

A Driving Simulator Study Examining Phone Dialing with an iPhone vs. a Button Style Flip-Phone

Bryan Reimer¹, Bruce Mehler¹, Birsen Donmez^{1,2}, Silviu Pala³, Ying Wang¹, Nan Zaho¹,
Kirsten Olson¹, John Wenzel¹ & Joseph F. Coughlin¹

¹MIT AgeLab & New England University Transportation Center, Cambridge, MA, USA

²University of Toronto, Department of Mechanical and Industrial Engineering, Toronto, Canada

³Denso International America, Southfield, MI, USA

A simulation study compared 36 young adult drivers' task completion time, eye behavior, and driving performance while dialing a flip-phone with tactile pushbuttons and an iPhone which provides a touch screen interface. Participants who often use a traditional manual button phone completed the dialing task faster when using the flip-phone compared to touch screen users using the iPhone. Females using the flip phone had the highest percentage of time spent with eyes on the road. Females were also less likely to exhibit glances greater than 2 seconds in duration with both phone types and particularly with the flip-phone. Some advantages may exist in a traditional tactile manual interface in terms of the percentage of time drivers kept their eyes on the road.

INTRODUCTION

Traditional cellular phones and smartphones are designed as portable, convenient communication devices. The use of cellular technology in automobiles creates situations where drivers are required to balance driving, a primary task with a high degree of learned behavior (Ranney, 1994) and a responsibility to public safety (Evans, 1991), with the seemingly simple activities involved with the use of cellular technologies, e.g., dialing, conversing, texting, browsing the web, etc. Drivers appear to have limited awareness of the level of distraction involved with phone use (Horrey & Lesch, 2008) and are not curtailed by the perceived risks (Walsh, White, Hyde, & Watson, 2008).

Numerous studies have assessed the impact of cell phones on the driving task (see Collet, Guillot, & Petit, 2010a; 2010b for a comprehensive review). A meta-analysis by Horrey and Wickens (2006) suggests that the act of conversing over a phone impacts a driver's behavior. Although some studies have reported on the effects of manual dialing (Angell et al., 2006; Reed & Green, 1999), research has largely focused on comparing manual phone dialing with voice dialing systems (Angell, et al., 2006; Ranney, Harbluk, & Noy, 2005; Schreiner, 2006). Overall, research has shown an increase in collision risk and fatality with phone use (Redelmeier & Tibshirani, 1997; Violanti, 1997, 1998; Violanti & Marshall, 1996; Wilson & Stimpson, 2010).

Studies have begun to appear assessing the behavior of drivers with smartphone applications (Basacik, Reed, & Robbins, 2012; Lee & Cheng, 2010) and text messaging (Drews, Yazdani, Godfrey, Cooper, & Strayer, 2009; Hoffman, Lee, McGehee, Macias, & Gellatly, 2005; Hosking, Young, & Regan, 2009). For instance, Drews et al. (2009) and Hosking et al. (2009) investigate changes in driver behavior with T9 (predictive text) based input, while Hoffman et al., (2005) investigates the effect of number of lines, page presentation and message control on glance behavior.

Manually dialing a phone is a visual-manipulative interaction that is relatively distinct from the primarily cognitive demands of voice dialing and conversing. Holding the phone once a call has been initiated through either manual or voice dialing, can be considered an added manipulative demand. Although voice dialing features continue to evolve, many phone calls and other textual interactions with cellular technologies continue to rely on manual interactions with the keypad. In recent years, there has been a rapid transition from cellular devices with T9 based input systems and QWERTY keyboards containing tactile buttons to touch screens. Yet there appears to be little research available that provides information regarding the extent to which different interface types may impact a driver's ability to enter information into the device.

In a small driving simulation experiment involving 18 participants, Samuel, Pollatsek and Fisher (2011) investigated glance behavior of frequent and infrequent text messagers that were distributed across two different phone conditions, a QWERTY keypad (Blackberry) and touchpad (iPhone). Although the mean number of glances over 2 seconds was larger for touchpad text messaging, there were no statistical differences in glance behavior. A larger scale effort involving 100 participants (Ranney, Baldwin, Parmer, Martin, & Mazzae, 2011) assessed the distraction potential of contact selection, dialing, and text messaging using a QWERTY keypad (Blackerry) and touchpad (iPhone). Task performance and driving performance were reported by phone type, but not eye glance measures. Target detection was lower during contact selection, car-following delay greater during 10-digit dialing, and the standard deviation of lane position higher during text messaging with the iPhone. In this paper, we aim to extend on this effort by exploring differences in task completion time, drivers' glance behavior, and driving performance during 10-digit dialing with both a manual button style and a touch screen phone interface.

METHODS

Participants

Recruitment methods and experimental content were approved by MIT's institutional review board. The sample was intentionally drawn from a younger age group (20-29 years) likely to have a high proportion of individuals with extensive experience using cell phones. Participants were required to be active, experienced drivers, defined as driving 3 or more times a week and having held a valid driver's license for 3+ years. Additionally, they needed to demonstrate a safe operating history by reporting a driving record free of accidents for the past year. The participant group was considered to be relatively healthy compared to an unscreened community sample based on self-report and specified health exclusion criteria including: major medical illness resulting in any hospitalization in the past 6 months, any neurological problems, treatment for a mental disorder, or regular use of a range of medications (e.g., anti-convulsant, immunosuppressive, cytotoxic, anti-depressant, anti-anxiety, anti-psychotics; medications for major medical conditions such as cancer, hypertension; medication to control heart rate or that causes drowsiness). Eye glasses were set as an exclusion criterion due to the use of eye tracking metrics as a primary dependent variable. Participants were drawn from community volunteers in the greater Boston area who responded to online, print advertisements, or referrals. Compensation of \$60 was provided.

Apparatus

The study was conducted in a driving simulator consisting of a fixed base, full cab Volkswagen New Beetle situated in front of an 8' by 6' (2.44m by 1.83m) projection screen positioned 76" (1.93m) in front of the mid-point of the windshield. This provided approximately a 40 degree view of the virtual world at a resolution of 1024 x 768 pixels. Graphical updates were generated at a minimum frame rate of 20 Hz by STISIM Drive version 2.08.02 (Systems Technology, Inc., Hawthorne, CA) based upon a driver's interaction with the steering wheel, brake and accelerator. Force feedback was provided through the steering wheel and auditory feedback consisting of engine noise, cornering sounds, and brake noise was played through the vehicle's sound system. Audio tasks and instructions were also provided through the vehicle sound system. Driving performance data were captured at 10 Hz. A FaceLAB® 5.0 eye tracking system (Seeing Machines, Canberra, Australia) recorded data at up to 60 Hz. Two video cameras, one mounted in front and one behind and to the side of the driver, captured images of the participant's face and hands to monitor general behavior and interaction with the cell phones. Validation work has established high correspondence between this simulator configuration and on-road behavior in the allocation of visual attention in interactions with visual manipulative human

machine interfaces (HMIs) (Wang et al., 2010) and cognitive demands (Reimer & Mehler, 2011).

The simulation scenario consisted of a divided highway with two lanes in each direction plus a 2 foot (0.61 m) shoulder on each side of the roadway. Lane width was 15 feet (3.62 m) and posted speed limit was 65 mph (104.6 km/h). Typical traffic events on the virtual highway included passing vehicles, lane changes, and slow downs. The average traffic density in the virtual scenario was set at 23 vehicles/mile (14.3/km). Average traffic speed for vehicles in the left lane was set equal to the posted speed limit of 65 mph (104.6 km/h) and 5 mph slower (96.5 km/h) for the right lane.

Two types of phone interfaces were tested: a flip-style phone with tactile buttons that the participant had to physically depress to engage (Samsung Model SCH-A670) and a smart phone with a touch screen representation of the same keyboard layout (original Apple iPhone).

Design and Measures

The assessment of phone dialing took place within the context of a broader study of HMI interactions that included two other real-world (face valid) task types, manipulation of the vehicle radio and address entry in a navigation system. Two surrogate tasks, each of which presented multiple levels of visual demand, were also examined. To control for order effects, tasks were presented in a design in which the three face valid tasks were presented in positions 1, 3 and 5 and the two surrogate tasks in positions 2 and 4. The ordering of the face valid tasks across the three positions was fully counterbalanced so that an equal number of participants experienced each of the six possible task orders. Similarly, half of the participants experienced one surrogate task first and half the other. Within the phone task period, half of the participants were presented the flip-phone first and the other half the iPhone. Finally, distribution of males and females across the various combinations was equal.

Procedure

Participants read and signed an informed consent, eligibility was verified by interview, and a questionnaire that included items related to cell phone experience and type usage was completed. Participants were asked to provide a 10 digit phone number that they knew well that would be used for dialing during the phone task. This was done to avoid demand due to the participant needing to hold an unfamiliar number in memory or direct their gaze at a visual display of a number.

When participants were seated in the simulator, the driver's seat and steering wheel were adjusted so that the individual was comfortable and their eyes and mouth nominally visible for the recording and eye tracking cameras. An eye tracking head model was then created. Recorded audio instructions described the simulator and provided the following guidance and incentive: "During the study, you will receive a monetary award for performing the tasks while you continue driving the simulator. While performance on the

tasks is important, you should balance driving safety while you attempt to complete the tasks, just as you would when driving a real car. Since in the real world you cannot disregard the traffic code, you may be penalized \$2 for every ticket you receive and \$5 for any collision.” These instructions were intended to encourage a realistic balance between secondary task engagement and driving safety and reinforced guidance that was provided in the informed consent form where it was specified that the monetary award for performing the secondary tasks could be up to \$10. In actuality, all participants received equivalent compensation.

A short drive of 2.65 miles (approximately 5 minutes) followed to provide initial familiarization with the simulator. Participants then received instructions to pull over to the side of the highway and stop the car. Recorded instructions concerning the first task set were presented along with supplemental training by a research associate to ensure that participants were able to complete the tasks. Participants then resumed driving. Two minutes after highway speed was obtained, the subsequent 2 minutes were used as a single task reference period. Thirty seconds following, recorded instructions for the first task were presented. Once all tasks for a task set were complete, 2 minutes of single task driving followed, and then the participants were again instructed to pull to the side of the highway and stop. Training for the next task set then began. This same procedure was followed until all task sets were presented. At a stop point approximately mid-way through the simulation session, participants were offered the opportunity to briefly exit the simulator.

During the training portion of the phone task set, participants practiced placing phone calls on both the flip and touch screen phone by dialing their self-selected familiar 10-digit number. When not in use, the phones were located in cup holders in the center console between the driver and front passenger seats. The flip-phone was always stored in the open position. During training, participants were instructed to pick-up a phone from the cup-holder, enter the 10-digit number, press SEND (CALL for the iPhone), then press the END button (CANCEL for the iPhone) to cancel the call, and then to return the phone to the cup-holder. Following the procedure established in the CAMP study (Angell, et al., 2006), participants were trained to say the word ‘done’ after pressing the SEND/CALL button to indicate that they had completed the task. ‘Done’ was used as a marker for determining the end point for timing task duration. The prompt to initiate a task consisted of the recorded instructions, “Your task is to make the phone call using the xxxx phone” where xxxx was either “touch screen” or “flip”. A 2 second pause followed and then the prompt, “Begin”. Participants placed 2 phone calls with one type of phone followed by 2 phone calls with the other. There were 1 minute spacing intervals between the completion of one phone call and the prompt to initiate the next.

Data Reduction & Analysis Periods

Eye data were processed following ISO standards (ISO 15007-1, 2002; ISO 15007-2, 2001) and counts of glances

greater than 1.6 and 2 seconds were computed. The percentage of time spent with eyes on the road was computed as a ratio of glance duration to the roadway by the total length of glance data derived for a task, i.e., computed ignoring saccadic movements and eye closure. Task time was computed by two independent research associates. Differences in scoring were resolved by the lead author.

For continuous random variables, repeated measures ANOVAs were conducted with the following independent variables: cell phone condition, self-reported cell-phone type use, and gender.

RESULTS

Sample Characteristics

Participants were enrolled and run through the protocol until all cells in the counterbalanced design described previously were equally filled with usable cases. The final analysis sample consisted of 36 subjects, half male and half female. Participants ranged in age from 21 to 29 years with an average of 24.6 years (SD = 2.3). The average age of male and female participants was not statistically different, $F(1,34) = .74, p = .40$.

Twenty three out of the 36 participants (64%) reported often using a cell phone that required using a touch screen. The remaining 13 (36%) reported often using a cell phone that did not require interacting with a touch screen. The proportion of men to women who were touch screen phone users (11:12) was not significantly different from the proportion who were non-touch screen phone users (7:6) ($\chi^2(1) = .12, p = .73$).

Time on Task and Error Rates

A total of 5 participants made an error (out of their two trials) when dialing on the flip phone. Only one participant made an error while dialing on the iPhone.

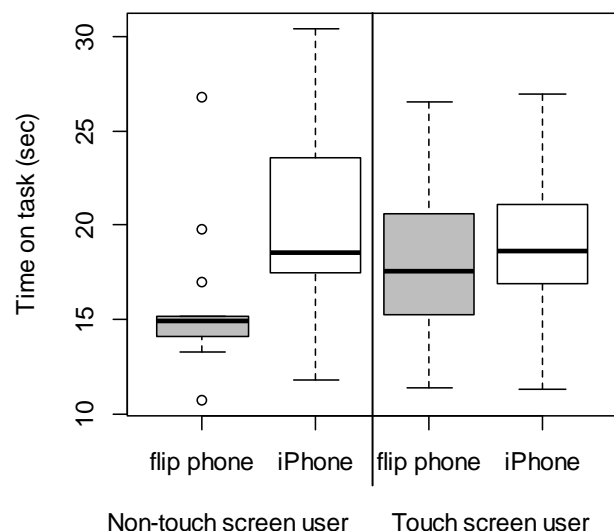


Figure 1. Time on task (sec).

For time on task (Figure 1), there was a significant interaction effect of cell-phone condition and self-reported cell-phone type use ($F(1,34)=6.46$, $p=.02$). Follow-up comparisons revealed that non-touch screen users spent less time on task when they were using the flip phone compared to when they were using the iPhone (mean difference=4.28 sec, $t(34)=3.84$, $p=.0005$). Further, this group of users were also faster on the flip phone than touch-screen users on the iPhone (mean difference=3.05 sec, $t(34)=2.09$, $p=.04$). No other statistically significant differences were found.

Visual Attention

Percentage of time spent with eyes on the road (Figure 2) revealed a significant gender and cell-phone type interaction ($F(1,29)=5.16$, $p=.03$). Females using the flip phone had the highest percentage of time spent with eyes on the road (vs. males flip phone: 16%, $t(29)=3.74$, $p=.0008$; vs. males iPhone: 15%, $t(29)=4.44$, $p=.0001$; vs. females iPhone: 9%, $t(29)=4.66$, $p<.0001$).

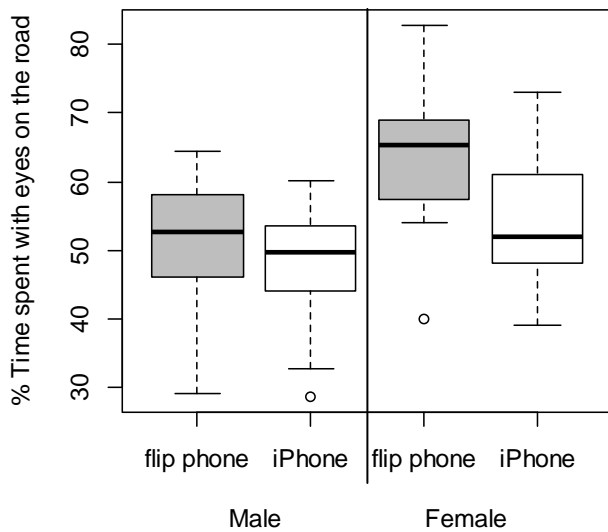


Figure 2. Percentage of time spent with eyes on the road.

Table 1. Frequency of drivers who had at least one glance to the device longer than 2 seconds. (Note: Six cases, 4 male and 2 female, with poor eye quality across one or more data points were not included in this analysis.)

Gender	Phone Type	Had at least one glance to the device greater than 2 sec.	
		Yes	No
Male	Flip phone	8	6
	iPhone	8	6
Female	Flip phone	1	15
	iPhone	6	10

Table 1 shows the number of drivers within the two gender groups whose glances to the devices (flip phone and iPhone) did or did not exceed 2 sec. A logistic regression

model was built to predict the tendency to make glances greater than 2 sec. The model was fitted using PROC GENMOD in SAS 9.1, with the specifications of logit link function and binomial distribution. Repeated measures were accounted for using GEE.

Females using the flip phone had the lowest odds of exhibiting glances greater than 2 seconds (vs. males flip phone: $\chi^2(1)=6.61$, $p=.01$; vs. males iPhone: $\chi^2(1)=6.61$, $p=.01$; vs. females iPhone: $\chi^2(1)=4.94$, $p=.03$).

Driving Performance

There was a significant interaction effect of cell-phone condition and self-reported cell-phone type use with average driving speed ($F(2,68)=3.15$, $p=0.049$) (Figure 3). There was no significant difference between the speed of two user groups within the reference period ($p>.05$). Non-touch screen users slowed down both when they were dialing on the flip phone and the iPhone (reference vs. flip phone: 1.43m/s, $t(68)=4.53$, $p<.0001$; reference vs. iPhone: 1.27m/s, $t(68)=4.00$, $p=.0002$). Touch screen users were found to slow down only when they were dialing on the flip phone (reference vs. flip phone: 0.63m/s, $t(68)=2.64$, $p=.01$). When the two groups were compared, it was found that the flip phone users maintained slower speeds when they were dialing (both flip phone and iPhone) compared to the speeds adopted by the touch screen users in all three experimental conditions (i.e., flip phone, iPhone, and reference) ($p<.05$).

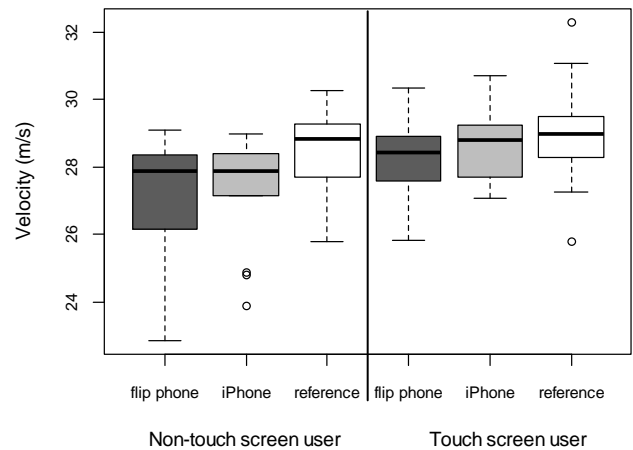


Figure 3. Driving speed (velocity in m/s).

DISCUSSION

Due to the limited research currently available comparing a tactile phone dialing interface with what is likely to soon become the dominant hand-held phone interface in the car, the touch screen, the intent of this study was largely exploratory rather than focusing on testing specific hypotheses regarding possible advantages of one interface type over the other. There was, however, some expectation that advantages might be seen with the traditional button interface due to the lack of

tactile feedback from touch screens; this could be expected to be more of an issue when driving as opposed to more stationary situations since there is an expectation that an individual will need to shift their visual attention back and forth between the phone and the roadway to maintain control of the vehicle.

Perhaps most striking in this study is the finding that those young adults who often use the more traditional manual button interface completed the dialing task significantly faster when using this form of HMI compared to when they used a touch screen, and compared to frequent touch screen users when they used either interface. In this sample, more frequent use of a touch screen interface offered little advantage in working with this form of HMI in terms of completion time under simulated driving conditions.

Some additional advantage is suggested for traditional manual interface in terms of the percentage of time drivers kept their eyes on the road. This effect is most apparent in the female participants who were found to keep their eyes directed toward the roadway an average of 64% of the time using the flip-phone vs. 54% using the touch screen interface. The apparent tactile interface advantage appears in spite of the fact that 67% of the females reported often using a touch screen cell phone for dialing. Complementary behavior is also seen in the fewer number of female participants exhibiting one or more long (2 second) duration glances with the flip-phone as compared to female participants with the iPhone, and male participants with both phone types.

Compensatory behavior, in the form of slowing vehicle speed under secondary task demand, was somewhat greater among frequent manual button interface users, regardless of which HMI type they used during the simulation. It is less clear as to the extent to which this can be considered an advantage since slowing one's normal travel speed may impede traffic and potentially increase risk of collisions with low observant drivers behind the individual dialing the phone. Conversely, this could suggest that these participants were more aware of the demand and safety implications of dialing and were actively attempting to compensate. This suggests that there may be some value in developing a better understanding of what behavioral / psychological linkages may exist between phone style usage and selected driving behavior variables.

While this comparison focuses specifically on handheld phones, the findings may have relevance to the study of the movement from largely tactile HMIs in the automobile to touch screen HMIs. Future work will need to more broadly assess performance with these interfaces. Time on task, gaze data, and driving performance measures all appear to show some sensitivity to the differing demands of the two interface types.

ACKNOWLEDGEMENTS

Support for various aspects of this research was provided by Denso Automotive, the United States Department of Transportation's Region One New England University Transportation Center and the Santos Family Foundation.

REFERENCES

- Angell, L., Auflick, J., Austria, P. A., Kochhar, D., Tijerina, L., Biever, W., et al. (2006). Driver Workload Metrics Task 2 Final Report. Washington, DC: U.S. DOT National Highway Traffic Safety Administration.
- Basacik, D., Reed, N., & Robbins, R. (2012). Smartphone use while driving: a simulator study. Berkshire, UK: Transport Research Laboratory.
- Collet, C., Guillot, A., & Petit, C. (2010a). Phoning while driving I: a review of epidemiological, psychological, behavioural and physiological studies. *Ergonomics*, 53(5), 589-601.
- Collet, C., Guillot, A., & Petit, C. (2010b). Phoning while driving II: a review of driving conditions. *Ergonomics*, 53(3), 602-616.
- Drews, F. A., Yazdani, H., Godfrey, C. N., Cooper, J. M., & Strayer, D. L. (2009). Text messaging during simulated driving. *Human Factors*, 51(5), 762-770.
- Evans, L. (1991). *Traffic safety and the driver*. New York: Van Nostrand Reinhold.
- Hoffman, J. D., Lee, J. D., McGehee, D. V., Macias, M., & Gellatly, A. W. (2005). Visual sampling of in-vehicle text messages: effects of number of lines, page presentation, and message control. *Transportation Research Record: Journal of the Transportation Research Board*, 1937, 22-30.
- Horrey, W. J., & Lesch, M. (2008). Assessing the awareness of performance decrements in distracted drivers. *Accident Analysis & Prevention*, 40(2), 675-682.
- 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. G., Young, K. L., & Regan, M. A. (2009). The effects of text messaging on young drivers. *Human Factors*, 51(4), 582-592.
- ISO 15007-1. (2002). Road vehicles - Measurement of driver visual behaviour with respect to transport information and control systems - Part 1: Definitions and parameters. Geneva, Switzerland: International Standards Organization.
- ISO 15007-2. (2001). Road vehicles - Measurement of driver visual behaviour with respect to transport information and control systems - Part 2: Equipment and procedures. Geneva, Switzerland: International Standards Organization.
- Lee, W., & Cheng, B. (2010). Comparison of portable and onboard navigation system for the effects in real driving. *Safety Science*, 48(10), 1421-1426.
- Ranney, T. A. (1994). Models of driving behavior: A review of their evolution. *Accident Analysis & Prevention*, 26(6), 733-750.
- Ranney, T. A., Baldwin, G. H. S., Parmer, E., Martin, J., & Mazzae, E. N. (2011). Distraction effects of manual number and text entry while driving. Washington, DC: U.S. DOT National Highway Traffic Safety Administration
- Ranney, T. A., Harbluk, J. L., & Noy, Y. I. (2005). Effects of voice technology on test track driving performance: implications for driver distraction. *Human Factors*, 47(2), 439-454.
- Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine*, 336(7), 453-458.
- Reed, M. P., & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialing task. *Ergonomics*, 42(8), 1015-1037.
- Reimer, B., & Mehler, B. (2011). The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation. *Ergonomics*, 54(10), 932-942.
- Samuel, S., Pollatsek, A., & Fisher, D. L. (2011). Texting while driving: Evaluation of glance distributions for frequent/infrequent texters and keypad/touchpad texters. *Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Lake Tahoe, CA, 424-432.
- Schreiner, C. S. (2006). The effect of phone interface and dialing method on simulated driving performance and user preference. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, San Francisco, CA, 2359-2363.
- Violanti, J. M. (1997). Cellular phones and traffic accidents. *Public Health*, 111, 423-428.
- Violanti, J. M. (1998). Cellular phones and fatal traffic collisions. *Accident Analysis and Prevention*, 30(4), 519-524.
- Violanti, J. M., & Marshall, J. R. (1996). Cellular phone and traffic accidents: an epidemiological approach. *Accident Analysis and Prevention*, 28(2), 265-270.
- Walsh, S. P., White, K. M., Hyde, M. K., & Watson, B. (2008). Dialling and driving: Factors influencing intentions to use a mobile phone while driving. *Accident Analysis & Prevention*, 40(6), 1893-1900.
- Wang, Y., Reimer, B., Mehler, B., Lammers, V., D'Ambrosio, L. A., & Coughlin, J. F. (2010). The validity of driving simulation for assessing differences between in-vehicle informational interfaces: a comparison with field testing. *Ergonomics* 53(3), 404-420.
- Wilson, F. A., & Stimpson, J. P. (2010). Trends in fatalities from distracted driving in the United States, 1999 to 2008. *American Journal of Public Health*, 100(11), 2213-2219.