

AWARENESS OF PERFORMANCE DECREMENTS DUE TO DISTRACTION IN YOUNGER AND OLDER DRIVERS

William J. Horrey, Mary F. Lesch, & Angela Garabet
Center for Safety Research
Liberty Mutual Research Institute for Safety
Hopkinton, MA USA
E-mail: william.horrey@libertymutual.com

Summary: Although many studies have documented the performance decrements associated with driver distractions, few have examined drivers' awareness of these distraction effects. The current study purports to measure how well-calibrated drivers are with respect to their own performance when distracted. Forty drivers completed a series of tasks on a hand-held or hands-free cell phone while driving an instrumented vehicle around a closed test track. Subjective estimates of performance decrements were recorded and compared to actual decrements observed on multiple measures of driving performance. Although their driving performance suffered in dual-task conditions, drivers were generally not well-calibrated to the magnitude of the distraction effects ($r = -.38$ to $.16$). In some cases, estimates of distraction were opposite of the observed effects (i.e., smaller estimates of distraction corresponded to larger performance deficits). There were some age and gender differences. We discuss the implications of these findings for potential mitigation strategies for distracted driving.

Over the past several years, numerous studies have documented the effects of engaging in cell phone conversations or other distracting tasks while driving. In general, these studies show that distracted drivers have slowed responses to critical traffic events or to discrete stimuli and are more likely to miss external events such as a changing traffic light, among other effects (Alm & Nilsson, 1994; Hancock, Lesch & Simmons, 2003; Horrey & Wickens, 2006; McKnight & McKnight, 1993; Strayer & Johnston, 2001). Much of the focus in such studies has been on detecting the presence and measuring the magnitude of performance decrements for different types of in-vehicle tasks or device configurations. However, to date there has not been much examination of drivers' awareness of these distraction effects.

The perception or awareness of distraction effects may influence drivers' decisions or their willingness to engage in distracting activities while on the road. For example, drivers that are not calibrated with respect to the magnitude of distraction effects may engage in activities because they don't realize their performance is compromised. In an earlier study, Lesch and Hancock (2004) found that a priori ratings of confidence in dealing with distracting tasks were not related to actual performance while distracted for female drivers, suggesting that some drivers may not be aware of performance decrements while distracted. In their study, younger and older drivers performed a memory and recall test while navigating an instrumented vehicle around a closed test track (some data is presented by Hancock et al., 2003). On some trials, drivers were presented with a visual prompt for the memory task that occurred in conjunction with a changing traffic light. Measures of stopping performance were assessed and compared against confidence

ratings. Overall, there was little association between the performance and confidence in female drivers; however, the disparities were greatest for older drivers.

In the current study, we wished to establish how well-calibrated drivers are with respect to distraction effects—that is, whether drivers' estimates of the magnitude of distraction reflect actual performance decrements on several driving tasks. We expand on the results from Lesch and Hancock (2004) using a distraction task that did not require visual processing. Furthermore, we were interested in whether subjective estimates of distraction and actual distraction varied as a function of phone type (hand-held, hands-free). Drivers in this study drove an instrumented van on a closed test track while performing a continuous mental arithmetic task administered over a hand-held or hands-free cell phone. Subjective measures of distraction effects were recorded and compared to actual performance on multiple measures of driving performance.

METHODS

Participants

Forty drivers, divided into two age groups, were recruited through advertisements in local newspapers. The younger driver group consisted of 20 drivers between the ages of 18 and 34 ($M = 21.8$ yrs, $SD = 4.3$). The older driver group consisted of 20 drivers between the ages of 55 and 82 ($M = 64.1$ yrs, $SD = 7.7$). Males and females were (approximately) balanced across the two age groups. All had normal or corrected-to-normal visual acuity. Drivers were paid \$20 for each hour of participation.

Materials

The experiment was conducted on a 0.5-mile (0.8 kilometer) closed-loop test track. The two-lane track was delineated to allow for continuous driving. A signalized intersection, located at the end of a straightaway, was controlled through track-based infrared sensors and vehicle-based GPS and speed information.

The instrumented vehicle was a 2002 Ford Windstar minivan, outfitted with several sensors and computers. A PC computer rack, mounted directly behind the driver, controlled and coordinated various aspects of the data acquisition and stimulus presentation. Auditory stimuli for the phone tasks were presented via handheld phone (Avaya model 3626) or through speakers mounted behind the driver (Sony model SRS-T100PC).

Five pace clocks were positioned at various points around the track. These 0.5-m (20-inch) diameter clocks were mounted on 1.8 m (6 foot) towers and placed adjacent to the drivers' lane. The bottom half of the clock was green and the top half was red. The arrow indicator moved around the clock at a constant rate and completed the full rotation every 12 seconds, on average (varied across clocks).

Procedure

At the start of the two-hour session, drivers completed an informed consent form, tests for visual functioning and several questionnaires. Drivers were then introduced to the safety features of the instrumented van and given several minutes of practice to familiarize themselves with the handling of the vehicle, the driving tasks, and the track.

For all conditions, drivers were allowed to select their speed (to an upper limit of 48 kph [30 mph]). When approaching the pace clocks, drivers were told to adjust their speed, either by accelerating or braking, in order to pass the clock when the arrow indicator was in the green portion of the clock (and not the red). In doing so, they were told to avoid bringing the vehicle to a full stop and exceeding the speed limit. On a random subset of trials (37.5%), the traffic light at the signalized intersection changed from red to green. On these occasions, drivers were instructed to bring the vehicle to a complete stop as quickly as they could and to try to stop before they reached a stop line marked by two traffic cones.

During the experimental blocks, drivers performed a concurrent phone task—a variation of the Paced Auditory Serial Addition Task (PASAT; Gronwall, 1977). Numbers between 1 and 9 were randomly presented every 7 seconds, either through the hand-held phone or through the vehicle speakers (hands-free). The task involved adding two consecutive numbers and responding verbally. Drivers were given a small monetary bonus for each correct response (5 cents per correct response, up to \$20 over the whole experiment).

Drivers completed three blocks of 8 laps, each lasting approximately 15-minutes. In two blocks, drivers performed the PASAT (once using the hand-held phone, once using the hands-free system). In the third block, drivers performed the driving tasks alone, with no PASAT. Drivers also completed a short, 3-minute baseline block for the PASAT, performed while the vehicle was stationary. The order of blocks was counterbalanced across driver.

Measures of pace clock accuracy (number of correct clock events out of 45 per block), braking response time (from the onset of the traffic light change) and stopping errors (failing to stop or failing to stop before the stop line) were recorded. In between blocks, drivers completed a modified NASA TLX (Hart & Staveland, 1988) that included subjective ratings of performance for all of the relevant driving sub-tasks (pace clocks, stopping task) and for the PASAT.

RESULTS

Driving Performance

Prior to examining drivers' calibration of distraction effects, we examined the dual-task costs in driving performance associated with the cell phone tasks (i.e., a manipulation check). In general, we used a mixed ANOVA with the variables of Task, Age and Gender (with the exception of the stopping errors, which were not normally distributed; for these we employ non-parametric tests). There were no significant effects of age and gender on any of the measures of driving performance.

Table 1 shows the mean performance scores across dependent measure and task condition. There was a significant effect of task condition across all measures of performance: pace clock accuracy ($F(2,68) = 27.1, p < .01$); braking response time ($F(2,60) = 5.8, p < .01$); and stopping errors ($\chi^2_F(2) = 16.0, p < .001$). The hand-held and hands-free PASAT conditions yielded poorer performance on all metrics than in the single-task driving baseline (all p 's $< .01$); however, in general, there were no differences between hand-held and hands-free cell phones, replicating previous results (e.g., Strayer & Johnston, 2001; Horrey & Wickens, 2006).

Table 1. Summary of driving performance data by task condition

Task Condition	Pace Clock Accuracy (%)	Brake Response Time (s)	Stop Light Errors (%)
Baseline	82.2 (1.2)	0.88 (0.02)	26 (6)
Hand-held	71.1 (1.6)	0.96 (0.03)	48 (6)
Hands-free	73.9 (1.2)	0.94 (0.03)	48 (5)

Note. Standard error appears in parentheses.

Calibration to Distraction Effects

In order to examine drivers' calibration to the effects of in-vehicle distraction, we first calculated estimated and actual distraction effects from the subjective performance ratings and from actual performance along the various measures of driving performance. Subjective estimates of distraction were expressed as a % difference between self-rated performance in the driving baseline block and self-rated performance in the distraction block (the baseline condition was used as a reference in order to control for any systematic biases in ratings). Actual distraction effects were expressed as the % difference in performance from the baseline block to the dual-task block. Thus, negative values denoted an actual or estimated loss in performance for the distraction condition relative to baseline, whereas a positive value indicated a gain in performance.

Next, we examined the relationship between subjective estimates of distraction effects and actual distraction effects. If drivers are well-calibrated, we would expect a positive correlation between the estimated and actual distraction effects. That is, larger estimations of distraction effects would correspond to larger observed performance decrements. However, if drivers are not well-calibrated to the magnitude of distraction effects, then there would be no such relationship or a negative relationship may exist.

The results from a correlational analysis are shown in Table 2. Overall, there were no significant relationships between estimates of distraction effects and actual performance decrements, lending support to the notion that drivers were not well-calibrated to the distracting effects of a concurrent in-vehicle task. In the hands-free condition, there was a significant relationship for stopping errors ($r = -.38$); however, this relationship is in the negative direction. Here, drivers that estimated the smallest performance decrements were actually exhibiting the largest ones! The confidence intervals around the correlations in Table 2 suggest that, in general, the poor

calibration of drivers to the effects of distraction were consistent across phone type and dependent measure.

An examination of the different driver groups revealed some interesting findings. For example, older male drivers were actually well-calibrated to the magnitude of distraction effects for the stopping task (stopping errors, $r = .71$; brake RT, $r = .82$). In contrast, younger males showed strong associations in the opposite direction (stopping errors, $r = -.91$; brake RT, $r = -.65$, the latter result was marginally significant). That is, young male drivers that thought they were doing better were actually doing worse than others. In general, female drivers did not exhibit any significant relationships between estimated and actual performance loss (Lesch & Hancock, 2004).

Table 2. Correlations between the magnitude of estimated and actual distraction effects (the difference between baseline and distraction conditions)

Driver Group	Δ Brake Response Time		Δ Pace Clock Accuracy		Δ Stop Light Errors	
	Held	Free	Held	Free	Held	Free
Overall	.06 (-.28, .39)	-.02 (-.33, .31)	.16 (-.18, .46)	.10 (-.22, .40)	.11 (-.23, .42)	-.38* (-.62, -.07)
Young Males	-.65 (-.93, .10)	-.33 (-.79, .38)	-.07 (-.70, .62)	.30 (-.41, .78)	.05 (-.64, .69)	-.91** (-.98, -.66)
Older Males	.82** (.34, .96)	.45 (-.26, .84)	.40 (-.36, .84)	-.22 (-.74, .48)	.30 (-.46, .80)	.71* (.15, .93)
Young Females	.07 (-.62, .70)	.13 (-.58, .73)	.42 (-.34, .85)	.51 (-.18, .86)	.34 (-.37, .80)	-.61 (-.91, .09)
Older Females	.38 (-.38, .83)	-.04 (-.65, .60)	-.23 (-.77, .52)	-.29 (-.78, .41)	-.29 (-.80, .46)	-.56 (-.88, .11)

Note. 95% confidence intervals are shown in parentheses. * $p < .05$, ** $p < .01$.

DISCUSSION

In the current study, younger and older drivers were asked to complete a phone task on a hand-held or hands-free cell phone while navigating a closed test-track in an instrumented vehicle. Compared to baseline driving, we observed dual-task decrements on all measures of driving performance, replicating previous results (Alm & Nilsson, 1994; Hancock et al., 2003; McKnight & McKnight, 1993; Strayer & Johnston, 2001).

To examine drivers' calibration to distraction effects, we compared drivers' subjective estimates of distraction with actual performance decrements, based on a number of measures of driving performance. The results suggest that, for the most part, drivers are not well-calibrated to the distracting effects of a hand-held or hands-free cell phone conversation. Across all measures of performance, subjective estimates of distraction effects were not related to the actual magnitude of the distraction. In some cases, driver's estimates of distraction were opposite of what was observed in actual performance. That is, drivers that estimated the smallest (or no) distraction effects exhibited the largest ones. In general, a disconnection between performance and awareness was observed across driving measure and phone type.

Analysis of the driver groups, though limited by a small sample size, showed some differences. For female drivers, there were no significant relationships across all measures and phone types (also as shown by Lesch & Hancock, 2004); however, there were some mixed results for males. For example, on some measures young and older males revealed a nearly equal but opposite relationship between estimated and actual distraction effects. Younger males, in particular, were poorly calibrated, suggesting that this may be an important group for targeted remediation.

The willingness to engage in distracting activities may be a function of drivers' perceptions of performance decrements. As such, drivers may engage in distracting activities simply because they do not realize that their performance is degraded or they may be overconfident in their skills and their ability to deal with distractions while behind the wheel (Wogalter & Mayhorn, 2005).

Strategies for improving drivers' calibration to distraction effects include both driver- and technology-based approaches. For the former, training to recognize or attend more closely to their driving activities may help drivers' determine when their performance is below "baseline." One question that remains is whether drivers would continue to perform in-vehicle tasks if they were perfectly calibrated to the distraction effects. In other words, the issue could become one of risk tolerance and the degree to which drivers are willing to take on the additional risk of the in-vehicle task.

Some technological innovations are intended to monitor driver performance and behaviors in real time and provide alerts when the system infers that a distracting activity is inappropriate (e.g., Donmez, Boyle, & Lee, in press). Thus, these systems have some authority in deciding when the distraction effects of in-vehicle activities are too severe. However, discrepancies between the driver and the system's estimated level of distraction could create some conflicts. For example, a driver's mental model of the system may be violated if warnings are issued for events which the driver does not believe to be problematic. The consequences of such a mismatch could include system disuse deriving from reduced trust in the system (Lee & Moray, 1994; Parasuraman & Riley, 1997) or worse (Sarter & Woods, 1995). Providing the raw data underlying the system functions may be a key consideration.

ACKNOWLEDGMENTS

The authors are grateful to Peter Teare, Eric Jones, Richard Holihan, and Ed Correa for their upkeep and maintenance of the instrumented van and related equipment. We also thank Yueng-hsiang Huang and Michael Blair for their helpful comments regarding an earlier version of this paper.

REFERENCES

- Alm, H. & Nilsson, L. (1994). Changes in driver behaviour as a function of handsfree mobile phones—A simulator study. *Accident Analysis and Prevention*, 26, 441-451.
- Ashley, S. (2001). Driving the info highway. *Scientific American*, 285, 52-58.
- Brookhuis, K.A., de Vries, G., & de Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention*, 23, 309-316.

- Donmez, B., Boyle, L.N., & Lee, J.D. (In press). Safety implications of providing real-time feedback to distracted drivers. *Accident Analysis and Prevention*.
- Gronwall, D.M.A. (1977). Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual and Motor Skills*, 44, 367-373.
- Hancock, P.A., Lesch, M., & Simmons, L. (2003). The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis and Prevention*, 35(4), 501-514.
- Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North-Holland, 139-183.
- 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.
- Lee, J.D. & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, 40, 153-184.
- Lesch, M.F. & Hancock, P.A. (2004). Driving performance during concurrent cell-phone use: Are drivers aware of their performance decrements? *Accident Analysis and Prevention*, 36, 471-480.
- McKnight, A.J. & McKnight, A.S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention*, 25, 259-265.
- Parasuraman, R. & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230-253.
- Sarter, N.B. & Woods, D.D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors*, 37(1), 5-19.
- Strayer, D.L. & Johnston, W.A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science*, 12(6), 462-466.
- Wogalter, M.S. & Mayhorn, C.B. (2005). Perceptions of driver distraction by cellular phone users and nonusers. *Human Factors*, 47(2), 455-467.