ELSEVIER

Accident Analysis and Prevention 40 (2008) 1282–1293

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

A meta-analysis of the effects of cell phones on driver performance

Jeff K. Caird [a,*], Chelsea R. Willness [b], Piers Steel [c], Chip Scialfa [b]

[a] Cognitive Ergonomics Research Laboratory, University of Calgary, 2500 University Dr. N.W. Calgary, Alberta T2N 1N4, Canada
[b] Department of Psychology, University of Calgary, 2500 University Dr. N.W. Calgary, Alberta T2N 1N4, Canada
[c] Haskayne School of Business, University of Calgary, 2500 University Dr. N.W. Calgary, Alberta T2N 1N4, Canada

Received 11 September 2007; received in revised form 2 January 2008; accepted 30 January 2008

Abstract

The empirical basis for legislation to limit cell phones while driving is addressed. A comprehensive meta-analysis of the effects of cell phones on driving performance was performed. A total of 33 studies collected through 2007 that met inclusion criteria yielded 94 effect size estimates, with a total sample size of approximately 2000 participants. The dependent variables of reaction time, lateral vehicle control, headway and speed and the moderating variables of research setting (i.e., laboratory, simulator, on-road), conversation target (passenger, cell phone) and conversation type (cognitive task, naturalistic) were coded. Reaction time (RT) to events and stimuli while talking produced the largest performance decrements. Handheld and hands-free phones produced similar RT decrements. Overall, a mean increase in RT of .25 s was found to all types of phone-related tasks. Observed performance decrements probably underestimate the true behavior of drivers with mobile phones in their own vehicles. In addition, drivers using either phone type do not appreciably compensate by giving greater headway or reducing speed. Tests for moderator effects on RT and speed found no statistically significant effect size differences across laboratory, driving simulation and on-road research settings. The implications of the results for legislation and future research are considered.

Keywords: Cellular or mobile phones; Meta-analysis; Driver performance; Experimental methods

1. Introduction

Many governments, institutions and corporations are interested in the safety implications of cell phone use while driving. Approximately 44 countries have banned the use of handheld phones while driving (Cellular-News, 2007; Sundeen, 2005). In the U.S., passed legislation to date has been limited to three states (New York, Washington DC and California). Newfoundland is the only province in Canada that has passed handheld legislation. Numerous corporations and institutions have also developed policies that prohibit employees from using cell phones while driving. Many legislative bodies are in the process of considering a variety of restrictions on mobile phones. However, do the data justify these and similar prohibitions?

Research on cell phones and driving has been the focus of a large body of literature, especially over the past decade. The first study to specifically investigate the impact of cell phones on driving was published by Brown et al. (1969). Today, the volume of research on all aspects of driving with cell phones is about a meter in width. However, the research on the dangers of using mobile phones while driving is not consistent, with some arguing that the results represent an overreaction based on faulty methods (Shinar et al., 2005). Policy makers need a firm foundation upon which to make decisions. If driving while on a cell phone is truly dangerous, then legislation prohibiting their use should be more widely enacted. If the use of mobile phones while driving proves to have negligible effects, legislation that channels limited law enforcement time toward policing requires convergent and rigorous empirical support. Given this, we need to summarize what is known.

Unfortunately, most previous efforts to review the research regarding cell phones and driving have been traditional qualitative literature reviews (e.g., Caird and Dewar, 2007; Goodman et al., 1997; McCartt et al., 2006; National Transportation Safety Board, 2003). Though a popular methodology, there are limitations in their ability to synthesize data as well as examine in a systematic manner complex relations among the variables of interest (Hunter and Schmidt, 1990). Rather, meta-analysis is recommended.

Meta-analysis avoids some of the limitations of the standard literature review. It improves researchers' ability to statistically

* Corresponding author. Tel.: +1 403 220 5571; fax: +1 403 282 8249.
E-mail address: jkcaird@ucalgary.ca (J.K. Caird).

0001-4575/$ – see front matter
doi:10.1016/j.aap.2008.01.009

combine the results of many studies as well as reconcile conflicting findings through the examination of moderators and mediators (Rosenthal and DiMatteo, 2001). Importantly, meta-analysis enables a larger number of research dimensions to be included in the analysis. Meta-analysis encourages finding all available studies and determining the extent to which attributes of the set of studies support central hypotheses. Consequently, the purpose of this paper is to meta-analytically estimate the true effects of cell phone interaction on driver performance. First, we consider the findings of two previous meta-analyses on cell phones and driving as well as how our present paper builds upon their limitations.

1.1. Past meta-analyses on cell phones and driving

Two previous meta-analyses of cell phones and driving performance have been conducted, and these have served to inform researchers and policy makers. Horrey and Wickens (2006) analyzed 23 studies and Caird et al. (2004) synthesized 22 studies, both covering the literature through 2004. Common to both studies was a detrimental effect of cell phone use on reaction time and smaller effects on lateral control.

However, the research interest in cell phones and driver distraction has continued with intensity since 2004, providing a sufficient foundation for further updating the literature. Aside from considering a larger body of literature, the present meta-analysis systematically codes a wider range of effects than these previous meta-analyses. Reaction time to events, lateral and longitudinal vehicle control, and glance behavior are the general categories of measures that have been used in studies on cell phones and driving. Other important variables, such as speed and headway, were not included in previous analyses and results elsewhere are not entirely consistent.

For instance, knowing whether drivers compensate (e.g., reducing speed or increasing headway) has been hypothesized and observed by a number of researchers (e.g., Ålm and Nilsson, 1995), but other researchers have not found driver compensation (Ishida and Matsuura, 2001). Time or distance headway, or how close a driver gets to a lead vehicle while engaged in a distraction task, indicates that drivers who are more engaged in a conversation may come closer to the lead vehicle (Ranney et al., 2004). To offset losses of perceived safety when using a cell phone, drivers may engage in adaptive or compensatory behaviors, such as increasing their following distance (Ålm and Nilsson, 1995) and decreasing their speed (Ålm and Nilsson, 1994). In general, the pattern of results suggests that drivers adjust their headway when using a handheld phone, but not with a hands-free phone. Thus, inclusion of longitudinal control measures should provide a more complete picture of the trade-offs drivers make when absorbed to varying degrees by cell phone interaction.

In addition to driver compensation, other variables may mitigate the effect of cell phones on driving performance. First, cell phones and driving can be examined in a variety of experimental contexts, such as on the road and in driving simulators. Horrey and Wickens (2006) compared the results from field and simulation studies and found that effect sizes were similar, with field studies potentially showing greater costs. The simulation category included laboratory tasks where no tracking was required (e.g., Consiglio et al., 2003). We argue that a separate category is required to distinguish between true driving simulation studies and laboratory tasks, as the latter may not adequately impose a sufficient steering load on the driver. For example, Caird et al. (2004) found that laboratory tasks produced a larger effect size estimate than either field or simulator studies. In addition, laboratory studies may be relatively free of confounding variables that mask effects with additional experimental error. A more thorough investigation of this moderator variable has a number of methodological implications.

Second, conversation with a passenger may or may not be analogous to conversation on a cell phone. Passengers may adjust when they speak to the driver because they are attuned to what is going on in traffic and with the road geometry. To date, most studies have found that passengers do not seem to regulate their conversation with the driver (Gugerty et al., 2004) and both kinds of conversations are a source of distraction (Horrey and Wickens, 2006). However, a number of additional studies have appeared and the question of which source of distraction imposes a greater load on the driver (i.e., cell phone or passenger) requires reconsideration.

Third, Shinar et al. (2005) argues that the ‘phone tasks’ used by experimenters to approximate conversations are not representative of real conversations. In contrast, other researchers used naturalistic conversation where participants select from a list of topics to discuss with an experimenter prior to entering the testing sessions (e.g., Rakauskas et al., 2004). The experimental tasks, Shinar argues, are more likely to impose a higher cognitive workload on a driver. However, Horrey and Wickens (2006) found higher effect size estimates for conversation tasks than for experimental tasks, which is the opposite of Shinar’s (2005) results. A re-consideration of this issue with more studies may reconcile this discrepancy.

Fourth, the pattern of effects for older drivers is also somewhat equivocal. The impact of cell phones on older drivers appears to be more detrimental in some studies (Brookhuis et al., 1991), but not others (Ålm and Nilsson, 1995). Synthesis of effects based on available studies serves to clarify this important individual difference.

Finally, a number of important statistical issues have not been sufficiently addressed by these earlier meta-analyses on cell phone use while driving, such as zero-coding of effect sizes and failsafe-*N* estimates. For example, McCartt et al. (2006, p. 96) critiques the results of an earlier version of Horrey and Wickens (2006), not only for failing to include studies that included other driving measures (besides RT and tracking), but also for inadequately accounting for publication bias, where specific results have been systematically excluded. For example, many studies simply do not report non-significant effects. If these are ignored, it biases the results upwards, making the average effect size larger than it should be. Alternatively, we could estimate these non-significant effects as zero. Unfortunately, this simply reverses the bias and leads to an underestimate of the average effect size (Pigott, 1994). Both Horrey and Wickens (2006) and Caird et al. (2004) used this second approach to code non-

Table 1
List of studies and effects coded included in the meta-analysis

No.	Study	Phone type	Dependent measures	Conversation target	Setting	Conversation type
1	Ålm and Nilsson (1994)	HF v. Base	Lat. pos., speed	Cell	Sim.	Cognitive
2	Ålm and Nilsson (1995)	HF v. Base	RT	Cell	Sim.	Cognitive
3	Amado and Ulupinar (2005)	HF v. Base	RT	Cell, Passenger	Sim.	Cognitive
4	Beede and Kass (2005)	HF v. Base	RT, Lat. pos., speed	Cell	Lab.	Cognitive
5	Brookhuis et al. (1991)	HF v. Base	RT, Lat. pos.	Cell	On-road	Cognitive
6	Burns et al. (2002)	HH, HF, Base	RT	Cell, Passenger	Sim.	Cognitive
7	Burns et al. (2003)	HF v. Base	RT	Cell	Sim.	Cognitive
8	Chisholm et al. (2006)	HF v. Base	RT. Lat. pos., speed	Cell	Sim.	Naturalistic
9	Consiglio et al. (2003)	HF, HH, Base	RT	Cell, Passenger	Lab.	Naturalistic
10	Cooper et al. (2003)	HF v. Base	RT, speed	Cell	On-road	Cognitive
11	Graham and Carter (2001)	HF v. Base	RT, Lat. pos.	Cell	Lab.	Cognitive
12	Gugerty et al. (2004)	HF v. Base	RT	Cell, Passenger	Lab.	Cognitive
13	Horberry et al. (2006)	HF v. Base	Speed	Cell	Sim.	Cognitive
14	Horswill and McKenna (1999)	HF v. Base	RT, Head	Cell	Lab.	Cognitive
15	Irwin et al. (2000)	HH v. Base	RT	Cell	Lab.	Naturalistic
16	Kircher et al. (2004)	HH, HF, Base	RT, Lat. pos., Head, speed	Cell	Sim.	Cognitive
17	Laberge et al. (2004)	HF v. Base	RT, Lat. pos., speed	Cell, Passenger	Sim.	Cognitive
18	Lui and Lee (2005)	HF v. Base	Speed	Cell	On-road	Cognitive
19	Lui and Lee (2006)	HF v. Base	Lat. pos., speed	Cell	On-road	Cognitive
20	McCarley et al. (2004)	HF v. Base	RT, Lat. pos.	Cell	Lab.	Naturalistic
21	McPhee et al. (2004)	HF v. Base	RT	Cell	Lab.	Cognitive
22	Parkes and Hööijmeijer (2001)	HF v. Base	RT	Cell	Sim.	Cognitive
23	Patten et al. (2004)	HH, HF, Base	RT	Cell	On-road	Cognitive
24	Rakauskas et al. (2004)	HF v. Base	RT, Lat. pos., speed	Cell	Sim.	Naturalistic
25	Ranney et al. (2004)	HF v. Base	RT, Lat. pos., Head, speed	Cell	Sim.	Cognitive
26	Shinar et al. (2005)	HF v. Base	Lat. pos., speed	Cell	Sim.	Naturalistic, cognitive
27	Strayer and Drews (2004)	HF v. Base	RT, Head, speed	Cell	Sim.	Naturalistic
28	Strayer et al. (2003)	HF v. Base	RT, Head	Cell	Sim.	Naturalistic
29	Strayer and Johnston (2001)	HH, HF, Base	RT	Cell	Lab.	Naturalistic
30	Uno and Hiramatsu (2000)	HF v. Base	Speed	Cell	On-road	Cognitive
31	de Waard et al. (2001)	HH v. Base	Lat. pos., speed	Cell	Sim.	Cognitive
32	Waugh et al. (2000)	HH v. Base	Speed	Cell	On-road	Cognitive
33	Woo and Lin (2001)	HH v. Base	RT	Cell	Sim.	Naturalistic

Note: HH is handheld, HF is hands-free, Base is baseline, RT is reaction time, Head is headway, Lat. pos. is lateral position, Cell is cellular or mobile telephone, Sim. is driving simulator and Lab. is laboratory.

significant differences, meaning that true effects are likely larger than previously reported.

The preferred methodology for dealing with unreported effects is simply to minimize the issue as much as possible. By contacting the study authors directly, they can often provide the needed information (Lipsey and Wilson, 2001), which was the approach taken here. Furthermore, we also calculated failsafe-*N*, a measure of our results' stability. It determines how many "file drawer" studies (i.e., non-reported results) would be necessary to invalidate our findings.

2. Methods

2.1. Literature search

A comprehensive search for references, which was conducted over the span of approximately 4 years, used a variety of techniques including querying databases (e.g., SafetyLit, http://www.safetylit.org, PSYCinfo), cross-referencing bibliographies from reviews (e.g., Goodman et al., 1997; National Transportation Safety Board, 2003), index searches of proceedings (e.g., Human Factors and Ergonomics Society, Transportation Research Board), backtracking references from articles, conversations with numerous colleagues, contacting editors and authors, and searching the Internet. Acquisition of published and lower-circulation publications, such as technical reports was integral to the search process. Frequently, studies appeared in multiple sources, such as a proceedings paper, a technical report and/or a journal. Variations of this publication progression are common in the human factors literature. Either the most complete form of a study (e.g., technical report) or the highest quality outlet (i.e., a peer reviewed journal) was used. Publication redundancies were carefully eliminated, and statistical results were cross-checked across versions.

After the elimination of publication redundancies, a total of 106 studies were obtained that were published from 1969 to 2007. A number of additional studies did not meet independent and dependent variable criteria for inclusion. For the dependent variable, studies needed to measure: reaction time, headway, speed or lateral control.

Of the remaining studies, many failed to adequately report methodological details and statistical information including null results or, most commonly, critical post hoc tests. For example, authors frequently report only main effects tests but not critical follow-up tests. To address this shortcoming, approximately 30 authors were contacted to obtain additional statistical and methodological information. Approximately half of these provided the requested information. This left us with 33 studies or 31% of the original set.

2.2. Coding procedures

A total of 33 studies contributed 94 effects size entries. A complete list of studies and coded categories is found in Table 1 and their citations are listed in the references. Entries were double-coded and discrepancies were reconciled through reviewing the original study and analytical discussion. The coding categories of the study were: (i) phone type, (ii) dependent measures, (iii) research setting, (iv) type of conversation and (v) conversation target.

The first of these coding categories, phone type, was either handheld or hands-free. The second category, dependent variable, consisted of: reaction time (RT), lateral position, headway and speed. RT included measures of responses to stimuli, such as peripheral signals typical of secondary task probes (Patten et al., 2004) and events, such as the sudden emergence of a pedestrian (Laberge et al., 2004). Four studies specifically compared younger and older drivers on RT, and the mean age of older driver samples ranged from 64 to 69 years of age (Ålm and Nilsson, 1995; McCarley et al., 2004; McPhee et al., 2004; Strayer and Drews, 2004). Lateral vehicle control included standard deviation of lane position (SDLP) and similar measures of deviation within a lane. Headway included the distance to a vehicle in front of a driver in time or distance. Speed was operationalized as mean velocity. The third category, the research setting, was categorized into laboratory, simulation and on-road. The most commonly used laboratory task was tracking, an approximation of steering. Driving simulation had approximations of steering and the visual traffic environment and varied from the National Advanced Driving Simulator (NADS), which is the most advanced driving simulator in the world (Ranney et al., 2004), to low-fidelity simulators. The on-road category included studies that were conducted on test tracks and on actual roadways, frequently using instrumented vehicles.

The fourth category, conversation target, was coded as either passenger or non-passenger. Studies where someone other than the driver was seated in the vehicle and engaged in the conversation was coded as "passenger". In other studies, the driver conversed with someone who was not in the vehicle, often an experimental confederate who guided the conversation in topic and pace.

The fifth category, conversation type, was coded as either an information-processing or experimental task or a naturalistic conversation. A number of information processing tasks have been used including word and mathematics games. For instance, the paced serial addition task or PASAT (Baddeley et al., 1985) has been used in a number of cell phone studies to approximate conversation (e.g., Brookhuis et al., 1991). Protocols for naturalistic conversations typically have participants select a number of topics from a list to be discussed during the course of the experimental drives (e.g., Rakauskas et al., 2004).

2.3. Meta-analytic method

Statistical values (i.e., p, F, t) were extracted from the identified studies and converted to correlation effect sizes (i.e., r_{es}). Effect sizes were only analyzed where at least three coefficients were available, which is the typical minimum standard (Lipsey and Wilson, 2001). Where the combined effect size is based on fewer studies, the value can be unstable. All computations were performed using *MetaExcel* (Steel, 2006), except for the moderator analysis, which was conducted using SPSS (v.14). Aside from summarizing the weighted mean effect size (i.e., r_c and Cohen's d), the following additional analyses were conducted.

Confidence intervals (CoI) indicate the precision with which we are able to calculate the mean effect size. *Credibility intervals* (CrI) is an estimate of generalizability, reflecting the range of effect sizes that likely contain any *specific* situation. The larger the CrI, the more likely there are significant moderator effects. Consequently, we can have small confidence intervals, indicating we know the mean effect size precisely, but wide credibility intervals, indicating that specific effect sizes may differ widely from situation to situation (Whitener, 1990; Steel, 2007).

To account for situation-specific effects, weighted least squares (WLS) regression is recommended (Steel and Kammeyer-Mueller, 2002). Each case is weighted by the inverse of sampling error, providing significantly more accurate results than alternative meta-analytic moderator techniques. Finally, we calculated failsafe-*N*. Failsafe-*N* calculations estimate the number of null results, when combined with the current set of effects, that would be needed to reduce a meta-analytic co-efficient to ±.10 (i.e., a small correlation with lower practical significance; Rosenthal, 1995).

3. Results

As per Table 2, results of the meta-analyses are organized into mean effect sizes for reaction time, lateral positioning, headway and speed. Moderator variables – which included research setting, conversation target and conversation type – appear in Tables 3 and 4. To cast the results in terms of relevant values (Lipsey and Wilson, 2001), Table 5 reports an analysis of 26 studies that specifically examined reaction time (RT) increases while using cell phones. Finally, failsafe-*N* estimates are presented in Table 6.

3.1. Mean effect sizes

3.1.1. Reaction time

Table 2 lists the number of effect sizes (k) combined to calculate the mean effect size, the number of participants (N), two expressions of effect size (i.e., r_c and Cohen's d), as well as the lower (L) and upper (U) limits of the credibility (CrI) and confidence (CoI) intervals (see previous section for description). Because both r_c and d provide redundant information, the analysis will focus on r_c as the summary statistic.

The effect on RT when talking on either a handheld (HH; $r_c = .546$) or hands-free (HF; $r_c = .460$) phone was equivalent to the mean effect size of the two previous meta-analyses. The

Table 2
Meta-analyses of the effects of cell-phone use on driving performance variables

Variable	*k*	*N*	r_c	*d*	95% CrI		95% CoI	
					L	U	L	U
Reaction time								
Hands-free vs. baseline	37	1263	.460	1.03	.10	.82	.38	.53
Handheld vs. baseline	5	170	.546	1.29	−.17	1.00	.17	.92
Hands-free vs. handheld	3	118	.063	.12	.00[a]	.00[a]	−.02	.14
Younger vs. older drivers	4	140	.515	1.19	.06	.97	.26	.77
Lateral positioning								
Hands-free vs. baseline	15	456	−.152	.31	−.98	.68	−.39	.09
Headway								
Hands-free vs. baseline	5	224	.176	.36	−.39	.74	−.14	.49
Speed								
Hands-free vs. baseline	17	495	.230	.47	−.34	.80	.06	.40
Handheld vs. baseline	5	160	.394	.85	.00[a]	.00[a]	.26	.52
Hands-free vs. handheld	3	128	.135	.27	−.54	.81	−.39	.66

Note: *k*, Number of samples; *N*, total sample size; r_c, weighted mean correlations corrected for reliability; *d*, Cohen's *d*-effect size transformed from r_c; CrI, credibility interval; CoI, confidence interval.

[a] A .00 credibility interval reflects zero residual variance after accounting for sampling error variance.

Table 3
Moderator analyses for the effects of hands-free cell phone use on reaction time

Variable	*k*	*N*	r_c	*d*	95% CrI		95% CoI		R^2	*F*
					L	U	L	U		
Research setting										
Simulator vs. on-road	18	559	.422	.93	.11	.73	.32	.53	.140	3.419
Laboratory vs. simulator	14	528	.489	1.12	.18	.80	.38	.60	.027	.836
On-road vs. laboratory	5	176	.491	1.12	−.06	1.00	.19	.80	.055	.990
Conversation target										
Cell phone	28	853	.470	1.06	.11	.83	.38	.56	.007	.217
Passenger	7	306	.468	1.06	.15	.79	.31	.63		
Conversation type										
Cognitive	21	845	.417	.92	.09	.74	.32	.51	.035	1.024
Naturalistic	9	260	.480	1.09	.07	.89	.30	.66		

Note: * $p < .05$. *k*, Number of samples; *N*, total number of data points; r_c, weighted mean correlations corrected for reliability; *d*, Cohen's *d*-effect size transformed from r_c; CrI, credibility interval; CoI, confidence interval.

direction of the effect size difference indicated that slower reaction times (RT) resulted from conversing on either type of phone than in baseline conditions. The values reported for HH versus baseline studies ranged from .54 to .90, with the exception of one outlier reporting −.03. Study effect size outliers may represent imperfect methodological quality or measurement artifacts (Huffcutt and Arthur, 1995). In general, the influence of this, and outliers in subsequent analyses, is intended to illustrate the

Table 4
Moderator analyses for the effects of hands-free cell phone use on driving speed

Variable	*k*	*N*	r_c	*d*	95% CrI		95% CoI		R^2	*F*
					L	U	L	U		
Research setting										
Simulator	11	337	.338	.72	−.11	.79	.16	.51	.226	4.08
On-road	5	122	.074	.15	−.49	.63	−.27	.42		
Conversation type										
Cognitive	13	391	.227	.46	−.38	.83	.02	.43	.029	.425
Naturalistic	3	84	.140	.28	.04	.24	−.13	.41		

Note: *$p < .05$. *k*, Number of samples; *N*, total number of data points; r_c, weighted mean correlations corrected for reliability; *d*, Cohen's *d*-effect size transformed from r_c; CrI, credibility interval; CoI, confidence interval.

Table 5
Mean reaction time increases (i.e., drive with distraction—baseline drive), standard deviation of difference means, number of studies and number of participants

Variable or condition	Mean increase in reaction time (s)[a]	Standard deviation (s)	Number of studies	Number of participants
All distraction tasks	.25	.28	26	1170
Age				
Younger drivers	.19	.26	5	83
Older drivers	.46	.56	5	59
Task				
Handheld phone	.21	.16	5	157
Hands-free phone	.18	.29	16	518
Cognitive task	.33	.39	10	292
Conversation	.14	.13	11	348
Dial/enter number	.36	.12	6	511
Converse with passenger	.20	.13	3	84
Listen to radio/other	.05	.03	3	88
Event or stimuli				
BRT[b], lead vehicle brakes	.36	.42	7	630
BRT, light change at intersection	.18	.19	5	504
BRT, pedestrian	.19	.09	3	472
RT, peripheral detection	.20	.15	3	124
BRT/RT, abstract S-R[c]	.19	.26	3	68
RT, simple	.17	.12	6	146

[a] Means are averaged by study.
[b] Brake reaction time is analogous to perception response time (PRT).
[c] S-R is stimulus-response.

influence of specific values on the mean effect. By removing this study (i.e., −.03), the meta-analytic coefficient increases to .772 with variance accounted for increasing from 8% to almost 18%.

Consistent with our similar meta-analytic findings for HH and HF were three studies that compared them directly. The effect sizes of HF versus HH phone studies did not differ appreciably from one another ($r_c = .063$), both having similar decrements on driver reaction time.

Four studies tested differences in performance between older and younger drivers. As might be expected, older drivers were slower reacting to events and stimuli while conversing than younger drivers ($r_c = .515$).

Table 6
Failsafe-*N* estimates for meta-analytic effect sizes

Variable	*r*-Original	Failsafe-*N*
Reaction time		
Handsfree vs. baseline	.460	153.59
Handheld vs. baseline	.546	27.39
Handsfree vs. handheld	.063	n/a
Younger vs. older drivers	.520	19.92
Lateral positioning		
Handsfree vs. baseline	.152	7.96
Headway		
Handsfree vs. baseline	.176	3.87
Speed		
Handsfree vs. baseline	.230	22.93
Handheld vs. baseline	.394	16.31
Handsfree vs. handheld	.135	n/a

Note: *r*-Original, Meta-analytic correlation generated in current study (see Table 2); failsafe-*N*, number of unpublished papers with an average correlation of zero required to equal *r*-criterion. For all variables above, *r*-criterion, ±.10 (proposed "true score" correlation due to publication bias); *r*-failsafe-*N*, .00 (estimated average value for unpublished studies).

3.1.2. Lateral positioning

Studies were coded in a positive direction if driver performance in the HF condition was more variable than in the baseline condition, and negative when baseline was more variable than the HF condition. Too few studies with a HH phone condition measured lateral control to be meta-analyzed. The results indicate a negative corrected value of −.152, which means that participants in the HF condition were more variable in their lateral performance than those in the baseline condition. The relative size of the mean effect size indicates that measures of lateral control or lane tracking had minimal impact on driver performance while conversing on the phone, but across studies these effects were quite variable. The coefficients compiled for testing lateral control in the hands-free versus baseline condition range from .73 to −.72. A meta-analysis can synthesize inconsistent literatures, and correct for many statistical artifacts. However, the 15 studies analyzed yielded minimal reconciliation of essentially contradictory results.

3.1.3. Headway

A similar coding strategy was used for headway as was used for lateral positioning. A positive value indicated that in the HF condition a greater headway was adopted, whereas a negative value indicated that baseline performance yielded greater headway. Participants who used a HF phone allowed somewhat greater headway than in the baseline condition. However, the effect size was relatively small ($r_c = .176$) and confidence intervals crossed zero. By removing the outlier study (−.22) from this analysis, the effect size of .176, which accounts for 17%

of the variance, increased to .413 with 97% variance accounted for. The influence of the outlier study masks the larger headway seen in the other two studies in response to being on a hands-free phone.

3.1.4. Speed

After RT, mean speed was the second most commonly used measure across studies. A higher speed in the baseline condition was coded as positive, whereas a negative value indicated that using a cell phone resulted in driving faster. The positive values for HH and HF phones compared to baseline indicated that use of either type of phone resulted in slower speeds compared to baseline. The comparison between HH and HF indicated slightly slower speeds for those driving with a handheld phone in that the mean effect size difference for HF versus baseline ($r_c = .230$) was less than the difference between HH versus baseline ($r_c = .394$). Drivers with handheld phones appear to reduce their speed slightly more than those with hands-free phones. Three studies directly compared HH and HF phone use and speed-related driving performance. Kircher et al. (2004) found that HF slowed more than HH, and Patten et al. (2004), who contributed two effect sizes, indicated HH slowed more than HF. The net effect size ($r_c = .135$) favors the contribution with two studies.

3.2. Moderator variables

Using WLS regression, the moderator analyses are summarized in Table 3 for RT and in Table 4 for speed. Only HF cell phones were analyzed as there was consistently insufficient numbers of primary studies testing the effects of HH phones. No moderator analyses were significant.

Reaction time was not significantly moderated by either conversation target or type of conversation, indicating that talking with a passenger or another person on a cell phone has approximately equal costs on RT performance. Similarly, cognitive tasks and naturalistic conversations resulted in comparable RT performance decrements. Cognitive and naturalistic conversations differed slightly, but not significantly, in their effect on speed such that cognitive tasks produced slightly greater speed reductions by drivers.

This absence of significant differences between conditions may be because statistical power is an issue here—there was a relatively small number of coefficients available in each case.

Given this issue of statistical power, we expect that research setting may be an exception. For both RT and speed, the contrast between simulator and on-road research generated effects of $R^2 = .140$ and .226, respectively. Both also approached statistical significance, with $p = .079$ for RT and $p = .063$ for speed. Consequently, hands-free phone conversation may result in greater speed reductions during simulation studies than on-road. Calibration of perceived speed in driving simulators is frequently more difficult than actual driving due to the lack of binocular depth cues, among other visual information. As a result, drivers may have been less attentive to their speed maintenance while conversing. However, this effect only approached significance.

3.3. In-depth RT analyses

The primary purpose of the in-depth RT analysis was to make available to researchers and practitioners as many studies as possible in a form that was tangible (Lipsey and Wilson, 2001). The resulting categorized results may benefit researchers, accident reconstruction experts, modelers and policy makers who are seeking specific data about cell phone tasks or stimuli (Olson and Farber, 2003). To provide estimates of the impact of various independent variables on RT while using a cell phone, 26 studies, which partially overlapped with studies included in the meta-analysis, were analyzed in greater depth. Inclusion in this analysis was dependent on the availability of both baseline and RT means from text, tables or figures. A number of categories were coded including the age of participants (e.g., younger, older drivers), the task that the driver had to perform while driving (e.g., handheld, hands-free, conversation, dialing, listening), and the event or stimuli that the driver had to respond to while driving (e.g., lead vehicle braking, intersection yellow light, pedestrian incursion). Means were calculated based on study and not weighted by the number of participants within a study. This was done to constrain the contribution of certain laboratories where multiple cell phone studies used RT (Lipsey and Wilson, 2001; Rosenthal and DiMatteo, 2001). Results of the in-depth RT analysis are shown in Table 5.

The mean increase in seconds of RT over baseline, the standard deviation of this mean, the number of studies included in the calculation and the number of participants from the studies are listed. The independent variable categories of age, task and event or stimuli compose entry groupings. Overall, the average increase in RT for all tasks was .25 s. Younger drivers were less affected by cell phone tasks (.19 s) than older drivers (.46 s). Handheld (.21 s) and hands-free (.18 s) phones both increased RT. The use of a cognitive task to approximate a cell phone conversation had a higher RT cost (.33 s) than naturalistic conversation (.14 s) compared to baseline measures. Dialing or entering a number also had a relatively higher RT cost (.36 s). With the exception of lead vehicle braking (.36 s), other events or stimuli exacted approximately the same RT increase (.20 s). Overall, reaction time decrements varied somewhat depending on age, task, event or stimuli.

3.4. Failsafe-N estimates

The results do not appear to be susceptible to publication bias. As per Table 6, many of the estimates in the current study are quite robust. For example, about 154 additional null-effect studies would be required to reduce the difference in RT between talking on a hands-free phone and a baseline condition.

Conversely, some of the other estimates, specifically where there were few studies available to meta-analyze, are less resistant to the file drawer problem. The relative stability of the effect sizes is also evident in the results where outliers were considered. For example, lateral positioning and headway effect sizes may be reduced to a marginal level of practical significance with the addition of eight and four studies with null results, respectively. However, cell phone conversations do not appear to affect lateral vehicle control or induce drivers to increase following distance to a great extent anyway. Overall, the number of studies required to reduce the RT or speed effect sizes to a marginal level is sufficiently large that future and file-drawer studies are not likely to affect these results.

The influence of publication bias on results and the use of failsafe-*N* estimates to set limits on the influence of unpublished insignificant results (Rosenthal, 1995) is debatable. For instance, Scargle (2000) goes so far as to conclude that "fail-safe file-drawer (FSFD) analysis is irrelevant because it treats the inherently biased file drawer as unbiased and gives grossly wrong estimates of the size of the file drawer" (p. 102). However, publication of results in this domain such as technical reports, conference proceedings and journal articles reduces the likelihood that there are other studies out there sitting in file drawers. In addition, we made extensive efforts to contact authors to include as many additional studies as practically possible. Based on the variety of publication outlets available and efforts to contact authors to include additional studies, the influence of the file drawer problem on our reported results is probably less of an issue than in other meta-analysis domains.

4. Discussion

The purpose of this study was to synthesize a section of the overall body of literature on driving performance and cell phones so that a number of convergent results could be clearly articulated, and that inconsistent findings reported in primary studies may be resolved. By providing more precise estimates of driver performance with cell phones, legislation, public policy and future research are expected to benefit. The results presented encompass a larger body of research, improve on the statistics used in previous meta-analyses and examine a number of new analyses. Specific results are prioritized and implications for legislation and future research considered.

Foremost, cell phone conversation while driving increases reaction time to events and stimuli. This result is unequivocal in this meta-analysis and is consistent with previous meta-analyses in this literature (Caird et al., 2004; Horrey and Wickens, 2006). The meta-analytic estimates for the effects on RT range from .460 to .546. The total mean increase in RT analysis was .25 s for all manipulations in the in-depth RT analysis. Further, the addition of either potentially excluded (i.e., file drawer) or future studies with null results is unlikely to change this result.

Handheld and hands-free phones produced similar mean RT effect sizes, such that the meta-analytic estimates for each type of phone were essentially the same (see Table 2). Common sense and urban legend suggest that hands-free phones may be safer to use while driving, but handheld and hands-free phones produced similar performance decrements. This result has important implications for legislative restrictions and vehicle and device manufacturers of hands-free in-vehicle systems.

Restrictions on cell phone use appear to be only loosely informed by science and have almost been exclusively directed at handheld mobile phones (Sundeen, 2005). However, hands-free and handheld phones produce comparable crash risks (McEvoy et al., 2005) and similar driver reaction time decrements (see Tables 2 and 5). Extensions to existing legislation and broader new legislation that encompasses hands-free devices seems to logically follow from epidemiological and driver performance research. However, the pragmatic issue of how a police officer will observe hands-free use is problematic. Hands-free device use also has the potential to be confused with passenger conversation, which society regards as acceptable. Similar costs to RT were found for conversations over the phone and with passengers (see Table 4), and interactions with passengers are a primary category of distraction-related crashes (Caird and Dewar, 2007; Chen et al., 2000; Stutts et al., 2001). That hands-free and passenger distractions represent a crash risk will be difficult to change from the prevailing societal view of acceptability.

McCartt et al. (2006) argues that cell phone restrictions may only be partially successful, because compliance with cell phone restrictions is transient (McCartt and Geary, 2004). After legislation restrictions are introduced, observed rates of cell phone use while driving drops. Over time cell phone use increases, but does not necessarily return to pre-legislation levels. In an analysis of the effectiveness of the Washington DC restrictions, McCartt and Hellinga (2007) found that enforcement of cell phone bans produced longer term reductions in use.

Do drivers compensate for potential RT decrements? Drivers with handheld phones appear to reduce their speed slightly more than those with hands-free phones. The mean effect size differences for HF versus baseline ($r_c = .230$) and HH versus baseline ($r_c = .394$) are interesting from a driver adaptation viewpoint. However, the number of studies where direct comparisons between HH and HF phones were made is insufficient to conclude that drivers compensate to the imposed cell phone use demands. As a working hypothesis, drivers may be more aware of the potential safety threats imposed by handheld phones and reduce their speed to a greater degree than those conversing on hands-free phones. The physical presence of the phone in the hand may provide a reminder for the driver to compensate. More research is needed that compares HH and HF phones directly.

Cell phone conversations did not affect headway and lateral control measures appreciably. Drivers do not seem to consistently increase their following distance while on hands-free phones. The relative size of the mean effect indicates that measures of lateral control or lane tracking had minimal impact on driver performance while conversing on the phone, but that across studies measures were quite variable. Headway was affected by a single outlier, whereas lateral control results were entirely inconsistent. This may be perhaps due to the inclusion of a number of measurements into this category. The number of studies available was relatively small for headway and further research with more consistent measurement is recommended.

The results of this meta-analysis most likely represent optimistic estimates of actual performance decrements that drivers experience in their own vehicles. Drivers know that they are being observed and most likely perform to the best of their abilities given the experimental demands. Yet despite these more optimal performance conditions, RT decrements were still common. Actual driving behavior is likely to reflect behaviors that are worse than those observed in the studies examined here (Evans, 2004).

The meta-analysis and the in-depth RT analysis of older and younger driver performance provide convergent evidence for older drivers having slower RT performance while using cell phones. The mean effect size difference between younger and older RT was .515, which indicates older driver performance was affected to a greater degree than younger drivers. Four studies investigated the RT of older drivers while they used cell phones (Ålm and Nilsson, 1995; McCarley et al., 2004; McPhee et al., 2004; Strayer and Drews, 2004). None of the samples of older drivers from these studies exceeded 75 years of age, and thus the performance effects calculated here for older drivers probably underestimate the actual decrements.

Novice driver performance with cell phones has also received limited research focus (Chisholm et al., 2006; Greenberg et al., 2003; Shinar et al., 2005) and could not be meta-analyzed. Restrictions of cell phone use during a graduated licensing (GDL) period are being considered by many jurisdictions (Sundeen, 2005) based on a recommendation of the National Transportation Safety Board (2003). However, a solid empirical basis for this recommendation that is based on multiples studies is not yet available.

One obvious measure of cell phones and driving is crashes. A limited number of driving performance studies report collisions (Chisholm et al., 2006; Strayer and Drews, 2004). An in-depth investigation of five deaths and two injuries by the National Transportation Safety Board (2003) illustrates how using a cell phone contributed to the probable cause of the crash. Failures to detect hazards, underestimation of vehicle gaps and crashes were not possible to address in this meta-analysis and require future consideration.

For reaction time and speed moderator analyses, on-road studies produced similar effects to those from laboratory and simulation settings. These results are somewhat surprising because researchers often assume that laboratory tasks, for instance, have little or no resemblance to driving and thus are not necessarily generalizable. The results of our moderator analysis do not support the widely held belief that a single research paradigm or method is necessarily better than others. The most common form of this argument is that only on-road studies can yield valid and reliable results (Carsten and Brookhuis, 2005; Klauer et al., 2006). Each method has a number of strengths and weaknesses that impact internal and external validity. At least for cell phone conversations, the pattern of results is convergent across methods. Methods used by inexperienced researchers can result in low research quality, and our review of almost all studies in this domain indicates that laboratory, simulation and on-road methods each have a distribution of studies that range from low to exceptional quality. In other words, no single method has necessarily cornered the research market on either good or bad quality. However, study research quality could be coded in the future (Wortman, 1994). Bayesian meta-analysis (Louis and Zelterman, 1994) may further clarify effect size differences across methods.

Based on the conversation target analysis, talking with either a passenger or a person on a cell phone had approximately equal costs on RT performance. Naturalistic conversation tasks have been used more frequently in the literature since 2003 than in the past. However, Shinar et al. (2005) argue that mobile phone conversations, as used by researchers, are not necessarily indicative of real conversations because cognitive tasks impose a higher workload on drivers than ordinary talking. These authors found that cognitive tasks used by researchers had a higher mental workload and performance cost than naturalistic conversation. In contrast to this assertion, the moderator analyses in the current analysis indicate that word and number cognitive tasks appear to produce no greater or lesser cost to RT performance than naturalistic conversations.

5. Future research

In a broader context, driver distraction is the transient redirection of attention from the task of driving to any thought, activity, event or object (Caird and Dewar, 2007), and in addition to mobile phones, includes in-vehicle information systems, such as navigation and entertainment systems (Angell et al., 2006). For instance, visual tasks that require the driver to look away from the roadway to interact with a device affect measures of lateral vehicle control (e.g., standard deviation of lane position) and glance behavior (e.g., percent time off road). When one's eyes are off the road and one's hands are off the wheel, a common result is a degradation of hazard detection and lateral vehicle control (Chisholm et al., 2008; Horrey et al., 2006). Dialing and reaching for a mobile phone in a purse or briefcase are relevant examples of visual and physical distraction imposed by these specific cell phone tasks (Fuse et al., 2001; Tijerina et al., 1996). In contrast, talking on the phone imposes a non-uniform cognitive workload that is dependent on emotional intensity and task demands. Consequently, there is a different pattern of decrements associated with conversing. Conversation may constrain the breadth of search (e.g., Recarte and Nunes, 2000, 2003), the depth of attentional processing with an area of interest (e.g., Strayer et al., 2003), and increase RT to hazards (Tables 2 and 5). Visual scanning of peripheral sources of information, such as side and rearview mirrors and the speedometer is reduced while cognitively engaged, whereas more glances are made to the road centre, perhaps to preserve vehicle control (Recarte and Nunes, 2000, 2003). Lateral control is less affected.

Surprisingly, a research knowledge gap exists with respect to cell phone dialing and handheld phones. Handheld phones are used by the overwhelming majority of drivers (Glassbrenner, 2005), even when legislation does not permit their use while driving (McEvoy et al., 2005). Based on this meta-analysis, research has overwhelmingly focused on hands-free use ($k = 37$) and much less so on handheld phones ($k = 5$). We do not necessarily know why researchers focused on the former and not the latter.

The original scope of this meta-analysis was intended to include the effects of dialing on driving performance. However, only two studies had sufficient statistical information to include in an analysis (Chisholm et al., 2006; Lamble et al., 1999). A number of authors were contacted, but many did not respond to our requests and as a result a dialing task meta-analysis could not be performed. Increased lane variability and lane excursions have been observed. In a naturalistic study, removal of the hands from the steering wheel was more likely while talking on a cell phone, dialing or answering (Stutts et al., 2005). Dialing has been found to affect lateral vehicle control due to biomechanical interference with steering (Brookhuis et al., 1991; Tijerina et al., 1996). Dialing numbers on a handheld also requires drivers to look at the device before each number or button sequence is pushed and when initiating a call. As the number of digits to be dialed increases, the number of glances to the phone also increases. Glance frequency and glance duration to dial a number affects drivers' capability to scan for hazards and respond to them (Green, 2007). Based on the above, future empirical and meta-analytic studies may wish to focus on the effects of dialing (or visual distractions) on driving performance.

Glance or eye movement measures while using cell phones could not be synthesized because of measurement variability across studies; typical measures include fixation duration, fixation frequency, percent eyes off road time, horizontal and vertical gaze variability, pupil diameter and proportion of glances to specific areas of interest (AOI) (e.g., mirrors, roadway, hazards). Data quality from video based records, "automated" eye tracking systems and the use of dissimilar measures are issues that need to be resolved (ISO, 2002; Green, 2007). As additional results are published on glance behavior, combining the effect sizes across studies would be useful for understanding the basic mechanisms by which performance declines while conversing, dialing and answering. The hypothesis that holding a handheld phone to one side of the face may limit scanning and detection in that visual field on the same side has not been sufficiently explored.

Meta-analysis can also be applied to odds ratios (Woodward, 2005), such as those used to estimate crash risk when using a cell phone (McEvoy et al., 2005). Combining effect sizes from epidemiological studies, given that there are so few of sufficient quality, will be important for establishing a conclusive estimate of crash risk. In general, the presence of a handheld or hands-free phone increases crash risk by about four times (McEvoy et al., 2005; Redelmeier and Tibshirani, 1997).

Because of the perceived importance of driver distraction, the threshold of acceptance to journals appears to be lower for studies addressing the impact of cell phones or other distractions on driving. Remarkably, many prominent journals, including this one, do not uniformly require authors to provide complete statistical or methodological information. Editors, reviewers and authors should be required to correct these oversights. Calls for greater statistical accountability of published studies have gone unheeded despite the laments of researchers (e.g., Horrey and Wickens, 2006; Rosenthal, 1995). Determining that single degree of freedom statistical tests and null results are reported is important for meta-analysts who seek to resolve conflicting results or provide clear outcomes concerning contentious issues. Providing sufficient details about experimental manipulations and precise measurement definitions rarely meet the criterion of replication. To replicate a study, sufficient detail must be provided such that it could be repeated by carefully following the method description. Methodological and statistical details are usually the first and second locations where authors trim information to meet word and page limits of conference proceedings and journals. We argue that the long-winded introductions and speculative interpretations should be cut first. However, doing so will require authors to consider the parsimony of their rationale and interpretations. Given that many traffic safety descriptions and solutions have the potential to either save or take lives, meta-analyses may provide a more accurate estimation of the scope of problems or the relative effectiveness of treatments.

Acknowledgments

We would like to thank the numerous authors who responded to our requests for additional statistical analyses and clarification of statistical tests. We are grateful to Jenn Nicol for assistance double-coding the studies and to Bill Horrey for a valuable critique of the manuscript. This research was supported by contracts and grants to the first author initially from Canadian Automobile Association (CAA), Foundation for Traffic Safety and the University of California at Berkeley/PATH. Geoff Ho, Alison Smiley, Chip Scialfa and the first author wrote a report based on a meta-analysis of 22 studies (Caird et al., 2004), which also appears as an extended abstract in the Proceedings of the 3rd International Conference on Human Factors in Driving Assessment, Training and Vehicle Design in Portland, Maine. A new scan of the literature and substantial new coding and analysis resulted in the present manuscript, which was supported by the University of California at Berkeley/PATH and AUTO21 Network of Centres of Excellence (NCE).

References[1]

*#Ålm, H., Nilsson, L., 1994. Changes in driver behaviour as a function of handsfree mobile phones: a simulator study. Accident Analysis and Prevention 26 (4), 441–451.

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

#Al-Tarawneh, B.S., Cohen, W.J., Trachtman, D., Krauss, D.A., Bishu, R.R., 2004. The effect of hands-free cellular telephone conversation complexity on choice response time in a detection task. In: Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, Human Factors and Ergonomics Society, Santa Monica, pp. 2223–2227.

*Amado, S., Ulupinar, P., 2005. The effects of conversation on attention and peripheral detection: is talking with a passenger and talking on the cell phone different? Transportation Research: Part F 8, 383–395.

Angell, L., Auflick, J., Austria, P.A., Kochar, D., Tijerina, L., Biever, W., Diptiman, T., Hogsett, J., Kiger, S., 2006. Driver Workload Metrics Task 2 Final Report (Rep. No. DOT HS 810 635). National Highway Traffic Safety Administration, Washington DC.

Baddeley, A., Logie, R., Nimmo-Smith, I., Brerefon, N., 1985. Components of fluent reading. Journal of Memory and Language 24, 119–131.

*#Beede, K.E., Kass, S.J., 2005. Engrossed in conversation: the impact of cell phones on simulated driving performance. Accident Analysis and Prevention 38 (2), 415–421.

[1] Studies that are preceded by an "*"appear in the meta-analysis. Studies that were included in the in-depth RT analysis are denoted by a number of pound sign (#).

*Brookhuis, K.A., de Vries, G., de Waard, D., 1991. The effect 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.

*#Burns, P.C., Parkes, A., Burton, S., Smith, R.K., Burch, D., 2002. How dangerous is driving with a mobile phone? Benchmarking the impairment to alcohol (Tech. Rep. TRL 547). Transport Research Laboratory, Crowthorne, UK.

*#Burns, P.C., Parkes, A.M., Lansdown, T.C., 2003. Conversations in cars: the relative hazards of mobile phones. In: Proceedings of the International Ergonomics Association Meeting, International Ergonomics Association, Seoul, Korea (CD-ROM).

Caird, J.K., Dewar, R.E., 2007. Driver distraction. In: Dewar, R.E., Olsen, R. (Eds.), Human Factors in Traffic Safety, second ed. Lawyers & Judges Publishing, Tucson, AZ, pp. 195–229.

Caird, J.K., Scialfa, C., Ho, G., Smiley, A., 2004. The Effects of Cellular Telephones on Driving Behaviour and Accident Risk: Results of a Meta-Analysis. Canadian Automobile Association (CAA), Foundation for Traffic Safety, Edmonton, Alberta.

Carsten, O., Brookhuis, K., 2005. The relationship between distraction and driving performance: towards a test regime for in-vehicle information systems. Transportation Research: Part F 8, 75–77.

Cellular-News, 2007. Countries that ban cell phone use while driving. http://www.cellular-news.com/car_bans/. Accessed December 19, 2007.

Chen, L.-H., Baker, S.P., Braver, E.R., Li, G., 2000. Carrying passengers as a risk factor for crashes fatal to 16- and 17-year-old drivers. Journal of the American Medical Association 28 (12), 1578–1582.

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

*#Chisholm, S., Caird, J.K., Teteris, E., Lockhart, J., Smiley, A., 2006. Novice and experienced driver performance with cell phones. In: Proceedings of the 50th Annual Human Factors and Ergonomics Meeting, Human Factors and Ergonomics Society, Santa Monica, CA, pp. 2354–2358.

*#Consiglio, W., Driscoll, P., Witte, M., Berg, W.P., 2003. Effect of cellular telephone conversations and other potential interference on reaction time in a braking response. Accident Analysis and Prevention 325, 495–500.

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

#Cremades, J.G., Smith, D.W., Zhang, J.J., 1997. Hand-held car phones and driving performance: projected risks and implications for prevention. Journal of Health Education 28 (2), 98–102.

*de Waard, Brookhuis, K.A., Hernandez-Gress, N., 2001. The feasibility of detecting phone-use related driver distraction. International Journal of Vehicle Design 26 (1), 85–95.

Evans, L., 2004. Traffic Safety. Science Serving Society, Bloomfield Hills, MI.

Fuse, T., Matsunaga, K., Shidoji, K., Matsuki, Y., Umezaki, K., 2001. The cause of traffic when drivers use car phones and the functional requirements of car phones for safe driving. International Journal of Vehicle Design 26 (1), 48–56.

Glassbrenner, D., 2005. Driver cell phone use in 2005—overall results (Rep. No. DOT HS 809 967). National Highway Traffic Safety Administration, Washington DC.

Goodman, M.J., Bents, F.D., Tijerena, L., Wierwille, W., Lerner, N., Benel, D., 1997. An Investigation of the Safety Implications of Wireless Communications in Vehicles (DOT HS 806-635). National Highway Transportation Safety Administration, Washington DC. http://www.nhtsa.dot.gov/people/injury/research/wireless/.

*#Graham, R., Carter, C., 2001. Voice dialing can reduce the interference between concurrent tasks of driving and phoning. International Journal of Vehicle Design 26 (1), 30–47.

Green, P., 2007. Where do drivers look while driving (and for how long)? In: Dewar, R.E., Olsen, R. (Eds.), Human Factors in Traffic Safety, second ed. Lawyers & Judges Publishing, Tucson, AZ, pp. 57–82.

Greenberg, J., Tijerina, L., Curry, R., Artz, B., Cathey, L., Grant, P., Kochhar, D., Kozak, K., Bloomer, M., 2003. Evaluation of driver distraction using an event detection paradigm. In: 82nd Annual Transportation Safety Board Meeting, Transportation Research Board, Washington DC (CD-ROM).

*Gugerty, L., Rakauskas, M., Brooks, J., 2004. Effects of remote and in-person verbal interactions on verbalization rates and attention to dynamic spatial scenes. Accident Analysis and Prevention 36, 1029–1043.

#Hancock, P.A., Hashemi, L., Howarth, H., Ranney, T., 1999. The effects of in-vehicle distraction on driver response during a critical driving maneuver. Transportation Human Factors 1 (4), 295–309.

#Hancock, P.A., Lesch, M., Simmons, L., 2003. The distraction effects of phone use during a crucial driving maneuver. Accident Analysis and Prevention 35, 510–514.

*Horberry, R., Anderson, J., Regan, M.A., Triggs, Y.J., Brown, J., 2006. Driver distraction: the effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. Accident Analysis and Prevention 38 (1), 185–191.

*Horswill, M.S., McKenna, F.P., 1999. The effect of interference on dynamic risk-taking judgments. British Journal of Psychology 90, 189–199.

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

Horrey, W.J., Wickens, C.D., Consalus, K.P., 2006. Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. Journal of Experimental Psychology: Applied 12 (2), 67–78.

Huffcutt, A.I., Arthur, W., 1995. Development of a new outlier statistic for meta-analytic data. Journal of Applied Psychology 80 (2), 327–334.

Hunter, J.E., Schmidt, F.L., 1990. Methods of Meta-Analysis: Correcting Error and Bias in Research Findings. Sage, Newbury Park, CA.

International Standards Organization, 2002. Road vehicles – Measurement of Driver Visual behaviour with Respect to Transport Information and Control Systems – Part 1: Definitions and Parameters, ISO committee Standard 15007-1. International Standard Organization, Geneva, Switzerland.

*#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.

#Ishida, T., Matsuura, T., 2001. The effect of cellular phone use on driving performance. IATSS Research 25 (2), 6–14.

*Kircher, A., Vogel, K., Tornros, J., Bolling, A., Nilsson, L., Patten, et al., 2004. Mobile telephone simulator study (Rep. No. 969A). Swedish National Road and Transport Research Institute, Linköping, Sweden.

Klauer, S.G., Dingus, D.R., Neale, T.A., Sudweeks, J., Ramsey, D.J., 2006. The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic study data (Rep. No. DOT HS 810 594). National Highway Traffic Safety Administration, Washington DC.

*#Laberge, J., Scialfa, C., White, C., Caird, J.K., 2004. The effect of passenger and cellular phone conversations on driver distraction. Transportation Research Record 1899, 109–116.

#Lamble, D., Kauranen, T., Laakso, M., Summala, H., 1999. Cognitive load and detection thresholds in car following situations: safety implications for using mobile telephone (cellular) telephones while driving. Accident Analysis and Prevention 31, 617–623.

Lipsey, M.W., Wilson, D.B., 2001. Practical Meta-Analysis. Sage Publications, Thousand Oaks, CA.

Louis, T.A., Zelterman, D., 1994. Bayesian approaches to research synthesis. In: Cooper, H., Hedges, L.V. (Eds.), Handbook of Research Synthesis. Russell Sage Foundation, New York, pp. 411–422.

*Lui, B.-S., Lee, Y.-H., 2005. Effects of car-phone use and aggressive disposition during critical driving maneuvers. Transportation Research: Part F 8, 369–382.

*Lui, B.-S., Lee, Y.-H., 2006. In-vehicle workload assessment: effects of traffic situations and cellular telephone use. Journal of Safety Research 37 (1), 99–105.

*McCarley, J.S., Vais, M.J., Pringle, H., Kramer, A.F., Irwin, D.E., Strayer, D.L., 2004. Conversation disrupts change detection in complex traffic scenes. Human Factors 46 (3), 423–436.

McCartt, A.T., Geary, L.L., 2004. Longer term effects of New York State's law on drivers' handheld cell phone use. Injury Prevention 10, 11–15.

McCartt, A.T., Hellinga, L.A., Bratiman, K.A., 2006. Cell phones and driving: review of research. Traffic Injury Prevention 7, 89–106.

McCartt, A.T., Hellinga, L.A., 2007. Longer term effects of Washington DC, law on drivers' hand-held phones. Traffic Injury Prevention 8 (2), 199–204.

McEvoy, S.P., Stevenson, M.R., McCartt, A.T., Woodward, M., Haworth, C., Palamara, P., Cercarelli, R., 2005. Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study. British Medical Journal 331, 428.

*McPhee, L., Scialfa, C., Dennis, W., Ho, G., Caird, J.K., 2004. Age differences in visual search for traffic signs during a simulated conversation. Human Factors 46 (4), 674–685.

National Transportation Safety Board, 2003. Ford Explorer Sport collision with Ford Windstar Minivan and Jeep Grand Cherokee on Interstate 95/495 near Largo, Maryland February 1, 2002 (Rep. No. NTSB/HAR-03/02). National Transportation Safety Board, Washington DC, http://www.ntsb.gov/publictn/2003/HAR0302.pd.

Olson, P.L., Farber, E., 2003. Forensic Aspects of Perception and Response, second ed. Lawyers and Judges Publishing, Tucson, AZ.

*#Parkes, A.M., Hööijmeijer, V., 2001. Driver situation awareness and car phone use. In: Proceedings of the 1st Human-Centered Transportation Simulation Conference. University of Iowa, Iowa City, IA (ISSN 1538-3288).

*#Patten, C.J.D., Kircher, A., Östlund, J., Nilsson, L., 2004. Using mobile telephones: cognitive workload and attention research allocation. Accident Analysis & Prevention 36, 341–350.

Pigott, T.D., 1994. Methods for handling missing data in research synthesis. In: Cooper, H., Hedges, L.V. (Eds.), Handbook of Research Synthesis. Russell Sage Foundation, New York, pp. 163–175.

*Rakauskas, M., Gugerty, L., Ward, N.J., 2004. Effects of cell phone conversations with naturalistic conversations. Journal of Safety Research 35, 453–464.

*#Ranney, T., Watson, G., Mazzae, E.N., Papelis, Y.E., Ahmad, O., Wightman, J.R., 2004. Examination of the distraction effects of wireless phone interfaces using the National Advanced Driving Simulator-Preliminary report on freeway pilot study (Pre. No. DOT 809 737). National Highway Transportation Safety Administration, East Liberty, OH.

Recarte, M.A., Nunes, L.M., 2000. Effects of verbal and spatial imagery tasks on eye fixations while driving. Journal of Experimental Psychology: Applied 6, 31–43.

Recarte, M.A., Nunes, L.M., 2003. Mental workload while driving: effects on visual search, discrimination, and decision making. Journal of Experimental Psychology: Applied 9 (2), 119–137.

Redelmeier, D.A., Tibshirani, R.J., 1997. Association between cellular-telephone calls and motor vehicle collisions. The New England Journal of Medicine 336 (7), 453–458.

Rosenthal, R., 1995. Writing meta-analytic reviews. Psychological Bulletin 118 (2), 183–192.

Rosenthal, R., DiMatteo, M.R., 2001. Meta-analysis: recent developments in quantitative methods for literature reviews. Annual Review of Psychology 52, 59–82.

Scargle, J.D., 2000. Publication bias: the file drawer problem in scientific inference. Journal of Scientific Exploration 14 (1), 91–106.

*Shinar, D., Tractinsky, N., Compton, R., 2005. Effects of practice, age, and task demands, on interference from a phone task while driving. Accident Analysis and Prevention 37, 315–326.

Steel, P., 2006. MetaExcel (Version 10.0) [Computer software]. Retrieved from http://www.ucalgary.ca/(steel/procrastinus/meta/meta.html.

Steel, P., 2007. The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. Psychological Bulletin 133 (1), 65–94.

Steel, P., Kammeyer-Mueller, J., 2002. Comparing meta-analytic moderator search techniques under realistic conditions. Journal of Applied Psychology 87, 96–111.

*#Strayer, D.L., Drews, F.A., 2004. Profiles in driver distraction: effects of cell phone conversations on younger and older drivers. Human Factors 46 (4), 640–649.

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

*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), 1–5.

Stutts, J., Feaganes, J., Reinfurt, D., Rodgman, E., Hamlett, C., Gish, K., et al., 2005. Driver's exposure to distractions in their natural driving environment. Accident Analysis and Prevention 37, 1093–1101.

Stutts, J.C., Reinfurt, D.W., Staplin, L., Rodgman, E.A., 2001. The Role of Driver Distraction in Traffic Crashes. AAA Foundation for Traffic Safety, Washington DC. http://www.aaafoundation.org/projects/index.cfm?button=distraction.

Sundeen, M., 2005. Cellphones and Highway Safety: 2005 State Legislative Update. National Conference of State Legislatures, Denver, CO. http://www.ncsl.org/programs/transportation/cellphoneupdate05.htm.

Tijerina, L., Kiger, S., Rockwell, T., Tornow, C., 1996. Heavy vehicle driver workload assessment, Task 7A: in-car text message system and cellular phone use by heavy vehicle drivers on the road (Rep. No. DOT HS 808 467 7A). National Highway Traffic Safety Administration, Washington DC.

#Tokunaga, R.A., Hagiwara, T., Kagaya, S., Shimojyo, A., 2000. Cellular telephone conversation while driving: effects on driver reaction time and subjective mental workload. Transportation Research Record 1724, 1–6.

*Uno, H., Hiramatsu, K., 2000. Effects of auditory distractions on driving behavior during lane change course negotiation: estimation of spare mental capacity as a index of distraction. JSAE Review 21, 219–224.

*Waugh, J.D., Glumm, M.M., Kilduff, P.W., Tauson, R.A., Symth, C.C., Pillalamarri, R.S., 2000. Cognitive workload while driving and talking on a cellular phone or to a passenger. In: Proceedings of the IEA/HFES 2000 Congress, Human Factors and Ergonomics Society, Santa Monica, CA, pp. 6-276–6-279.

Whitener, E.M., 1990. Confusion of confidence intervals and credibility intervals in meta-analysis. Journal of Applied Psychology 75, 315–321.

*Woo, T.H., Lin, J., 2001. Influence of mobile phone use while driving. IATSS Research 25 (2), 15–19.

Woodward, M., 2005. Epidemiology: Study Design and Data Analysis, second ed. CRC Press, Boca Raton, FL.

Wortman, P.M., 1994. Judging research quality. In: Cooper, H., Hedges, L.V. (Eds.), Handbook of Research Synthesis. Russell Sage Foundation, New York, pp. 97–109.