Journal of Safety Research 40 (2009) 7–12

Contents lists available at ScienceDirect

Journal of Safety Research

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

Dissociation between driving performance and drivers' subjective estimates of performance and workload in dual-task conditions

W.J. Horrey *, M.F. Lesch, A. Garabet

Liberty Mutual Research Institute for Safety, Hopkinton, MA, USA

ARTICLE INFO

Available online 24 January 2009

Keywords:
Driver distraction
Dual-task
Dissociation
Subjective workload and performance
Younger and older drivers

ABSTRACT

Introduction: The current study measured how concurrent driving and in-vehicle activities of different levels of engagement varied in terms of performance and subjective estimates of demand and performance. *Method:* In this test track study, 41 younger and older drivers completed a series of cognitive tasks while driving an instrumented vehicle. One task involved an engaging guessing game where drivers tried to guess the identity of an object. The other task involved a simple mental arithmetic task. *Results:* We observed some dissociation between drivers' performance and their subjective reports. For instance, drivers tended to estimate their performance as better for the more engaging guessing task than the arithmetic task, though their performance was actually worse. At the same time, subjective estimates of workload across the two tasks did not vary in the dual-task condition even though they did in the single-task baseline conditions, suggesting that drivers failed to account for the added demands in dual-task situations. *Conclusions:* We discuss the implications of these findings for driver safety. *Impact on Industry:* Crashes due to distraction can carry tremendous costs for employers, in terms of injury, disability, and loss of potentially productive work years, whether these crashes occur on or off the job.

1. Introduction

Today, drivers are facing more and more possible in-vehicle activities as new embedded and portable devices and technologies become available. While these devices may afford drivers enhanced productivity, there are obvious safety concerns to the extent that these devices and related activities detract attention from the driving task. There are a number of studies that suggest that driver distraction and inattention are contributing factors in a significant percentage of crashes and near-misses (Klauer, Neale, Dingus, Ramsey, & Sudweeks, 2005; Ranney, Mazzae, Garrott, & Goodman, 2000; Sussman, Bishop, Madnick, & Walter, 1985).

The performance of in-vehicle tasks depends on a number of factors—resource demand, task structure, and attention switching and allocation—all of which will influence the effectiveness of time-sharing for multiple tasks (Wickens & Hollands, 2000). In general, task performance requires a certain amount of mental resources in order to complete or meet the task demands (Kahneman, 1973). The term 'resources' carries the connotation that they are both limited in nature and can be allocated to different tasks (Wickens, 2002). The amount of resources required for a given task will depend on task difficulty and the effort invested. For very difficult tasks, or for multiple tasks, there may be insufficient resources to accomplish some or all of the tasks and performance may suffer as a consequence. In general, several studies that compare in-vehicle non-visual tasks of varying complexity in the driving context have shown that the more demanding tasks result in greater performance decrements (Briem & Hedman, 1995; Irwin, Fitzgerald, & Berg, 2000; McKnight & McKnight, 1993; Patten, Kircher, Ostlund, & Nilsson, 2004; Rakauskas, Gugerty, & Ward, 2004).

The level of engagement associated with an activity may also be an important factor in determining the driver's resource allocation policy. That is, drivers may devote more attentional resources to activities that are more engaging. In educational research, the term "engagement" has been used to refer to students' cognitive investment, active participation, and emotional engagement in specific learning tasks and is assessed using cognitive criteria (e.g., the extent to which students attend to and expend mental effort in a task), behavioral criteria (e.g., extent to which students are making active responses), and affective criteria (e.g., emotional reactions to the learning task; see Chapman, 2003). In another view (e.g., Pintrich & De Groot, 1990), engagement is reflected by the type of cognitive strategy used in performing a task: shallow (e.g., rehearsal) or deep (e.g., elaboration). In the area of knowledge acquisition and skill development, level of interest is considered central in determining how we select and persist in processing certain types of information in preference to others (e.g., Hidi, 1990). In general, research in the domain of driver distraction has not yet addressed these aspects of engagement. For example, many studies have employed information processing tasks, such as mental arithmetic or logical reasoning, as proxies for conversation (see Haigney & Westerman, 2001; Brookhuis, de Vries,

* Corresponding author. Tel.: +1 508 497 0237; fax: +1 508 435 0482.
E-mail address: william.horrey@libertymutual.com (W.J. Horrey).

0022-4375/$ – see front matter © 2009 National Safety Council and Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsr.2008.10.011

& de Waard, 1991; Patten et al., 2004; McKnight & McKnight, 1993; Briem & Hedman, 1995). In general, these tasks tend to rely on working memory and relatively shallow processing of the information (Pintrich & De Groot, 1990). It is possible that these tasks underestimate the true hazards of distracting activities (Caird, Scialfa, Ho, & Smiley, 2004; Horrey & Wickens, 2006). Other studies have employed more conversation-like tasks, such as a question and answer approach in which the experimenter asks the driver a series of scripted questions (e.g., Irwin et al., 2000; Rakauskas et al., 2004; McKnight & McKnight, 1993). These conversations are more naturalistic, however they can be difficult to measure and quantify in a controlled experiment (in addition to difficulties in controlling the discussion content across individual participants).

1.1. Current Study

The current study sought to assess the effects of a competing in-vehicle task on driving performance (both actual and perceived) among younger and older drivers when tasks varied by level of engagement. As noted above, tasks that are cognitively engaging may draw more attentional resources and thus have greater implications for safe driving. We also wished to explore actual performance in the context of the drivers' subjective assessments of workload demands and task performance.

2. Methods

2.1. Participants

Forty-one participants were recruited for the study, divided into younger (N=21, *M*=22.7 yrs, *SD*=4.8) and older (N=20, *M*=63.8 yrs, *SD*=7.3) age groups. Both groups were approximately balanced by gender. On average, younger drivers had 6.4 years of driving experience (*SD*=4.8) and drove 20,300 km per year (*SD*=17,050). Older drivers had 45.7 years of driving experience (*SD*=8.7) and drove 25,950 km per year (*SD*=13,350). All drivers had normal or corrected-to-normal far visual acuity (min. 20/40) and normal color vision. All participants had a valid U.S. Drivers License and owned and operated a cell phone. Participants were recruited from advertisements in local papers and paid $20 for each hour of participation, plus an additional $20 performance bonus.

2.2. Materials

2.2.1. Apparatus

This study was conducted on a 0.8 km two-lane closed-loop track, delineated for continuous travel. Five pace clocks, used in a speed control task, were positioned adjacent to the driver's lane (left side) at various points around the track. Each clock face was 0.5 m in diameter and was mounted on a 1.8 m tower. The bottom half of the clock was green and the top half was red and the arrow hand rotated clockwise at approximately 12 seconds per rotation. Along a straight section of the track, there was a signalized intersection (traffic light), which was controlled by track- and vehicle-based sensors.

The instrumented vehicle was a 2002 Ford Windstar minivan, equipped with sensors and computers to record vehicle position, speed, and driver inputs, among other variables. Auditory messages were played through a set of speakers mounted behind the driver.

2.2.2. Secondary Tasks

There were two in-vehicle tasks: an addition task and a guessing task. For the addition task, we used the Paced Auditory Serial Addition Task (PASAT) which has been used previously in studies on driver distraction (e.g., Brookhuis et al., 1991; Patten et al., 2004) and is representative of many information processing tasks relying on mental rehearsal and working memory (i.e., relatively low level of engagement). A random number between 1 and 9 was presented every 7 seconds and drivers were required to remember this number, add it to the following number, and say the result aloud. Thus, the task involved working memory, mental arithmetic, and vocal responses.

The guessing task, based on the Twenty Questions Test (TQT; Mosher & Hornsby, 1966, Kafer & Hunter, 1997), fulfills several of the criteria for engagement outlined by others (Chapman, 2003; Pintrich & De Groot, 1990; Hidi, 1990). This task requires that drivers spontaneously generate questions in order to determine the identity of an object in as few yes-or-no questions as possible. Thus, this task utilizes problem-solving skills, strategic planning, information integration, and working memory in a more game-like context. Because there is a back-and-forth information exchange, the task is more like a conversation than the mental arithmetic task. Also, the TQT has an additional advantage over earlier Q&A approaches in that performance can be measured and quantified. Target objects were selected on the basis of a pilot test with seven independent raters. They were from one of three categories (fruit, vegetable, or land mammal) and represented the most generic form of the object rather than a particular type (e.g., dog versus a particular breed). For this task, the experimenter specified the category (using a prerecorded message) before the driver began asking questions. None of the objects were repeated over the course of the experiment and the objects were randomized for each person. After the driver correctly guessed the object, or after they indicated that they gave up on an object, the next object category was specified. The experimenter responded to the questions by pressing buttons on a small keypad that would generate a prerecorded 'yes' or 'no' response over the vehicle speakers.

2.2.3. Procedure

At the start of the two-hour session, drivers completed an informed consent form. Vision was tested with a Titmus Vision Tester (Titmus Optical Inc., Chester, VA) and drivers were excluded if their far visual acuity was below 20/40 or if they were color blind. Drivers then completed a demographic and driving history questionnaire as well as scales comparing their driving skills and abilities to the average U.S. driver (based on Horswill, Waylen, & Tofield, 2004).

Following the completion of the questionnaires, drivers were taken outside to the track, instructed in the operation of the van, and given several practice laps around the track. The driving task included lane keeping, speed control, and speeded response to the traffic light. For the lane keeping task, drivers were instructed to keep their vehicle in the center of their lane throughout the trial. For the speed control (pace clock) task, drivers were told to speed up or slow down during the approach to a clock such that they would pass by the clock when the arrow was in the green portion of the clock (i.e., avoid passing in the red portion). They were also instructed to avoid coming to a complete stop and to avoid exceeding 48 kph. For the traffic light, the light changed from green to red on a random subset (38%) of trials. To discourage drivers from running the light, there was no yellow light sequence. When the light changed, drivers had to bring the vehicle to a complete stop as quickly as they could and before reaching a stop line marked by two traffic cones. The timing of the light change was short (between 4.5 and 5.5 seconds before the intersection), requiring a quick response.

The study consisted of five experimental blocks: three driving blocks and two non-driving blocks. Before the experimental blocks, drivers were given training and practice on the two tasks. Each of the driving blocks consisted of 8 laps around the test track and lasted approximately 15-minutes. In the baseline driving block, participants performed the driving task alone. In the remaining two driving blocks, participants also performed a secondary task while navigating the course (the addition task or the guessing task). To encourage completion, drivers were offered a financial incentive for each correct response—a 5-cent bonus for each correct addition response and $1 bonus for each object correctly guessed (up to a maximum of $20 over the whole experiment); however, all participants received the full bonus amount regardless of performance. In two non-driving blocks, participants performed either the addition or the guessing task while the vehicle was parked on the side of the track. Short breaks were provided between blocks and the order of blocks was counterbalanced across driver. The addition task was also executed in a separate block using a hand-held phone; however, these results are presented elsewhere (Horrey, Lesch, & Garabet, 2008).

Throughout the driving blocks, lane position, pace clock errors, braking response time (to the traffic light), and stopping errors were recorded. In-vehicle task performance was gauged by the percent correct in the addition task and response rate (number of correct items per unit time) for the guessing game. After each block, a modified NASA TLX (Hart & Staveland, 1988) was used to assess drivers' ratings of mental demand, physical demand, time pressure, frustration, and mental effort. Additionally, drivers rated their driving performance on the lane keeping, pace clock, and stop light tasks as well as their performance on the in-vehicle tasks.

At the end of experiment, participants rated each in-vehicle task on a continuous scale on dimensions of difficulty (anchors: incredibly difficult, incredibly easy), distraction (anchors: incredibly distracting, not at all distracting), enjoyment (anchors: not enjoyable at all, perfectly enjoyable), and engagement (anchors: not at all engaging, incredibly engaging). They also completed a sensation seeking scale (Zuckerman, Kolin, Price, & Zoob, 1964) and driving behavior questionnaire (after Parker, Reason, Manstead, & Stradling, 1995). Finally, drivers were thanked and remunerated for their participation.

3. Results

Prior to the analysis of performance data, we wished to establish critical differences between the two tasks (Guessing and Addition), based on several dimensions. Following these analyses, the performance data and subjective ratings were analyzed according to the 3×2 mixed design, with the within-subject variable Condition (Baseline, Addition, Guess) and the between-subject variable Age (Younger, Older). Variance in the degrees of freedom across the analyses was due to occasional equipment failure and subsequent data loss encountered on some trials. In these cases, the analyses were completed with list-wise deletion of cases. Prior to these analyses, we examined whether there was an effect of gender on any of our measures. As there were no main effects of gender, nor any significant interactions, we did not include this variable in subsequent analyses.

3.1. Task Comparison

As noted above, we contrasted the two tasks (Guessing and Addition) on several dimensions. First, as expected, the guessing task was considered more engaging than the addition task ($F(1,39)=3.7$, $p=.06$). Second, in the in-vehicle task baseline blocks, drivers rated the guessing task as significantly more demanding than the addition task, based on the composite workload rating ($t(40)=6.7$, $p<.001$). With respect to the workload sub-scales, drivers rated the guessing task higher than the addition task on mental effort ($t(40)=6.1$, $p<.001$), mental demands ($t(40)=7.4$, $p<.001$), frustration ($t(40)=3.9$, $p<.001$), and time pressure ($t(40)=4.9$, $p<.001$). There was no difference in ratings of physical demands, though none was expected ($t(39)=1.4$, $p=.18$). Moreover, the tasks did not vary in terms of drivers' ratings of enjoyment ($F(1,39)=1.0$, $p=.34$); however, further analysis yielded a significant Age×Condition interaction ($F(1,39)=5.0$, $p=.03$). Younger drivers reported that the addition task was less enjoyable than the guessing game ($t(20)=2.5$, $p=.02$); whereas there was no significant difference in the enjoyment ratings for the tasks for older drivers ($t(19)=.81$, $p=.43$).

Thus, the guessing task fulfills several of the criteria for engagement outlined by others (Chapman, 2003; Pintrich & De Groot, 1990; Hidi, 1990). In addition to the structural differences between the tasks (mental rehearsal / working memory versus problem-solving skills, strategic planning, information integration, and working memory), the guessing task was perceived as demanding more attentional resources and involving added affective investment. Finally, the guessing task was perceived to be more engaging than the addition task.

3.2. Driving Performance

A MANOVA revealed an overall effect of Condition across three measures of driving performance: brake response time, pace clock accuracy, and lane keeping ability ($F(6,128)=13.6$, $p<.001$). A follow-up model confirmed that there was an overall difference between the two different secondary tasks as well ($F(3,31)=2.9$, $p=.05$). As such, the general pattern of results with performance being highest in the baseline condition, followed by the addition task, and then the guessing task was preserved across all measures. Below, we describe the univariate statistics, which are summarized in Table 1.

Brake response time was measured from the moment the traffic light changed until the initial deflection of the brake pedal by the driver. There was a significant main effect of condition in brake RT (Table 1), with the fastest responses in the baseline condition, followed by the addition task and then the guessing task as shown in Fig. 1a. There was no significant effect of Age or a significant interaction.

Errors in the pace clock task included passing the clock when the arrow was in the green portion, bringing the vehicle to a complete stop, and speeding excessively in an attempt to pass the clock. Performance was expressed as a percent correct. With respect to pace clock accuracy, there was a significant main effect of condition. Drivers made fewer mistakes in the baseline condition, followed by the addition and the guessing conditions (see Fig. 1b). As with the other measures, there was no significant main effect of Age or a significant interaction.

Lane position was sampled along three straight sections of the track. Prior to analysis, these data were smoothed using a simple moving average with a window of 30 samples (1 second). The cumulative samples in each block were used to determine the variability in lane keeping. The analysis revealed a significant main effect of Condition in drivers' lane keeping ability, as measured by the standard deviation in lane position. Performance was better in the baseline condition (i.e., drivers' lane keeping was less variable), followed by the addition condition. Lane keeping performance was worst in the guessing condition (Fig. 1c). The main effect for Age was not significant, nor was the interaction (see Table 1).

Stopping errors included failure to comply with the light (i.e., running a red light) or stopping after the stop line. Because these data were not normally distributed, we used a series of non-parametric tests to examine stopping errors. A series of Wilcoxon signed ranks t-tests for paired samples showed that drivers made more errors in the

Table 1
ANOVA results for various dependent measures related to driving performance

Source	Brake Response Time			Pace Clock Accuracy			Variability in Lane Keeping		
	df	MSE	F	df	MSE	F	df	MSE	F
Between-subjects									
Age (A)	1	0.04	0.6	1	53.3	0.4	1	<0.001	0.06
Error	37	0.06		35	129.0		37	0.01	
Within-subjects									
Task (T)	2	0.09	7.9*	2	1347.9	35.7*	2	0.01	10.4*
T x A	2	0.01	0.5	2	38.0	1.0	2	0.001	0.8
Error	74	0.01		70	37.7		74	0.001	

Note. * p<.001.

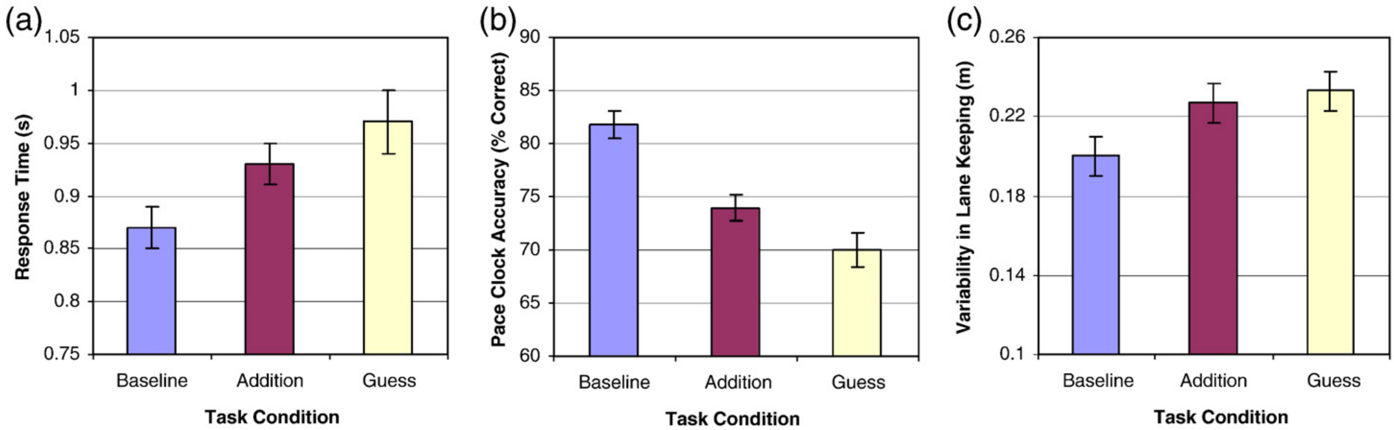

Fig. 1. (a) Braking response time to stop light; (b) accuracy on pace clock task; and (c) variability in lane keeping performance. Standard error bars are shown.

addition (M=48) and guessing (M=42) conditions compared to baseline (M=26; $Z(38)$=-3.6, p<.001; $Z(39)$=-2.9, p=.004; respectively). There was no significant difference between the two distraction conditions ($Z(39)$=-0.4, p=.66). Additionally, a Mann-Whitney test did not reveal a significant effect of Age ($Z(40)$=-0.1, p=.93).

To summarize, across all measures of driving performance, there were dual-task costs associated with performing the in-vehicle activities, replicating previous results (Briem & Hedman, 1995; Irwin et al., 2000; McKnight & McKnight, 1993; Patten et al., 2004; Rakauskas et al., 2004). In general, concurrent performance of the guessing task yielded poorer performance on the driving tasks.

3.3. In-vehicle task performance

We also observed decrements in in-vehicle task performance, relative to single task baseline. For the guessing task, response rates declined from 0.53 correct items per minute (single-task) to 0.39 items per minute in the dual-task condition ($t(40)$=2.3, p=.03). Likewise, accuracy in the addition task declined from 97.1% in the baseline block to 89.5% in the dual-task blocks ($t(40)$=6.4, p<.001). Thus, we note that all tasks (both driving-related and in-vehicle) were adversely impacted by concurrency.

In order to compare the two tasks, we analyzed the percent change scores. That is, for each person we calculated the percent difference between the dual- and single-task blocks for each in-vehicle activity. These change scores were analyzed with an ANOVA. There were no significant main effects (Age, Condition) or any interactions. That is, the magnitude of the decline did not vary as a function of task type, suggesting that the driving task did not differentially impact the performance of the in-vehicle tasks ($F(1,36)$=0.03, p=.87).

3.4. Subjective Ratings of Workload Demand and Performance

Subjective (composite) workload scores and performance ratings were also measured following each experimental block. As noted above, drivers rated the guessing task as significantly more demanding than the addition task in single-task baseline conditions. In blocks where driving took place, there was a main effect of Condition ($F(2,76)$=71.4, p<.001), with higher workload ratings in both dual-task blocks compared to single-task driving. Interestingly, there was no difference in workload between the guessing and addition tasks while driving ($t(40)$=.93, p=.36), despite differences between single-task conditions. While there was no main effect of age ($F(1,38)$=1.9, p=.18), there was a significant interaction ($F(2,76)$=5.5, p=.006) with younger drivers rating the addition task as more demanding than the older drivers. Both groups rated the guessing task equivalently in terms of workload.

With respect to subjective performance ratings, drivers rated their performance across all driving tasks (lane keeping, pace clocks, and stopping task) as better in the baseline condition, followed by the guessing condition and then, finally, the addition task (MANOVA results; $F(6,148)$=7.3, p<.001; see Fig. 2). Univariate statistics are shown in Table 2. With respect to in-vehicle task performance, drivers thought they performed better on the addition task than the guessing task while driving ($F(1,36)$=9.4, p=.004). There were no effects of age, nor any significant interactions. However, when we examine subjective estimates of performance in dual-task blocks compared to

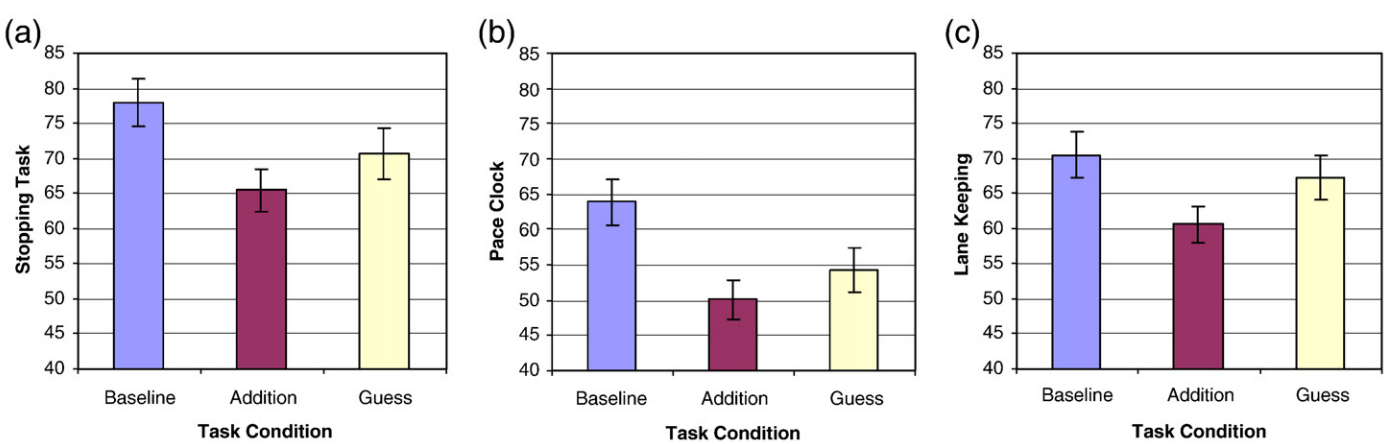

Fig. 2. Subjective performance ratings on the (a) stopping task; (b) pace clock task; and (c) lane keeping task. Standard error bars are shown.

Table 2
ANOVA results for subjective ratings of driving performance

Source	Stopping Task			Pace Clock Accuracy			Variability in Lane Keeping		
	df	MSE	F	df	MSE	F	df	MSE	F
Between-subjects									
Age (A)	1	578.3	0.5	1	19.5	0.02	1	413.7	0.5
Error	38	1200.1		38	863.8		38	824.7	
Within-subjects									
Task (T)	2	1567.0	17.0*	2	1996.1	13.5*	2	1013.8	7.9*
T x A	2	0.3	0.004	2	36.5	0.2	2	82.4	0.6
Error	76	92.4		76	147.6		76	128.5	

Note. * p<.001.

baseline conditions, we observed some interesting differences. Although overall performance on the guessing task was rated lower than the addition task in the driving blocks, the estimated performance in the dual-task guessing block was higher than single-task block, whereas the single-task addition performance was considered to be higher than the dual ($t(34)=3.2$, $p=.003$).

In summary, the guessing game was rated as higher in workload than was the addition task; however, this task difference was not preserved in the dual-task blocks. That is, when the driving and in-vehicle tasks were combined, drivers did not consider the guessing game to carry more workload, though it did in single task conditions. With respect to subjective estimates of performance, drivers rated their driving performance as higher in the guessing condition than in the addition condition. Although they rated their in-vehicle task performance as better in the addition task, compared to single-task baseline performance, drivers tended to rate themselves as improved relative to baseline in the guessing condition compared to the addition task where they thought that their performance declined in the dual-task block.

4. Discussion

Given the advent of new activities and devices that drivers may perform while driving, it is important to understand the implications for safety. Driver distraction and inattention has been shown to be a contributing factor in a significant percentage of crashes and near-misses (Klauer et al., 2005; Ranney et al., 2000; Sussman et al., 1985). The effectiveness of time-sharing for multiple concurrent tasks depends on resource demand, task structure, and attention switching and allocation (Wickens & Hollands, 2000). The current study sought to examine how driving performance and drivers' subjective estimates of performance and workload were impacted by two tasks of varying degrees of engagement. Tasks that are cognitively engaging may draw more attentional resources and have greater implications for safety.

The results from the current study showed that the more engaging task, the guessing game, yielded greater performance deficits on measures of driving performance than did the less engaging, addition task. We also observed decrements on the in-vehicle task (addition or guessing) while drivers navigated the track compared to a single task baseline. While these results are comparable to other studies employing tasks of varying complexity (Briem & Hedman, 1995; Irwin et al., 2000; McKnight & McKnight, 1993; Patten et al., 2004; Rakauskas et al., 2004), an examination of the subjective estimates of workload and performance revealed some intriguing results.

4.1. Dissociations Between and Within Ratings and Performance

The analysis of the subjective estimates of workload and performance yielded some dissociations. First, although drivers rated the guessing task as more demanding than the addition task (in the single task baseline), they did not rate the guessing task as more demanding when it was combined with the driving task compared to the addition task (e.g., Rakauskas et al., 2004). That is, drivers did not fully account for the added workload from the guessing task as evidenced by the equivalent degree of perceived workload in dual-task blocks—a result that was clear in the WL ratings of single-task blocks. It is possible that drivers failed to account fully for the added demands of the guessing task in the dual-task situation, instead opting for a default level of workload for both tasks. For example, Vidulich and Wickens (1986) found a disconnection between ratings of workload and performance. They speculated that their continuous tracking task (akin to the driving task in the current study) dominated the dual-task ratings such that the differences between the intermittent secondary tasks were obscured / not reflected. Alternatively, the observed workload assessments could reflect that drivers were fully invested in the task sets. Given the differences in tasks (from baseline), both possibilities would yield the pattern of results that was observed: that there were greater performance decrements for the guessing task compared to the addition task. The observed dissociation between similar levels of workload yielding different levels of performance has been described elsewhere.

Importantly, there were differences between actual performance and the subjective estimates of performance. Drivers rated their own driving performance, following the experimental block, as higher for the guessing task compared to the corresponding addition block. However, and as noted in the above section, performance in the guessing task was actually impaired relative to the addition block, suggesting that there is dissociation between drivers' estimates of the impact of the engaging task and their actual performance. Vidulich and Bortolussi (1988) also found higher performance in a speech condition that was associated with lower ratings. Why would drivers rate their performance as better on the guessing game compared to the addition task when their performance is actually worse? Previous work suggests that drivers are not aware of their own performance decrements due to distraction (Lesch & Hancock, 2004). It is possible that tasks that are more engaging might systematically bias performance estimates towards more favorable outcomes. For example, drivers tend to be overconfident in their skills and abilities (e.g., Wogalter & Mayhorn, 2005; Horswill et al., 2004). Given that more attentional resources are tied-up in the engaging activities (guessing task), the performance assessment may more adequately represent drivers' pre-existing biases and overconfidence in their skills as opposed to an online estimate of performance that requires attentional resources itself Yeh & Wickens, 1988).

4.2. Limitations & Implications

While normal aging is often associated with degraded performance on many tasks, such as response time (e.g., Salthouse, 1996), we did not observe any age-related performance deficits on the current measures. The wide range of ages in each group could have contributed to greater heterogeneity in measured responses, leading to decreased statistical power associated with an age effect. We also note that, because the experiment was carried out on a closed test track, we could not replicate real world driving conditions, thus limiting the generalizability of our results. That being said, lane keeping, speed control, and response to traffic controls are representative of many of the coordinated tasks required for driving.

In many situations, drivers themselves decide how and when to perform in-vehicle (and potentially distracting) activities (Lerner, 2005). These decisions likely are influenced by drivers' perceptions of the combined resources demanded by the tasks as well as their level of performance. To the extent that drivers' perceptions are erroneous, there are concerns for safety—especially if drivers elect to perform concurrent activities when it is clearly inappropriate to do so. In the real world, drivers may be much more engaged in certain in-vehicle activities—an outcome that might not be fully realized in laboratory-based information processing tasks or by scripted conversations. Given the dissociations observed in the current study, task engagement may be of particular relevance, both from a safety perspective and from a methodological perspective. More research on tasks of varying degrees of engagement in both laboratory and naturalistic settings is merited.

5. Impact on Industry

Today, it is important to understand how new in-vehicle tasks affect drivers' performance as well as how they impact drivers' perceptions of their own performance. Commercial drivers often have to deal with additional in-vehicle tasks as part of their jobs (e.g., communicating with dispatch, using navigation devices). As discussed above, degraded driving performance can contribute to road traffic crashes—an outcome that impacts business through injury, disability, and loss of potentially productive work years. Off-the-job crashes or crashes during workplace commutes also carry tremendous costs for employers. For example, NHTSA estimates that 80% of the health-related fringe benefit costs of motor-vehicle crashes are incurred off-the-job. Furthermore, off-the-job crash injuries accounted for more than 92% of employer health care costs (NHTSA, 2003). While current U.S. data systems such as the Fatality Analysis Reporting System (FARS) and General Estimates System (GES) do not consider crashes that occur during a commute to be work-related, Boufous and Williamson (2006), using police crash records and workers compensation data from Australia, found that roughly 75% of work-related traffic crash casualties occurred during a workplace commute. Unfortunately, the average length of workplace commutes is on the rise. In 2000, over 97 million single-occupant vehicles were used in daily commutes to and from work in the United States—a 35 million increase from the number in 1980 (TRB, 2006).

Acknowledgements

The authors are grateful to Peter Teare, Eric Jones, Richard Holihan, and Ed Correa for their maintenance and upkeep of the instrumented van and other equipment. We also thank Mary Jane Woiszwillo for her assistance with the secondary tasks. William Shaw and Jia-hua Lin provided helpful comments and suggestions on earlier drafts of this manuscript.

References

Boufous, S., & Williamson, A. (2006). Work-related traffic crashes: A record linkage study. *Accident Analysis and Prevention*, *38*(1), 14–21.
Briem, V., & Hedman, L. R. (1995). Behavioral effects of mobile telephone use during simulated driving. *Ergonomics*, *38*(12), 2536–2562.
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.
Caird, J. K., Scialfa, C. T., Ho, G., & Smiley, A. (2004). *Effects of Cellular Telephones on Driving Behaviour and Crash Risk: Results from a Meta-Analysis.* Edmonton, AB: CAA Foundation for Traffic Safety.
Chapman, E. (2003). Alternative approaches to assessing student engagement rates. *Practical Assessment, Research & Evaluation*, *8*(13) Retrieved May 22, 2008 from http://PAREonline.net/getvn.asp?v=8&n=13
Haigney, D., & Westerman, S. J. (2001). Mobile (cellular) phone use and driving: A critical review of research methodology. *Ergonomics*, *44*(2), 132–143.
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* (pp. 139–183). Amsterdam: North-Holland.
Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, *60*, 549–571.
Horrey, W. J., Lesch, M. F., & Garabet, A. (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.
Horswill, M. S., Waylen, A. E., & Tofield, M. I. (2004). Drivers' ratings of different components of their own driving skill: A greater illusion of superiority for skills that relate to accident involvement. *Journal of Applied Social Psychology*, *34*(1), 177–195.
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*, *90*, 1130–1134.
Kafer, K. L., & Hunter, M. (1997). On testing the face validity of planning/problem-solving tasks in a normal population. *Journal of the International Neuropsychological Society*, *3*, 108–119.
Kahneman, D. (1973). *Attention and Effort.* Englewood Cliffs, NJ: Prentice-Hall.
Klauer, S. G., Neale, V. L., Dingus, T. A., Ramsey, D., & Sudweeks, J. (2005). Driver inattention: A contributing factor to crashes and near-crashes. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 1922–1926). Santa Monica, CA: Human Factors and Ergonomics Society.
Lerner, N. D. (2005). Deciding to be distracted. *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (pp. 499–505). Iowa City, IA: University of Iowa.
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*(3), 471–480.
National Highway Traffic Safety Administration [NHTSA]. (2003). *The Economic Burden of Traffic Crashes on Employers (Report No. DOT HS 809 682).* Washington DC: Author.
McKnight, A. J., & McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention*, *25*(3), 259–265.
Mosher, F. A., & Hornsby, J. R. (1966). On asking questions. In J. S. Bruner (Ed.), *Studies in Cognitive Growth* (pp. 86–102). New York: John Wiley & Sons.
Parker, D., Reason, J. T., Manstead, A. S. R., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics*, *38*(5), 1036–1048.
Patten, C. J., Kircher, A., Ostlund, J., & Nilsson, L. (2004). Using mobile telephones: Cognitive workload and attention resource allocation. *Accident Analysis and Prevention*, *36*(3), 341–350.
Pintrich, P. R., & De Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33–50.
Rakauskas, M. E., Gugerty, L. J., & Ward, N. J. (2004). Effects of naturalistic cell phone conversations on driving performance. *Journal of Safety Research*, *35*(4), 453–464.
Ranney, T. A., Mazzae, E., Garrott, R., & Goodman, M. J. (2000). *NHTSA driver distraction research: Past, present, and future [On-line].* Available: www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/welcome.htm
Salthouse, T. A. (1996). Processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.
Sussman, E. D., Bishop, H., Madnick, B., & Walter, R. (1985). Driver inattention and highway safety. *Transportation Research Record*, *1047*, 40–48.
Transportation Research Board. (2006). *Commuting in America III: The Third National Report on Commuting Patterns and Trends (NCHRP Report 550/TCRP Report 110).* Washington, DC: National Academies, TRB.
Vidulich, M. A., & Bortolussi, M. R. (1988). A dissociation of objective and subjective workload measures in assessing the impact of speech controls in advanced helicopters. *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1471–1475). Santa Monica, CA: Human Factors Society.
Vidulich, M. A., & Wickens, C. D. (1986). Causes of dissociation between subjective workload measures and performance: Caveats for the use of subjective assessments. *Applied Ergonomics*, *17*, 291–296.
Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*, 3rd ed. Upper Saddle River, NJ: Prentice Hall.
Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*, 159–177.
Wogalter, M. S., & Mayhorn, C. B. (2005). Perceptions of driver distraction by cellular phone users and nonusers. *Human Factors*, *47*(2), 455–467.
Yeh, Y. Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors*, *30*(1), 111–120.
Zuckerman, M., Kolin, E. A., Price, L., & Zoob, I. (1964). Development of a sensation-seeking scale. *Journal of Consulting Psychology*, *28*, 477–482.

William Horrey is currently a Research Scientist in the Center for Behavioral Sciences at the Liberty Mutual Research Institute for Safety in Hopkinton, MA. His work focuses primarily on transportation human factors and issues in driver inattention. He received his Ph.D. in engineering psychology from the University of Illinois at Urbana-Champaign in 2005.

Mary Lesch holds a Ph.D. in cognitive psychology from the University of Massachusetts - Amherst and has spent more than 15 years conducting research in the areas of language processing and working memory. Currently, she is a Research Scientist in the Center for Behavioral Sciences at the Liberty Mutual Research Institute for Safety in Hopkinton, MA where her research focuses on the role of cognitive failures in accident causation.

Angela Garabet is a currently a Research Specialist at the Liberty Mutual Research Institute for Safety providing expertise and support in the Center for Behavioral Sciences. She holds a MASc in Systems Design Engineering from the University of Waterloo.