944**	**4.12**	**<0.0001**
Reading.Long	*** Time**	0.012	0.028	944	0.42	0.671
Reading.Medium		0.042	0.028	944	1.49	0.138
Entry.Long		**0.165**	**0.028**	**944**	**5.81**	**<0.0001**
Entry.Medium		-0.201	0.105	944	-1.92	0.056
Entry.Short		0.031	0.105	944	0.30	0.767
Reading.Long	* Driver Risk Low	-0.202	0.105	944	-1.92	0.055
Reading.Medium		-0.101	0.105	944	-0.96	0.335
Entry.Long		-0.013	0.105	944	-0.12	0.905
Time	* Driver Risk Low	-0.017	0.048	944	-0.35	0.729
Entry.Medium		-0.040	0.040	944	-0.99	0.321
Entry.Short		**-0.105**	**0.040**	**944**	**-2.62**	**0.009**
Reading.Long	*** Driver Risk Low*Tim**	0.012	0.040	944	0.29	0.771
Reading.Medium		-0.003	0.040	944	-0.08	0.937
Entry.Long		**-0.137**	**0.040**	**944**	**-3.42**	**0.001**

RANDOM EFFECTS	StdDev	Corr
(Intercept)	0.136	
Time	0.101	-0.785
Task Demand j	0.168	
Task Demand j'	0.168	0.571
Without Traffic	0.157	
With Traffic	0.157	0.766
Residual	0.148	

MODEL FIT	df	AIC	logLik
At convergence	34	-467.882	267.941
Null	2	1376.528	-686.2642
		L.Ratio	**p-value**
		1908.41	<0.0001

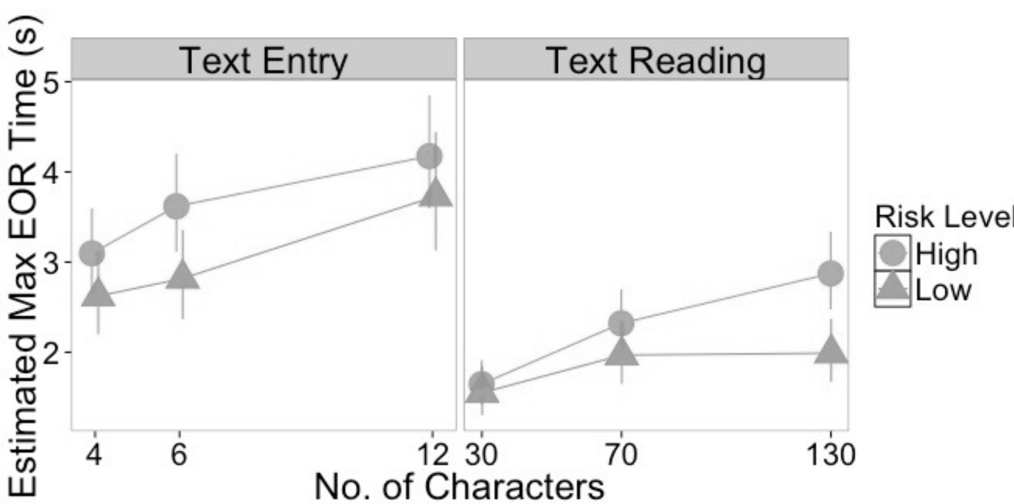

Figure 22. Estimated maximum EOR time by task type, text length, and driver risk levels (Experiment 2)

The variable *time* was significant for all text entry tasks, but only for text reading tasks with medium length. The maximum EOR time increased approximately 10% over each day of experiment for medium and long text entry tasks, and about 7% for short text entry tasks (Figure 23). The maximum EOR time remained similar over time for text reading tasks, and there was no significant interaction effect between time and driver risk levels found in the model.

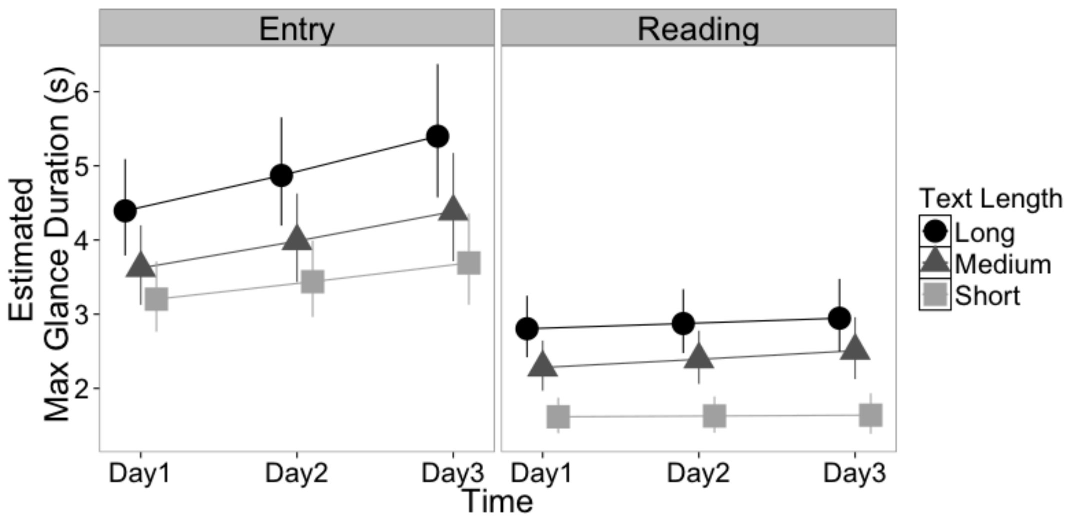

Figure 23. Estimated maximum EOR time by task demands and time (Experiment 2)

Figure 24 shows the estimated max EOR time for individual drivers and shows the three-way interaction among driver risk levels, task demands, and time. It can be seen from the figure that high and low risk drivers had relatively similar adaptive behavior over time when performing text reading tasks. However, high-risk drivers had significantly longer max EOR when performing long text entry tasks, and it increased significantly over time. Additionally, high-risk drivers also much larger between-subject variations on max EOR when performing text entry tasks, especially with long text, compared to low-risk drivers.

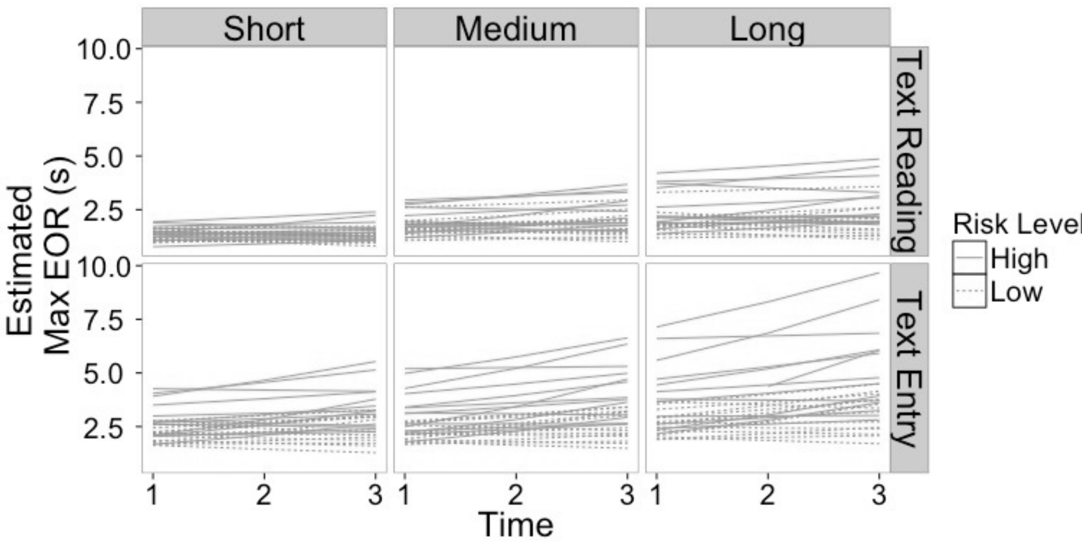

Figure 24. Three-way interactions between driver risk levels, task demands, and time on max EOR (Experiment 2)

Hypothesis 2.1:Proportion of EOR Time

The proportion of EOR time on average ranged from 54.4% to 65.8% for short to long text entry tasks, and 59.6% to 69.6% for short to long text reading tasks. With traffic condition on average had 2% less EOR time compared to without traffic condition. There was also a slight increase of the proportion of EOR time over time across task and traffic conditions.

Figure 25 shows the density plots of proportion of EOR time by driver risk levels, task types, and over time. It can be seen from the plot that the proportion of EOR time for high risk drivers skewed more toward the right on Day3 compared to the previous two days. However, this is not true for the low risk group drivers.

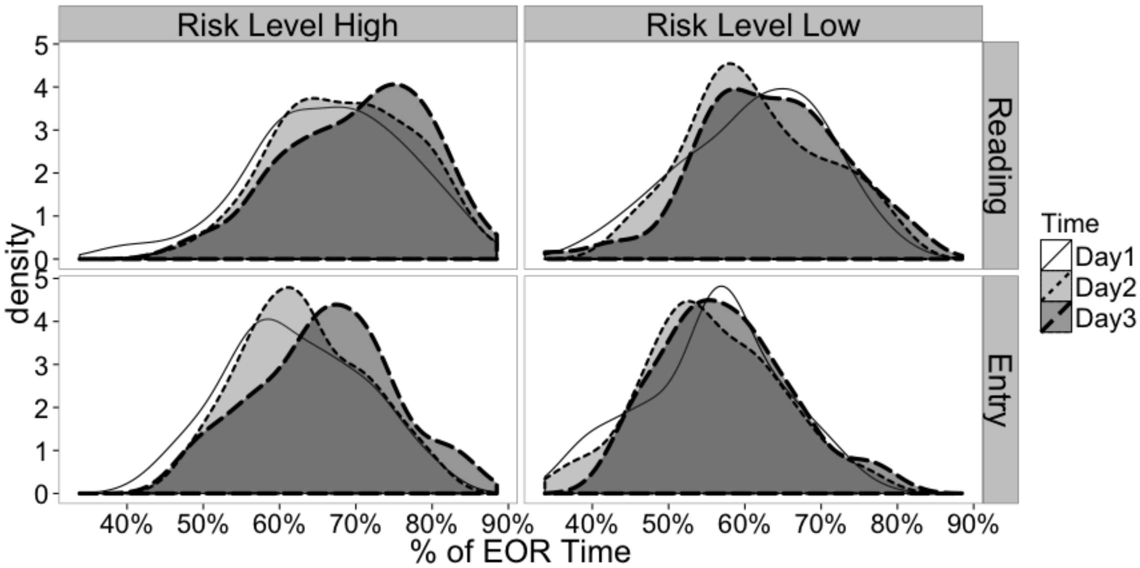

Figure 25. Density plot of proportion of EOR time by driver risk group, task type, and time (Experiment 2)

The random coefficient model showed significant main effects of time, task demands, traffic condition, gender, and driver risk levels (Table 14). However, no significant interaction effects between time and driver risk levels were found in the model. Specifically, the proportion of EOR time was estimated to increase 1.4% over each day of experimentation. Driving that encountered ambient traffic reduced the proportion of EOR time by 2%. Male drivers spent 5.7% more time looking at the IVIS screen, and high-risk drivers spent 5% more time looking at the IVIS. In addition, the estimates of proportion of EOR time for different task demands were similar as the results from Experiment 1.

Table 13. Random coefficient model results for proportion of EOR time (Experiment 2)

FIX EFFECTS		Coef. Est.	Std.Error	DF	t-value	p-value
(Intercept)		0.575	0.0218	960	26.38	<0.0001
Time		0.014	0.0050	960	2.83	0.0048
Entry.Short		-0.053	0.0074	960	-7.08	<0.0001
Reading.Medium		0.067	0.0074	960	9.06	<0.0001
Entry.Medium	**(vs. Reading.Short)**	0.000	0.0074	960	-0.02	0.9824
Reading.Long		0.100	0.0074	960	13.41	<0.0001
Entry.Long		0.061	0.0074	960	8.23	<0.0001
With Traffic (vs. Without Traffic)		-0.020	0.0060	960	-3.31	0.0010
Male (vs. Female)		0.057	0.0203	25	2.83	0.0091
Driver Risk Low (vs. Driver Risk High)		-0.050	0.0203	25	-2.47	0.0207

RANDOM EFFECTS	StdDev	Corr	MODEL FIT	df	AIC	logLik
(Intercept)	0.050		At convergence	18	-3091.141	1563.57
Time	0.030	-0.880	Null	2	-1737.578	870.79
Task Demand j	0.040				**L.Ratio**	**p-value**
Task Demand j'	0.040	0.739			1385.563	<0.0001
Without Traffic	0.030					
With Traffic	0.030	0.500				
Residual	0.040					

Hypothesis 2.2: SDLP

The SDLP was similar as found in Experiment 1 for each IVIS condition. In addition, driving “with” and “without” ambient traffic had similar SDLP across IVIS task conditions. The SDLP appeared to be slightly reduced from day 1 to day 3 of the experiment not only for both text entry and text reading tasks, but for driving only condition as well (Figure 26). High-risk drivers had slightly larger SDLP when performing text entry tasks than low risk drivers, but had very similar SDLP distributions for text reading tasks and driving-only condition.

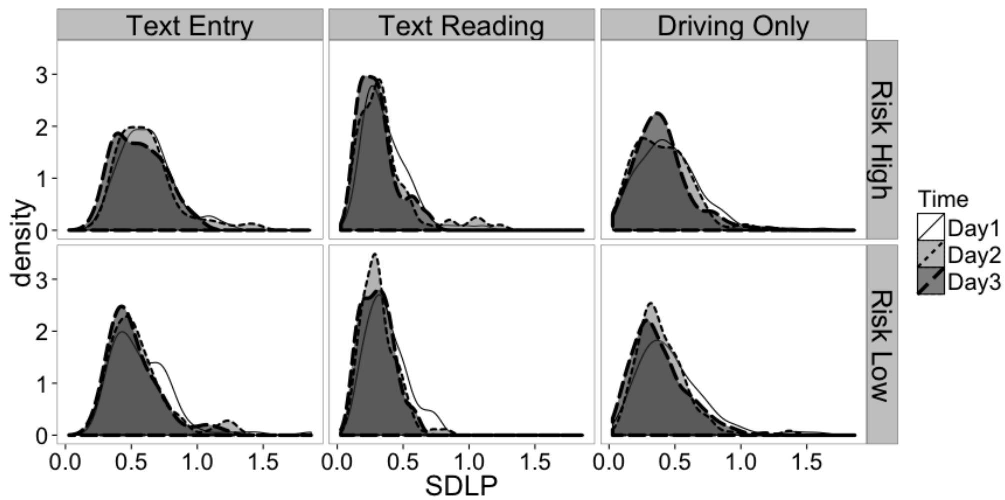

Figure 26. Density of SDLP by task demands, time, and driver risk levels (Experiment 2)

The random coefficient model showed a significant main effect for time, task demand, and task duration. Short and medium text reading had slightly better or similar SDLP compared to the driving-only condition with same durations; text entry tasks had approximately 13% to 20% larger SDLP compared to driving-only condition with same durations (Figure 27). The SDLP was estimated to decrease by about 3% for each day of experiment, independent of IVIS tasks (Figure 28). Driver risk levels and traffic conditions were not found significant in the model. No other main effects or interaction effects were found as well.

Table 14. Model estimates for (log transformed) SDLP (Experiment 2)

FIX EFFECTS		Coef. Est.	Std.Error	DF	t-value	p-value
(Intercept)		-2.085	0.068	1895	-30.48	<0.0001
Time		-0.031	0.014	1895	-2.19	0.0283
Reading.Short	**vs. Driving Only**	-0.072	0.036	1895	-1.99	0.0467
Entry.Short		0.125	0.032	1895	3.89	0.0001
Reading.Medium		0.020	0.033	1895	0.60	0.5481
Entry.Medium		0.188	0.032	1895	5.79	<0.0001
Reading.Long		0.091	0.032	1895	2.84	0.0046
Entry.Long		0.154	0.035	1895	4.35	<0.0001
log(Duration)		0.519	0.019	1895	27.82	<0.0001
RANDOM EFFECTS	**StdDev**	**Corr**	**MODEL FIT**	**df**	**AIC**	**logLik**
(Intercept)	0.227		At convergence	15	1442.99	-706.49
Time	0.050	-0.296	Null	2	3090.77	-1543.39
Task Demand j	0.087				**L.Ratio**	**p-value**
Task Demand j'	0.087	0.633			1673.78	<0.0001
Residual	0.335					

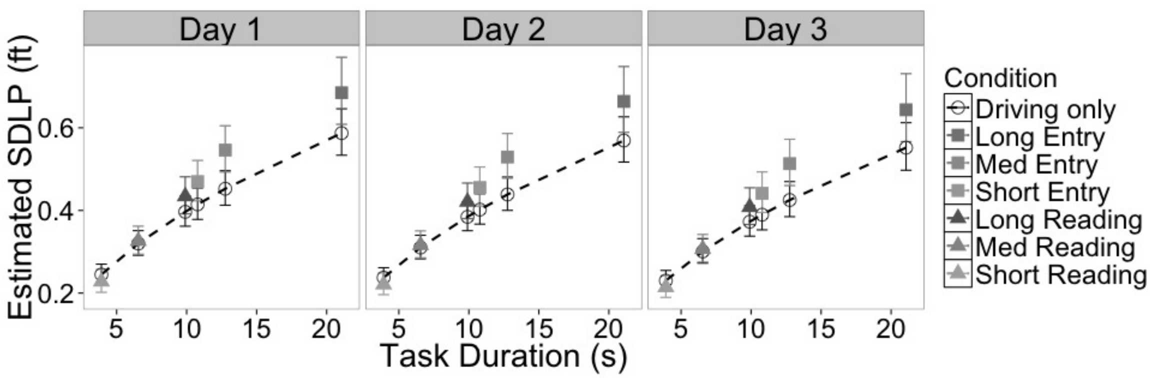

Figure 27. Estimated SDLP by task demands (Experiment 2)

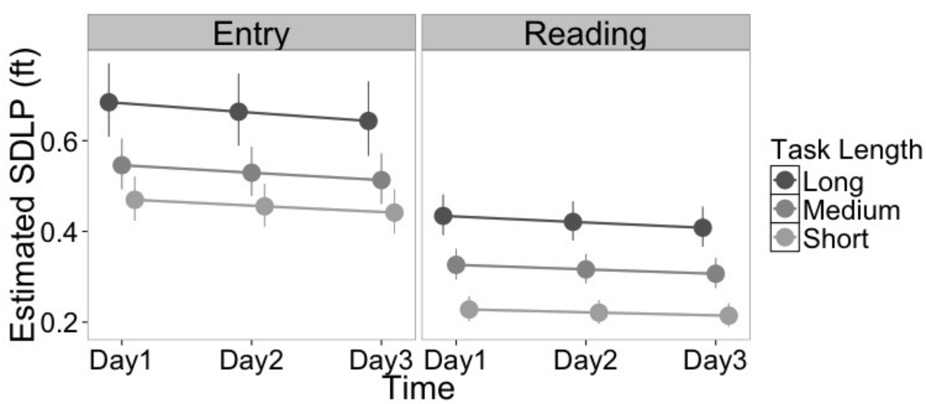

Figure 28. Estimated SDLP by time (Experiment 2)

Hypothesis 2.2: Standard Deviation of Time Headway

The standard deviation of time headway was slightly smaller for the with traffic condition than the without traffic condition across IVIS task types (diff = - 0.037 s), and larger for low-risk drivers compared to high-risk drivers (diff = 0.031 s). Figure 29 shows the density plots for the standard deviation of time headways by task type, traffic, and time. It appears that over time the distribution did not change for driving only condition, but slightly skewed to the left for the text reading and entry tasks.

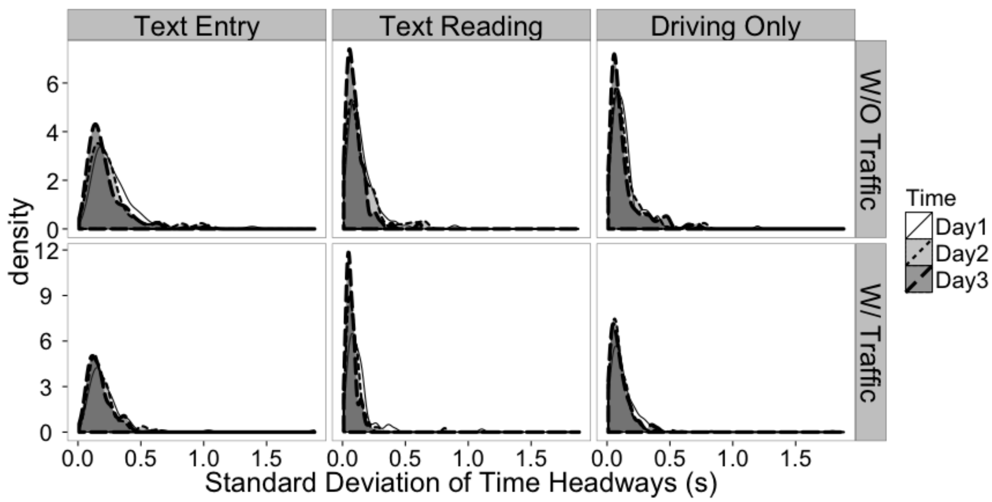

Figure 29. Density plot for standard deviation of time headways by task demands, traffic, and time (Experiment 2)

The regression model suggested that traffic, task demands, and task duration had significant main effects on the standard deviation of time headways (Table 16). There was also a significant interaction effect between time and task demands. However, driver risk levels, gender, and age were not significant in the model. Specifically, the model suggested that the standard deviation of time headway was about 27% larger when there was no ambient traffic. When controlled for task duration, the time headway was significantly more varied for all IVIS task conditions compared to driving-only condition on day 1, except the long text entry tasks (Figure 30). Over time, the time headway variation did not significantly change for driving-only condition and text entry tasks. However, the time headway variation significantly reduced approximately 12% to 20% per experimentation day for text reading tasks (Figure 31).

Table 15. Model estimates for standard deviation of time headway (Experiment 2)

FIX EFFECTS		Coef. Est.	Std.Error	DF	t-value	p-value
(Intercept)		-2.882	0.105	1888	-27.48	<0.0001
Reading.Short	**vs. Driving Only**	**0.346**	**0.102**	**1888**	**3.41**	**0.0007**
Entry.Short	**vs. Driving Only**	**0.555**	**0.099**	**1888**	**5.58**	**<0.0001**
Reading.Medium	**vs. Driving Only**	**0.575**	**0.100**	**1888**	**5.74**	**<0.0001**
Entry.Medium	**vs. Driving Only**	**0.507**	**0.100**	**1888**	**5.09**	**<0.0001**
Reading.Long	**vs. Driving Only**	**0.561**	**0.100**	**1888**	**5.64**	**<0.0001**
Entry.Long	**vs. Driving Only**	0.093	0.105	1888	0.89	0.3759
Time		-0.034	0.035	1888	-0.96	0.3377
With Traffic (vs. Without Traffic)		**-0.241**	**0.036**	**1888**	**-6.69**	**<0.0001**
Duration		**0.065**	**0.002**	**1888**	**27.34**	**<0.0001**
Reading.Short	*** Time**	**-0.190**	**0.045**	**1888**	**-4.22**	**<0.0001**
Entry.Short	*** Time**	-0.025	0.045	1888	-0.56	0.5774
Reading.Medium	*** Time**	**-0.161**	**0.045**	**1888**	**-3.56**	**0.0004**
Entry.Medium	*** Time**	-0.010	0.045	1888	-0.23	0.8208
Reading.Long	*** Time**	**-0.094**	**0.045**	**1888**	**-2.08**	**0.0379**
Entry.Long	*** Time**	0.085	0.045	1888	1.87	0.0616

RANDOM EFFECTS	StdDev	Corr
(Intercept)	0.393	
Time	0.159	-0.747
Task Demand j	0.174	
Task Demand j'	0.174	0.818
Without Traffic	0.218	
With Traffic	0.218	0.739
Residual	0.435	

MODEL FIT	df	AIC	logLik
At convergence	24	2545.90	-1248.95
Null	2	4545.30	-2270.65
		L.Ratio	**p-value**
		2043.398	<0.0001

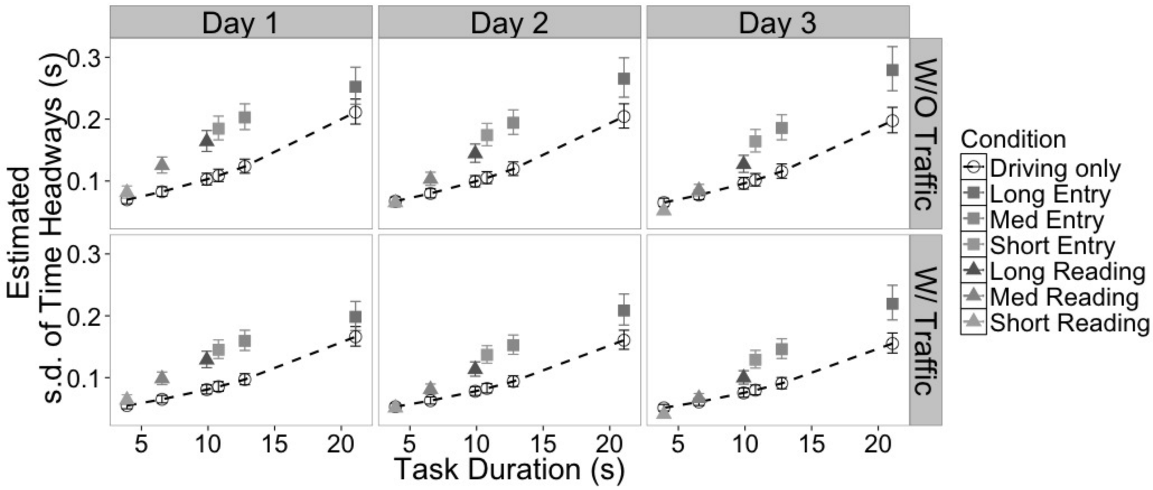

Figure 30. Estimated standard deviation of time headways by traffic and task demands (Experiment 2)

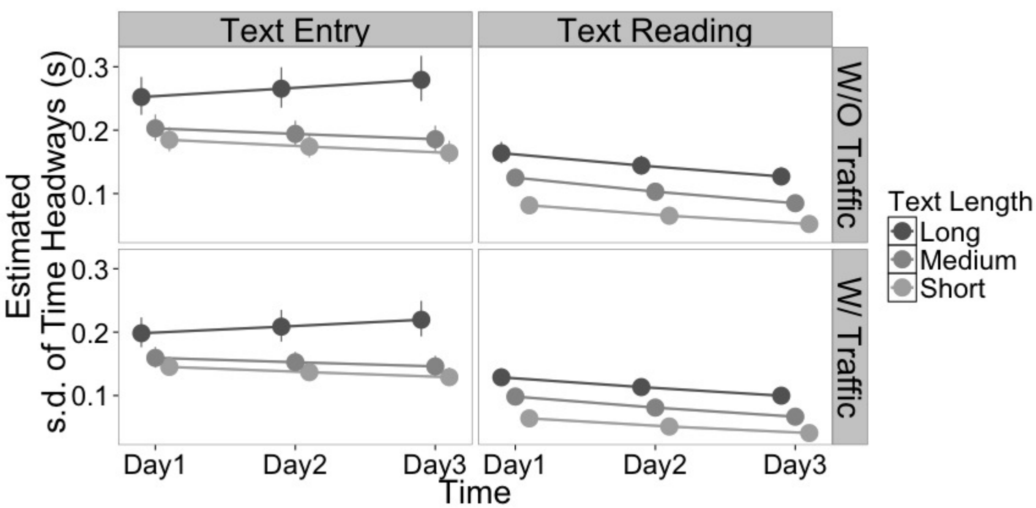

Figure 31. Estimated standard deviation of time headways by time (Experiment 2)

Discussion

The purpose of Experiment 2 was to examine low and high risk drivers' adaptive behavior over time and under different driving demands. For this purpose, a similar driving simulator study with the same IVIS tasks was conducted with 28 young drivers from 18 to 30 years old.

The glance behavior data from both Experiment 1 and Experiment 2 revealed large individual differences, which suggests different driver sub-groups. An examination of individual differences was conducted, but there was a great deal of variation, making it difficult to observe and see any generalizable outcomes. Cluster analysis was considered a reasonable method to classify drivers based on their driving habits, opinions, and behavior. By clustering drivers with similar risk levels together, it makes it easier to understand their common characteristics. The results of cluster analysis suggested that there were two clusters of drivers exist with low and high risk levels, determined by their driving habits and opinions on driving safety. The high risk drivers tend to think that driving fast, aggressively, and disobey traffic lights under certain situations were not as dangerous, and they also slightly more frequently did so. The regression models further showed that the two clusters of drivers had different adaptation on glance behavior. Specifically, the two clusters of drivers had similar maximum EOR time when performing short and medium text reading tasks, but the high risk drivers had significantly longer EOR time when performing more visually demanding tasks compared to low risk drivers. This suggests that low risk drivers were able to better compensate for the increase risk by maintaining relatively short off-road glances, even the task demand increased. However, even though high risk drivers had longer and more proportion of EOR time, they did not have worse vehicle control performance in terms of lateral and longitudinal vehicle control. This may indicate that these high risk drivers were also better at multitasks, and thus they were able to look longer off the road without having worse vehicle control performance.

Interestingly, both high and low risk drivers considered typing addresses and using cell phone that require off-road glances as somewhat or extremely dangerous while driving. Yet, the high risk drivers still had substantially long glances while interacting with IVIS. Atchley, Atwood, and Boulton (2011) found similar results in their survey study, which showed that young drivers' perceived risk of texting has little or no influence on their actual choices on texting while driving.

Cluster analysis is only one way of characterizing individual differences on risk taking. Drivers' risk levels may in fact, be expressed as a continuous measure instead of a discrete one (i.e., high and low), and follow a multi-dimensional normal or log normal distribution. In this case, grouping drivers into sub-groups using cluster analysis may be as arbitrary as cutting the distribution using percentiles. Further studies may measure drivers' perceptions on risks using continuous measures instead of Likert scales, and use as a covariate when modeling driver performance to have a more refined relationship between risk taking and performance.

The results from Experiment 2 also demonstrated that random coefficient models could provide a better estimation on individual performance when large individual differences exist, compared to marginal models such as GEE. When individual differences are large, the population average estimation may underestimate the performance of a few high-risk drivers. Different drivers adapt their behavior in different ways, and the subject-specific estimation can be more accurate when modeling individual driver's performance.

Risk compensation behavior was found when drivers performing IVIS tasks under the "with traffic" driving scenario. When there was ambient traffic presented on the road, both high and low risk drivers tend to have shorter off-road glances and take more proportion of time looking on the road during IVIS tasks. They also had smaller time headway variation under the "with traffic" scenario. This is somewhat consistent with Tsimhoni and Green (2001) study, which showed that drivers had shorter glances off the road when driving on curve road than on straight road. This positive adaptation suggested that drivers realized the increased risk of performing

IVIS tasks when driving demand was higher (i.e. there was ambient traffic around), and thus, tried to compensate this risk by looking more on the road and keep relatively stable headway distances. However, for high risk male drivers, the maximum EOR time could still be as long as 4 seconds when typing a long word when there was traffic, which could be a potential danger for both the driver and the other roadway users.

This experiment also examined driver behavioral adaptation over time, and the results showed that for both high and low risk drivers, their maximum EOR time increased over time when performing text entry tasks but not text reading tasks. In addition, the proportion of EOR time increased over time for both text entry and reading tasks. This suggests a negative behavioral adaptation: when drivers performing more visually demanding tasks such as long text entry, they would have even longer EOR time over time, which potentially further increased their risk on the road. However, it needs to be noted that drivers did not have worse vehicle control performance over time even with the increased off-road glances. In fact, their lateral and longitudinal control ability slightly improved over time. Similar findings on vehicle control performance improvement were also found in previous studies (Chisholm et al., 2008; Shinar et al., 2005). This indicates that after using the IVIS for a period of time, drivers were able to adapt to the tasks so that they could look away longer from the road and pay more attention to IVIS but without sacrificing their vehicle control performance.

However, it also needs to be noted that the appeared better lateral and longitudinal control ability with long glances does not mean safer driving or suggest reduced risks associate with using IVIS. Looking away the roadway for too long may cause the drivers missing sudden events on the road, such as pedestrian crossing, lead vehicle braking, etc. Strayer et al. (2011) showed that drivers were unable to better detect sudden events even after getting more familiar with a cellphone task over time. The authors did not examine drivers' glance behavior over time in the same study. However, this dissertation appears to support their results in that drivers were found to have longer glances off the road over time, which may explain why their reaction time for sudden events could not be improved. With the absence of sudden roadway events, drivers' reaction time, which is an important measure of distraction and driving safety, could not be examined in this dissertation to further validate this hypothesis.

A limitation of this experiment is that each driver only experienced the IVIS tasks in three sessions over one week. The current analysis assumed linear relationship between drivers' glance behavior and vehicle control performance change with time, but the effect of time is likely to be nonlinear over a longer term. It may be interesting in further studies to examine drivers' adaptive behavior over months and even years after using IVISs.

One seemly controversial finding from this experiment is that over time, drivers' off-road glance duration increased but their SDLP and standard deviation of time headways decreased. One possible reason for this is that after practice, drivers were able to control the vehicle better, and also became more familiar with the IVIS tasks. This might help them to gain more confidence on multitasking and thus felt safer to have longer glances off the road. The next chapter will discuss and test this hypothesis by comparing drivers' subjective ratings on their driving performance to their actual vehicle control performance and glance behavior over time.

CHAPTER 5: RELATION BETWEEN DRIVERS' SUBJECTIVE RATING, VEHICLE CONTROL PERFORMANCE, AND GLANCE BEHAVIOR OVER TIME

This chapter examines the relationship among drivers' subject ratings of the change in their performance, and the change in their actual vehicle control performance and glance behavior over time. Correlation studies among the three measurements were performed. It was hypothesized that drivers' subjective ratings on their performance positively correlated with their actual vehicle control performance but negatively correlated with their glance behavior change.

Subjective Ratings on Performance Change

At the end of the third session of the experiment 2 drivers were asked to fill out an end of study questionnaire (Appendix 2) where they were asked to rate their driving performance on Day 3 compared to Day 1. Specifically, they were asked to rate whether performing text entry, text reading, and driving only under both with or without traffic scenarios respectively, were "much worse", "somewhat worse", "the same", "somewhat better", or "much better" comparing to Day 1.

Approximately 90% of the participants considered themselves performed somewhat or much better under the text entry tasks without traffic scenario; 86% drivers thought they performed at least somewhat better under the text entry tasks with traffic scenario (Figure 32). Slightly fewer drivers thought their driving performance was improved when performing text reading tasks (82% for "without traffic" condition, and 75% for "with traffic" condition). Additionally, approximately 36% drivers thought their driving performance was the same under the driving only scenarios.

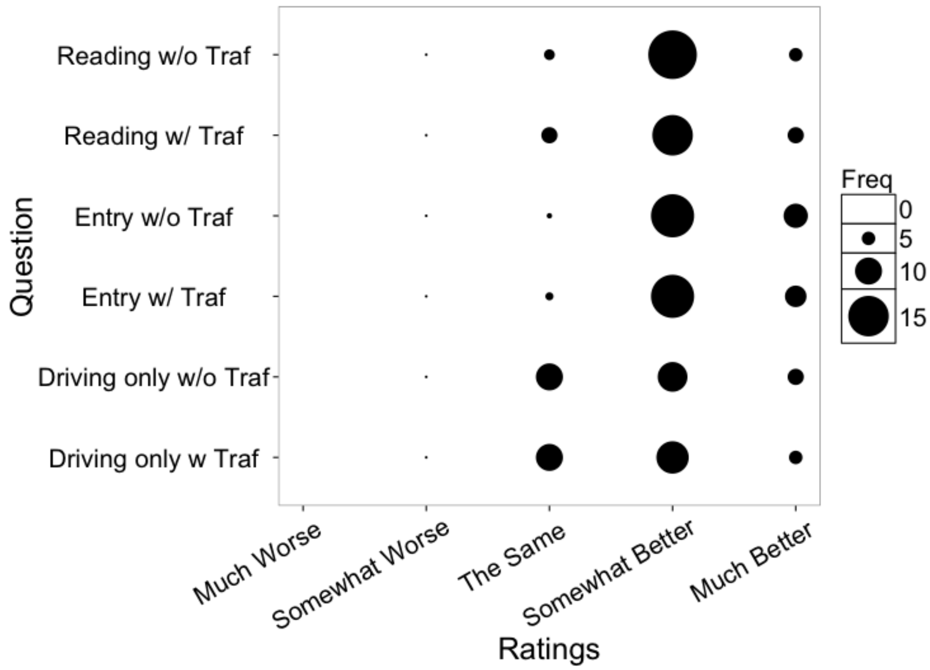

Figure 32. Subjective ratings on driving performance improvement over time

Correlation Study

Changes in drivers actual vehicle control performance and glance behavior from Day 1 to Day 3 were obtained using the random effect estimations for each driver from the random coefficient model in Chapter 4. From the random coefficient models, the slope of change on the outcome variables (i.e., SDLP, standard deviation of time headway, maximum and proportion of EOR time) over time under certain task and driving demand were able to be calculated by extracting the random and fixed effect estimates of the variable time, task demands, and traffic conditions, as well as the intercept. For SDLP and standard deviation of time headway, a negative slope suggests an improved performance over time, while for maximum and proportion of EOR time a positive slope suggests a worse glance behavior over time.

Pearson correlation tests were performed to test the correlation between the slope of change on the vehicle control performance as well as glance measurements and drivers' subjective ratings, under text reading with traffic, text reading without traffic, text entry with traffic, and text entry without traffic conditions (i.e. total four conditions). Given that the slope of change were similar for short, medium, and long text reading or entry tasks, the average of the three estimates under each of the four conditions was used in the correlation tests.

The results showed that the subjective ratings had a moderate negative correlation with the slope of change on SDLP for text entry and text reading “with” traffic, and a small correlation with text reading “without” traffic condition (Figure 33). The Pearson correlations were -0.56 ($p = 0.0024$), -0.45 ($p = 0.0171$), and -0.36 ($p = 0.0574$), respectively. Since a negative slope of change on SDLP suggests improvement on vehicle control performance, these negative correlations actually suggest a positive correlation between drivers’ subjective rating and their actual improvement on vehicle control performance. No correlations were found between the subjective ratings and the slope of change on the standard deviation of time headway.

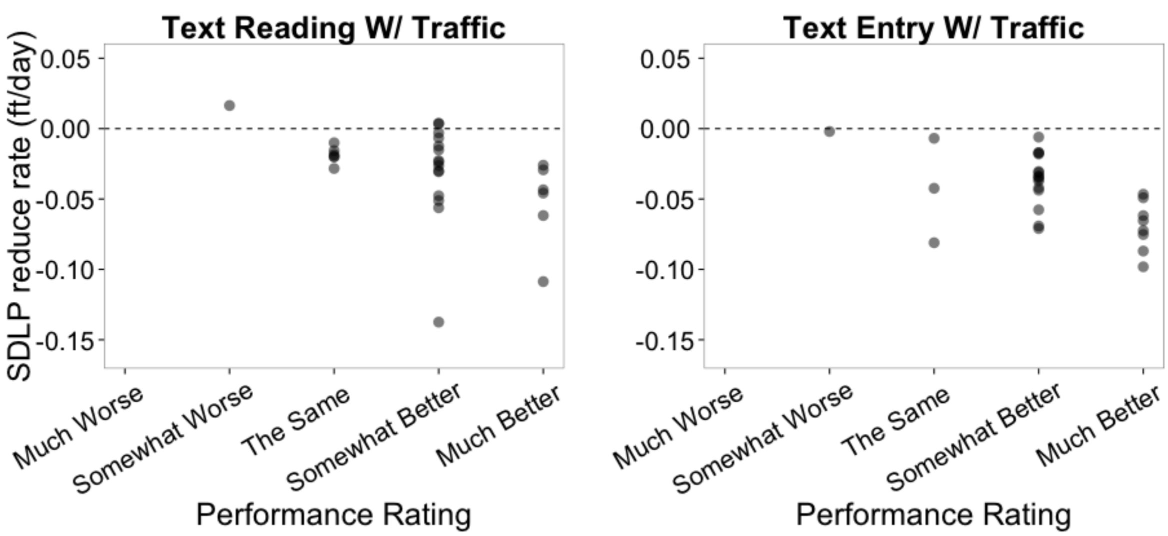

Figure 33. Correlation between the slope of change of SDLP and driver’s subjective ratings

Figure 34 shows the relation between the slope of change on maximum EOR time and drivers’ subjective ratings. For most drivers, their slope of change on the maximum and proportion of EOR time was positive over time, which suggested a worse glance behavior.

However, no significant correlations were found between the slope of change on the maximum and proportion of EOR time and the subjective ratings.

Figure 34. Correlation between the slope of change on maximum EOR time and the subjective ratings

Discussion

This chapter used the data from Experiment 2 and aimed to understand the relationship between drivers' subjective ratings on their driving performance change over time, and their actual vehicle control performance and glance behavior change over time. The results suggested that drivers' perceived performance improvement were somewhat consistent with their actual vehicle control improvement. However, no significant correlations were found between their subjective ratings and glance behavior over time.

Most drivers had significantly improved vehicle control performance (e.g., SDLP), and they also realized they were getting better. However, most of them had longer and worse glance behavior, which was not considered in their self-rating of driving performance. This indicates that drivers' assessment on their performance while using IVIS was based on vehicle control ability (especially lateral control ability), but not glance behavior. Therefore, this finding suggests that drivers may not realize the danger of having long glances off the road.

Another possible explanation of the inconsistence between drivers' self-rating and glance behavior is that there may be biases associated with drivers' judgments. Drivers might not be able to well calibrate their actual performance change, or distinguish the relativity small differences in their glance durations. Future studies may examine drivers' calibration and discrimination abilities on assessing their own performance change (Yaniv, Yates, & Smith, 1991), in order to further understand the relationship between drivers' perceived and actual behavioral adaptation.

CHAPTER 6: GENERAL CONCLUSIONS

This chapter summarizes the overall findings of the dissertation, discusses the contributions of the results to the research field and publications, as well as future research topics that relate to the current research aims.

Overall Findings

The main purpose of this dissertation is to examine drivers' behavioral adaptation related to IVIS use, and understand how drivers with different risk levels change their glance behavior and vehicle control performance due to different IVIS task demands, driving demands, and exposures over time. It was hypothesized that drivers may compensate the risk associated with the increased IVIS task demands and driving demands, but negatively adapt their glance behavior over time. Two simulator experiments were conducted for this purpose, and the key findings are summarized as follow.

1. Drivers tend to have longer glances off the road when the IVIS tasks were more visually demanding, such as typing longer words or reading longer phrases. However their glance durations did not linearly increase with the increased task demand. Rather, the rate of increase of their glance durations slowed down as the task demand increased. This may be considered as drivers' risk compensation strategy for the increased risk of performing highly demanding IVIS tasks.

2. Drivers also tend to compensate for the increased risk of performing IVIS tasks while there is ambient traffic on the road. Specifically, they tend to have shorter and less proportion of off-road glance time, as well as more stabilized headway distances.

3. Although drivers tried to compensate the increased risk by adjusting their glance behavior, their lateral and longitudinal vehicle control ability was still degraded as the task

demand increased. It was also found that the degraded vehicle control performance was mainly associated with the increased task completion time for tasks with higher demands. Thus, it is important to reduce the length of the IVIS tasks so they can be completed in shorter durations.

4. Over time, drivers tend to have longer glances toward the IVIS, which can be considered as a negative behavioral adaptation. However, their lateral and longitudinal vehicle control ability was not affected by the increased off-road glance duration, and even slightly improved. This indicates that drivers were able to better multitask over time and pay more attention to the IVIS without sacrificing their ability to control the vehicle.

5. Drivers exhibit quite large differences on their glance behavior. Two clusters of drivers were identified based on driving habits and safety opinions and classified as high and low risk drivers. High risk drivers were found to have longer off-road glances than low risk drivers when performing IVIS with higher demands, which suggested that they did not compensate risk as much as low risk drivers.

6. Although most drivers had worse glance behavior over time, they did not really realize it and actually considered themselves had improved performance.

Contributions

Intellectual Merit

Previous studies in the literature have mainly focused on driver behavioral adaptation for safety alert systems such as lane departure warning systems, and advanced driving assistant systems such as adaptive cruise control. This dissertation, however, shed lights on behavioral adaptation while interacting with regular IVIS. In particular, drivers do not use the system in the same way they use it on day 1, and there is a negative adaptation associated with glance behavior over time. Although they may have short glances off the road and appear to attend more to the roadway when they first interact with the system due to unfamiliarity, they may have much longer off-road glances as they get more familiar with the system and feel more comfortable to do so. This suggests that driver behavioral adaptation not only occurs for advanced in-vehicle technologies, but also for secondary tasks such as typing and reading using regular in-vehicle devices.

This dissertation also demonstrates how random coefficient modeling can be used for driver behavior studies. Random coefficient model is essentially the same as the linear mixed model or repeated measures ANOVA, which have actually been frequently used in the driving research domain. However, most studies only focus on interpreting the fixed effects and obtaining the "population-average" estimates, without examining the random intercept, slope and it their correlation. This dissertation found that enabling the random slopes in the regression model significantly reduced AIC values and improved model fitting. By using the random coefficient model and extracting the random effects, we can also obtain "subject-specific" driving performance estimates. This can be particularly useful when the driving performance is largely different among drivers and the "population-average" estimates do not well represent individual performance. The "subject-specific" estimates can also be used to predict performance from participants in the experiment, or any individuals who are not in the experiment but share similar characteristics, which can be useful for forecasting and personalized driver education or feedback systems.

Broader Impacts

Drivers' glance behavior is not consistent for different tasks, but rather changes for IVIS tasks with higher demands. This needs to be taken into account when designing IVIS tasks to ensure safe glances. For example, typing words with 12 characters can be too demanding as most drivers had off-road glances longer than 2 seconds while doing so. Additionally, it needs to be noted that different drivers may have dramatically different glance strategies while interacting with IVIS. Although most middle age and older drivers kept their EOR time within a reasonable range, some younger drivers tend to have much longer EOR time when performing long tasks. Therefore, this driver factor may also need to be considered when designing future in-vehicle systems and driving assistant systems. Additionally, when examining glance behavior, a number of studies as well as NHTSA Distraction Guidelines chose to use the mean EOR time as the measurement for glance. However, crashes are rare events and may not be well represented by the mean of the glance distribution (Horrey & Wickens, 2007). Rather, it is the tail of the distribution, with those extreme glances placing the greatest crash risk on the drivers while engaged in an IVIS. This dissertation used the maximum EOR time rather than mean EOR time to model the influence of IVIS tasks, and revealed that extreme glances do exist when drivers interact with IVIS. System designers and safety researchers need to be aware of this extreme behavior when designing and testing new systems. This work has been published in the proceeding of the *7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (Peng, Boyle, Ghazizadeh, & Lee, 2013).

Drivers' adaptive behavior over time suggests that when testing drivers' glance behavior while using IVIS, their experience on IVIS should be considered as they may have longer EOR time after long-term use of the system. The results also suggest that it is important to examine drivers' glance behavior after long-termed use of the IVIS to fully understand the potential distraction and safety risks of the system. The work of the Experiment 2 of the dissertation will be submitted to *Human Factors*.

The benefit of text reading over text entry tasks as well as short over long tasks may largely depend on the task completion time. The results from Experiment 1 suggested that task completion time could be a key factor that affects driving safety while using IVIS, in spite of task types. The magnitude of distraction caused by IVIS tasks appears to reflect only the task length, but not the task type. This study observed that on average, drivers can complete reading tasks that contain up to 140 characters in about 10 seconds, but can only enter 4 letters in that same period using a touchscreen. This implies that reading tasks, which require less time to finish shows better performance than entry tasks. When text entry task is required, a better designed keypad or interface that can allow drivers to enter texts faster may help reduce distractions while driving. The findings relate to the association between driving performance and the IVIS task completion time has been accepted by *Transportation Research Part F* (Peng, Boyle, & Lee, 2014).

IVIS displays will continue to evolve and it is important to recognize its potential safety consequences. The first experiment of this dissertation showed one extreme situation that included very long text reading and text entry on a simple road that posed minimal driving demand. However, it is clear that driving in the real world does pose many other demands from within and outside the vehicle. Previous studies have shown that driving performance under different driving conditions can be very different. Driver distraction from cell phones or e-mail systems have greater impact on drivers in terms of reaction time, speed, and target detections, when in more demanding driving environments (J. D. Lee et al., 2001; Strayer, Drews, & Johnston, 2003; Strayer & Johnston, 2001). The second experiment of this dissertation also found similar findings: driving demands have significant impact on drivers glance behavior and longitudinal vehicle control performance while interacting with IVIS. This suggests that testing the driver distraction of IVIS should consider different driving demands. It may also shed lights on providing endpoints for the whole range of values that encompasses the potential distractions of IVIS under different driving situations.

Future Research

One of the limitations discussed in Chapter 5 is that this study design did not include any sudden events such as lead vehicle braking, and thus driving performance such as reaction time could not be examined. However, reaction time can be important in order to examine the distraction effect of IVIS use, especially when it involved long off-road glances. This study did not find any further degradation on drivers' lateral and longitudinal vehicle control over time although their EOR time increased, but it is possible that their reaction time to sudden events could have been further impacted with the increased EOR time. Thus, future studies should examine this effect to fully understand the potential safety impact on drivers' negative adaptation of glance behavior. Additionally, future studies may validate the results from these simulator studies on the real road, and with longer time period to test the longer-term use of IVIS.

REFERENCES

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis*. Beverly Hills: Sage Publications.

Anttila, V., & Luoma, J. (2005). Surrogate in-vehicle information systems and driver behavior in an urban environment: A field study on the effects of visual and cognitive load. *Transportation Research Part F, 8*, 121-133.

Atchley, P., Atwood, S., & Boulton, A. (2011). The choice to text and drive in younger drivers: Behavior may shape attitude. *Accident Analysis & Prevention, 43*(1), 134-142.

Blanco, M., Biever, W. J., Gallagher, J. P., & Dingus, T. A. (2006). The impact of secondary task cognitive processing demand on driving performance. *Accident Analysis & Prevention, 38*(5), 895-906. doi: http://dx.doi.org/10.1016/j.aap.2006.02.015

Boyle, L., Lee, J. D., Peng, Y., Ghazizadeh, M., Wu, Y., Miller, E., . . . Jenness, J. (2013). NHTSA Visual-Manual Driver Distraction Guidelines: Texting. Washington, DC: National Highway Traffic Safety Administration.

Brock, G., Pihur, V., Datta, S., & Datta, S. (2011). clValid, an R package for cluster validation. *Journal of Statistical Software, 25*(4).

Campbell, J., Carney, C., & Kantowitz, B. (1998). Human Factors Design Guidelines for Advanced Traveler Information Systems (ATIS) and Commercial Vehicle Operations (CVO). Seattle: Battelle Human Factors Transportation Center.

Chiang, D. P., Brooks, A. M., & Weir, D. H. (2004a). An experimental study of destination entry with an example automobile navigation system. *Society of Automotive Engineers Special Publication, SP-1593*.

Chiang, D. P., Brooks, A. M., & Weir, D. H. (2004b). On the highway measures of driver glance behavior with an example automobile navigation system. *Applied Ergonomics, 35*(3), 215-223.

Chisholm, S. L., Caird, J. K., & Lockhart, J. (2008). The effects of practice with MP3 players on driving performance. *Accident Analysis & Prevention, 40*, 704-713.

Deery, H. A. (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research, 30*(4), 225-236.

Dingus, T. A., Antin, J. A., Hulse, M., & Wierwille, W. W. (1989). Attentional demand requirements of an automobile moving-map navigation system. *Transportation Research Part A, 23*(4), 301-315.

Dingus, T. A., Hulse, M. C., Mollenhauer, M. A., Fleischman, R. N., McGehee, D., & Manakkal, N. (1997). Effects of age, system experience, and navigation technique on driving with an advanced traveler information system. *Human Factors, 39*, 177–199.

Donmez, B., Boyle, L., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. *Accident Analysis & Prevention, 39*(3), 581-590.

Donmez, B., Boyle, L., & Lee, J. D. (2008). Mitigating driver distraction with retrospective and concurrent feedback. *Accident Analysis and Prevention, 40*(2), 776-786.

Donmez, B., Boyle, L., & Lee, J. D. (2010). Differences in off-road glances: effects on young drivers' performance. *Journal of Transportation Engineering, 136*(5), 403-410.

Drews, F., Yazdani, H., Godfrey, C., Cooper, J., & Strayer, D. (2009). Text messaging during simulated driving. *Human Factors, 51*(5), 762-770.

Dudek, C. L. (1992). Guidelines on the Selection and Design of Messages for Changeable Message Signs: Federal Highway Administration Report

Dudek, C. L. (1997). NCHRP Synthesis 237: Changeable Message Signs. *Transportation Research Board*.

Ferguson, S. A. (2003). Other high-risk factors for young drivers: how graduated licensing does, doesn't, or could address them. *Journal of Safety Research, 34*(1), 71-77.

Fisher, D. L., Laurie, N. E., Glaser, R., Connerney, K., Pollatsek, A., & Duffy, S. A. (2002). Use of a fixed-base driving simulator to evaluate the effects of experience and PCbased
risk awareness training on drivers' decisions. *Human Factors, 44*(2), 287-302.

Forbes, N. L. (2009). *Behavioral adaptation to in-vehicle navigation systems.* (PhD), University of Nottingham.

Formann, A. K. (1984). *Die Latent-Class-Analyse: Einführung in Theorie und Anwendung*. Weinheim: Beltz.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brownb, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis and Prevention*, 185–191.

Horrey, W., & Wickens, C. (2007). In-vehicle glance duration: distributions, tails, and model of crash risk. *Transportation Research Record, 2018*(1), 22-28. doi: 10.3141/2018-04

Hosking, S., Young, K., & Regan, M. (2006). The effects of text messaging on young novice driver performance: Monash University Accident Research Center.

JAMA. (2004). Japan Automobile Manufacturers Association (JAMA) Guideline for In-Vehicle Display Systems, Version 3.0 *August*.

Jamson, H. A., & Merat, N. (2005). Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. *Transportation Research Part F: Psychology and Behaviour, 8*(2), 79-96.

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

Labiale, G. (1996). Complexity of in-car visual messages and driver's performance. In A.G. Gale et al. (Ed.), *Vision in Vehicles* (pp. 187-194). Bron Cedex, France: INRETS.

Lansdown, T. C. (2009). Individual differences during driver secondary task performance: verbal protocol and visual allocation findings. *Accident Analysis & Prevention, 34*(5), 655-662.

Lee, J., Roberts, S. C., Hoffman, J. D., & Angell, L. (2012). Scrolling and driving: how an MP3 player and its aftermarket controller affect driving performance and visual behavior. *Human Factors, 54*(2), 250-263.

Lee, J. D. (1997). *A functional description of ATIS/CVO systems to accommodate driver needs and limits.* Mahwah, NJ: Lawrence Erlbaum Associates.

Lee, J. D., Caven, B., Haake, S., & Brown, T. (2001). Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors, 43*, 631-640.

Lee, J. D., Regan, M., & Young, K. (2008). What Drives Distraction? Distraction as a Breakdown of Multilevel Control. In M. A. Regan, J. D. Lee & K. L. Young (Eds.), *Driver distraction: Theory, effects, and mitigation* (pp. 31-40). Boca Raton, FL: CRC Press, Taylor & Francis Group.

Maciej, J., & Vollrath, M. (2009). Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accident Analysis & Prevention, 41*(5), 924-930. doi: http://dx.doi.org/10.1016/j.aap.2009.05.007

McGehee, D. V., Raby, M., Carney, C., Lee, J. D., & Reyes, M. L. (2007). Extending parental mentoring using an event-triggered video intervention inrural teen drivers. *Journal of Safety Research, 38*, 215-227.

Merat, N., Lai, F., & Jamson, S. L. (2011). The comparative merits of expert observation, subjective and objective data in determining the effects of in-vehicle information systems on driving performance. *Safety Science, 49*(2), 172-177. doi: http://dx.doi.org/10.1016/j.ssci.2010.07.005

Metz, B. (2009). *Worauf achtet der Fahrer? Steuerung der aufmerksamkeit beim Fahren mit visuellen Nebenaufgaben.* (Ph.D.), Würzburg: Julius-Maximilians Universität.

Metz, B., Schomig, N., & Kruger, H.-P. (2011). Attention during visual secondary tasks in driving: Adaptation to the demands of the driving task. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(5), 369-380. doi: http://dx.doi.org/10.1016/j.trf.2011.04.004

Mooi, E., & Sarstedt, M. (2011). Chapter 9 Cluster Analysis *A Concise Guide to Market Research*. Heidelberg: Springer.

OECD. (1990). Behavioral adaptation to changes in the road transport system. Paris: OECD.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences, 106*(37), 15583-15587. doi: 10.1073/pnas.0903620106

Owens, J. M., McLaughlin, S. B., & Sudweeks, J. (2011). Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis & Prevention, 43*(3), 939-947.

Peng, Y., Boyle, L., Ghazizadeh, M., & Lee, J. D. (2013). *Factors affecting glance behavior when interacting with in-vehicle devices: implications from a simulator study.* Paper presented at the the 7th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Bolton Landing, New York.

Peng, Y., Boyle, L., & Lee, J. D. (2014). Reading, Typing, and Driving: How interactions with in-vehicle systems degrade driving performance. *Transportation Research Part F.*

Ranney, T., Baldwin, S., Parmer, E., Martin, J., & Mazzae, E. (2011). Distraction Effects of Manual Number and Text Entry While Driving. Washington, DC: National Highway Traffic Safety Administration.

Ranney, T., Baldwin, S., Parmer, E., Martin, J., & Mazzae, E. (2012). Distraction Effects of In-Vehicle Tasks Requiring Number and Text Entry Using Auto Alliance's Principal 2.1B Verification Procedure. Washington, D. C.

Romesburg, H. C. (2004). *Cluster analysis for researchers*. Lulu Pr.

Rouzikhah, H., King, M., & Rakotonirainy, A. (2013). Examining the effects of an eco-driving message on driver distraction. *Accident Analysis & Prevention, 50*, 975-983. doi: http://dx.doi.org/10.1016/j.aap.2012.07.024

Sagberg, F., Fosser, S., & Setermo, I. A. F. (1997). An investigation of behavioural adaptation to airbags and antilock brakes among taxi drivers. *Accident Analysis & Prevention, 29*(3), 293-302.

Schieber, F., Holtz, A., B., S., & McCall, R. (2008). *Analysis of Visual Demands of In-Vehicle Text Displays Reveals an Age-Related Increase in Time Needed to Reallocate Attention to the Road.* Paper presented at the Human Factors and Ergonomics Society Annual Meeting, New York City, New York.

Schomig, N., Metz, B., & Kruger, H.-P. (2011). Anticipatory and control processes in the interaction with secondary tasks while driving. *Transportation Research Part F, 14*, 525-538.

Shinar, D., Tractinsky, N., & Compton, R. (2005). Effects of practice, age, and task demands, on interference from a phone task while driving. *Accident Analysis & Prevention, 37*, 315-326.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of Applied Psychology, 9*(1), 23–32.

Strayer, D. L., Jason, M. W., & Drews, F. A. (2011). Cognitive distraction while multitasking in the automobile. *Psychology of Learning and Motivation: Adances in Research and Theory, 54*.

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

Tsimhoni, O., & Green, P. (2001). *Visual demand of driving and the execution of display-intensive in-vehicle tasks.* Paper presented at the Human Factors and Ergonomics Society 45th Annual Meeting, Minneapolis, Minnesota.

Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: Speech recognition versus a touch-screen keyboard. *Human Factors, 46*.

Twisk, J. W. R. (2004). Longitudinal data analysis. A comparison between generalized estimating equations and random coefficient analysis. *European Journal of Epidemiology, 19*, 769-776.

Vaa, T. (2007). Modelling Driver Behaviour on Basis of Emotions and Feelings: Intelligent Transport Systems and Behavioural Adaptations *Modelling Driver Behaviour in Automotive Environments* (pp. 208-232).

Vashitz, G., Shinar, D., & Blum, Y. (2008). In-vehicle information systems to improve traffic safety in road tunnels. *Transportation Research Part F: Traffic Psychology and Behaviour, 11*(1), 61-74. doi: http://dx.doi.org/10.1016/j.trf.2007.07.001

Verbeke, G., Molenberghs, G., & Rizopoulos, D. (2010). Chapter 2 Random Effects Models for Longitudinal Data *Longitudinal Research with Latent Variables* (pp. 38): Springer.

Wakefield, J. (2013). *Bayesian and Frequentist Regression Methods*: Springer.

Ware, J. H., & Laird, N. M. (1982). Random-effecs models for longitudinal data. *Biometrics, 38*(4), 963-974.

Wierwille, W. W., & Tijerina, L. (1998). Modelling the relationship between driver in-vehicle visual demands and accident occurrence. In A. G. Gale, I. D. Brown, C. M. Haslegrave & S. P. Taylor (Eds.), *Vision in Vehicles VI*. North-Holland, Amsterdam.

Wikman, A. S., Nieminen, T., & Summala, H. (1998). Driving experience and time sharing during in-car tasks on roads of different widths. *Ergonomics, 41*(3), 358-372.

Williams, A. F. (2003). Teenage drivers: patterns of risk. *Journal of Safety Research, 34*, 5-15.

Wolfe, J. M., & Pashler, H. (1988). Visual search. *Attention.* London: University College London Press.

Yaniv, I., Yates, J. F., & Smith, J. K. (1991). Measures of discrimination skill in probabilistic judgment. *Psychological Bulletin and Review, 110*(3), 611–617.

Zwahlen, H. T., Adams, C. C. J., & Debald, D. P. (1987). *Safety aspects of CRT thouch panel controls in automobiles.* Paper presented at the The Second Interation Conference on Vision in Vehicles, Nottingham, U.K.

APPENDIX 1: PRE-STUDY DRIVER QUESTIONNAIRE EXPERIMENT TWO

As part of this study, it is useful to collect information describing each participant. The following questions ask about your basic demographic information, health, driving frequency and patterns. Please read each question carefully. If something is unclear, ask the research assistant for help. Your participation is voluntary and you may skip any questions that you do not want to answer.

Demographic Information

1) What is your age? ________
2) What is your gender? ❒ Male ❒ Female
3) Do you wear contact lenses? ❒ Yes ❒ No
4) Do you wear glasses when driving? ❒ Yes ❒ No
5) What is your highest level of education?

❒ Some high school or less	❒ 4-year college degree
❒ High school Diploma	❒ Master's degree
❒ Some college	❒ Professional degree
❒ 2-year college degree/trade school	❒ Doctorate degree

Driving History

6) How old were you when you started to drive (including driving before licensure)? ________
7) At what age did you get your first driver's license (not just a permit)? _________
8) How often do you drive? (Check the most appropriate category)
   - ❒ Less than once weekly
   - ❒ At least once weekly
   - ❒ At least once daily
9) How many miles did you drive last week? ___________
10) How many **moving violations** have you had in the last **3 years**: ________
    a. If **more than zero**, how many were in the **past year**: ________
11) How many **vehicular crashes** have you been involved in during the **last 3 years**: ________
    a. If **more than zero**, how many of them were considered **'your fault'**: ________
12) How many **vehicular crashes** have you been involved in during the **past year**: ________
    a. If **more than zero**, how many of them were considered **'your fault'**: ________

Driving Habits and Opinions

13) While driving, how often do you…

	Never	Rarely	Sometimes	Mostly	Always
a. Wear a seat belt	☐	☐	☐	☐	☐
b. Use your horn	☐	☐	☐	☐	☐
c. Tailgate the car in front of you	☐	☐	☐	☐	☐
d. Use angry or insulting gestures toward other drivers	☐	☐	☐	☐	☐
e. Drive through the intersection when you see a yellow light	☐	☐	☐	☐	☐
f. Drive at least 10 mph over the speed limit	☐	☐	☐	☐	☐
g. Turn without signaling	☐	☐	☐	☐	☐
h. Not come to a full stop at a stop sign	☐	☐	☐	☐	☐

14) How would you classify these driving actions…

Driving Actions	Extremely Dangerous	Somewhat Dangerous	Neither	Somewhat Safe	Extremely Safe
a. Cutting in front of other cars	☐	☐	☐	☐	☐
b. Switching back and forth between lanes to drive through traffic	☐	☐	☐	☐	☐
c. Pass a school bus while it's red lights are flashing	☐	☐	☐	☐	☐
d. Driving the speed limit in adverse weather (e.g., snow, heavy rain)	☐	☐	☐	☐	☐
e. Driving 10 mph over the speed limit	☐	☐	☐	☐	☐
f. Run a red light if nobody is around	☐	☐	☐	☐	☐
g. Drive soon after drinking alcohol or using recreational drugs	☐	☐	☐	☐	☐
h. Disregard the speed limit late at night or early in the morning	☐	☐	☐	☐	☐
i. Using cell phone while driving	☐	☐	☐	☐	☐
j. Typing address on a GPS device while driving	☐	☐	☐	☐	☐

Vision and Hearing

15) What type of prescription glasses or contact lenses will you be wearing for today's study? (Check only one)

- ❒ None (Go to question 15)
- ❒ Single Lens Glasses
- ❒ Bifocals
- ❒ Trifocals
- ❒ Contact Lenses

16) What type of vision problem do you have? (Check only one)

- ❒ Near-sighted - can only see items that are near without glasses
- ❒ Far-sighted - can only see items that are far away without glasses
- ❒ Near and Far sighted - cannot see items that are near or far without glasses
- ❒ Other, explain____

17) Do you have any known hearing problem?

- ❒ No
- ❒ Yes

18) Do you currently use a hearing aid?

- ❒ No
- ❒ Yes

Motion sickness

19) How often do you experience motion sickness? (Circle only one)

0 1 2 3 4 5 6 7 8 9 10

Never Always

20) How severe are your symptoms when you experience motion sickness (Circle only one)

0 1 2 3 4 5 6 7 8 9 10

None Severe

21) Have you taken any medication in the past 48 hours?

	Visit 1	Visit 2	Visit 3
No	❒	❒	❒
Yes	❒	❒	❒
If yes, please list all			

22) Have you consumed any alcohol or other drugs in the past 24 hours?

	Visit 1	Visit 2	Visit 3
No	❒	❒	❒
Yes	❒	❒	❒
If yes, please list all			

APPENDIX 2: END OF STUDY QUESTIONNAIRE EXPERIMENT TWO

The following questions ask about your simulator driving experience during these three visits. Please read carefully and check one and only one box that best expresses your experience.

1) In general, how demanding did you find the following scenarios

	Scenarios	Extremely	Somewhat	Not at all
a.	Performing text ENTRY tasks WITHOUT traffic	☐	☐	☐
b.	Performing text ENTRY tasks WITH traffic	☐	☐	☐
c.	Performing text READING tasks WITHOUT traffic	☐	☐	☐
d.	Performing text READING tasks WITH traffic	☐	☐	☐
e.	Driving only (no tasks) WITHOUT traffic	☐	☐	☐
f.	Driving only (no tasks) WITH traffic	☐	☐	☐

2) Rate your TODAY's driving performance in terms of safety in the following scenarios

	Scenarios	Extremely Dangerous	Somewhat Dangerous	Neither	Somewhat Safe	Extremely Safe
a.	Performing text ENTRY tasks WITHOUT traffic	☐	☐	☐	☐	☐
b.	Performing text ENTRY tasks WITH traffic	☐	☐	☐	☐	☐
c.	Performing text READING tasks WITHOUT traffic	☐	☐	☐	☐	☐
d.	Performing text READING tasks WITH traffic	☐	☐	☐	☐	☐
e.	Driving only (no tasks) WITHOUT traffic	☐	☐	☐	☐	☐
f.	Driving only (no tasks) WITH traffic	☐	☐	☐	☐	☐

3) Rate your driving performance in the following scenarios TODAY as compared to DAY 1…

	Scenarios	Much Worse	Somewhat Worse	The Same	Somewhat Better	Much Better
a.	Performing text ENTRY tasks WITHOUT traffic	☐	☐	☐	☐	☐
b.	Performing text ENTRY tasks WITH traffic	☐	☐	☐	☐	☐
c.	Performing text READING tasks WITHOUT traffic	☐	☐	☐	☐	☐
d.	Performing text READING tasks WITH traffic	☐	☐	☐	☐	☐
e.	Driving only (no tasks) WITHOUT traffic	☐	☐	☐	☐	☐
f.	Driving only (no tasks) WITH traffic	☐	☐	☐	☐	☐