ATLAS CENTER

Advancing Transportation Leadership and Safety

RESEARCH

Investigating the Roles of Touchscreen and Physical Control Interface Characteristics on Driver Distraction and Multitasking Performance

www.ATLAS-Center.org

Investigating the Roles of Touchscreen and Physical Control Interface Characteristics on Driver Distraction and Multitasking Performance

Report: ATLAS-2015-08

Thomas K. Ferris and Youngbo Suh
Texas A&M Engineering Experiment Station (TEES)

and

Jeffrey D. Miles
Texas A&M Transportation Institute (TTI)

University of Michigan
2901 Baxter Rd., Room 124
Ann Arbor, MI 48109-2150

and

Texas A&M Transportation Institute
College Station, TX 77843-3135

January 2016

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ACKNOWLEDGMENT

The authors would especially like to thank Christine Yager, formerly of Texas A&M Transportation Institute (TTI), for her experimental design contributions and assistance in the early phases of this project. We would also like to thank Ko-Ching (Katrina) Wu for her efforts in experimental software development; Sungeun Cho for assisting with experimental hardware development and pilot testing; Austin Fivash, Kyle Kingsbury, Ben Starnes, and Adrian Contreras for managing the experimental reconfiguration activities during data collection; Shiyan Yang for conducting data collection; Sandra Stone and Nada Trout for coordinating recruitment and scheduling of participants; Christie Havemann for her administrative support in purchasing experimental equipment; and several other TTI researchers who volunteered to share resources that were necessary for this project, especially those in TTI's Visibility Research group. Finally, we would like to thank technical monitor Joel Cooper and project coordinators Beth Jakubowski, Barb Lorenz, Robert Wunderlich, and other ATLAS representatives for project management and technical support.

Technical Report Documentation Page

1. Report No. ATLAS-2015-08	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Investigating the Roles of Touchscreen and Physical Control Interface Characteristics on Driver Distraction and Multitasking Performance		5. Report Date January 2016 6. Performing Organization Code
7. Author(s) Thomas K. Ferris, Youngbo Suh, and Jeffrey D. Miles		8. Performing Organization Report No.
9. Performing Organization Name and Address Texas A&M University Texas A&M Transportation Institute College Station, Texas 77843-3135		10. Work Unit no. (TRAIS) 11. Contract or Grant No. DTRT13-G-UTC54
12. Sponsoring Agency Name and Address Advancing Transportation Leadership and Safety (ATLAS) Center 2901 Baxter Rd., Room 124, Ann Arbor, MI 48109-2150 USA		13. Type of Report and Period Covered 14. Sponsoring Agency Code
15. Supplementary Notes Supported by a grant from the U.S. Department of Transportation, OST-R, University Transportation Centers Program		

16. Abstract

This study aimed to assess the potential of driver distraction, task performance, orientation of attention, and perceived workload in a multitasking situation involving interaction with touchscreen interface, compared to physical interface. Authors conducted a real-driving experiment focusing on qualities of synthetic feedback produced from the interfaces mounted on the center console, when a driver engages in an input task with the interfaces while maintaining visual attention and awareness on the road and the roadside. Participants drove a vehicle along a straight double-lane route prescribed on the closed-course track in Texas A&M University–Riverside Campus. The results revealed that the lack of both auditory and vibro-tactile feedback in touchscreen interaction led to significant degradation in input task performance, compared to a natural haptic feedback from the physical interface. The insufficiency also significantly deteriorated drivers' ability in detecting and promptly responding to objects designed to suddenly appear on the roadside, as opposed to the physical interface or touchscreen interface with synthetic feedback. Perceived workload appeared to be not significantly affected by the difference in the interface characteristics. The findings emphasized the impact of synthetic feedback on multitasking performance as previous studies did, and highlighted its potential on regulating visual attention resource on which awareness is based. Designers of user interface or policy makers concerning with driving safety will benefit from the findings.

17. Key Words Touchscreen interaction, Synthetic feedback, Multitasking performance, Driver distraction, Awareness, Visual Attention		18. Distribution Statement Unlimited	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 35	22. Price

TABLE OF CONTENTS

List of Figures ii

List of Tables ii

Introduction 1

Method 4

Participants 4

Apparatus 5

Procedure 10

Measures 12

Results and Discussion 14

Input Task 14

Acknowledgment Task 17

Speed Maintenance Task 18

Perceived Workload 19

Conclusions 20

References 23

Appendix—Data Collection Materials 27

LIST OF FIGURES

Figure 1. Overview of the Experimental Environment. 4

Figure 2. TTI Fleet Vehicle—2006 Toyota Highlander. 5

Figure 3. Placing Traffic Delineation Barrels and the Signage. 5

Figure 4. Two Types of the Signage (Left—Regular One, Right—Rotated One). 6

Figure 5. Comparison of Touchscreen Keypad (TK) (Left) and Physical Keypad (P) (Right). 7

Figure 6. Comparison of Touchscreen Keypad (TK) (Left) and Large Touchscreen Keypad (L) (Right). 7

Figure 7. An Illustration of a Graphical Instruction for the Input Task. 8

Figure 8. All Sequences for the Input Task. 9

Figure 9. A Screenshot from the Recorded Videos. 9

Figure 10. Illustration of Signage Detection in the Acknowledgment Task (from [a] to [d]). 11

Figure 11. Descriptive Statistics in the Input Task—(a) Input Accuracy (%) (Top-Left), (b) Input Efficiency (Count) (Top-Right), (c) Input Time (ms) (Bottom-Left), and (d) Input Awareness Error (Count) (Bottom-Right). Error Bars Represent Standard Deviation, and Red Arrows Represent Significant Difference. 15

Figure 12. Descriptive Statistics in the Acknowledgment Task—(a) Detection Rate (%) (Left) and (b) Response Promptness (%) (Right). Error Bars Represent Standard Deviation, and Red Arrows Represent Significant Difference. 17

Figure 13. Descriptive Statistics in (a) Speed Maintenance (s^2) and (b) Perceived Workload. Error Bars Represent Standard Deviation. 19

LIST OF TABLES

Table 1. Types of Feedback and Interface Settings. 8

Table 2. Order of Interface Settings. 10

INTRODUCTION

Breakthroughs in the development of input and display technologies offer promising new options for vehicle manufacturers in efforts to improve the driver's experience and satisfaction. However, these technologies should be thoroughly evaluated in a driving context to understand how interface design features may also negatively impact driver performance and safety. Otherwise, features that were first designed to support humans may actually result in increased safety risk, for example, by becoming a source of driver distraction (Petzoldt et al., 2014). Vehicle manufacturers employ touchscreens for control of secondary functions (such as adjustment of climate or audio system settings), and one of the pioneering manufacturers prominently features a 17-inch touchscreen interface on a center console with no physical control elements. Touchscreens can support more virtual real estate and a much greater degree of flexibility in display and control than can physical interfaces. A transition from physical controls such as button or knob to touchscreen systems gains the advantage of more expression in the visual qualities of the controls. However, the nature of the haptic (touch) and auditory cues associated with control interactions also changes dramatically. This feedback is inherent to well-designed physical controls; users can feel buttons press and knobs turn, can hear the click of activation, and the rich feedback these cues provide make it easy to interact with them while only minimally requiring visual resources. Touchscreens must recreate the haptic and auditory feedback cues synthetically in the forms of artificial click sounds and vibrations, impoverished forms of feedback compared to that naturally present in the physical controls. Consequently, touchscreen interactions may require more visual resources as compensation for the impoverished haptic/auditory feedback and may require diverting more visual and attentional resources to the touchscreen, and thus away from the driving task. This can lead to greater time with eyes off the road, more safety-critical glances away from the road, and ultimately more serious concerns for driving safety. Considering the potential utility of touchscreen interfaces for drivers, the full extent of positive and negative implications of employing touchscreens must be better understood to emphasize the benefits while minimizing the imposed risks.

In-vehicle secondary tasks are known to cause problems with driver distraction and increased cognitive and/or manual workload, threatening driver safety significantly (Alm and Nilsson, 1997; Briem and Hedman, 1995; Heenan et al., 2014; Kass et al., 2007; Iqbal et al., 2010; Kutila et al., 2007; Petzoldt et al., 2014; Ryu et al., 2013; Sethumadhavan, 2011). Human factors researchers have investigated the potential of synthetic feedback to assist drivers (Donmez et al., 2007; Ratwani et al., 2008; Sethumadhaven, 2011; Van Erp and Van Veen, 2004), generally showing that performance in user interaction while driving is supported by introducing synthetic feedback to existing interfaces. For example, Van Erp and Van Veen (2004) showed how synthetic visual and vibro-tactile feedback could produce a significant reduction in response time for in-vehicle interactions. Few studies have specifically investigated touchscreen interfaces in the driving environment, with fewer exploring the impact of synthetic

feedback on touchscreen interaction (Lee and Spence, 2008; Pitts et al., 2010; Pitts et al., 2012a,b). Among these, Lee and Spence (2008) investigated the full spectrum of synthetic feedback and proposed that the more synthetic feedback could lead to greater reductions in reaction time and subjective workload for a secondary task. Pitts et al. (2012a) also reported that vibro-tactile feedback contributed to a reduced task completion time and perceived task difficulty. Although these previous studies revealed some benefits of synthetic feedback in touchscreen interaction, little is known about the effects of synthetic feedback on drivers' ability to properly attend to the roadway and driving task. With initial evidence that a touchscreen can demand more of a driver's visual attention and exacerbate the negative effects of a secondary task (Crandall and Chapparro, 2012), this study examined the ability to promptly respond to a safety-critical object on the road as a function of interface type (touchscreen or physical interface), control button size, and combinations of synthetic feedback. The findings can suggest how design features of touchscreens (and physical interfaces) affect the orientation of visual attention, as well as the driving and secondary task performance implications.

One way to determine the effects on roadway awareness and driving safety comes from the literature on situation awareness (SA). Endsley (1988) proposed one of the more prominent conceptualizations of SA, defining three SA levels: 1) perception of stimuli (Level 1 SA), 2) overall comprehension of the current state (Level 2 SA), and 3) projection of the state to the future (Level 3 SA). The field of aviation studies employed the conceptualization at first, then studies regarding ground vehicle used it extensively. Measurement of SA while driving usually benefits from a simulator, as the measurement involves "pausing" a simulated world and administering a questionnaire about the current context in the simulated world (Ma and Kaber, 2007). In a real vehicle, it is difficult to measure any of the three levels of SA as the real world cannot as easily be "paused". Nonetheless, the conceptualization of SA is a useful way to assess driver safety, with several sources of evidence linking a failure in developing Level 1 SA directly to increased risk of safety issues (e.g., Newcomb, 2012). Previous studies have been primarily interested in measuring Level 1 SA (Hyman et al., 2010; Nasar et al., 2007; Nasar and Troyer, 2013) by using time and accuracy measures in detecting inserted probes. This method not only represents one of the more "viable options" for measuring SA quantitatively in a real vehicle, but also has been found to be sensitive to secondary tasks and other manipulations common to the driving environment (Jones and Endsley, 2004; Walker et al., 2008).

Most if not all previous driving studies concerned with the performance and safety effects of interface feedback have been conducted in simulator environments (Lee and Spence, 2008; Pitts et al., 2010; Pitts et al., 2012 a, b). The richness of perception in a real vehicle, inclusion of vehicle forces, and the credibility of safety risks are significant factors that warrant study of interface feedback in a real vehicle environment. In contrast, driving simulators enable researchers to assess the negative implications of driver distraction due to secondary task, in safety-related contexts like a car crash, but without a significant increase in physical risk to

participants (Medeiro-Ward et al., 2013). Although operating a real vehicle in an experiment imposes more practical constraints to ensure safety, it provides participants with "an adequate level of realism" that a driving simulator cannot provide (National Highway Traffic Safety Administration (NHTSA), 2013). For example, NHTSA (2013) asserted that force feedback from the steering wheel can substantially affect natural hand motion and control input, and feedback should parallel pedal input in the task of managing vehicle speed. In addition, imperfections in the roadway and engine vibrations introduce low- and high-frequency vibrations that represent natural noise in the driving environment, and make secondary interaction tasks more difficult. Whereas simulated driving is most demanding of visual resources, driving in a real vehicle involves a richer multisensory experience, so conclusions about the effects of multisensory feedback can best be evaluated in a real driving environment.

The assumption tested in this study is that awareness on the road, measured by time and accuracy measures in reacting visual probes in the driving environment, will be affected by the richness of synthetic feedback in a secondary task interface, with the best multitask performance supported overall with a physical interface. Also, the statistical significance would be observed in performance measures, attention orientation, and perceived workload. To some extent, this study is inspired by previous studies that showed how touchscreen-based phones degraded texting-while-driving performance to a greater extent than did phones with physical buttons (Crandall and Chapparro, 2012; Reimer et al., 2014; Young et al., 2014). The current study built on this earlier work by exploring the driving performance and safety effects of various types of synthetic feedback on a touchscreen, comparing to those effects when using a physical interface, and assessing the impact of secondary task feedback on awareness on the road with visual probes that represented safety-critical visual cues.

METHOD

This study involves a real-world driving experiment to examine the impact of synthetic feedback on drivers interacting with in-vehicle controls. As shown in Figure 1, participants drove a fleet vehicle from Texas A&M Transportation Institute (TTI) along the prescribed course as instructed to perform a multitask consisting of 1) an input task with a 3×3 keypad equipped with different types of synthetic feedback and 2) an acknowledgement task toward the signage placed on the right side of the road. A smartphone mounted above the instrument cluster guided the input task, and its screen graphically presented a sequence of inputting. A table touchscreen realized one of the keypads used in this study, designed to produce a set of sensory feedback. The other keypad was a physical numeric one. An acknowledgement task required participants to make a verbal response as soon as they noticed the signage placed intermittently within lines of traffic delineation barrels (TDBs) and obscured such that they were not visible until within a certain distance. In addition to this verbal response, participants had to change lanes or continue in the current lane according to what the acknowledged sign indicated. They tried their best to keep the vehicle speed at 20 mph.

Figure 1. Overview of the Experimental Environment.

Participants

29 drivers (male: 12, female: 17) participated in this study. The average of age was 27.3 (StDev: 4.85), and it had been approximately 11.2 years (StDev: 5.82) since they were firstly licensed to drive. Although three of them had few (3 to 4) years of driving experience, their level of driving experience sufficed to not exhibit a significant effect on peripheral search on the roadside while driving (Crundall et al., 1999; 2002). Only two participants reported that their own vehicle has a touchscreen interface, and five have used a vehicle-mounted touchscreen before. All participants typically drove their own vehicles at least 5 days in a week. Their normal

or corrected visual acuity were at least 20/50 and a separate test confirmed that participants did not have color-anomalous vision. Compensation was $40 for approximately 1.5 hours of participation in the study.

Apparatus

The vehicle used in this experiment was 2006 Toyota Highlander, the base model, shown in Figure 2.

Figure 2. TTI Fleet Vehicle—2006 Toyota Highlander.

Participants drove the vehicle in the closed-course track in Texas A&M University–Riverside Campus, which is a repurposed airport runway. The track was built with paved road tiles. The course was a straight, two-lane road outlined on both sides with rows of TDBs, as shown in Figure 3. Each roadside had forty-four TDBs, and spacing between TDBs was 60 feet with a few exceptions to accommodate cross-road intersections (no other vehicles were allowed on the course during data collection, so the intersections were considered part of the straightaway). In accordance with the prescribed route, participants drove the length of the course, then turn around and double back along the same roadway, which resulted in 5,200 feet of straight driving with periodic lane changing. Considering the 20 mph speed recommendation, participants completed each down and back lap in approximately 3 minutes.

Figure 3. Placing Traffic Delineation Barrels and the Signage.

Figure 3 also shows the signage devised for this study (in the back of the truck). These signs were rotatable incomplete circles printed on transparent plastic boards. Each plastic board was affixed to a delineation post, and the board was rotated such that it appeared in one of the two orientations illustrated in Figure 4. The left orientation (C) signified a need to initiate the lane-change and a verbal response, and the orientation shown on the right (U) required a verbal acknowledgment of the sign only, without a lane change. The height of the post was the same height as the TDBs approximately, so the TBD could obscure the post and the attached sign. The signage came in two sizes: 10-inch and 6-inch diameter shapes. Eight pairs of 10 and 6-inch signs were on each roadside, alternating large and small shapes. Every sign was 5 feet away from the nearest TDB, and the number of TDBs between each sign was randomly chosen but was at least three.

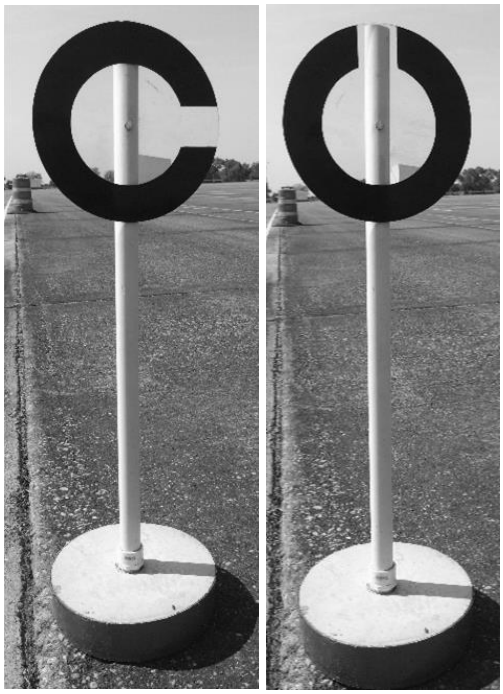


Figure 4. Two Types of the Signage (Left—Regular One, Right—Rotated One).

As the interfaces, the input task used numeric keypads designed for this study. The keypads came in three types; 1) touchscreen keypad (TK), 2) large touchscreen keypad (L), and 3) physical keypad (P). TK and L were an Android application running in Samsung Galaxy Note 10.1 (GT-P7510), and P was Targus Wireless Numeric Keypad (AKP11US) affixed to a hardboard just as large as the size of the tablet, in order for the overall area of all interfaces to be maintained equally, as shown in Figure 5. The appearance looked like a layout of 3×3 button array, in addition to enter and delete button at the bottom. TK was identical to P in size and layout but distinct in interface feedback. Whereas TK could produce several types of synthetic feedback, P exhibited its own haptic feedback by nature. The unnecessary buttons in P were taped and obscured.

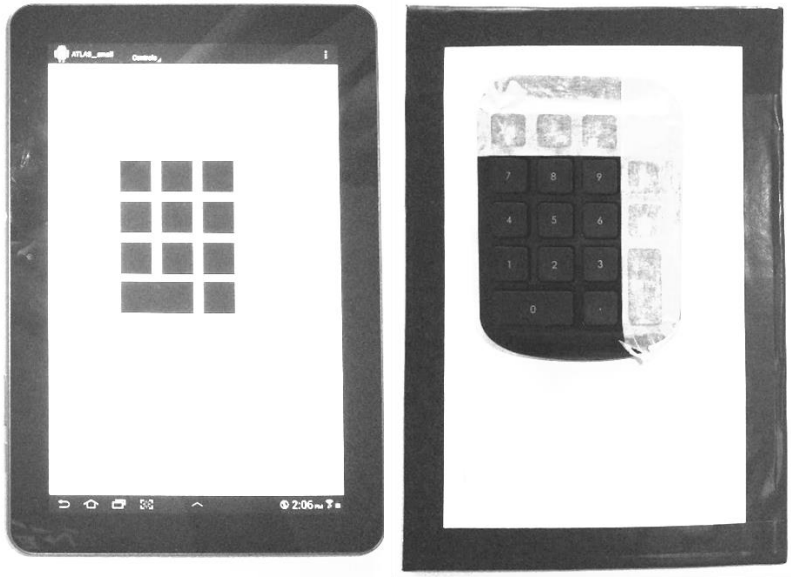


Figure 5. Comparison of Touchscreen Keypad (TK) (Left) and Physical Keypad (P) (Right).

Compared to TK, L produced an equal set of synthetic feedback and provided participants with presumably more convenience due to the larger button size, as shown in Figure 6, in order to find out whether or not the size of the layout made a difference in the collected data.

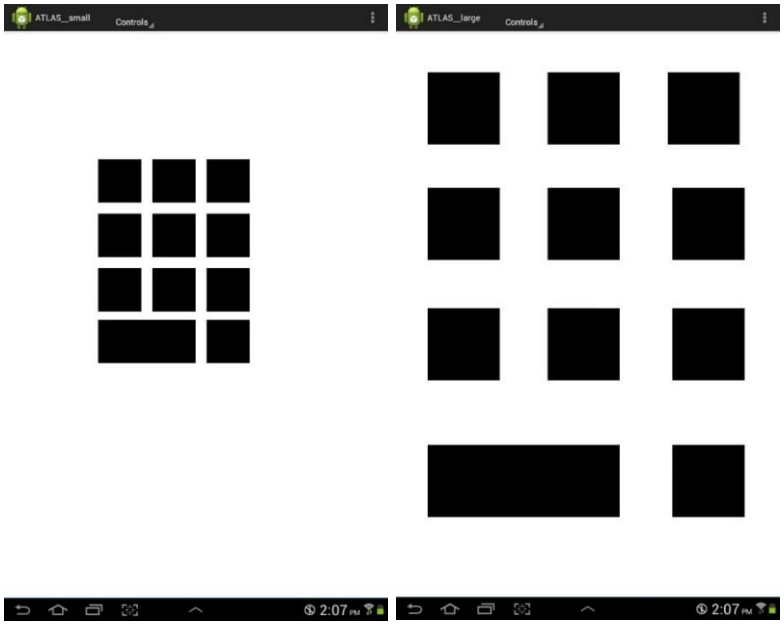


Figure 6. Comparison of Touchscreen Keypad (TK) (Left) and Large Touchscreen Keypad (L) (Right).

TK and L generated synthetic feedback as any button is touched, without delay. Initially equipped with a form of visual feedback, the interfaces could additionally provide auditory and/or vibro-tactile feedback in accordance with experimental settings, in order for participants

to experience multisensory feedback in the touchscreen, in addition to using the physical control interface. As described in Table 1, six types of interface settings were arranged. As visual feedback, participants were able to see the touched black (RGB: #000000) button blinks yellow (RGB: #FFFF00) instantly for one time. Auditory feedback let participants hear a keystroke sound from the tablet's default settings. Vibro-tactile feedback triggered the tablet to vibrate for 500 milliseconds (ms). The intensity of feedback such as screen brightness, volume, and vibration intensity was maximum, in accordance with the tablet's software and hardware constraints. All participants reported that they were able to perceive the three types of feedback in the pre-experiment training.

Table 1. Types of Feedback and Interface Settings.

Category	Type	Description
Interface setting	V	Touchscreen keypad with visual feedback only
	VA	Touchscreen keypad with visual and auditory feedback
	VT	Touchscreen keypad with visual and vibro-tactile feedback
	VAT	Touchscreen keypad with visual, auditory, and vibro-tactile feedback
	L	Large touchscreen keypad with visual, auditory, and vibro-tactile feedback
	P	Physical keypad
Synthetic Feedback	Visual	The color of a touched button instantly blinks yellow from black
	Auditory	A clicking sound from the tablet's default settings is played
	Vibro-tactile	The tablet vibrates for 500 ms

A smartphone, iPhone® 4S, was mounted above the instrument cluster and inside the primary visual area of drivers, so as to present a graphical instruction of how to conduct the input task. As shown in Figure 7, the instruction depicted the required input sequence for button presses in an input task completed on either the touchscreen interface or on a similarly mounted physical button interface. The sequence starts at the diamond and ends at the arrow. In all sequences, the number of buttons to press was 5, and the travel distance of the sequences was 4 vertical/horizontal moves.

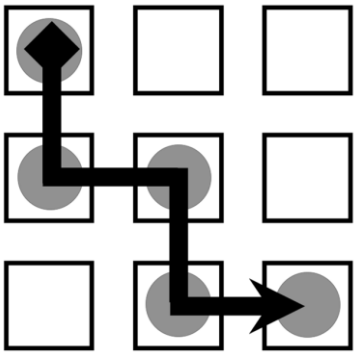


Figure 7. An Illustration of a Graphical Instruction for the Input Task.

Figure 8 shows 90 distinct sequences used in the experiment. Instructed to drive six laps of the prescribed course in total, participants could see 15 distinct sequences in one lap. The smartphone presented each set of 15 sequences repeatedly until finishing the corresponding lap.

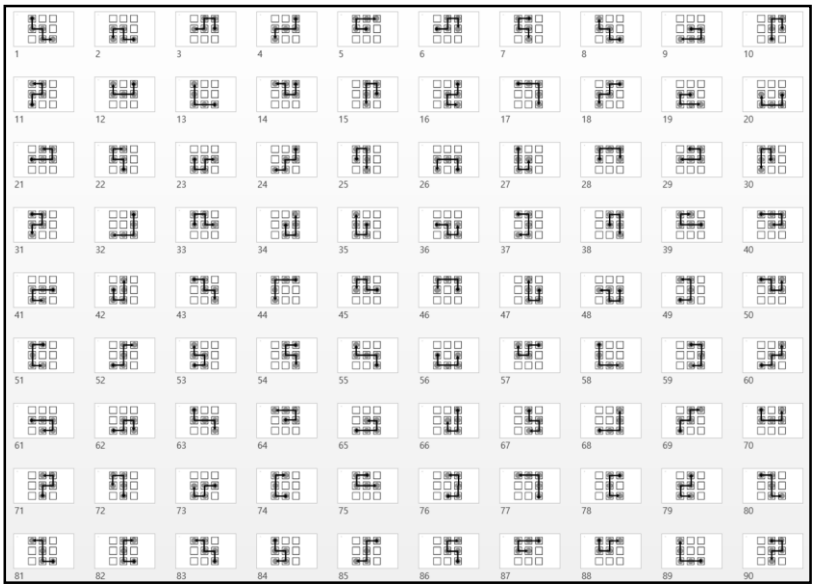


Figure 8. All Sequences for the Input Task.

A Logitech WebCam C920 HD Pro was used for video/audio recording. The camera was hung on the ceiling of the interior of the instrumented vehicle, which enabled us to capture the driver's interactions with the shape display and a view of the course track, as shown in Figure 9. The resolution was 1920×1080 at 30 frame per second (fps), although the observed frame rate was approximately 23 fps. Recorded audio provided confirmation of the verbal response and its timing.

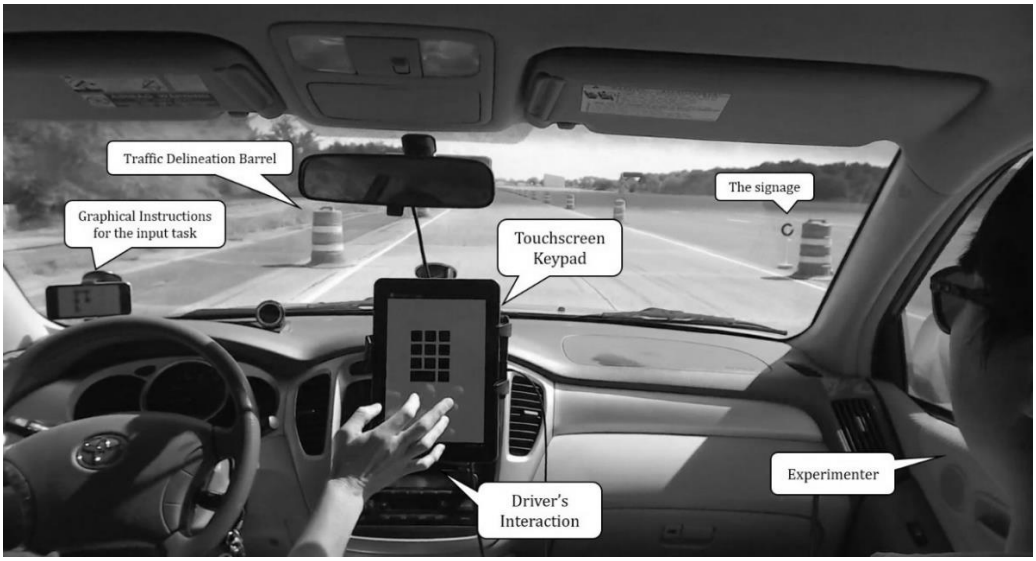


Figure 9. A Screenshot from the Recorded Videos.

Procedure

A training session facilitated the data collection to avoid safety risks due to low-level skills in in-vehicle interaction (Jahn et al., 2009) and to eliminate the learning effect as possible. Participants drove three laps while conducting the experimental tasks—one lap consisted of two lengths, as the route went forth and back on the straight course. They conducted the input task in the first length and the acknowledgment task in the second length, in the first lap. The both tasks were conducted simultaneously in the second and third lap. At the beginning of each length in the latter two laps, the experimenter switched the interface settings including VAT (touchscreen keypad with visual, auditory, and vibro-tactile synthetic feedback), L, and P, as stated in Table 2. During the training, the experimenter notified them every time the vehicle speed went over or below 20 mph, as they needed to be familiar with a sense of the desired speed. In the main experiment, however, the experimenter tried not to interfere with them in regard to speed management, unless the traveling speed introduced a safety risk according to the experimenter's judgment. Participants had to equally prioritize the acknowledgment task, the input task, and speed management, as the experiment advised that the three are equivalently important.

The input task required participants to press, in order, the sequence of five buttons shown in whichever of the patterns from Figure 8 was currently on display on the smartphone. After entering their input on the experimental interface (touchscreen or physical buttons), participants tapped on the smartphone screen, which then advanced to display the next pattern. Participants tried to assure accurate entry of the sequence (they used a delete/clear all button when an input error occurred) and to enter as many correct input sequences as possible during indicated stretches of the experimental course. Six interface settings were prepared for the input task, and each was given in each of six experimental laps. Table 2 shows the initial order of the interface settings, but the experimenter changed the order based on the participant number given when recruited. The first interface setting had the initial order number matching the remainder of the participant number when divided by 6, and the rest of the interface settings were determined sequentially based on the initial order. For example, when the 15th participant came, the participant number became 15, and the remainder should be 3. The first interface to use in experimental lap #1 would then be VT, and the order was VT-VAT-L-P-V-VA.

Table 2. Order of Interface Settings.

Initial Order	Setting
1st	V
2nd	VA
3rd	VT
4th	VAT
5th	L
6th	P

The acknowledgment task required verbally reporting each observed sign placed intermittently between the TDBs and performing the proper lane change response (or properly remaining in the current lane). The experimenter instructed participants to verbalize the response as soon as noticing the signage, as the signage was not visible at distance and become detectable for a very short period of time. As shown in Figure 10 (a), although the vehicle was approaching the TDB, the participant was not able to tell the presence of the signage until sight lines allowed the sign to be unveiled (see Figure 10 [b]) and become detectable (see Figure 10 [c]). When traveling at 20 mph, each sign was visible for less than two seconds (see Figure 10 [d]). When participants observed the C-shape orientation (left side of Figure 4), they would then perform a lane change (change from right lane to left lane, or vice versa). Participants learned a few suggested ways to refer to the signs (C or U; change or unchanged/no change) for the verbal response, but they could use their own words as long as the response was accurate and speeded. After each lap, experimental assistants trailed behind the experimental vehicle and randomly reconfigured the signage to minimize any advantage by remembering orientations of signs from previous runs.

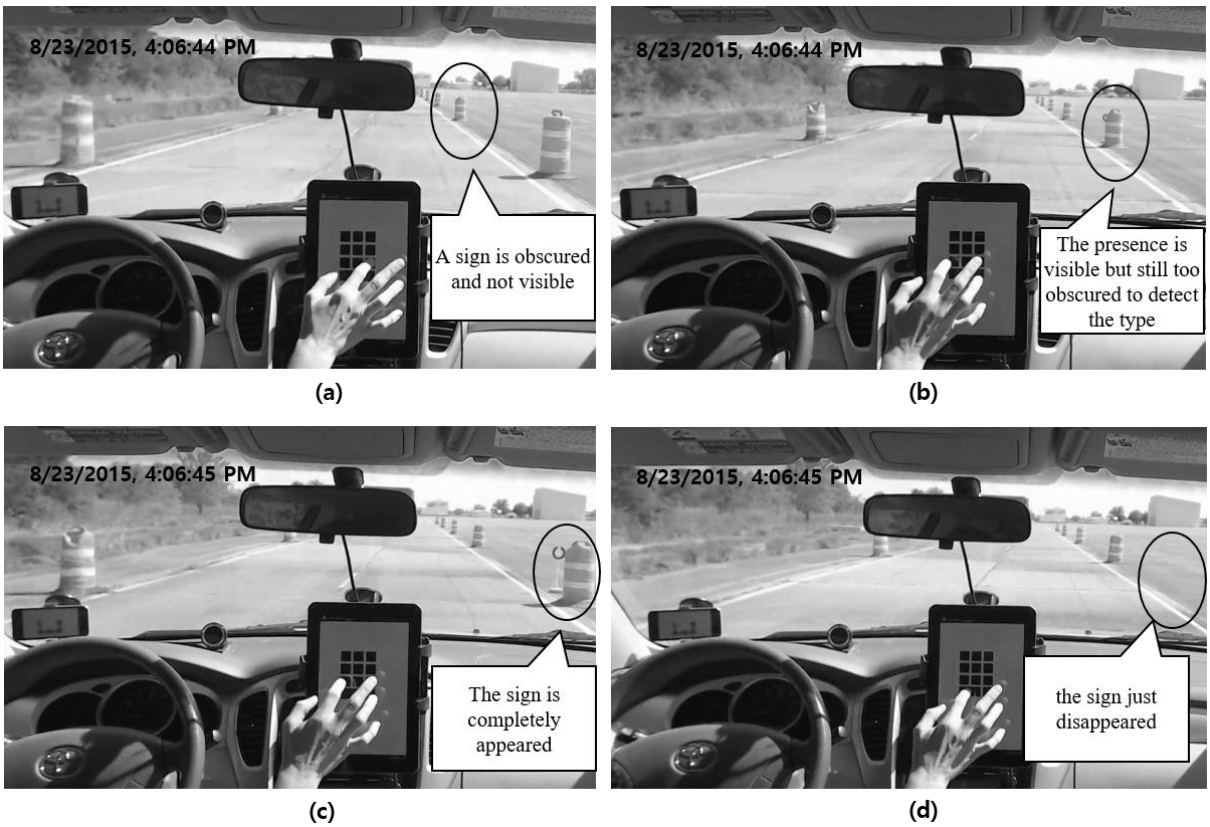


Figure 10. Illustration of Signage Detection in the Acknowledgment Task (from [a] to [d]).

As another indication of driving SA, participants conducted the task of maintaining vehicle speed at 20 mph throughout the driving. Before entering each length of the course, they had about 200 feet to reach the desired speed. They did not receive any notification to slow down to

20 mph even when over, unless the experimenter judged that safety risks were developing due to the speed violation.

Measures

Measures for each experimental task facilitated data collection that involves the interface settings as the primary within-subject independent measure.

Input Task

Four variables accounted for performance in the input task:

1. Input accuracy (%) was defined as the percentage of correctly entered sequences over all entered sequences in one lap. Deleted inputs were not included.
2. Input efficiency was defined as the quantity of all correctly entered sequences in one lap. Participants could continue to renew the sequences and to input them as long as driving.
3. Input time (ms) was defined as the time in average to enter one correct sequence in one lap. The Android application that hosted the input task acquired the timestamps in ms for inputting the 3×3 button patterns as well as enter and delete.
4. Input awareness error was defined as the quantity of entered sequences with more or less than 5 inputs in one lap. If a participant entered a sequence after inputting insufficient or excessive numbers of button, it would imply that he/she was not fully aware of what he/she just inputted at that moment.

Acknowledgment Task

Three variables accounted for performance in the acknowledgment task but the statistical analysis used only two of them:

1. Detection rate (%) was defined as the percentage of correctly initiated responses over all signs in one lap. The confirmation of the response was aided by the recorded videos and audios.
2. Response time (RT) was defined as a count of the number of frames in the recorded videos from the moment that the verbal response was made to the moment that the barrel on the same side as the vehicle but most near to the acknowledged signage disappeared in the windshield. RT needed to be scaled, as the obscured object placed on the right became visible earlier when driving in the left lane than in the right lane. To handle this difference, response promptness, a scaled measure based on the maximum and the minimum RT to the signage was proposed as below.
3. Response promptness (%) was defined as below, which assessed performance in temporal availability to respond, compared to the largest possible durations of availability in the experimental setting.

$$\text{Response Promptness (\%)} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \frac{\mathrm{RT_i} - \mathrm{MIN.\,RT_i}}{\mathrm{MAX.\,RT_i} - \mathrm{MIN.\,RT_i}} \times 100$$

N: Total number of signs to be acknowledged, $\mathrm{RT_i}$: response time to sign i of the participant
$\mathrm{MAX.\,RT_i}$: response time to sign i, observed as maximum in all participants
$\mathrm{MIN.\,RT_i}$: response time to sign i, observed as minimum in all participants

Speed Maintenance Task

Two variables accounted for performance in the speed maintenance task but the statistical analysis used only one of them:

1. Lap time was defined as the time (in seconds) in the recorded video measured from the moment that the first barrel passes the windshield to the moment that the last barrel does the same in each half lap. Lap time developed speed maintenance as below.
2. Speed maintenance (s^2) was defined as below and assessed the time difference between the ideal lap time (s) and the actual lap time (s). The ideal lap time was 89 seconds, considering that the vehicle speed was instructed to be 20 mph, and that each length of the course was approximately 2,600 feet. Squaring was introduced to highlight the deviation.

$$\text{Speed Maintenance } (s^2) = (89 - LT_1)^2 + (89 - LT_2)^2$$

LT_1: Lap time(s) in the first half of one lap, LT_2: Lap time(s) in the second half of one lap

Perceived Workload

After completion of each experimental lap, participants received a questionnaire form to record their subjective ratings of perceived workload following the method specified for calculating NASA-Task Load Indices (Hart and Staveland, 1988). Six ratings scales assessed mental/physical/temporal demand, performance, effort, and frustration following each multitasking interface condition listed in Table 2. The ratings were then averaged for each participant in each interface condition.

RESULTS AND DISCUSSION

The collected data were statistically analyzed with the interface settings as the primary within-subject independent measure. The analyses produced models based on analysis of variance (ANOVA), involving the measures as response variable and verified their variance homogeneity by Bartlett's test. When an ANOVA model indicated statistical significance, multiple comparison was performed. Some of the models did not exhibit normality of residuals. Despite the violation of the normality assumption, the false positive rate would not be affected by the violation because both variance homogeneity and treatment independency were satisfied and the sample size was large enough (Harwell et al., 1992; Lix et al., 1996). Bartlett's test verified variance homogeneity for those models with residuals normality, followed by Tukey's test for multiple comparison. When the normality was not found, Levene's test was used to test for the variance homogeneity, and multiple comparison was adjusted by Bonferroni method that is known to be more conservative than Tukey's test and preferable in controlling for type I error (Maxwell, 1980). Although the collected data had 29 replications, a few of them were disqualified in some of the models due to a technical problem, which will be pointed out in the analyses of the acknowledgment task and the speed maintenance task. All statistical analyses used a significance level of alpha = 0.05.

In addition to the quantitative measures, participants commented their overall impression of the tasks with the interface settings after the completion of the experiment. Their comments also involved whether or not they could recognize the differences among the interface settings in supporting the tasks. Most participants reported that the tasks were demanding and they were too busy to think or do anything other than the tasks. Many if not most expressed that the P (physical keyboard) conditions were better for supporting multitask performance and driving safety, compared to the various touchscreen conditions. Often this was because participants could rest/anchor their hands on P and still have the tactile sensation of the keys. Participants also commonly mentioned that having synthetic feedback when using the touchscreen was helpful, and felt less confident about their performance when the touchscreen did not provide them with auditory and/or vibro-tactile feedback. Their preferences regarding the interface settings were varied. Some found it more annoying when they were not in favor of the selected settings, though no evidence claimed that task performance related to such annoyance.

Input Task

Figure 11 shows the descriptive statistics of each measure in the input task. Each bar depicts mean values, and error bars represent standard deviation. Developing the ANOVA models, preliminary analysis revealed a deviation from normality in every model from the input task. Nonetheless, Levene's test verified variance homogeneity from the models, and the recorded videos showed that the treatments were applied independently. 29 replications were included.

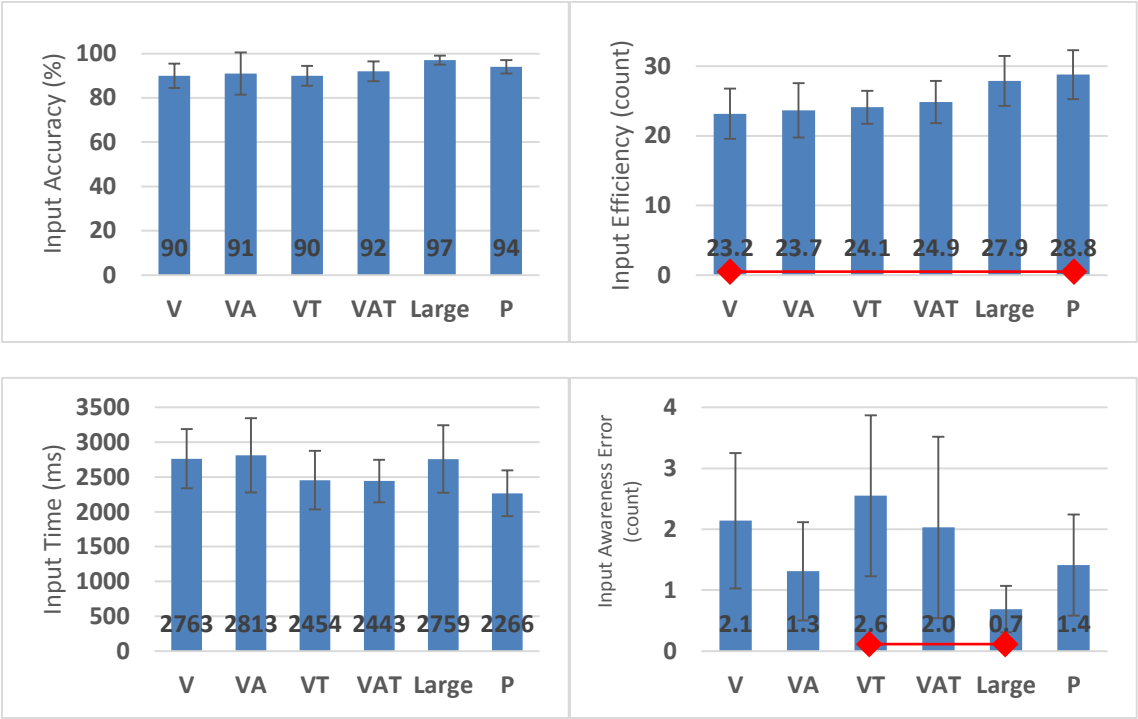


Figure 11. Descriptive Statistics in the Input Task—(a) Input Accuracy (%) (Top-Left), (b) Input Efficiency (Count) (Top-Right), (c) Input Time (ms) (Bottom-Left), and (d) Input Awareness Error (Count) (Bottom-Right). Error Bars Represent Standard Deviation, and Red Diamond Arrows Represent Significant Difference.

Input Accuracy

For input accuracy, although the difference in means between the worst, V (touchscreen keypad with visual feedback only), and the best, L (large touchscreen keypad with visual, auditory, and vibro-tactile feedback), was 7 percent; the ANOVA model did not show significance in the interface settings. Levene's test identified variance homogeneity ($F_{5, 168}$ = 1.3154, p-value = 0.2598). An alternative analysis regarding a count of the number of delete button presses was also conducted to find a different angle on input accuracy, as delete could mean an occasion where participants made a mistake or at least they thought they did so. However, the number of delete presses did not show a significance in a statistical sense. Indeed, the frequency with which participants entered delete to deal with input errors seemed to correspond more so with perturbations caused by the bumpiness of the road than the interface settings. As the worst performance reached 90 percent accuracy, as shown in Figure 11(a), this implies that interface characteristics did not substantially (or significantly) impacted input accuracy, as the task was relatively simple and self-paced. The task may suffer from a ceiling effect that did not allow much room to improve the accuracy with richer feedback. Potentially the 7 percent difference between best and worst accuracies could approach significance with a more difficult or visually demanding input task.

Input Efficiency and Input Time

In input efficiency, the ANOVA model showed the significance ($F_{5, 168} = 3.5429$, p-value = 0.0045), and Levene's test identified variance homogeneity ($F_{5, 168} = 0.7243$, p-value = 0.6061). The lack of the normality led to conduct multiple comparisons with Bonferroni adjustments, which revealed a significant difference between P and V (p-value = 0.0270). Differences between other pairs were not significant, suggesting that not having auditory or vibro-tactile feedback in touchscreen can degrade input efficiency, compared to the physical interface. By definition, degradation in input efficiency means that more time will be needed for the same number of sequences to be input. In input time, the ANOVA model did not show the significance (Levene's test identified variance homogeneity [$F_{5, 168} = 1.8274$, p-value = 0.1100]), and yet Figure 11 (c) shows V required more time than P by a half second on average to enter one input sequence. The time difference is noteworthy, as a half second with a secondary task equates to more than 50 feet of travel distance at normal highway speeds (75 mph), which is a significant distance to cover under distracting conditions. Also, the difference could scale upward when the secondary task becomes more complex or complicated than the artificial one used in this study.

Input Awareness Error

Figure 11 (d) shows that L achieved the lowest input awareness error. The ANOVA model showed the significance in input awareness error ($F_{5, 168} = 2.9978$, p-value = 0.0133), and the Bonferroni method identified the pair of L and VT (touchscreen keypad with visual and vibro-tactile feedback) as significant (p-value = 0.0140). A speculation that the button size is a more prominent contributor to input awareness error was possible, considering that the significance was not found between the interface settings with the same-sized buttons.

One possible reason for the best performance with L is that the buttons were spatially more distinct and easier to accurately contact the intended button. Another reason came from the recorded videos showing that the size was large enough to force drivers to use their forearm in the finger interaction whereas other interface settings did not necessarily require participants to do so. The greater manual task load in these cases may have led to an increase in cognitive resources invested and better performance according to the Yerkes-Dodson Law, which explains how simple tasks can show improved performance with an increase in task load (Yerkes and Dodson, 1908). Understanding why VT, not V, exhibited the worst performance in input awareness error needed further explanation, as VT had not only the same visual feedback as V but also a vibratory cue that was assumed to help. The recorded video enlightened us in some extent, showing some participants inputting each button faster than the duration of the vibration (= 0.5 second). According to the follow-up interviews, they experienced that "the tablet just kept vibrating" throughout entering each input sequence, so a vibration presented for durations of this magnitude may not be as helpful in supporting nonvisual awareness of button input, and may actually hinder awareness when visual cues alone are sufficient. However, participants who felt

less confident in using V that did not include auditory or vibro-tactile feedback glanced at the interface more frequently. The increased visual attention on the input task may have naturally supported more speeded performance than the conditions with richer multisensory feedback, perhaps at the cost of the acknowledgment and speed management tasks. The statistical analysis in the acknowledgment task supported this further in the next section, as the task necessitated constant visual attention on the road and the worst performance appeared in V.

Acknowledgment Task

Figure 12 shows the descriptive statistics of each measure in the acknowledgment task. Each bar depicts the mean of each measure, and error bars represent standard deviation. The data from the acknowledgment task produced two ANOVA models. Although one showed a violation of the normality assumption, both models exhibited variance homogeneity. The recorded videos suggested identified treatment independency. Due to a technical problem, the quality of several recorded scenes were too poor to confirm the verbal responses of a few participants. These data were removed from analyses, resulting in 27 participants' worth of data for the model of detection rate and 25 for the model of response promptness.

The ANOVA model showed significance of the interface settings ($F_{5, 156} = 6.7049$, p-value = 0.0001) on detection rate. The normality violation was found, and Levene's test verified variance homogeneity ($F_{5, 156} = 0.8186$, p-value = 0.5382). As shown in Figure 12 (a), V showed the lowest detection rate, 85 percent. With the interface settings involving vibro-tactile feedback (VT and VAT), physical keyboard (P) or with larger buttons (L), the richer synthetic feedback caused detection rate to go up significantly, to as high as 95 percent. Multiple comparisons by Bonferroni adjustment found the significance in several pairs including L vs. V, P vs. V, VAT vs. V, and VT vs. V (p-value = 0.0021, 0.0004, 0.0001, and 0.0002, respectively). All interface settings except for VA showed significantly higher detection rate than V.

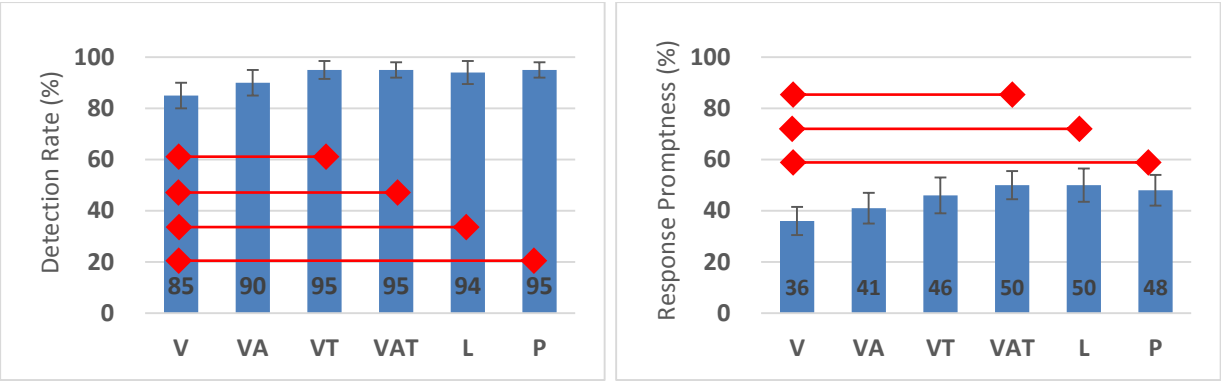


Figure 12. Descriptive Statistics in the Acknowledgment Task—(a) Detection Rate (%) (Left) and (b) Response Promptness (%) (Right). Error Bars Represent Standard Deviation, and Red Diamond Arrows Represent Significant Difference.

As Figure 12 (a) and Figure 12 (b) showed a similar tendency, the statistical analysis for response promptness found a consistent result. The ANOVA model in response promptness showed the normality of residuals by Shapiro-Wilk test (W = 0.9862, p-value = 0.1417) and variance homogeneity by Bartlett's test (χ^2 = 2.4212, df = 5, p-value = 0.7883). The impact of the interface settings on response promptness showed statistical significance ($F_{5, 144}$ = 5.2200, p-value = 0.0002), and Tukey's test revealed significant differences between the treatments such as L vs. V, P vs. V, and VAT vs. V (p-value = 0.0011, 0.0040, and 0.0013, respectively).

Considering that good performance in the acknowledgment task required more visual attention be paid to the roadside, richer nonvisual synthetic feedback may be critical for supporting visual awareness of the roadway while secondary in-vehicle interaction tasks are performed. Also noteworthy, performance with regard to detection rate and response promptness measures did not get *worse* with rich-feedback touchscreen interfaces compared to the physical button interface, which provides arguably the richest forms of feedback. This finding is one of the most important of this study and represents the most impactful contribution that can be used to inform guidelines for in-vehicle interface design. NHTSA is collecting inputs from manufacturers and researchers and planned to release interface guidelines in a three-phase process. They published the Phase I Guideline in 2013. The Phase I Guideline detailed very comprehensive concerns in driver distraction due to in-vehicle electronic devices, including those that employ touchscreen interfaces, and also offered practical suggestions to deal with these concerns. However the specific roles of synthetic feedback types are not discussed for their impact on distraction potential, driving safety, and performance. The findings of this study suggest a reason for modifications. For example, the Phase I Guideline pointed out that "more than six button or key presses during a single task" could disturb driving safety, whereas this study revealed how a six-keypress sequence negatively affected drivers' visual attention allocation as a function of feedback configuration.

Speed Maintenance Task

Figure 13 (a) shows speed maintenance as a function of interface settings. No significant statistical affect was found, Levene's test verified variance homogeneity ($F_{5, 156}$ = 0.1656, p-value = 0.9748). Considering the interface settings affected the above visually demanding tasks, the insignificance in the speed maintenance task implied that maintaining the 20 mph could be supported by not only seeing the speedometer or the outside but also the multisensory feedback provided by the realistic driving setting. As the experiment employed a real vehicle, participants felt true acceleration and deceleration forces. The real in-vehicle interface also enabled them to be given force feedback from the acceleration pedal to regulate the speed. In addition, since the course track was built with equally sized road tiles, the roadway produced regular bumps that drivers could feel and hear. The frequency of the bump sounds provided as auditory display enabling them to perceive when the vehicle sped up or slowed down. This is just one example of how the richer multisensory feedback of the real driving environment can change the demand on

visual resources compared to in simulator settings. Because competition for the engaged sensory channels and other information processing resources is one of the primary factors affecting human abilities to multitask (Wickens, 2002), this may illustrate the extent to which feedback variables may differentially affect performance in simulated and real environments.

Perceived Workload

The ANOVA model for perceived workload suggested the interface settings did not showed a significance. Shapiro-Wilk test verified residual normality (W = 0.9925, p-value = 0.5054), and Bartlett's test verified variance homogeneity (χ^2 = 5.5134, df = 5, p-value = 0.3565). It is possible that synthetic feedback was not effective in lowering workload or that the experimental tasks were not complex enough to be eased by synthetic feedback. The latter is promising, given that participants' average age was 27.3 and had experience in driving for over 11 years in average. That is to say, the participant demographics included mostly younger drivers (so no age-related perceptual or cognitive deficits) with considerable familiarity and experience in regulating driving behaviors in a multitasking situation. Still, further researches are needed if synthetic feedback could show the significance in perceived workload when a multitasked driver is unskilled or older. However in Figure 13(b), the highest mean report of subjective workload was shown in V, which at least partially mirrors expectations.

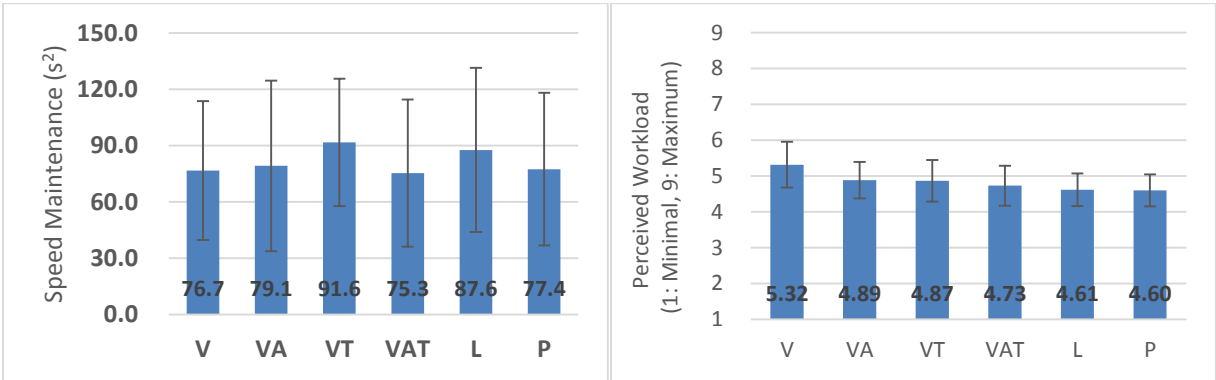


Figure 13. Descriptive Statistics in (a) Speed Maintenance (s^2) and (b) Perceived Workload. Error Bars Represent Standard Deviation.

CONCLUSIONS

This study involved a closed-course driving experiment in a real vehicle to investigate the impact of touchscreen and physical interface characteristics on driver distraction and multitasking performance in a rich and representative multisensory environment. The distracting secondary input task required drivers to orient visual attention frequently, and multitasking performance involved visual-manual interaction, probe detection, and vehicle control. A touchscreen interface does not support *better* performance than a physical button interface, however with richer synthetic feedback through visual, auditory, and vibro-tactile cues, a similar level of multitask performance and safety can be achieved. Particularly, the addition of vibro-tactile feedback seemed to make a larger positive impact than auditory feedback, and the combination of visual and vibro-tactile cues (sometimes without auditory cues) approached levels of performance found with physical keyboards. In addition, the size of the touchscreen control elements appeared to affect the measures of interest, but further research is needed to determine any potential interaction effects between button size and feedback richness.

The results revealed the significant effects of the interface characteristics across multiple measures of performance and safety. For input efficiency, a significant difference was found between the touchscreen and the physical interface *unless* the touchscreen interaction supported both auditory and vibro-tactile feedback. Performance in input awareness error was harder to draw conclusions from, but larger-sized buttons on an interface, requiring larger movements and more visual resources for spatial orientation, would enhance performance in specifically the device-interaction task, but perhaps at the cost of visual resources on the roadway. Additionally, the timing and duration of nonvisual (auditory and/or vibro-tactile) cues may warrant additional investigation. The 0.5-second duration of vibro-tactile feedback cues may have misled participants in discerning between inputs and was not necessarily helpful in supporting all tasks. Detection rate and response promptness were the most indicative measures of driver attention and the extent to which a driver may be distracted in a multitasking situation. Both measures consistently exhibited how performance deteriorated when the input task used a touchscreen without both auditory and vibro-tactile feedback. The interface settings did not significantly affect performance in speed maintenance, which may indicate that at the slow vehicle speeds in this study it is an insensitive measure. Or it may partially reflect the performance benefit drivers experience in the real driving environment, with rich multisensory feedback including force feedback, g-force, and auditory noise. The result can be informative as an illustration that captured the effect of naturalistic factors when employing a real vehicle instead of a driving simulator.

Driving is frequently a multitasking operation, and technological advances are available to support better and safer in-vehicle secondary task interaction. In addition to exploring these benefits, it is critical to understand the implications of these tasks for attention orientation, especially with regard for the demand on visual resources. The findings of this study suggest that

touchscreen interfaces with richer synthetic feedback can support significantly more effective and safer multitasking while driving by minimizing the negative implications of distracted driving. The findings also demonstrated that synthetic feedback could alleviate the performance deterioration due to the transition from physical button interfaces to more advanced touchscreens, as is the apparent trend in vehicle manufacturing. Further research may explore more dimensions of feedback, such as the duration and intensity, as well as more complex secondary interaction tasks that can further improve the realism of the experimental setting to more closely match that of the everyday driving experience.

REFERENCES

Alm, H. and L. Nilsson (1995). “The effects of a mobile telephone task on driver behaviour in a car following situation.” *Accident Analysis and Prevention*, 27(5) 707–715.

Briem, V. and L. R. Hedman (1995). Behavioural effects of mobile telephone use during simulated driving. *Ergonomics*, 38(12), 2536–2562.

Crandall, J. M. and Chaparro, A. (2012). “Driver distraction: effects of text entry methods on driving performance.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1693–1697.

Crundall, D., Underwood, G., and Chapman, P. (1999). “Driving experience and the functional field of view.” *Perception*, 28(9), 1075–1087.

Crundall, D., Underwood, G., and Chapman, P. (2002). “Attending to the peripheral world while driving.” *Applied Cognitive Psychology*, 16(4), 459–475.

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

Endsley, M. R. (1988). “Design and Evaluation for Situation Awareness Enhancement.” *Proceedings of the Human Factors Society Annual Meeting*, 97–101.

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

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., and Olds, C. C. (1992). “Summarizing Monte Carlo results in methodological research: The one-and two-factor fixed effects ANOVA cases.” *Journal of Educational and Behavioral Statistics*, 17(4), 315–339.

Heenan, A., Herdman, C. M., Brown, M. S., and Robert, N. (2014). “Effects of Conversation on Situation Awareness and Working Memory in Simulated Driving.” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56(6), 1077–1092.

Hyman, I. E., Boss, S. M., Wise, B. M., McKenzie, K. E., and Caggiano, J. M. (2010). “Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone.” *Applied Cognitive Psychology*, 24(5), 597–607.

Iqbal, S. T., Ju, Y.-C., and Horvitz, E. (2010). “Cars, calls, and cognition: investigating driving and divided attention.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1281–1290.

Jahn, G., Krems, J. F., and Gelau, C. (2009). “Skill acquisition while operating in-vehicle information systems: Interface design determines the level of safety-relevant distractions.” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 51(2), 136–151.

Jones, D. G. and Endsley, M. R. (2004). “Use of real-time probes for measuring situation awareness.” *The International Journal of Aviation Psychology*, 14(4), 343–367.

Kass, S. J., Cole, K. S., and Stanny, C. J. (2007). “Effects of distraction and experience on situation awareness and simulated driving.” *Transportation Research Part F: Traffic Psychology and Behaviour*, 10(4), 321–329.

Kutila, M. H., Jokela, M., Mäkinen, T., Viitanen, J., Markkula, G., and Victor, T. W. (2007). “Driver cognitive distraction detection: Feature estimation and implementation. Proceedings of the Institution of Mechanical Engineers,” *Part D: Journal of Automobile Engineering*, 221(9), 1027–1040.

Lee, J.-H. and Spence C. (2008). “Assessing the benefits of multimodal feedback on dual-task performance under demanding conditions.” *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*, 1, 185–192.

Lix, L. M., Keselman, J. C., and Keselman, H. J. (1996). “Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test.” *Review of Educational Research*, 66(4), 579–619.

Ma, R. and Kaber, D. B. (2007). “Situation awareness and driving performance in a simulated navigation task.” *Ergonomics*, 50(8), 1351–1364.

Maxwell, S. E. (1980). “Pairwise multiple comparisons in repeated measures designs.” *Journal of Educational and Behavioral Statistics*, 5(3), 269–287.

Medeiros-Ward, N., Cooper, J.M., and Strayer, D.L. (2014). “Hierarchical control and driving.” *Journal of Experimental Psychology: General*, 143(3), 953–958.

Nasar, J., Hecht, P., and Wener, R. (2007). “‘Call if You Have Trouble’: Mobile Phones and Safety among College Students.” *International Journal of Urban and Regional Research*, 31(4), 863–873.

Nasar, J. L. and Troyer, D. (2013). “Pedestrian injuries due to mobile phone use in public places.” *Accident Analysis and Prevention*, 57, 91–95.

National Highway Traffic Safety Administration (2013). *Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices*, Federal Register, 78(81).

Newcomb, A. (2012, May 13). “Texting while walking banned in New Jersey town.” *ABC News Blogs*. 5/13/2012. Retrieved from http://gma.yahoo.com/blogs/abc-blogs/texting-while-walking-banned-jersey-town-152720795--abc-news-topstories.html.

Petzoldt, T., Bellem, H., and Krems, J. F. (2014). “The Critical Tracking Task A Potentially Useful Method to Assess Driver Distraction?” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56(4), 789–808.

Pitts, M. J., Burnett, G., Skrypchuk, L., Wellings, T., Attridge, A., and Williams, M. A. (2012a). “Visual–haptic feedback interaction in automotive touchscreens.” *Displays*, 33(1), 7–16.

Pitts, M. J., Burnett, G. E., Williams, M. A., and Wellings T. (2010). “Does haptic feedback change the way we view touchscreens in cars?” *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, ACM.

Pitts, M. J., Skrypchuk, L. Wellings, T., Attridge, A., and Williams, M. A. (2012b). “Evaluating user response to in-car haptic feedback touchscreens using the lane change test.” *Advances in Human-Computer Interaction 2012*, 1–13.

Reimer, B., Mehler, B., and Donmez, B. (2014). “A study of young adults examining phone dialing while driving using a touchscreen vs. a button style flip-phone.” *Transportation Research Part F: Traffic Psychology and Behaviour*, 23, 57–68.

Ratwani, R. M., Andrews, A. E., Sousk, J. D., and Trafton, J. G. (2008). “The effect of interruption modality on primary task resumption.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 52(4), 393–397.

Ryu, J. B., Sihn, Y. K., and Yu, S. B. (2013). “Assessment of risky driving caused by the amount of time focused on an in-vehicle display system.” *International journal of automotive technology*, 14(2), 259–264.

Sethumadhavan, A. (2011). “In-vehicle technologies and driver distraction.” *Ergonomics in Design: The Quarterly of Human Factors Applications*, 19(4), 29–30.

Van Erp, J. B. and Van Veen, H. A. (2004). “Vibro-tactile in-vehicle navigation system.” *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(4), 247–256.

Walker, G. H., Stanton, N. A., and Young, M. S. (2008). “Feedback and driver situation awareness (SA): A comparison of SA measures and contexts.” *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(4), 282–299.

Wickens, C.D. (2002). “Multiple resources and performance prediction.” *Theoretical Issues in Ergonomics Science*, 3(2): 159–177.

Yerkes, R.M. and Dodson, J.D. (1908). “The relation of strength of stimulus to the rapidity of habit formation.” *Journal of Comparative Neurology and Psychology*, 18, 459–482.

Young, K. L., Rudin-Brown, C. M., Patten, C., Ceci, R., and Lenné, M. G. (2014). “Effects of phone type on driving and eye glance behaviour while text-messaging.” *Safety Science*, 68, 47–54.

APPENDIX—DATA COLLECTION MATERIALS

Attached #1—Perceived Workload Questionnaire

Attached #2—Demographic and Driver Background Questionnaire

Attached #3—Consent Form

Attached #4—Flyer

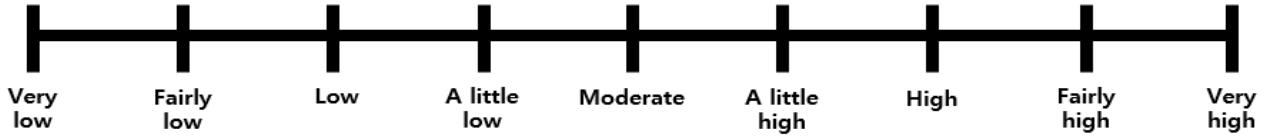
Attached #1

Subject #: _____ Task code: _____

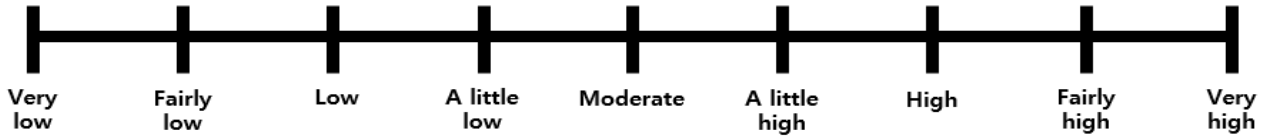
Perceived Workload Questionnaire

<INSTRUCTION>
Please circle one of the predefined qualities in each question.

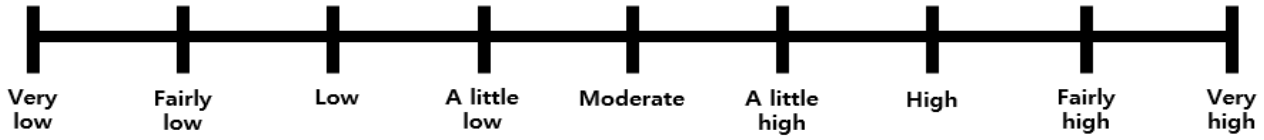
1. MENTAL DEMAND—How mentally demanding was the task?



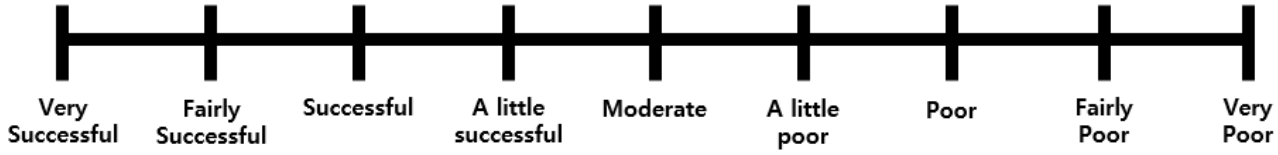
2. PHYSICAL DEMAND—How physically demanding was the task?



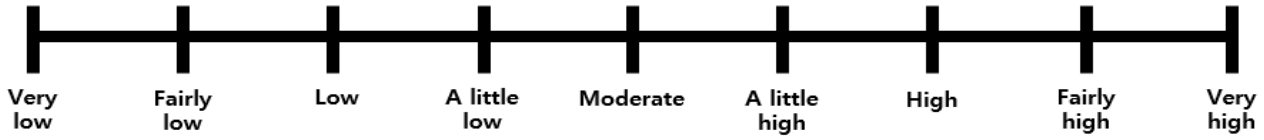
3. TEMPORAL DEMAND—How hurried or rushed was the pace of the task?



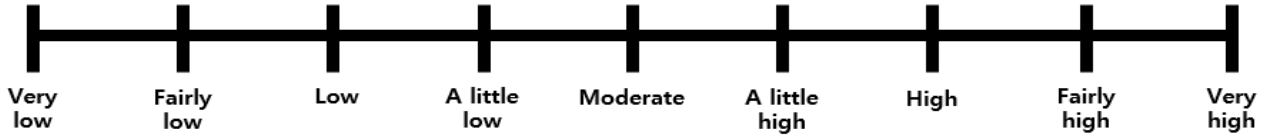
4. PERFORMANCE—How successful were you in accomplishing what you were asked to do?



5. EFFORT—How hard did you have to work to accomplish your level of performance?



6. FRUSTRATION—How insecure, discouraged, intimated, stressed and annoyed were you?



Attached #2

Subject #: _____

Demographic and Driver Background Questionnaire

<INSTRUCTION>
Please provide your response accordingly

1. Gender: Male / Female

2. Age: _____

3. How long have you been licensed to drive? _____ year(s)

4. How often do you drive? _____ day(s) per week

5. Current Vehicle: _____ (year) _____ (make) _____ (model)

6. Does your current vehicle have a touchscreen-based dashboard? Yes / No

7. Have you ever owned or regularly driven a vehicle with a touchscreen-based dashboard?

Yes / No

Attached #3

CONSENT FORM

Evaluation of Drivers' Performance and Awareness while Driving

Introduction

The purpose of this form is to provide you with the information that may affect your decision as to whether or not to participate in this experiment. If you decide to participate in this study, this form will be used to record your consent.

In this experiment, you will be asked to drive safely in a multitasking situation. You will be instructed to conduct lane-change tasks along the prescribed course track while conducting two secondary tasks. Data collection consists of your performance of the driving task, the secondary tasks, subjective ratings of perceived workload, and physiological measures. Additionally, you will be asked to fill out a form pertaining to general background questions.

You can participate in this experiment if you are satisfying all of the following criteria:

- A valid Texas driver license with no less than 5 years of driving experience.
- Normal or corrected-to-normal visual and auditory acuity (glasses/contacts/hearing aids are ok).
- Age between 20 and 40.
- Able to speak, write, and read English.
- No known color vision deficiency, and no physical conditions that would limit your ability to drive safely while conducting the secondary tasks.

What will I be asked to do?

1. Please read through and sign this form if you have not done so.
2. Familiarizing yourself with the three input interfaces specially designed for this experiment:
    - A smartphone screen and requires your visual attention, mounted on the instrument cluster.
    - Two input interfaces. One is implemented on a physical keypad, and the other is on a touchscreen tablet. Both will be mounted on the center console, and you will interact with either of the two at a time.
3. Training session. You will be instructed to drive along the prescribed course track for training.
4. Wearing two physiological monitors to record heart rate and skin conductance level. They are not supposed to cause any pain, though we recommend you report any discomfort to the experimenter(s). You cannot participate in this experiment without wearing them.
5. Conducting experimental sessions involving the driving task with two secondary tasks:
    - Driving task—performing lane-change on the prescribed track while maintaining 25 mph.
    - Secondary task A—making verbal acknowledgement for signage placed alongside the track.
    - Secondary task B—inputting data by pressing buttons on the interfaces mounted on the center console.
6. Rating your subjective workload for each experimental session, based on NASA-Task Load Index (NASA-TLX) scale. The experimenter will instruct you how to do so.
7. Filling a form for general background questions.

You are encouraged to take a break between sessions. In total, the entire duration of the above procedure is expected to be no more than 2 hours.

What are the risks involved in this study?

You will be asked to drive in a distracted state. The risk is expected to show minor increase over minimal, as you will operate an actual vehicle though the course track. The track is reserved only for this experiment and closed to others. It is not likely that you will see other vehicles running while you are on the track. You are strongly advised to maintain 25 mph and not to exceed the speed limit, or you will be notified by the experimenter (with the potential to end the experiment). The experimenter will be in the

front passenger seat throughout the experiment, paying visual attention to the road ahead as well as the surroundings. He/she will ask you for caution if the vehicle starts to deviate from the prescribed path or lane, which will be evidenced by visual judgment. If you did not get back on the track immediately, he/she will instruct you to stop the secondary tasks, pull over the vehicle, and restart the trial. Any objects within 15 yards from the course track will be cleared, except for the experimental signage as well as delineation devices, which will be mainly made up of plastic.

If you experience stress or discomfort that might make you feel unsafe, you can stop the experiment and leave at any moment. If there is any situation that the experimenter can deem that your safety is at risk (evidenced by vocalizations, facial expressions, or body language that are normally associated with discomfort), he/she will stop the experiment for further assessment of the situation (with the potential to end the experiment). In addition, there is a slight risk of discomfort when wearing the physiological measurement devices, but the ability to take frequent breaks and remove the devices should minimize this discomfort. If you are feeling uncomfortable for any reason, please do not hesitate to let the experimenters know, and keep in mind you can leave the experiment at any time.

If you suffer any injury as a result of taking part in this research study, please understand that nothing has been arranged to provide free treatment of the injury or any other type of payment. In case of an accident or medical emergency, appropriate emergency medical services will be called. However, neither TTI nor Texas A&M University will assume financial responsibility for any medical costs incurred by participants due to participation in this study. You are encouraged to review your medical insurance coverage to ensure it is adequate.

What are the possible benefits of this study?

There are no foreseeable direct benefits to you. The benefits to society include a greater understanding of the characteristics of user interaction that support multitask performance. This is especially important when human operators conduct safety-critical tasks. The results of this study can help inform product designers, in order to best support human information processing abilities in data-rich environments.

Do I have to participate?

No. Your participation is voluntary. You may decide not to participate or to withdraw at any time without your current or future relations with Texas A&M University being affected.

Will I be compensated?

Once each participant has completed the driving task, we will return to an office; following a post-drive interview, the participant will be compensated a one-time payment of up to $40.00 cash. If either the participant or the experimenter (due to unforeseen circumstances) chooses to stop the experiment at any time after the training but before the end of the study, the participant will receive compensation at a rate of $20 per hour, rounded up to the nearest 10-minute interval.

Who will know about my participation in this study?

Unless you inform another on your own, only the study team will know about your participation in the study. The signed consent forms will be kept in a locked cabinet in the TAMU Human Factors and Cognitive Systems Laboratory. Data files will be anonymized with generic and unique subject labels and will not in any way be linked to subject identifiers. Data files will only be accessible by the study team. Video/audio recordings collected from this experiment will be analyzed for the purpose of research and not be used in public. All data will be kept until data analysis have been completed, or for at most seven years, and then erased.

Who do I contact with questions about the study?

If you have questions regarding this study, you may contact Dr. Thomas Ferris, (979) 458-2340, tferris@tamu.edu.

Who do I contact about my rights as a research participant?

For questions about your rights as a research participant, to provide input regarding research, or if you have questions, complaints, or concerns about the research, you may call the Texas A&M University Human Subjects Protection Program office by phone at 1- 979-458-4067, toll free at 1-855-795-8636, or by email at irb@tamu.edu.

Will photos, video or audio recordings be made of me during the study?

The researchers will make an audio and/or video recording during the study so as to analyze further in the event that an action or intention is unclear or otherwise not captured by performance data. If you do not give permission for the audio/video recording to be obtained, you cannot participate in this study.

_____ I give my permission for audio/video recordings to be made of me during my participation in this study.

Signature for consent form

Please be sure you have read the above information and received satisfactory answers to any questions. You will be given a copy of this consent form for your records. By signing this document, you consent to participate in this study.

Signature of Participant: _____Date: _____

Printed Name: _____

Attached #4

Flyer

Hi [Name],

Human Factors and Cognitive System Laboratory of Industrial and Systems Engineering in Texas A&M University is seeking participants for a driving study that will test and evaluate in-vehicle interactions. The purpose of this study is to research drivers' performance in driving and interaction with the specially designed dashboard-mounted interface and to measure the potential for distraction, physiological response, and cognitive load while interacting with the interface. Data from this study will be kept confidential.

The expected duration will be no more than three hours including the resting time. You will be asked to drive an instrumented vehicle on a closed-course track at Texas A&M University Riverside Campus. No other vehicles will be running on the track while conducting this experiment.

Participants will be compensated one-time payment of up to $40.00 in cash for complete participation. If either the participant or the experimenter (due to unforeseen circumstances) chooses to stop the experiment at any time after the training but before the end of the study, the participant will receive compensation at a rate of $20 per hour, rounded up to the nearest 10-minute interval. You will be asked to sign for video/audio recordings obtained during the study. The recordings will be used for the purpose of research only and not be published.

Participants must:

- Possess a valid Texas driver license with no less than 5 years of driving experience
- Age between 20–40
- No color vision deficiency
- No physical disabilities, condition, pain, or injuries that interfere with driving as well as haptic practices
- Be able to speak, read, and write in English

When: [Dates], 2015
Duration: No more than 2 hours
Where: Texas A&M Transportation Institute, Texas A&M Riverside Campus

Pre-Registration is required.
To pre-register or for questions, please contact Youngbo Suh of the Texas A&M University at suh0816@tamu.edu. Please leave a message including your name and age, along with a preferred contact information. Please pass this along to anyone you know who may be interested.

Thank you,
Human Factors and Cognitive System Laboratory
Industrial and Systems Engineering, Texas A&M University