Accident Analysis and Prevention 43 (2011) 939–947

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Driver performance while text messaging using handheld and in-vehicle systems

Justin M. Owens*, Shane B. McLaughlin, Jeremy Sudweeks

Virginia Tech Transportation Institute, 3500 Transportation Research Plaza, Blacksburg, VA 24061, USA

ARTICLE INFO

Article history:
Received 20 July 2010
Received in revised form 8 November 2010
Accepted 24 November 2010

Keywords:
Driving
Text messaging
Distraction
Mobile phone

ABSTRACT

This study presents an evaluation of driver performance while text messaging via handheld mobile phones and an in-vehicle texting system. Participants sent and received text messages while driving with an experimenter on a closed-road course, using their personal mobile phones and the vehicle's system. The test vehicle was an instrumented 2010 Mercury Mariner equipped with an OEM in-vehicle system that supports text messaging and voice control of mobile devices via Bluetooth, which was modified to allow text message sending during driving. Twenty participants were tested, 11 younger (19–34) and 9 older (39–51). All participants were regular users of the in-vehicle system, although none had experience with the texting functions.

Results indicated that handheld text message sending and receiving resulted in higher mental demand, more frequent and longer glances away from the roadway, and degraded steering measures compared to baseline. Using the in-vehicle system to send messages showed less performance degradation, but still had more task-related interior glance time and higher mental demand than baseline; using the system's text-to-speech functionality for incoming messages showed no differences from baseline. These findings suggest that using handheld phones to send and receive text messages may interfere with drivers' visual and steering behaviors; the in-vehicle system showed improvement, but performance was not at baseline levels during message sending.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The use of in-vehicle electronic devices has risen dramatically in the past decade, posing a unique challenge for safety researchers. As these devices have become increasingly popular, lawmakers must decide how best to regulate their use while not unduly restricting tasks that drivers can safely perform. To identify tasks requiring regulation, research must be conducted to evaluate the wide range of devices currently being used and the tasks being performed.

In the past several years, the use of mobile phones for making calls while driving has been a topic of much research. Although researchers have yet to reach a consensus on the risk level of calling while driving (Drews et al., 2008; Klauer et al., 2006), the research landscape has been well established using driving simulators, on-road testing, and naturalistic observation (Horrey and Wickens, 2006). Recently, however, mobile phones have been increasingly used for text messaging in addition to making calls. The International Association for the Wireless Telecommunications Industry (CTIA) reports that text messaging increased by more than 23 times from 2005 to 2009, with nearly 1.6 trillion text messages currently sent annually (CTIA, 2010). Consequently, texting is becoming more common during driving; a recent American Automobile Association (AAA) survey found that 14.1% of all drivers, and 48.5% of drivers between the ages of 18 and 24, reported texting while driving (AAA, 2008).

While fewer studies have been conducted examining the safety aspects of text messaging while driving, several of these have suggested that it is equally, if not more, detrimental to the driving task than calling. Crisler et al. (2008) asked participants to drive winding roads in a simulator while performing various tasks including text messaging and using an MP3 player. They found that text messaging while driving led to degraded lanekeeping, speed modulation, and steering performance. Drews et al. (2009), also using a driving simulator, had participants follow a lead vehicle while having a free-form text-message conversation with a friend using their personal phone. They found that texting resulted in slower brake onset times, longer following distances, shorter minimum following distances (exhibiting increased variability in following distance), poorer lanekeeping, and more crashes during dual-tasking.

Hosking et al. (2009) used a driving simulator to study text messaging during varied driving situations, including vehicle-

* Corresponding author at: Virginia Tech Transportation Institute, Center for Automotive Safety Research, 3500 Transportation Research Plaza, Blacksburg, VA 24061, USA. Tel.: +1 540 231 1010; fax: +1 540 231 1555.

E-mail addresses: jowens@vtti.vt.edu (J.M. Owens), smclaughlin@vtti.vt.edu (S.B. McLaughlin), jsudweeks@vtti.vt.edu (J. Sudweeks).

0001-4575/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2010.11.019

following, pedestrian avoidance, and lane-changing tasks. Here, participants were trained on the use of a cell phone with a predictive-text input and were asked to answer simple questions that were preloaded on the phone. Participants were given 15 s to read the message, and were then prompted for a response. They found that when texting, participants spent more time looking inside the car, increased their following distance, missed more lane-change cues, and deviated from the lane more often; interestingly, there was no change in hazard detection performance. The authors suggest that this could be attributable to the higher rate of hazards present in the study than in the real world, which could raise participant expectancy; this is quite possibly the case, and is a factor worth noting in any study that includes enough hazard trials to reach statistical significance.

Although their data were collected before text messaging became prevalent, Klauer et al. (2006) found in their large-scale naturalistic driving study that dialing a handheld device increased crash or near-crash risk by roughly 2.8 times compared to baseline. Further, they found that taking the eyes away from the forward roadway for 2 s out of a 5 s span significantly increased crash risk. These results suggest that texting, a task requiring similar actions and visual–spatial attention to dialing, but with potentially longer duration, may pose an even greater risk of crash.

As the above discussion suggests, there has to this point been a lack of controlled on-road studies examining the effects of text messaging. While driving simulators are a flexible tool for investigating distracted driving, it is likely that participants do not behave exactly as they would on a physical roadway that possesses both the full kinematics of driving and a real sense of risk. Even on a closed road, a loss of vehicle control or a crash would have more significant consequences than in a simulator. Naturalistic studies are also of great value, as they are able to capture what drivers actually do on public roads, but they lack control over the circumstances under which drivers engage in secondary tasks. To the authors' knowledge, the study presented here is the first controlled, physical-road study of the effects text messaging while driving. In addition, this study was designed to compare handheld text messaging with an in-vehicle system for text message sending and receiving. This system combines audio text-to-speech playback of received messages and sending of preprogrammed messages accessed through steering wheel controls and a display at the top of the center stack. Integrated systems that allow control over mobile devices are becoming common on new vehicles, and they are worth evaluating to determine the extent to which they are able to mediate the potentially detrimental effects of secondary tasks.

Testing was performed on a closed-road course to minimize risk, and immediately followed an on-road portion that examined driver performance while dialing contacts, having phone conversations, and selecting music tracks using the in-vehicle system and drivers' own devices (Owens et al., 2010). This on-road study indicated that the in-vehicle system, using voice control for these tasks, showed improvements over handheld calling and music selection in task duration, glance behavior, steering measures, and subjective mental demand. Dependent measures in the current study included total task duration, total number of task-related interior glances (defined as fixations to the instrument cluster, steering wheel, center stack, or interior objects including handheld devices), total task-related interior glance duration and longest task-related single-glance duration, steering wheel position variance, maximum steering wheel velocity, number of large steering wheel reversals (>5°), and mental demand. Glance behavior, particularly glance duration away from the forward roadway, has previously been shown to be predictive of crash risk (Klauer et al., 2006; Wierwille and Tijerina, 1998). Increased steering variability, maximum steering rates, and steering reversal rate indicate decreased vehicular control and increased workload (Dingus, 1995; Macdonald and Hoffmann, 1980; McLaughlin et al., 2009; Wierwille and Gutmann, 1978).

2. Methods

2.1. Materials

2.1.1. Test vehicle

Testing was conducted using a 2010 Mercury Mariner equipped with the Ford SYNC system. This system uses a Bluetooth wireless connection to interface with a mobile phone to allow control through the vehicle steering wheel, center stack, and audio system. To send a text message, drivers use steering-wheel-mounted buttons (Fig. 1A) to cycle through a set of 15 pre-programmed messages shown in the vehicle message display (Fig. 1B); once the desired message is selected, the OK button is pressed to send the message. To receive an incoming text message, after a notification chime the driver presses the OK button twice, which prompts the system to read the message aloud through the car speakers using text-to-speech functionality. For this study, the SYNC system in this vehicle was modified by the manufacturer to allow the driver to send text messages while the vehicle was moving, and was connected to a mobile phone via Bluetooth. Text messages were remotely sent to participants via a Dell laptop using Google Voice; this resulted in near instantaneous message transmission.

The vehicle was instrumented with four cameras capturing views of the driver's face, the forward roadway, the instrument panel, and center stack; it also included cabin audio (Fig. 2). A vehicle Data Acquisition System (DAS) collected vehicle network data, including steering wheel movements, at 10 Hz.

2.1.2. Test track

Testing was conducted on an approximately 1.4 mile-long section of a closed two-lane road built to Department of Transportation (DOT) highway specifications. Two straight areas of the roadway were used for trials, one heading uphill and one heading downhill. A white 1999 Ford Contour served as a confederate vehicle on the roadway.

2.1.3. Mobile phones

Participants used their personal mobile phones for handheld texting tasks. These included 10 with standard 12-button numeric physical keypads, 6 with touch screen input, and 4 with full QWERTY keypads.

2.2. Participants

Area owners of vehicles equipped with the SYNC system were contacted and screened to identify regular users of the system. Potential participants were identified using a list of local owners provided by Ford, and were contacted by letter and telephone. Owners who used the system to make calls and play music three or more times a week and who were comfortable using their phone for text messaging were invited to participate. As text messaging while driving is not currently legal in Virginia, participants were not screened regarding their current use of text messages while driving. Further, as not all mobile phones are compatible with the text messaging features of the SYNC system, experience with these features was not a criterion for inclusion; in fact, no participants had used the system for text messaging prior to the study. Twenty-one participants were tested, although one had to be excluded due to equipment failure. The number and mean ages of the remaining 9 female and 11 male participants are shown in Table 1. Younger participants reported owning their current vehicle for a mean of 16.9 months, and older participants reported owning their vehicle

Fig. 1. Layout of vehicle system interface. (A) depicts the control buttons on the steering wheel; the arrow buttons select from options presented on a screen located at the top of the center stack (B), while the "OK" button confirms selections.

for a mean of 13.2 months. All participants signed an IRB-approved informed consent form, had valid U.S. driver's licenses, scored 20/40 or better on a Snellen visual acuity test, and could read a sample text message sent to their phone without difficulty. Participants were paid $20 per hour for their participation, plus an additional $5 to cover any fees incurred for text messages or cellular minutes used. The entire study, including the previously published on-road portion, took approximately 3 h.

Fig. 2. Camera views of driver (face is not blurred in actual video), forward roadway, and wide- and close-views of driver's hands and center stack.

Table 1
Number of participants (range; $\bar{x}$) by group.

	Female	Male
Younger	5 (19–26; 24)	6 (19–34; 27)
Older	4 (40–50, 44)	5 (39–51, 48)

Table 2
Text messages used for tasks.

	Send message	Receive message
Handheld Phone	I'm stuck in traffic.	What color is the sky on a sunny day?
	Call you later.	How many minutes are in an hour?
Vehicle System	I'm stuck in traffic.	What state are you in now?
	Call you later.	What color is fresh snow?

2.3. Texting tasks

Participants were asked to send and receive text messages using both their handheld phone and the in-vehicle system. As outgoing messages were pre-programmed into the in-vehicle system, two messages from this list (one near the beginning, one near the end) were chosen to match handheld and in-vehicle system modalities (Table 2). For both modalities, participants were told what message to send, after which the experimenter instructed them to begin. For message-receiving tasks, participants were sent text messages to either their handheld phone or the in-vehicle system. The participant initiated receiving the message when a notification sounded. As shown in Table 2, the messages were simple questions that were designed to ensure the participant attended to the message. For both modalities, the participant was instructed to read/listen to the message, then reply out loud.

2.4. Experimental design

This study was a $2 \times 2 \times 2$ mixed factorial design, with Age (Younger and Older) being the between-subjects factor and Task (Send and Receive) and Modality (Handheld and Vehicle System) being within-subjects. There were six laps driven overall, including practice and baseline laps, and this portion of the study took approximately 40 min occurring roughly 2 h into the overall study. Baseline measurement always started at the point where tasks were given on experimental laps; baseline duration was calculated post hoc to match the mean duration of all tasks, 17.2 s.

2.5. Dependent measures

Dependent measures were derived from the vehicle network, post hoc video reduction, and subjective report. Network variables included steering wheel position variance, maximum steering wheel velocity, and number of steering wheel reversals greater than 5°. Variables derived from video reduction included total task duration, total number of task-related interior glances, total task-related interior glance duration, and longest task-related single-glance duration. As discussed above, interior glances were defined as fixations directed to the instrument cluster, steering wheel, center stack, or interior objects (including handheld devices), but did not include glances to the experimenter. Video was reduced by assistants who had received 8–10 h of general eyeglance-reduction training and testing, in addition to project-specific training. All reduced eyeglance data were double-checked by the reduction lab supervisor. Participants also reported mental demand after each task, using a scale from the National Aeronautics and Space Administration Task Load Index (NASA TLX). They verbally rated the mental demand they experienced from the combination of driving and performing the secondary task from 1 (Very Low) to 7 (Very High); a paper copy of the scale was attached to the dashboard for reference. This scale was chosen to minimize the amount of time drivers would need to look away from the forward roadway and the amount of demand the scale itself would require.

2.6. Procedure

As discussed above, the texting portion of this study was conducted on a closed road following a public-road portion where participants dialed contacts, had phone conversations, and selected music tracks using their personal handheld devices and the in-vehicle system (Owens et al., 2010). Upon returning to the research facility from the on-road portion, participants were directed onto the closed road. They were then instructed on how to use the in-vehicle system to send and receive text messages, and were allowed to practice until they could perform each task with no experimenter help. To assist in learning how to send messages using the system, participants were given a list of all 15 pre-programmed messages available to send. For this study, sending text messages involved replying to a message that had been previously stored in both the in-vehicle system and the participant's phone. To reply to a message, drivers pressed the "OK" button on the steering wheel twice, used arrow buttons to cycle through the list of available messages until the desired message was found, then pressed the "OK" button twice again to complete the sending of the message. When a message was received by the system, it chimed and the information display indicated that a new message had arrived. To have the system read the incoming message aloud, participants pressed the "OK" button twice. Participants were also asked to read and reply to a text message received on their phone.

After participants were comfortable sending and receiving text messages using both their device and the in-vehicle system, they were directed to take a familiarization lap of the test track. Testing began after the participant returned to the starting point on the track. Participants were instructed to stay at or below 35 mph at all times. Each lap had two trials of the same condition, one per leg. Participants completed two counterbalanced conditions for the first two laps, a baseline lap for the third, and two more counterbalanced conditions for the final two laps. Task counterbalancing was achieved through the use of permutations; this resulted in 24 orders, although only 21 were used due to time constraints and the strict eligibility criteria. Tasks were always completed in the same locations on each lap, and after each task participants were asked to rate their mental demand. During the baseline lap, no tasks were given, but participants were asked to rate their mental demand at the task locations.

For tasks in which participants received text messages, these messages were generated and sent by an experimenter located in a control tower overlooking the test track. Video monitors in the control room allowed this experimenter to send messages when the test vehicle passed particular landmarks, ensuring that messages were always received in the same segment of roadway. For tasks where participants were asked to send messages, the in-vehicle experimenter instructed them to do so as these landmarks were passed. Participants were told in advance of each leg what message they would be asked to send to avoid confusion.

For safety reasons no other vehicles were permitted on the roadway during testing. However, a confederate vehicle entered the roadway with the test vehicle, and parked on a side street during the experiment to maintain the illusion that other vehicles could enter the road. During the initial practice lap, this confederate vehicle passed the test vehicle in the opposite direction before parking in the experimental location.

3. Results

The General Linear Model (GLM) in SAS 9.2 was used to analyze task duration, mental demand, number of task-related

Table 3
Summary of significant texting results.

	Total task Dur.	Mental demand	Total # of interior glances	Total interior glance Dur.	Longest interior glance Dur.	Steering wheel position variance	Max steering wheel velocity	Number of steering reversals
Modality								
$F(1,18)=$	96.71	95.72	76.56	81.96	7.40	17.87	69.14	38.71
$p<$	0.0001	0.0001	0.0001	0.0001	0.05	0.001	0.0001	0.0001
Partial η^2	0.84	0.84	0.81	0.82	0.29	0.50	0.79	0.68
Task								
$F(1,18)=$	90.15	75.35	130.35	109.29	41.28	11.60	17.48	43.87
$p<$	0.0001	0.0001	0.0001	0.0001	0.0001	0.01	0.001	0.0001
Partial η^2	0.83	0.81	0.88	0.86	0.70	0.40	0.49	0.71
Age								
$F(1,18)=$	6.88	–	–	7.83	8.39	16.58	10.63	10.39
$p<$	0.05			0.05	0.01	0.001	0.01	0.01
Partial η^2	0.28			0.30	0.32	0.48	0.37	0.37
M × T								
$F(1,18)=$	51.05	11.07	41.95	43.28	–	–	–	45.65
$p<$	0.0001	0.01	0.0001	0.0001				0.0001
Partial η^2	0.74	0.38	0.70	0.71				0.72
M × A								
$F(1,18)=$	6.15	–	–	5.13	–	13.56	9.36	11.99
$p<$	0.05			0.05		0.01	0.01	0.01
Partial η^2	0.25			0.22		0.43	0.34	0.40
T × A								
$F(1,18)=$	7.71	–	4.71	10.93	–	–	–	–
$p<$	0.05		0.05	0.01				
Partial η^2	0.30		0.21	0.38				
M × T × A								
$F(1,18)=$	5.18	–	5.21	7.12	–	–	–	5.91
$p<$	0.05		0.05	0.05				0.05
Partial η^2	0.22		0.22	0.28				0.25

interior glances, total task-related interior glance duration and mean longest single duration of task-related interior glances, steering variance, maximum steering wheel velocity, and number of steering wheel reversals greater than 5°. Post hoc tests using a Tukey correction for multiple comparisons were used to investigate interactions and to compare task conditions to Baseline. The between-subject ANOVA assumptions of normality and heterogeneity of variance were evaluated with residual plots. Although some deviations from these conditions were observed, transformations were not considered; rather, the analysis relied on the robustness of ANOVA to mild departure from these assumptions. As a univariate analysis was employed the additional assumption of sphericity was evaluated for the within-subject factors. For two level factors, the sphericity condition was met. Table 3 presents a summary of significant results and effect sizes, with specific interactions discussed below.

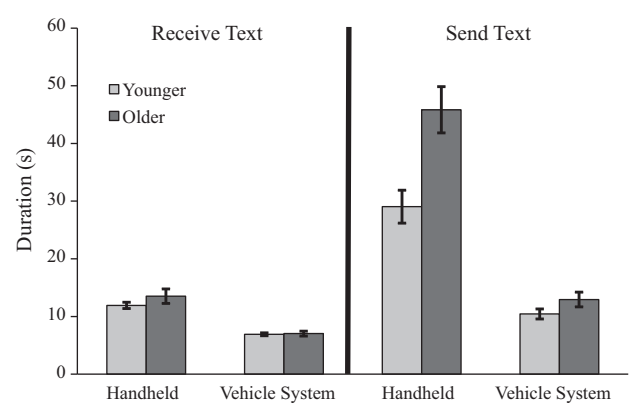

Fig. 3. Task durations (in seconds) for all conditions.

3.1.1. Total task duration

Fig. 3 presents mean task durations and standard error bars for all conditions. Here, the Handheld Send condition, with an overall mean duration of 37.4 s, had a significantly greater duration than all other tasks, $p < 0.0001$. Further, as the interactions with Age suggest, this difference was larger for older drivers. There was a marginally significant difference in task duration between Handheld Receive ($\bar{x} = 12.8$ s) and Vehicle System Receive ($\bar{x} = 7.1$ s), $p = 0.05$. No other tasks differed significantly from one another. Handheld text message sending took significantly longer than sending using the in-vehicle system, and longer than receiving messages via either method.

As the "Receive" messages between modalities were slightly different lengths, a post hoc analysis was conducted to see whether this disparity could account for the differences seen between Handheld and Vehicle System task durations. Expected mean message duration was calculated for each message/response pair using the vehicle system and established speaking rates (Robb and Gillon, 2007). From these, the mean expected duration for Handheld messages was 2.08 s, and the mean expected duration for vehicle system messages was 1.86 s, for a mean difference of 0.22 s. Confidence intervals were then constructed using the SAS CLM function, which showed a difference of 3.9 s between the upper 95% confidence limit of Vehicle System Receive Text duration and the lower 95% confidence limit of Handheld Receive Text duration. As this is larger than the predicted mean difference of 0.22 s, it can be concluded that differences in message length did not confound the observed differences between handheld and in-vehicle system text message receiving.

3.1.2. Mental demand

Mean subjective mental demand ratings and standard errors are presented in Fig. 4. It was found via post hoc Tukey comparisons that the Handheld Send condition ($\bar{x} = 5.25$ s) had higher demand than both Handheld Receive ($\bar{x} = 3.42$) and Vehicle System Send ($\bar{x} = 3.11$), $p < 0.0001$, and Handheld Receive had higher demand

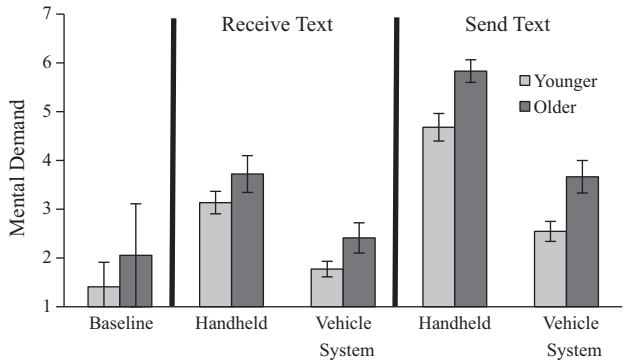

Fig. 4. Mental demand ratings for all conditions and baseline.

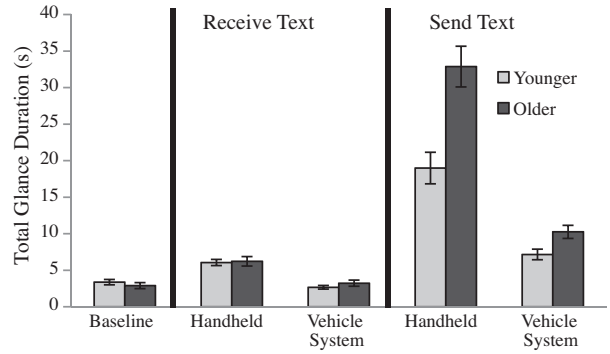

Fig. 6. Mean total task-related interior glance duration for all conditions and baseline.

than Vehicle System Receive ($\bar{x} = 2.09$), $p < 0.0001$. Further, all conditions except Vehicle System Receive had higher mental demand ratings than Baseline ($\bar{x} = 1.73$), $p < 0.0001$. Thus, all tasks except receiving texts via the in-vehicle system were rated to be more mentally demanding than driving alone, sending texts was rated as more mentally demanding than receiving texts, and handheld operation was rated as more demanding than using the in-vehicle system.

3.1.3. Glance measurements

This section presents three measurements of glance behavior. The mean total number of task-related interior glances and their standard errors for each condition are presented in Fig. 5. Handheld Sending ($\bar{x} = 17.45$) had significantly more interior glances than all other conditions ($p < 0.0001$), and the Vehicle System had more glances when Sending ($\bar{x} = 6.55$) than when Receiving ($\bar{x} = 3.06$), $p < 0.01$. During text-message Receiving, there was no difference between modalities. Both Sending tasks had more task-related interior glances than Baseline ($\bar{x} = 3.46$), $p < 0.05$. As suggested by the significant Age × Modality × Task interaction, older drivers took substantially more glances to the interior when sending a message with a handheld phone. To summarize, sending text messages using either handheld devices or the in-vehicle system increased the number of glances away from the road, although more than twice as many glances were taken when sending by hand than when using the in-vehicle system.

Fig. 6 presents the mean total task-related interior glance durations and standard errors for each task. Handheld Sending ($\bar{x} = 25.93$ s) had a longer total glance duration than all other conditions, $p < 0.0001$, and using the Vehicle System for Sending ($\bar{x} = 8.71$ s) resulted in longer total glance durations than using the system for Receiving ($\bar{x} = 3.06$ s), $p < 0.01$. Both Send conditions had longer total glance duration than Baseline ($\bar{x} = 3.14$ s), $p < 0.01$. The interactions with Age can be seen in the Sending tasks, where older drivers had longer glance durations for both modalities, particularly for Handheld. Sending messages required significantly more glance time than receiving messages or baseline, and Handheld Sending had much higher total task-related interior glance time than sending with the in-vehicle system.

The average longest task-related single-glance durations and standard errors for each condition are shown in Fig. 7. Handheld Sending ($\bar{x} = 2.72$ s) and Vehicle System Sending ($\bar{x} = 2.45$ s) had significantly longer glance durations than Handheld Receiving ($\bar{x} = 2.05$ s) and Vehicle System Receiving ($\bar{x} = 1.45$ s), respectively, $p < 0.05$. There was no significant difference either between the two Sending modalities or between the two Receiving modalities. Handheld Receiving and both Sending modalities had longer maximum glance durations than Baseline ($\bar{x} = 1.13$ s), $p < 0.001$. Again, the main effect of age can be seen, as older drivers had longer maximum glance durations for all conditions except baseline. Although the differences are smaller in magnitude than the previous glance variables, mean longest duration glance is a critical measure, as eyes-off-road time greater than 2 s has been shown to be associated with higher crash risk (Klauer et al., 2006). Here, both sending modalities and handheld receiving included, on average, at least one glance longer than 2 s. This will be discussed further below.

3.1.4. Steering measures

Fig. 8 presents steering wheel position variance and standard errors in degrees squared for all conditions. Here, Handheld Send ($\bar{x} = 2.91\,\text{deg}^2$) had greater variance than all other conditions, $p < 0.05$, and Handheld Receive ($\bar{x} = 2.18\,\text{deg}^2$) had greater variance than Vehicle System Receive ($\bar{x} = 1.46\,\text{deg}^2$), $p < 0.05$. In addition, Handheld Send was the only condition to have significantly more steering wheel position variance than Baseline ($\bar{x} = 1.66\,\text{deg}^2$),

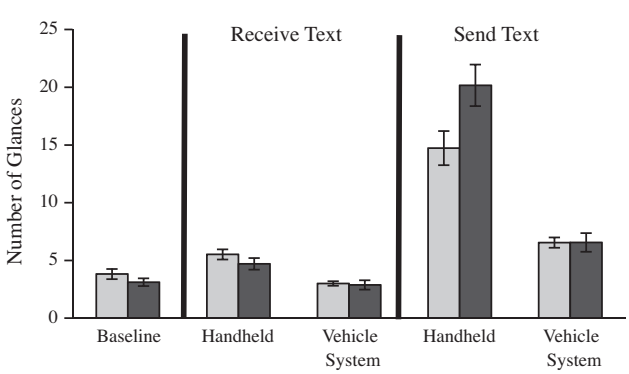

Fig. 5. Mean number of task-related interior glances for all conditions and baseline.

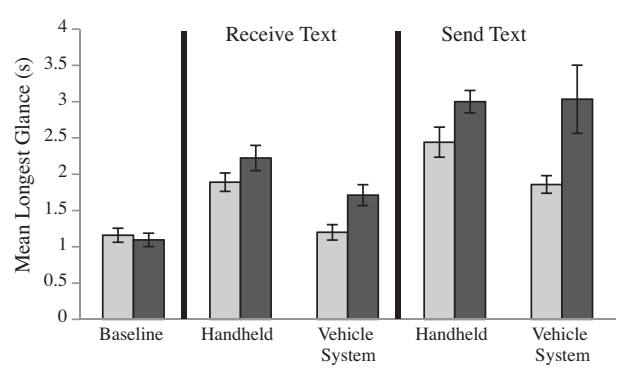

Fig. 7. Mean longest duration glance for all conditions and baseline.

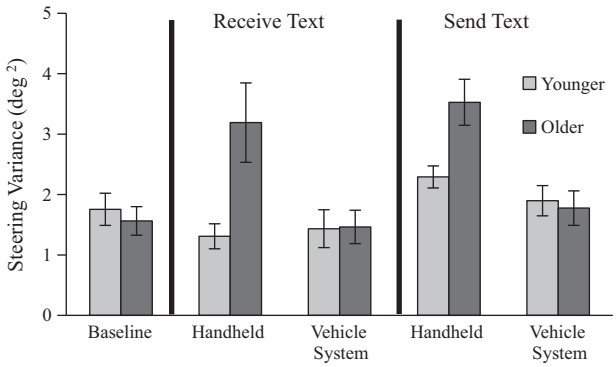

Fig. 8. Mean steering variance for all conditions and baseline.

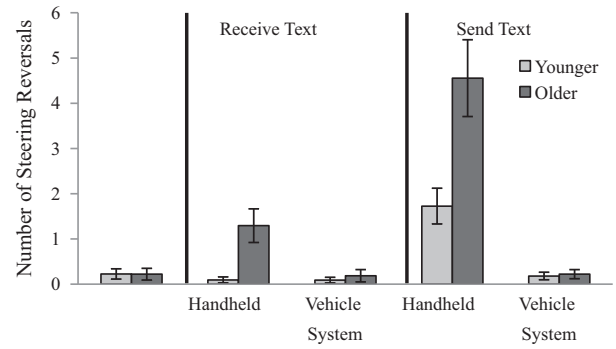

Fig. 10. Mean number of steering reversals greater than 5° for all conditions.

$p < 0.0001$. The interaction between Modality and Age can be clearly seen in the Handheld conditions, where older drivers had considerably degraded performance for both sending and receiving tasks. It appears, then, that handheld text messaging contributes to increased steering wheel position variance, particularly for older drivers, but that texting using the in-vehicle system did not increase steering variance above baseline.

Maximum steering wheel velocities and standard errors for all conditions (in deg/s) are shown in Fig. 9. Handheld Send ($\bar{x} = 22.78$ deg/s) had a higher maximum wheel velocity than all other conditions, $p < 0.001$. Handheld Receive ($\bar{x} = 17.18$ deg/s) had higher maximum wheel velocity than Vehicle System Receive ($\bar{x} = 10.02$ deg/s), $p < 0.0001$, and Vehicle System Send ($\bar{x} = 13.60$ deg/s) had higher maximum wheel velocity than Vehicle System Receive, $p < 0.05$. Both Handheld Send and Handheld Receive had higher maximum wheel velocity than Baseline ($\bar{x} = 11.86$ deg/s), $p < 0.01$. Again, older drivers had higher maximum steering velocities for handheld sending and receiving. These results suggest that handheld texting results in more rapid steering corrections than baseline or texting using the in-vehicle system, which is indicative of degraded control of vehicle position.

The mean number of steering wheel reversals greater than 5°, with standard error bars, is presented in Fig. 10. The Handheld Send condition ($\bar{x} = 3.14$) had more large steering reversals than all other conditions including Baseline ($\bar{x} = 0.22$), $p < 0.0001$ for all conditions. The interactions with age can be clearly seen, as older drivers had more steering reversals for handheld sending and receiving than younger drivers. This increase in large steering reversals suggests that drivers are less able to maintain their desired lane position while text messaging by hand, particularly when sending, and must make larger corrections when the vehicle drifts off course. As with other measures, the degradation is larger for older drivers.

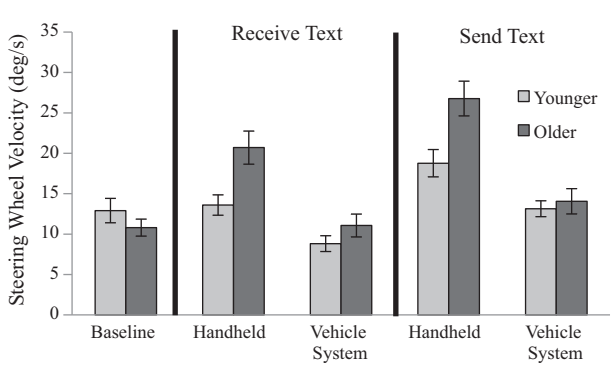

Fig. 9. Maximum steering wheel velocity for all conditions and baseline.

4. Discussion

In this study, two modalities of sending and receiving text messages were compared while participants drove on a closed road. For the Handheld conditions, participants used their personal cell phones; when using the Vehicle System, participants used a modified production Ford SYNC system that read incoming messages aloud and allowed sending of pre-programmed outgoing messages (note that the message-sending feature is currently locked out of production SYNC systems while driving over 3 mph.) All participants were skilled users of other functions of the system, although none had experience with the texting functions. Participants drove five laps of the closed road, completing two trials for each modality (Handheld and Vehicle system) and task combination (Sending and Receiving) and two trials of baseline driving.

Task duration analyses show that handheld texting took significantly longer than all other tasks, and that this difference was greater for older drivers, who spent an average of nearly 46 s sending a medium-length text message. All other factors aside, this suggests that handheld sending of text messages is a task that requires driver attention for a substantial amount of time, particularly for older drivers.

As would be expected, subjective ratings showed that all texting tasks had higher mental demand than baseline driving. Both younger and older drivers found handheld texting to be more demanding than using the in-vehicle system, and sending messages was overall more demanding than receiving messages. These findings suggest that drivers are aware of the increased task demands of text messaging, especially for sending messages by hand, which received a mean rating of 5.3 out of 7 (on a scale of Very Low to Very High demand).

Text messaging also resulted in substantial time spent looking to the interior of the vehicle. Sending text messages resulted in more task-related interior glances than baseline for both modalities, although there were far more interior glances when participants used their handheld phones (with a mean of 17.4 glances for handheld compared to 6.5 for the in-vehicle system). Sending messages using either modality and receiving messages with a handheld device were also associated with longer duration interior glance times than baseline or receiving messages using the in-vehicle system. Both handheld tasks and sending texts using the in-vehicle system had average longest-duration glances over 2 s; as Klauer et al. (2006) and Wierwille and Tijerina (1998) have shown, long glances away from the forward roadway are associated with higher crash likelihood. One factor to note when considering these results is that the message display was located on the top of the center stack in the test vehicle, while it varies in location among other vehicles. This high-mount location may have permitted some degree of monitoring of vehicle lane position or of the forward roadway. The extent to which drivers can accurately monitor traffic and the

road while performing other tasks, and how willing they are to appropriately adjust glance durations and task involvement, are open questions. Finally, the total task-related interior glance time was substantially higher for handheld sending than for all other conditions, with older drivers looking inside the vehicle a mean of 32.8 s for each text message, and younger drivers looking inside the vehicle a mean of 18.9 s. Although the in-vehicle system required less total eyes-interior time for sending messages, it was still significantly longer than baseline. To summarize the eyeglance data, sending text messages took drivers' eyes away from the forward roadway more often and for longer periods of time than receiving messages or baseline driving for both modalities, and handheld sending degraded measures considerably more than when the in-vehicle system was used.

This increased eyes-off-road time likely contributed to degraded vehicle control, as steering measures were likewise degraded while performing texting tasks, especially for older drivers. Handheld text-sending resulted in higher steering wheel position variance, and both handheld sending and receiving had higher maximum steering wheel velocities than baseline. Further, handheld sending resulted in more large steering wheel reversals than all other conditions and baseline driving. This implies that drivers were less able to maintain an intended vehicle heading while performing handheld text messaging, resulting in large, fast, course-correcting steering inputs.

5. Limitations

As stated above, participants were frequent users of the tested in-vehicle system, although not the specific texting functions. While this may have reduced learning effects, it is possible that this resulted in a sample population that may not be representative of the general public. Although the restriction to current users likely overlaps the population of drivers who are prone to text while driving (e.g., technologically inclined drivers), future studies may benefit from a more varied sample, and from a larger sample size. Further, as this study was conducted on a closed road as opposed to a public roadway, it is possible that driver behavior was not identical to behavior on open roads. For example, it is possible that participants felt there was less risk than may be present on open roads, and may have altered their behavior based on this. Nonetheless, it is feasible that the operation of a real automobile on a physical roadway elicits more "natural" behavior than is possible in a simulator, as both kinematic information and a true sense of risk exist even on a closed road. In addition, to keep the number of counterbalance permutations in the range of participants able to be recruited, the baseline condition always occurred on the third lap. This had the benefit of allowing all participants to drive a non-task lap at the same point in the study, but counterbalancing could have helped preclude any fatigue effects. This study also did not examine differences among text input methods for manual texting devices or examination of surprise events. The ability of drivers to react to surprise events while texting would be a valuable addition to the literature if conducted on a physical roadway, as several simulator studies have found impairments to surprise object detection when drivers are manipulating in-vehicle technology (Chisholm et al., 2008; Drews et al., 2009).

6. Conclusion

Overall, these results provide on-road support for previous findings that handheld text messaging presents significant safety concerns, especially for sending messages. Handheld message-sending resulted in far longer task durations, higher mental demand ratings, more frequent and longer glances away from the roadway, and degraded steering measures. Handheld receiving showed less performance decrement than sending, but still had higher mental demand, higher maximum steering wheel velocities, and longer task-related interior glance durations than baseline. The in-vehicle system showed improvement over handheld use, particularly for text message reception, where the text-to-speech function showed no difference from baseline for all variable measures. However, sending pre-determined messages through the system did result in higher mental demand and more frequent and longer glances to the vehicle interior than did baseline driving. These results are consistent with the part of this study investigating handheld and voice control (Owens et al., 2010). In that analysis, degradation of performance measures with handheld device operation was found for these same participants, but no performance degradation was found with voice control.

In nearly all cases, the older driver group exhibited considerably higher degradation of driving performance than did the younger group. This could be attributed to older adults' lack of familiarity with text messaging; however, all participants in this study reported being comfortable texting. It is also possible that the physiological effects of aging – including decreased visual capabilities, slowing of psychomotor functions, and decreased ability to multitask (Wood, 2002) – contributed to this general performance decrement; however, the older driver group tested here was not elderly, and it is unlikely that they experienced substantial age-related impairment. Future research would do well to further investigate the interplay between aging and technology, especially as technological advancements are often introduced in more expensive vehicles with an older demographic that may be less able or willing to engage with them.

Acknowledgements

Funding for this study was provided by The National Surface Transportation Safety Center for Excellence, and the test vehicle and system modification for testing was provided by the Ford Motor Company.

References

American Automobile Association, 2008. Cell Phones and Driving: Research Update, Retrieved March 18, 2010 from http://www.aaafoundation.org/pdf/CellPhonesandDrivingReport.pdf.

Chisholm, S.L., Caird, J.K., Lockhart, J., 2008. The effects of practice with MP3 players on driving performance. Accident Analysis and Prevention 40, 704–713.

Crisler, M.C., Brooks, J.O., Ogle, J.H., Guirl, C.D., Alluri, P., Dixon, K.K., 2008. Effect of wireless communication and entertainment devices on simulated driving performance Transportation Research Record. Journal of the Transportation Research Board 2069, 48–54.

CTIA – The Wireless Association, 2010. Wireless Quick Facts. Year-End Figures, Retrieved May 5, 2010 from http://www.ctia.org/media/index.cfm/AID/10323.

Dingus, T.A., 1995. Moving from measures of performance to measures of effectiveness in the safety evaluation of ITS products or demonstrations. In: Safety Evaluation of Intelligent Transportation Systems: Workshop Proceedings , Reston, VA, May 1–2, 1995.

Drews, F.A., Pasupathi, M., Strayer, D.L., 2008. Passenger and cell phone conversations in simulated driving. Journal of Experimental Psychology: Applied 14 (4), 392–400.

Drews, F.A., Yazdani, H., Godfrey, C.N., Cooper, J.J., Strayer, D.L., 2009. Text messaging during simulated driving. Human Factors 51 (5), 762–770.

Horrey, W.J., Wickens, C.D., 2006. Examining the impact of cell phone conversations on driving using meta-analytic techniques. Human Factors 48 (1), 196–205.

Hosking, S., Young, K.L., Regan, M.A., 2009. The effects of text messaging on young drivers. Human Factors 51 (4), 582–592.

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

Macdonald, W.A., Hoffmann, E.R., 1980. Review of relationships between steering wheel eversal rate and driving task demand. Human Factors 22 (6), 733–739.

McLaughlin, S., Hankey, J., Dingus, T., 2009. Driver measurement: methods and applications. In: Harris, D. (Ed.), Engineering Psychology and Cognitive Ergonomics: HCII 2009, LNAI 5639. Springer-Verlag, Berlin, Heidelberg, pp. 404–413.

Owens, J.M., McLaughlin, S.B., Sudweeks, J., 2010. On-road comparison of driving performance measures when using handheld and voice-control interfaces for mobile phones and portable music players. SAE International Journal of Passengers Cars – Mechanical Systems 3 (1), 734–743.

Robb, M.P., Gillon, G.T., 2007. Speech rates of New Zealand English- and American English-speaking children. Advances in Speech-Language Pathology 9 (2), 1–8.

Wierwille, W.W., Gutmann, J.C., 1978. Comparison of primary and secondary measures as a function of simulated vehicle dynamics and driving conditions. Human Factors 20 (2), 233–244.

Wierwille, W.W., Tijerina, L., 1998. Modelling the relationship between driver in-vehicle visual demands and accident occurrence. In: Gale, A.G., et, al. (Eds.), Vision in Vehicles VI. Elsevier, Amsterdam, pp. 233–243.

Wood, J.M., 2002. Aging, driving, and vision. Clinical and Experimental Optometry 85 (4), 214–220.