Pergamon

Journal of Safety Research 35 (2004) 453–464

www.elsevier.com/locate/jsr

www.nsc.org

Effects of naturalistic cell phone conversations on driving performance

Michael E. Rakauskas*, Leo J. Gugerty, Nicholas J. Ward

111 Church Street, Minneapolis, MN 55455, United States

Received 21 July 2003; received in revised form 11 March 2004; accepted 28 June 2004

Abstract

Problem: The prevalence of automobile drivers talking on cell phones is growing, but the effect of that behavior on driving performance is unclear. Also unclear is the relationship between the difficulty level of a phone conversation and the resulting distraction. **Method**: This study used a driving simulator to determine the effect that easy and difficult cell phone conversations have on driving performance. **Results**: Cell phone use caused participants to have higher variation in accelerator pedal position, drive more slowly with more variation in speed, and report a higher level of workload regardless of conversation difficulty level. **Conclusions**: Drivers may cope with the additional stress of phone conversations by enduring higher workloads or setting reduced performance goals. **Impact on Industry**: Because an increasing number of people talk on the phone while driving, crashes caused by distracted drivers using cell phones will cause disruptions in business, as well as injury, disability, and permanent loss of personnel.

Keywords: Driving; Distraction; Cell phone; Conversation; Simulator

1. Introduction

There is a growing trend in the United States for the use of mobile communication technology. By the beginning of 2004, more than 154 million people in the United States were cell phone subscribers (Cellular Telecommunications Industry Association, 2004). The portability of these phones has made their use a common occurrence among drivers. This raises concern about the increased risk of motor-vehicle collision associated with mobile phone use while driving (Redelmeier & Tibshirani, 1997). For example, Canadian research has suggested that the risk of having a serious crash involving injury to the driver is increased by 38% when a cell phone is being used (ITS America, 2002a). Moreover, a driver using a cell phone is 16% more likely to have caused the crash.

The etiology of this increased crash risk during cell phone use may involve a number of sources. First, cell phone users may behave differently than nonusers in the driving context even when not using the phone. For example, "cell phone users tend to have more violations for speeding, impaired driving, seat belt non-use, aggressive driving and lifestyle attitudes and personality" (ITS America, 2002a, ¶ 9). This suggests that the disposition of cell phone users predisposes them to greater crash risk (Response Insurance, 2003).

Second, the operation of the cell phone and engagement of conversation may cause impairment of driver performance. The primary task of the driver is to control the vehicle with respect to monitored hazards in the environment. This task demands motor control and cognitive resources (attention). A driver interacting with a cell phone must also apply control and attention to operate the phone and participate in the conversation. To the extent that this secondary task diverts critical control and attention resources away from driving, the performance of the driver may be

* Corresponding author. Tel.: +1 612 624 4614; fax: +1 612 624 8930.
E-mail address: mickr@me.umn.edu (M.E. Rakauskas).

0022-4375/$ - see front matter © 2004 National Safety Council and Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsr.2004.06.003

impaired with respect to the task objectives of driving: namely, mobility and safety.

Evidence of impairment while a driver is using a cell phone has been observed in several experimental studies that have been extensively reviewed (e.g., Goodman et al., 1997; McKnight & McKnight, 1993). This research looked at attention distraction with respect to the different stages of operating a cell phone, such as dialing (Kantowitz & Hanowski, 1996; Reed & Green, 1999), answering (Waugh et al., 2000), and holding the phone (Brookhuis, de Vries, & de Waard, 1991; Haigney, Taylor, & Westerman, 2000). However, the effect distraction has on impairment is not solely attributable to the manual resources applied to holding and operating a phone, or even the visual attention needed to locate and monitor the device (Strayer & Johnston, 2001). Studies using hands-free phones, which do not have a manual component and require only a limited visual demand to operate, have also shown driver impairment related to distraction (Alm & Nilsson, 1995; Briem & Hedman, 1995; Haigney et al., 2000; McKnight & McKnight, 1993; Parkes & Hooijmeijer, 2000; Strayer & Johnston, 2001; Strayer, Drews & Johnston, 2003). In this case, the cognitive effort needed to participate in a cell phone conversation is also an impairment factor. That is, "holding the phone isn't the main issue—thinking is" (ITS America, 2002b, ¶ 4).

Such impairment from cell phone use distraction is dangerous because of the driver's limited capacity to share task resources while adequately monitoring and controlling the safe path of the vehicle in the traffic environment. Distraction-related crashes result from the unexpected onset of a traffic hazard while the driver's attention is diverted to the demands of the cell phone interaction (Ranney, Mazzae, Garrott, & Goodman, 2000). As a result of this growing evidence of the crash risk associated with cell phone use, crashes and personal injury related to cell phone use are becoming significant liability issues and culpable drivers are being prosecuted for involvement in distraction-related crashes (Glater, 2002). Moreover, policymakers are beginning to consider the need for regulations that would govern the use of cell phones while driving (Sundeen, 2001). Many other countries have already established legislation to ban cell phones in some form, but only a few states in the United States have imposed bans (Cellular-News, 2003). As of 2001, only New York had a legislated ban on the use of (hand-held) cell phones while driving, although Massachusetts and Illinois ban cell phone use for bus drivers, and New Jersey prohibits the use of cell phones for provisional drivers. Several other states are currently debating legislation, but many others have either proposed bans that have subsequently failed to be legislated, or have moved responsibility for setting such bans to local city councils. Even so, new evidence suggests that such bans do not affect long-term behavior of drivers without sustained enforcement and publicity (Royal, 2003).

Nevertheless, to support more effective policy decisions and litigation, it is necessary to continue investigating the potential distractions from cell phone use that may impair driver performance and increase crash risk. This need becomes elevated as new transportation services, such as OnStar and 511 Traveler Information Services, become available and can be accessed from cell phones while driving. As the prevalence of cell phones expands and motivations to use them while driving increase, the research knowledge base must also expand to provide informed policy for the deployment and design of these devices and services.

To gain a more complete understanding of how attention is diverted away from driving when using a cell phone, it is necessary to look at how one's attention is utilized both by driving and during conversation tasks when using a cell phone. This requires the study of representative task demands using realistic driving situations and naturalistic conversations. In this regard, there is a concern that some of the previous research on driver distraction while using cell phones has used artificial tasks, such as mathematical or verbal tests, as content for simulated conversations. Mathematical tests have included computations or recognition of presented digits relative to a memorized numerical set (Hancock, Lesch, & Simmons, 2003). For verbal tests, some researchers have used a shadowing technique, which simply requires a subject to repeat a word that was just stated (Strayer & Johnston, 2001). Other research has employed a word generation task that only requires participants to create a word based on the last word stated (Strayer & Johnston, 2001), or to answer extremely simple questions (Irwin, Fitzgerald, & Berg, 2000; McKnight & McKnight, 1993). Some studies have used well-known verbal tasks that require listening to sentences, remembering elements of the sentences, and then repeating some of the words just heard in the correct sequence (Brown, Tickner, & Simmonds, 1969; Haigney et al., 2000; Waugh et al., 2000). Whereas such tasks may be practical to implement and quantify, they represent neither typical conversations nor the demands drivers engage in when using a cell phone.

Haigney and Westerman (2001) concluded that although structured verbal tasks have some advantages, more generalizable forms of conversation must be implemented in experimental settings to provide external validity to the research conclusions. For one, typical conversation includes elements of memory and emotional engagement. Moreover, conversations employed in research should extend across a range of complexity so that the conclusions relate to different contexts and applications. Since this study employed naturalistic conversation tasks and explicitly varied the difficulty of the conversations, other studies that employed varied conversation tasks to study their effects on driving are described next.

Some studies have attempted to use conversation tasks that are more natural in form and content. For example, Irwin et al. (2000) had participants hold a phone or not hold a phone and listen to weather information, answer simple one or two word questions, answer questions of greater depth, or answer questions that were meant to elicit an emotional response. The authors asked participants to keep their foot on a simulated gas pedal and to initiate a braking response with a separate brake pedal when a red light was activated. Even though this guidance-related braking response was the only part of the normal driving experience that was examined, the researchers showed that holding and listening to a phone significantly increased the reaction time (RT) of a braking response. While there was a difference between the control condition and using a phone, whether hand-held or hands-free, no significant difference in RT was found between the two types of phone. This finding suggests that merely listening to a conversation over a phone interferes with the processes of scene perception inherent when considering guidance actions, regardless of the difficulty of the conversation.

McKnight and McKnight (1993) also used a wide variety of verbal distracter tasks in an experiment examining how drivers' reactions changed when their attention was taxed. They had participants drive using a mock steering wheel and pedals while observing videotaped scenes. While "driving," participants were also asked to tune a radio, dial and place a call, converse casually with the test administrator, or verbally perform problem-solving exercises. These scenes were of situations warranting some type of vehicle control action and were filmed from the driver's viewpoint. Though the driving scenes did not change in response to the participant's use of the wheel, pedals, or turn signals, the experimenters counted any manipulation of these controls as a response to the scene. As expected, there were significantly fewer responses to scenes under conditions of radio tuning, placing a call, and both casual and intense conversation than in the no-distraction condition. Intense conversation was also found to be more distracting than easy conversation.

Whereas the preceding examples of research on cell phone distraction and driving can be commended for their use of secondary tasks that resemble the nature and range of natural conversations, they are nonetheless limited by the artificial nature of the primary task in the simulated driving environment. Thus, while the workload imposed by the conversation tasks may be considered realistic and representative, the results from the driving performance measures may have limited generalizability to actual driving.

The current research incorporated two types of natural conversation within a moderately realistic driving simulator. Participants drove through three 10-minute trials and were exposed to an unpredictable hazardous event in each of the trials. Performance was measured in various aspects of vehicle control and guidance as well as mental workload. Participants drove two trials while performing conversation tasks of two different difficulty levels and drove one trial without conversing. It was hypothesized that driving while engaged in the cell phone conversations would result in impaired primary task (driving) performance and higher mental effort. It was also hypothesized that the degree of impairment and effort would be greater when the participant was engaged in the more complex (demanding) conversation.

2. Method

2.1. Participants

The 24 participants (12 males, 12 females, mean age = 20.4 years, range = 18 to 32 years) were given monetary compensation for their participation. All participants had a valid drivers license and at least two years of driving experience (mean driving experience = 4.7 years). None of the participants had previous experience with this driving simulator.

2.2. Materials

2.2.1. Driving simulator

A GlobalSim Corporation driving simulator was used for the study. This simulator had a 150° (horizontal) by 40° (vertical) forward field of view and 50° (horizontal) by 40° (vertical) rear field of view available through a side-view mirror. The physics model built into the simulator simulated the dynamics of a four-cylinder Ford Taurus. The "car" body was a wooden car mock up, including a steering wheel, brake and accelerator pedals, side view mirror, and shifter (with park, reverse, and forward gears). Speed, in miles per hour (mph), was displayed at the bottom front of the participant's forward field of view.

2.2.2. Driving environment

The driving environment developed for the experiment consisted of two-lane roads in a rural setting. As shown in Fig. 1, the roads made a closed circuit with a crossroad intersection at the center controlled by a stoplight. There were stop signs at all of the "T" intersections around the outer ring road, positioned so that only traffic entering the ring road was signaled to stop. Ambient traffic was present throughout the environment and was controlled so that driving conditions remained constant for all participants. There were numerous cars parked on the side of the road to act as distracters for the following staged hazardous driving events:

- Pull Out: a parked car pulls out in front of the driver.
- Swerve: an oncoming car swerves in front of the driver.
- Run Red Light: an ambulance runs a red light in front of the driver.

Invisible triggers along the driving course initiated the hazardous events (or changes in ambient traffic conditions) when the driver passed over them.

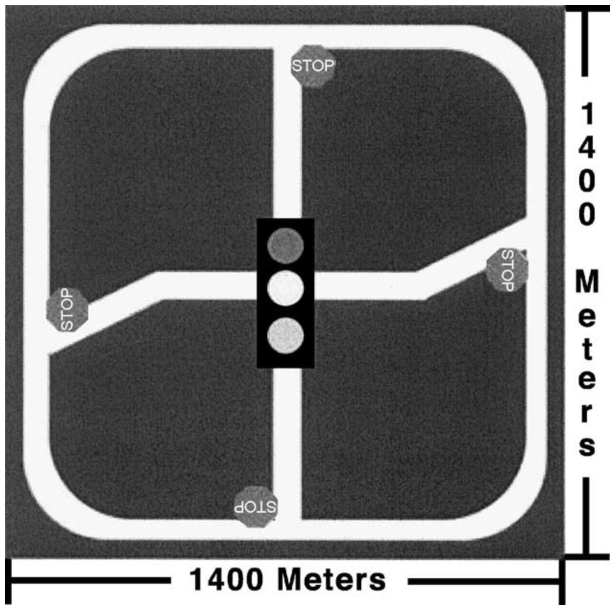

Fig. 1. Schematic of simulated driving environment.

2.2.3. Conversation questions

There were two difficulty levels of conversation questions (easy and difficult), which were validated through pilot testing. Potential questions were based on preliminary data collected from undergraduates at Clemson University and from *The Book of Questions* (Stock, 1985). These students were asked to list types of cell phone conversation topics they felt were and were not distracting to driving based on their own experiences. Special care was taken to avoid including questions that would interfere with visual/spatial tasks.

The complete list of questions was given to three pilot participants (mean age = 23.3 years) and the average time it took to answer each question was recorded. Then the questions were presented to another group of nine pilot participants (mean age = 24.3 years) who were asked to read each question and think about how they would answer. They then rated each question in terms of difficulty using a scale of 1 (easiest) to 5 (most difficult). Questions that had an average difficulty rating less than or equal to 2 were used as easy conversation questions (34 questions, mean rating = 1.74). Some examples of easy conversation questions include:

- What is your major? Why do you find that interesting?
- What do you plan on doing tomorrow afternoon?
- Are you free to meet me next Monday at 11:30 a.m.? If not, when are you free to meet?

Questions that had an average rating greater than or equal to 3 were used as difficult conversation questions (35 questions, mean rating = 3.40). Some examples of difficult conversation questions include:

- If a new medicine were developed that would cure arthritis but cause a fatal reaction in 1 percent of those who took it, would you release it to the public? Why/why not?
- Do you think that the world will be a better or worse place 100 years from now? In what ways? Give some examples.

The difference between the rated difficulties of the two groups of questions was statistically significant, t (67) = 21.58, $p < .001$, but the time required to answer each group was comparable (not significantly different). Thus, the perceived difficulty between the conversation levels was not confounded with the time required to respond.

For both levels of question difficulty, preplanned follow-up questions were asked, if applicable, to facilitate the elaboration of a response. The questions for a particular difficulty level were asked together during a continuous driving trial. That is, the two types of questions were blocked, not intermixed. For each difficulty level, questions were presented in two different random orders so that half of the participants experienced each order. The experimenter asked the next question on the list after the participant finished his or her response until the end of the trial.

2.3. Design

This study examined the effects of cell phone conversation on driving using a within-subjects design. There were three levels of conversation difficulty as an independent variable: (a) control (no conversation), (b) easy conversation task, and (c) difficult conversation task. Each participant performed three driving trials, with each trial assigned a level of conversation. Trials lasted approximately 10 minutes when driven at 45 mph, and each trial was broken into two paths separated by a break, thereby creating six total driving paths.

The hazardous events were embedded within the paths so that there was one event during each of the three trials. These occurred between the second and fourth minute of a 5-minute path. Each participant was exposed to each of the three hazardous events once during the entire experiment. As shown in Table 1, the ordering of the paths and the placement of the hazardous events were distributed across the three trials in two ordered sequences. The subjects were evenly divided between these sequences, so that the type of hazard event was not confounded with trial number (first,

Table 1
Path running order and hazard distribution over the three trials

Trial:	Order A	Order B
1st	1, 2(P)	5(R), 6
2nd	3(S), 4	3, 4(S)
3rd	5, 6(R)	1(P), 2

1–6=Paths 1 through 6.
P=Pull Out Hazard.
S=Swerve Hazard.
R=Run Red Light Hazard.

second, or third). The ordering of conversation level was counterbalanced across subjects to avoid confounding with trial number or type of hazardous event.

2.4. Measures

Measures were included for the mobility and safety objectives of the primary driving task. Measurements of speed maintenance, lane-keeping, and crash avoidance were computed for path sections that did not include intersections or curves. Speed and lane-keeping measures were recorded at a frequency of 4 Hz, and crash avoidance measures at 20 Hz.

2.4.1. Speed maintenance

The following variables were computed based on data recorded by the simulator:

- Accelerator position variability, based on the standard deviation of the accelerator pedal position (0 = released to 1 = fully depressed).
- Speed variability, based on the standard deviation of driving speed (mph).
- Average speed, based on mean driving speed (mph).

2.4.2. Lane position maintenance

The following variables were computed based on data recorded by the simulator:

- Steering offset, calculated as the standard deviation of the distance that the top-most point of the steering wheel moved from center (degrees, negative for left of center and positive for right of center).
- Mean lateral speed, calculated as the mean of the lateral distance that the participant's car traveled per second (feet per second).

2.4.3. Crash avoidance

The following variables were computed based on data recorded by the simulator:

- Collisions with other vehicles, recorded as a binary variable and presented as a percentage of total events.
- Reaction time (RT), calculated in seconds (s) as the time elapsed from the event trigger to the first occurrence of one the following three responses:
    1. An accelerator position equal to 0, indicating pedal release.
    2. A braking position greater than 0, indicating brake activation.
    3. A change in steering angle more than three standard deviations above or below the average steering angle of that participant on straight-aways. A change in steering angle of this magnitude was taken to indicate the initiation of a turning maneuver. Before applying this rule, steering response values were run through a digital, low-pass filter that removed changes greater than 2 Hz in frequency, which were thought to reflect noise rather than actual steering movements.

2.4.4. Mental workload

The Rating Scale of Mental Effort (RSME) was used to measure the self-reported perceptions of mental workload (Zijlstra, 1993). The RSME was presented as a single continuum on a sheet of paper with validated reference points along the scale (e.g., "Absolutely No Effort," "Some Effort," "Extreme Effort," etc.). This single-dimension scale has been found to have good sensitivity to both visual and mental workload (Verwy & Veltman, 1996).

2.5. Procedure

Participants completed demographic and general driving background questionnaires before being seated in the simulator. They were fitted with a headset (DC-Com Model 200 Portable Aircraft Intercom) through which they could hear simulator instructions and simulator road noises as well as communicate with the experimenter using an attached hands-free microphone. This headset system closely simulated a hands-free cell phone system. The forward field of view and the audio communicated through the headset were recorded for all driving portions of the experiment.

Participants were acclimated to the driving simulator through dictated instructions and by driving two practice sessions. The first practice session allowed participants to get used to the simulator controls. The second allowed them to become familiar with navigation and turn instructions from the simulator, consisting of automated visual and auditory cues. If a turn was missed, or an incorrect turn executed, participants were told by the simulator to stop, and the experimenter directed them back onto the correct path (this only occurred once during the experiment). After both practice sessions were completed, participants walked to a nearby table and completed the first workload test.

Participants were asked to maintain a speed at or near 45 mph. If participants drove above 50 mph, the speedometer text turned yellow. If they drove above 55 mph, the speedometer turned red and they heard a recorded message notifying them to slow down. Participants drove a given path with one or two turns. After each path was completed, participants walked to a nearby table to fill out the RSME.

In both dual-task conditions, the participant conversed over the headset with the experimenter who was located five feet behind the participant and out of the participant's field of view. So as not to affect the conversation, the experimenter could not see the front driving field of view or the driving demands put on the driver. Both the easy and difficult conversation tasks were performed continuously from the start to the finish of a trial.

2.6. Model for analysis

The driving performance variables consisted of dependent measures of speed maintenance (accelerator position variability, average speed, and speed variability) or lane keeping (steering offset and mean lateral speed). Each of the three driving trials consisted of two paths, and each path consisted of 7 to 10 straight road sections that were discontinuous (i.e., separated by turns). Each driving performance variable was first calculated over the data within each of the discontinuous road sections. Then, for each of the three trials, the variables were averaged across the road sections within that trial (averages were weighted based on the number of data points within each section).

To control for inflation of experiment-wide error due to the use of multiple driving performance variables to assess speed maintenance, a MANOVA was conducted to test whether conversation level (none, easy, or difficult) affected the multiple speed-maintenance variables in a similar fashion. A MANOVA was also conducted to assess the effect of conversation level on the multiple lane-keeping variables. Subsequent univariate contrast analyses (represented by planned t-tests) were used to test specific hypotheses regarding whether any conversation, easy or difficult, degraded driving more than no conversation (conversation effect) and whether difficult conversations degraded driving more than easy ones (difficulty effect). The contrasts were:

- Conversation effect: control (−1.0), easy conversation (0.5), difficult conversation (0.5).
- Difficulty effect: easy conversation (−1.0), difficult conversation (1.0).

A chi-square analysis was used to test for differences in the number of collisions between conditions. An alpha level of .05 was used for all analyses.

3. Results

3.1. Mental workload data

The difficult conversation questions were significantly more difficult to answer than the easy questions, as rated by the pilot participants. These difficulty ratings were made when the pilot participants were not driving, and thus focused exclusively on the conversation task. As a further check on the difficulty of the questions, the subjective reports of mental effort were compared for the three levels of conversation while driving in the experiment. These workload ratings reflected the overall effort required for both the driving and the conversation tasks. As predicted and as seen in Fig. 2, subjective mental effort was higher in the presence of a conversation relative to no conversation, but there was no significant difference in reported effort between the conversation levels while driving. Thus, as intended, the engagement of a conversation did increase driver effort to cope while driving. However, the design intent to create a low-level and a high-level of conversation difficulty was not sufficient to be reported subjectively in the context of also performing the driving task.

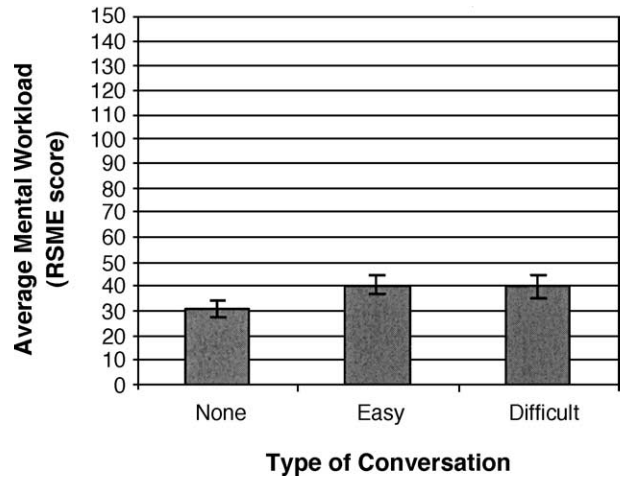

Fig. 2. Effect of conversation and level of difficulty on reported mental effort.

Although the questions were designed to reflect different difficulties, they were also intended to require the same response time so that both levels of conversation could have the same number of responses in a given time period (response rate). To verify this, the number of questions that participants answered during each 10-minute conversation trial was recorded. The pilot data showed that there was not a significant difference between the time participants took to answer the easy and difficult questions. As a check, the mean number of questions answered during each trial was compared with each difficulty level since the time allotted to answer the questions during the conversation trials was the same. The mean number of easy and difficult conversation questions answered, as well as the mean number by gender, is presented in Table 2. Consistent with the pilot study, a repeated measures ANOVA showed no significant difference between the two conversation conditions in terms of number of questions answered, $F\ (1, 22) = 0.26$. However, a significant between-subjects effect was present for gender, $F\ (1, 22) = 9.89$, $p = .005$, showing that males answered more questions than females and suggesting that males spent less time thinking about and answering each question.

3.2. Practice effects

When performing a within-subjects study, there is the chance that a participant's performance on later trials is affected by the knowledge and experience gleaned from earlier trials. To check for these effects, an analysis was conducted to compare participants' improvement in performance from the first to the third driving trials for all variables (except collisions). These contrasts compared the performance in the first trial (given a weight of −1) to the

Table 2
Number of questions answered in each conservation condition, mean and mean standard error

	Easy		Difficult	
	Mean	SE	Mean	SE
All participants:	19.83	1.03	20.42	0.99
Female:	17.58	1.35	17.92	1.32
Male:	22.08	1.30	22.92	1.10

performance in the last trial (given a weight of 1; the second trial was given a weight of 0).

Aside from RT to the hazard events, none of the variables showed significant practice effects (Table 3). RT to hazardous events did significantly improve with experience, t (23) = 2.42, p = .012. This suggests that participants became more vigilant to the hazardous conditions as the experiment progressed. However, the ordering of experimental conditions was counterbalanced. Thus, the practice effect for reaction time does not confound the conclusions regarding the effects of the independent variables, but it may limit the sensitivity of this task to the experimental conditions.

3.3. Speed maintenance (Mobility)

A MANOVA was conducted using accelerator position variability, speed variability, and average speed as one multivariate factor, and conversation condition as a repeated-measures factor. The effect of conversation level was significant, exact F (4, 20) = 2.91, p = .048. In addition, a separate contrast analysis was conducted for each of the hypothesized effects for the speed maintenance variables.

3.3.1. Conversation effects

As shown in Table 4, the presence of a cell phone conversation significantly influenced accelerator position variation, speed variability, and mean speed.

As shown in Fig. 3, the cell phone conversations significantly reduced performance in terms of: (a) increased accelerator position variability, (b) increased speed variability, and (c) reduced average speed. Although these changes were statistically significant, the practical size may be considered small and represent changes of approximately 0.5 to 1.0 mph in response to the cell phone conversation.

3.3.2. Difficulty effects

As shown in Table 4, the presence of different difficulties of cell phone conversation did not significantly influence accelerator position variation, speed variability, or mean speed.

3.4. Lane position maintenance (Safety)

A MANOVA was conducted using steering offset and mean lateral speed as one multivariate factor, and conversation condition as a repeated-measures factor. The effect of conversation level was not significant, exact F (4, 20) = 1.58, p = .219. In addition, a separate contrast analysis was conducted for each of the hypothesized effects for the lane position maintenance variables.

3.4.1. Conversation effects

As shown in Table 4, the presence of a cell phone conversation did not significantly influence steering variability or mean lateral speed.

3.4.2. Difficulty effects

As shown in Table 4, the presence of different difficulties of cell phone conversation did not significantly influence steering variability or mean lateral speed.

3.5. Crash avoidance (Safety)

There was no significant effect of the cell phone conversations in terms of reaction time to hazard events or percentage of events resulting in a collision (Table 4).

In summary, the most consistent effect was for the impairment of mobility measures in the presence of any cell phone conversation. The higher workload associated with

Table 3
Contrast analysis results of practice effects

Category	Dependent variable	Means and mean standard errors							
		Trial 1		Trial 2		Trial 3		Improvement from Trial 1 to Trial 3	
		Mean	SE	Mean	SE	Mean	SE	t	p
Speed Maintenance	Accelerator Variability	0.043	0.003	0.047	0.004	0.047	0.004	0.334	0.371
	Speed Variability	0.623	0.033	0.677	0.029	0.599	0.037	0.911	0.186
	Average Speed	44.65	0.492	44.67	0.407	45.77	0.453	1.486	0.075
Lane Position Maintenace	Steering Offset	2.14	0.260	2.31	0.399	1.94	0.153	0.601	0.277
	Mean Lateral Speed	0.084	0.006	0.087	0.005	0.086	0.005	0.278	0.392
Crash Avoidance	RT (standardized)	0.383	0.174	−0.001	0.204	−0.377	0.202	2.417	0.012*
	Collisions	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Mental Workload	RSME	38.56	4.087	37.48	4.128	35.75	4.158	0.621	0.270

df=23 for all of the above one-tailed t tests.

* Significant at .05 level.

Table 4
Mean, mean standard error, and contrast analysis results for all dependent variables

Category	Dependant variable	Means and mean standard errors						Contrast analyses			
		No conversation		Easy conversation		Difficult conversation		Conversation effects		Difficulty effects	
		Mean	SE	Mean	SE	Mean	SE	t	p	t	p
Speed Maintenance	Accelerator Variability	0.042	0.003	0.046	0.004	0.049	0.004	1.018	0.039*	1.308	0.102
	Speed Variability	1.279	0.079	1.438	0.070	1.530	0.066	2.436	0.012**	1.274	0.108
	Average Speed	45.62	0.398	44.51	0.526	44.93	0.434	2.306	0.015*	1.139	0.133
Lane Position Maintenance	Steering Varibility	1.920	0.268	2.408	0.398	2.052	0.130	0.856	0.201	1.154	0.130
	Mean Lateral Speed	0.090	0.006	0.084	0.005	0.083	0.004	1.018	0.160	0.234	0.409
Crash Avoidance	RT(standardized)	0.217	0.256	−0.042	0.198	−0.170	0.130	0.371	0.357	0.525	0.302
	Collisions	50.0%	0.104	48.5%	0.104	48.5%	0.104	1.093	0.143	0.000	0.500
Mental Workload	RSME	31.04	3.425	40.50	3.853	40.27	4.701	3.001	0.003**	0.059	0.477

df=23 for all of the above one-tailed t tests.
* Significant at .05 level.
** Significant at .01 level.

the cell phone conversations coincided with slower average speeds and larger speed variation. There was also some indication of increased accelerator position variability. In contrast, there were no significant effects of conversation difficulty on driving performance (Table 4). Thus, this study does not provide support to the hypothesis that level of conversation difficulty can impact driving performance. The main impairment effect occurs during any conversation, relative to driving without conversing.

4. Discussion

This study examined the potential distraction of cell phone operations based on naturalistic conversations while driving. Participants' workload ratings showed that the engagement of a conversation increased the reported effort to cope with the simultaneous task demands of driving. This is consistent with other research that indicated that the cell phone conversation imposed a workload demand

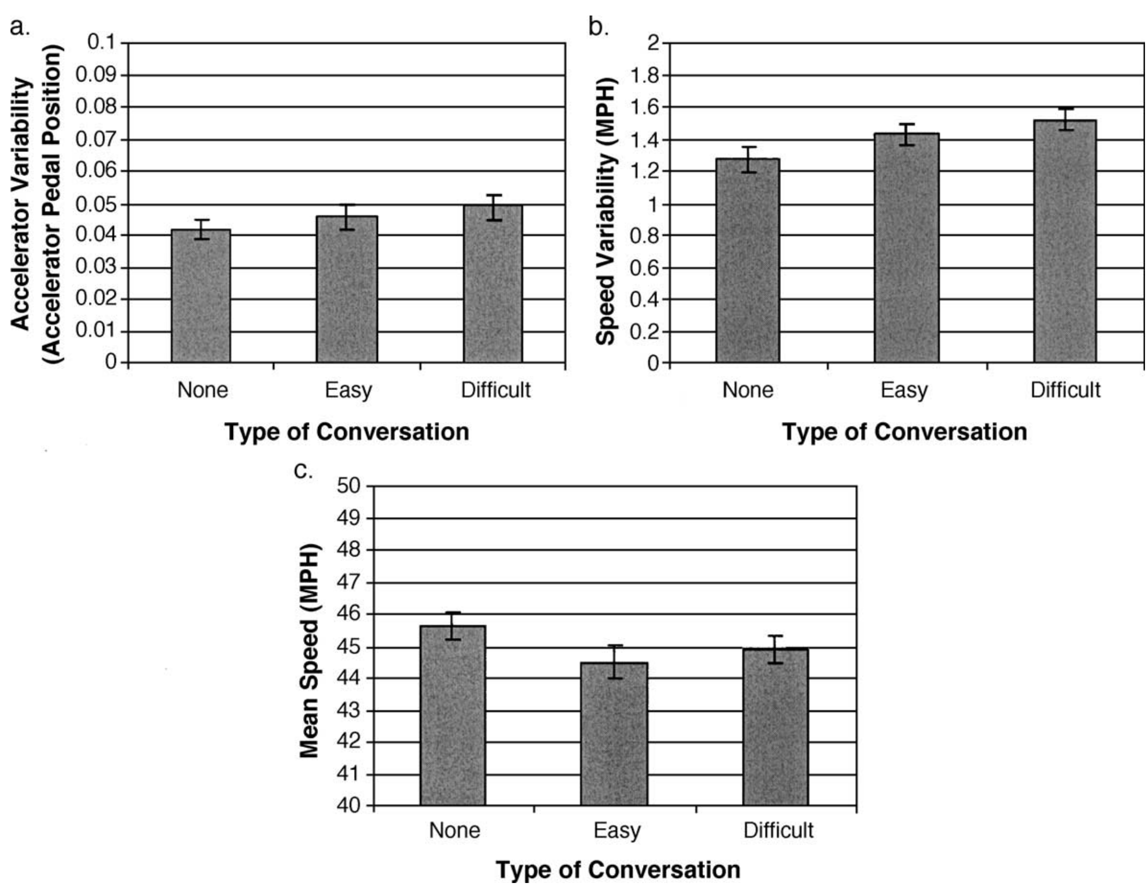

Fig. 3. Conversation effect for (a) accelerator position variability, (b) mean speed variability, and (c) average speed, with standard error bars.

irrespective of the nature of operating the phone: hand held or hands free (Matthews, Legg, & Charlton, 2003). This implies that resources must be allocated to process a conversation and time-share attention with the driving task even when no effort is required to manipulate the phone. However, the workload ratings gave no evidence that our manipulation of conversation difficulty affected the overall effort required for the dual task of driving while conversing. This may be a limitation of the method used to represent conversation difficulty, although pilot testing suggested that the two levels of conversation questions did differ in difficulty. Alternatively, our findings regarding the workload data may suggest that resource allocation is more sensitive to the initial engagement of attention to a conversation source than to the differentiation of complexity in the content of that conversation. This is supported by other research (Briem & Hedman, 1995; Irwin et al., 2000; McKnight & McKnight, 1993) that found effects for conversations, but little or no effect between verbal task difficulty levels.

Several outcomes may be expected as a result of the increased workload imposed on a driver who is operating a vehicle while engaged in a conversation. These can be surmised from theoretical models of the coping methods adopted by operators in the presence of a stressor, such as increased task demands (Mulder, 1986; Hockey, 1993). In these models, the term "state" refers to a profile of energy resources applied by the individual to process information and to select and execute responses (Mulder, 1986). For a given task, there will be a "target state" in terms of a resource distribution consistent with the task demands to achieve optimal task performance goals. This is a hypothetical state based on ideal task performance conditions. In actuality, a person will exhibit a "cognitive state" in response to stressors that are currently present in the task environment. The correspondence between the optimal and cognitive states will determine the extent to which the performance satisfies the task goals (Hockey, 1993).

The "stress response" refers to the form and extent of compensatory effort applied in the presence of the stressor to satisfy performance goals or performance impairment in the absence of an affective compensation strategy. Hockey (1986) contends that the form of stress response will be motivated to reduce the discrepancy between the stress states and the optimal state. These stress response strategies are applied in response to an error signal generated by the comparison of the optimal (target) state and the current cognitive state. These alternative strategies involve effort applied to: (a) changing the current cognitive stress state; (b) modifying the optimal target state by reducing performance goals; (c) removing or modifying the stressor in the environment; and (d) enduring the stress state rather than taking direct action.

In the context of driving while operating a cell phone and engaging in a conversation, the different driving performance metrics assessed in this study can be related to these proposed stress response strategies to cope with the increased workload. Given that the drivers could not select to disengage the conversation, they did not have the option to remove or modify this stressor (option 3). Moreover, although it was evident that the drivers reported greater effort in conjunction with cell phone conversations while driving, there was still evidence of performance impairment, suggesting that these drivers were not successful in changing their current cognitive state with respect to the optimal target state (option 1). Rather, the most consistent response of the drivers during cell phone conversations is congruent with the coping mechanisms of either changing the optimal target state (option 2) or enduring the stress state (option 4).

First, the driving performance degradations observed suggest that participants changed the optimal target state by reducing performance goals so that primary task demands were lowered. Notably, the performance goal for mobility (speed) was lowered such that drivers were satisfied to attain a slower average speed. Similarly, the related performance goal of controllability was relaxed such that larger variations in speed (and accelerator position) were tolerated. However, assuming that the intention of setting reduced goals will lower task demands such that the (modified) optimal state was consistent with the current cognitive state, then the drivers would not be expected to report subjective stress in terms of increased effort during the conversations. And yet, drivers did report increased stress (effort) despite the lower performance associated with reduced performance goals for mobility and controllability. This may suggest that the drivers either underestimated the required reduction in their goal setting, or perceived the environment and task context as imposing limits on the acceptability of the reductions.

Second, it is also possible to frame these same results in terms of enduring the combined task demands whereby the slower speed and increased speed variability are interpreted to be the direct results of attention being diverted to the cell phone conversation and insufficient resources being allocated to these primary task goals. Indeed, given that reported effort (stress) remained significantly higher during the cell phone conversations in conjunction with reduced performance, this is a more viable interpretation of the coping method that may be used by drivers using cell phones. Indeed, these two interpretations can be integrated by assuming that drivers attempt to tolerate the effects of stressors below some threshold, and then attempt to set reduced performance goals to sustain that (suboptimal) threshold (if the stressor cannot be removed or additional resources are not available).[1]

[1] To distinguish these explanations, it would be necessary to also measure self-assessed performance relative to the implicit goal setting by the driver. If performance is reduced objectively, but drivers report satisfactory performance, then it is possible that they are adopting a goal-setting strategy to cope with the stress of the combined task demands.

Both interpretations have safety implications. In terms of enduring the stress imposed by the additional effort needed to use a cell phone for a conversation while driving, crash risk may be increased as driver behavior becomes unstable and as attention resources are diverted away from road hazards. In terms of goal setting, the driver may underestimate the required resources such that the performance standard is still too high. In this case, the driver has to endure a (reduced) level of stress, but has accepted lower performance standards that can be related to increased risk. Even when performance goals are set at a reduced level to eliminate resource competition between driving and the use of the cell phone, the performance standard accepted may be significantly below that required for safety in the driving context.

Admittedly, there are some proponents who advocate the benefits of cell phones in the driving context, such as emergency assistance and traveler information services. Indeed, some of these advocates try to provide an economic argument that the value assigned to the use of a cell phone is more valuable to society than the total cost of traffic fatalities and injuries associated with the use of this technology (Harvard Center for Risk Analysis, 2002). Unfortunately, such conclusions are flawed when the persons that operate and value cell phones cause fatalities and injuries to other road users that place no value on these devices and unlimited value on their own lives. Any economic rationalization for a risk factor may not transfer from the aggregate level to the individual level, and may not be validly applied to persons who are victims rather than participants in that factor.

It should be noted that the major effects of cell phone conversations on driving performance in this study were related to the continuous task demands of longitudinal (speed) control. Impairment of this task can impact safety, especially in terms of crashes associated with rear-end scenarios. However, this study was not able to demonstrate these safety impairments directly, as conversations had little effect on crashes or reaction times to hazardous events. One factor affecting this result was the fact that participants experienced only one hazardous event per conversation condition. This was done so that participants would not be excessively vigilant toward the dangerous driving situations. Even this attempt at limiting the number of dangerous situations, however, was not effective since it was shown that the drivers became more alert to hazardous conditions as the experiment progressed, as evidenced by their faster RTs. As a result, the hazardous events may not have been sufficiently unexpected, critical, or complex enough to be sensitive to the time-sharing process of driving while conversing on the cell phone (Cooper et al., 2003). Another possible reason for the lack of sensitivity of our hazard-avoidance variables (e.g., RT) is that the low number of hazardous events in each condition probably reduced the reliability of our measurements of these variables. Other studies that have shown significant decrements in hazard avoidance during distracted driving have used multiple hazards in each condition in order to increase reliability (Lee, Craven, Haake, & Brown, 2001).

A limitation of our design was the exclusive use of young drivers (mean age = 20.4 years). For this reason, the results cannot readily generalize to populations of different ages. However, it is expected that older driver samples would have worse performance during these driving situations and conversation distractions. Similar findings have been demonstrated in previous studies where older drivers had higher braking reaction times (Alm & Nilsson, 1995), higher percentage of missed vehicle control responses (McKnight & McKnight, 1993), and worse lane keeping (Reed & Green, 1999) while conversing. Young drivers are an important group to study for a number of reasons beyond the fact that they may demonstrate better performance on these measures. Teenage drivers tend to have more crashes than older drivers (Evans, 1991) and drivers between the ages of 16 and 29 are more likely to use a cell phone while driving than all other age groups (Royal, 2003). Thus, understanding the performance of younger drivers, while not generalizable to all drivers, is worthwhile in its own right.

A word of warning must also be given regarding the nature of a simulator experiment. The driving simulator allows us to monitor specific details of driving attention that would be too dangerous to replicate in a real driving situation. Though the tasks and scenes were meant to accurately represent the true nature of driving, the complicated and intricate nature of driving can never completely be replicated in a simulator setting. In the same regard, Reed and Green (1999) found that performance not only on driving but also on secondary tasks was poorer in simulated situations than when driving on a real track. One example of this is that while talking on a phone, drivers made more frequent and larger corrections in steering offset while in the simulator. One explanation for these findings, and a further problem with simulators in general, is driver underload. This is defined as a situation where simulator participants do not pay attention to the primary task, since perceived risk is lower than in a real life situation with the same dangers.

However, it should be noted that Reed and Green also found comparable significant effects (of conversation and age on driving) for the simulator and for real driving. Thus, results from driving simulators have shown that they provide valid, yet somewhat exaggerated, assessments of the factors affecting driving performance measures. Unfortunately, there are few examples of experiments with cell phones conducted on roads or realistic test track settings (Hancock et al., 2003), and fewer still that have used naturalistic conversations in real traffic environments. This may be due to ethical considerations that preclude using high-traffic scenarios during on-road assessments of driving distraction.

5. Conclusions

Whereas other risk factors, such as fatigue, may be more prevalent in traffic crashes, distraction from cell phone use should also be a priority for policymakers (VCU, 2003). The findings from this study suggest that having a conversation using a hands-free cell phone while driving can cause decrements in the speed maintenance performance, while also leading to decreased average speeds. This study also showed that people think that talking on a cell phone while driving is more mentally demanding than driving without talking. In addition, this study suggests that regardless of the intensity of conversation, driving performance will be affected by this attentional distraction.

Future research should focus on the quantification of conversation complexity and dynamics and its relation to attention allocation and driver distraction. This could be supported by a research framework that combines cognitive models of driving and information processing, including conversation such that the potential distraction of all forms of information acquisition and processing (including cell phone conversations, dialogue with passengers, radio news, and other cognitive tasks) can be anticipated, interpreted, and investigated.

And while technology providers may propose design solutions to minimize the risks imposed by cell phones and other distractions inside vehicles, policymakers must ultimately decide on the fundamental purpose and objective of driving—safety or convenience. Until definitive conclusions to support policy can be made, the following may be a suitable compromise for the operation of cell phones while driving:

Having a cell phone in a car is a good idea. It can relieve anxiety about unavoidable delays and be lifesaving in an emergency. But safety-conscious drivers would be wise to avoid using it for causal conversations and should always pull off the road when dialing, talking, or answering it. (Brody, 2002, ¶ 23)

The driving decrements found in this study due to both types of conversation, coupled with the lack of any effect of conversation difficulty, reinforces this last point that even casual conversations during driving may be dangerous.

6. Impact on industry

It is important to understand how elements introduced into the driving environment affect one's ability to keep the proper amount of attention on the road and to maintain control of one's vehicle. Though evidence exists that cell phone use distracts attention away from driving, much more needs to be known about the amount of distraction that devices such as cell phones and navigation systems bring to driving. This is especially the case with cell phone technology because it is affordable, portable, and familiar to most everyone. Not only is this true for commuters and casual drivers, but also for the growing number of people who are on the road all day making deliveries, driving cargo cross-country, plowing snow, or otherwise working on the road who are also conversing on cell phones while driving. This said, it is easy to see that crashes caused by distracted drivers will not only cause a disruption of business, but also potential injury, disability, and permanent loss of personnel. Further, these effects are felt by both the distracted driver as well as other parties involved in an unfortunate driving incident, whether they were distracted or not. Therefore, quantifying such distraction is an important element in understanding what is distracting and learning how we can teach drivers to use this modern-day convenience responsibly.

References

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

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

Brody, J.E. (2002, October 1). *Cellphone: A convenience, a hazard or both?* Retrieved from *The New York Times*, January 30, 2004, http://www.nytimes.com/2002/10/01/health/01BROD.html.

Brookhuis, K. A., de Vries, G., & de Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention*, *23*(4), 309–316.

Brown, I. D., Tickner, A. H., & Simmonds, D. C. V. (1969). Interference between concurrent tasks of driving and telephoning. *Journal of Applied Psychology*, *53*(5), 419–424.

Cellular-News. (2003, March 12). *Countries that ban cell phones while driving.* Retrieved January 30, 2004, from http://www.cellular-news.com/car_bans/.

Cellular Telecommunications Industry Association. (2004, January 30). *Current U.S. Wireless Subscribers*. Retrieved January 30, 2004, from http://www.wow-com.com.

Cooper, P. J., Zheng, Y., Richard, C., Vavrik, J., Heinrichs, B., & Siegmund, G. (2003). The impacts of hands-free message reception/response on driving task performance. *Accident Analysis and Prevention*, *35*, 23–36.

Evans, L. (1991). Older-driver risks to themselves and to other road users. *Transportation Research Record*, *1325*, 34–41.

Glater, J. D. (2002). *Doing business by cellphone creates new liability issues. The New York Times.* December 3, 2002.

Goodman, M., Bents, F. D., Tijerina, L., Wierwille, W., Lerner, N., & Benel, D. (1997). *An Investigation of the Safety Implications of Wireless Communications in Vehicles*. Report No. DOT HS 808-635. Washington, DC: National Highway Traffic Safety Administration.

Haigney, D. E., Taylor, R. G., & Westerman, S. J. (2000). Concurrent mobile (cellular) phone use and driving performance: Task demand characteristics and compensatory processes. *Transportation Research - Part F*, *3*, 113–121.

Haigney, D. E., & Westerman, S. J. (2001). Mobile (cellular) phone use and driving: A critical review of research methodology. *Ergonomics*, *44*(2), 132–143.

Hancock, P. A., Lesch, M., & Simmons, L. (2003). The distraction effects of phone use during crucial driving maneuvers. *Accident Analysis and Prevention*, *35*, 501–514.

Harvard Center for Risk Analysis. (2002, December 2). *Updated study shows higher risk of fatality from cell phones while driving*. Retrieved January 30, 2004, from http://ww.hcra.harvard.edu/cellphones.html.

Hockey, G. R. (1986). Control model of state regulation. In G. R. Hockey (Ed.), *Energetics and human information processing*. Dordrecht: Martinus Nijhoff Publishers.

Hockey, G. R. (1993). Cognitive-energetical control mechanisms in the management of work demands and psychological health. In A. Baddeley, & L. Weiskrants (Eds.), *Attention: Selection, awareness and control*. Oxford: Clarendon Press.

Irwin, M., Fitzgerald, C., & Berg, W. P. (2000). Effect of the intensity of wireless telephone conversations on reaction time in a braking response. *Perceptual and Motor Skills*, *2000*(90), 1130–1134.

ITS America. (2002a, May 15). *Car cell phone use raises crash risk, studies find*. Retrieved January 30, 2004, from http://www.itsa.org/itsnews.nsf/key/8206.

ITS America. (2002b, July 10). *Cell phone use in cars causes tunnel vision*. Retrieved January 30, 2004, from http://www.itsa.org/itsnews.nsf/key/866A.

Kantowitz, B., & Hanowski, R. (1996). Simulator evaluation of heavy-vehicle workload: II. Complex secondary tasks. *Proceedings of the Human Factors and Erognomics Society 40th Annual Meeting* (pp. 877–881).

Lee, J., Caven, B., Haake, S., & Brown, T. (2001). *Speech-based interaction with in-vehicle computers: The effect of speech-based email on drivers' attention to the roadway. Human Factors*, *43*(4), 631–640.

Matthews, R., Legg, S., & Charlton, S. (2003). The effects of cell phone types on drivers subjective workload during concurrent driving and conversing. *Accident Analysis and Prevention*, *35*, 451–457.

McKnight, A. J., & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention*, *25*(3), 259–265.

Mulder, G. (1986). The concept and measurement of mental effort. In G. R. Hockey, et al. (Ed.), *Energetics and human information processing*. Dordrecht: Martinus Nijhoff Publishers.

Parkes, A., & Hooijmeijer, V. (2000). *The influence of the use of mobile phones on driver situation awareness. National Highway Traffic Safety Administration (NHTSA).* Internet forum on "The safety impact of driver distraction using in-vehicle devices." Retrieved July 5, 2000, from http://www.driverdistraction.org.

Ranney, T. A., Mazzae, E. Garrott, R., & Goodman, M. (2000). *NHTSA driver distraction research: past, present, and future*. Driver Distraction Internet Forum, United States Department of Transportation, National Highway Traffic Safety Administration (www_nrd.nhtsa.dot.gov/driver_distraction), Washington, DC: National Highway Traffic Safety Administration.

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.

Response Insurance (2003, July 1). *Even when not on a phone, cell phone users are more distracted when driving*. Retrieved January 30, 2004, from http://www.response.com/about_response/press_releases/distracted_driver_personality.asp.

Royal, D. (2003). Volume 1: Findings. *National survey of distracted and drowsy driving attitudes and behaviors, 2002.* National Highway Traffic Safety Administration (NHTSA) Report Number: DOT HS 809 566. Retrieved March 1, 2004, from http://www.nhtsa.dot.gov/people/injury/drowsy_driving1/survey-distractive03/index.htm.

Stock, G. (1985). *The book of questions*. New York: Workman Publishing.

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

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

Sundeen, M. (2001). Cell Phones and Highway Safety: 2001 State Legislative Update. *National Conference of State Legislatures: Environment, Energy, and Transportation Program.* Retrieved January 30, 2004, from http://204.131.235.67/programs/ESNR/2001cellph.htm.

VCU. (2003, March 31). Traffic and fatigue, not cell phones, are top driving distractions. *Highway and Vehicle Safety Report*, p. 3.

Verwy, W. B., & Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology*, *2*(3), 270–285.

Waugh, J. D., Glumm, M. M., Kilduff, P. W., Tauson, R. A., Smyth, C. C., & Pillalamarri, R. S. (2000). Cognitive workload while driving and talking on a cellular phone or to a passenger. *Proceedings from the IEA 2000/HFES 2000 Congress*. New Jersey: Prentice Hall.

Zijlstra, F. R. H. (1993). *Efficiency in work behaviour*. Unpublished doctoral dissertation. Technical University, Delft, The Netherlands.

Michael (Mick) Rakauskas obtained a M.S. in Human Factors, Applied Psychology from Clemson University, South Carolina, for his thesis on the effects of cell phone conversation difficulty level on driving performance, critical incident avoidance, and mental workload while in a driving simulator. He has done Usability Engineering as an intern at Microsoft in Redmond, WA, and has worked as a Human Factors Consultant for Veridian Engineering at Johnson Controls Inc. He is currently a Research Fellow at the ITS Institute HumanFIRST Program at the University of Minnesota.

Dr. Lee Gugerty obtained a Ph.D. in Experimental Psychology from the University of Michigan. He has done human factors work for NASA at the Johnson Space Center and for the Air Force Research Laboratory. He is currently an Associate Professor in the Psychology Department at Clemson University. His research focuses on situation awareness and navigation during real-time transportation tasks, and on software usability.

Dr. Nicholas Ward (M. Erg. S) obtained his Ph.D. in Human Factors psychology from Queen's University (Canada). He spent more than 8 years conducting Human Factors research in transportation safety at Loughborough University and the University of Leeds. His research focused on driver impairment, risk perception, and assessment of older drivers, as well as the design and evaluation of ITS applications. He is now the Director of the ITS Institute HumanFIRST Program at the University of Minnesota. This program operates a state-of-the-art driving simulator for a multidisciplinary team of researchers with the objective of understanding driver behaviour and designing interventions to improve traffic safety. He is also acting as the Technical Advisor for a European project on driver impairment (IMMORTAL), which includes specific tasks to develop assessment methods for at-risk drivers, including older populations and those with medical conditions (http://immortal.or.at/research_r1.php).