Using Capacity Analysis to Evaluate the Effect of MP3 Player Usage on Driver Behavior

Shannon C. Roberts[1], John D. Lee[1], Joshua D. Hoffman[2]
[1]University of Wisconsin-Madison, Madison, WI, USA
[2]John Deere Company, Moline, IL, USA

Objective: To develop measures of driver interaction with in-vehicle devices through capacity analysis. **Background:** Common measures of driver behavior focus on the mean values of measures that do not consider temporal distribution of attention. Capacity analysis can quantify the temporal profile of a drivers' capacity and the resulting task efficiency. **Methods:** Fifty participants performed search tasks on either an MP3 player or its aftermarket controller in a medium-fidelity driving simulator. A capacity analysis used a hazard function to describe tasks according to their probability of completion. **Results:** Capacity coefficients, calculated as the ratio of two integrated hazard functions, were not constant over time indicating that the probability of task completion is dynamic and must be described with more than a measure of the central tendency of distribution, such as mean task completion time. **Conclusion:** Along with other commonly used driving performance metrics, capacity analysis can reveal task completion dynamics that are often neglected with reaction time measures, providing insight into how drivers direct their attention away from the road to complete secondary tasks.

INTRODUCTION

Much research examines how activities, such as cell phone conversations, draw attention away from driving and influence crash risk (McCartt, Hellinga, & Bratiman, 2006; Regan, Lee, & Young, 2008). Common metrics used to measure distraction include driving performance, eye movement, and task performance. All of these metrics help identify whether and how certain device characteristics interfere with driving. Typically, these metrics summarize data across a task or driving event and do not reveal patterns or details about the temporal distribution of attention to secondary tasks and the consequence for drivers' ability to handle demanding situations.

However, if each task is considered a process that must be completed in a certain amount of time, capacity analysis can be used to measure temporal engagement (Wenger & Gibson, 2004). Capacity analysis considers how efficiently a person can process information, reflecting cognitive resources required for a task (Wenger & Townsend, 2000). In relation to driving, each task the driver engages in, whether it is checking the mirrors or tuning a radio station, requires a portion of the driver's capacity. Problems occur when secondary tasks, such as cellular phone conversations, rob cognitive capacity from the primary task of driving.

Capacity analysis provides a more comprehensive description of the response process than contained in aggregate measures such as mean task completion time. The analysis can compare processing capacity across experimental conditions based on task completion times. Examples of capacity analysis are found in facial recognition (Wenger & Townsend, 2000), multisensory interactions (Hugenschmidt, Hayasaka, Peiffer, & Laurienti, 2010; Patching & Quinlan, 2004), and epistemic actions (Maglio, Wenger, & Copeland, 2008). These analyses assess differences between conditions over the entire distribution, not just the central tendency.

Wenger and Townsend (2000) used a hazard function to show that a fine-grained analysis of response times can measure capacity and the resulting task efficiency. The hazard function describes the probability of finishing the task within the next instant. The integrated hazard function represents the cumulative probability of finishing the task by a certain time, given that it has not yet been completed. The capacity coefficient is the ratio of two integrated hazard functions, one corresponding to each task condition. If the ratio is greater than 1.0, the condition in the numerator has a greater integrated hazard function value and supports greater task efficiency than the condition in the denominator. A capacity coefficient of 1.0 implies both conditions allow similar task efficiency.

The purpose of this study was to apply hazard functions and the associated capacity analysis to driver interaction with an MP3 player. MP3 players are of particular interest because over 90% of new vehicles in the United States have an option for iPod connectivity (http://www.apple.com/ipod/car-integration/). Three previous simulator studies (Salvucci, Markley, Zuber, & Brumby, 2007; Chisholm, Caird, & Lockhart, 2008; Crisler, et. al, 2008) examined the effects of MP3 player interactions on driving performance. The studies indicated that searching through long playlists resulted in increased reaction times to braking events, decrements in lane keeping performance, and longer glances directed towards the MP3 player. The capacity analysis complements previous studies' analyses by providing a temporal profile of drivers' capacity and the resulting task efficiency.

METHODS

Experimental Setup

Participants. Fifty people (29 males, 21 females), ages 18-25, participated in the experiment. Each participant had a minimum of two years driving experience and had normal or corrected-to-normal vision. All participants had owned an MP3 player for at least four months and used an MP3 player while driving at least once a week.

Simulated Driving Environment. Participants drove a medium-fidelity driving simulator with a full vehicle cab and standard controls. The field of view was approximately 50 degrees horizontally. The driving environment was a straight, five-lane suburban roadway, which included two lanes in each direction and a center turning lane. Participants were instructed to drive at a speed of 45 mph.

Experimental Design. There were four independent variables: interface (two levels), in-vehicle task (five levels), roadway demand (three levels), and demand timing (three levels). Interface was the only between subjects variable and consisted of either using an MP3 player (a white, fifth generation, 30 GB iPod with color screen and video capability) or an aftermarket controller designed for the MP3 player (Harman/Kardon Drive + Play). The aftermarket controller displayed information on a separate screen and had a separate rotary controller that mimicked the MP3 player controls. In-vehicle task was a within subjects variable and was comprised of: a radio tuning task using seek buttons, an MP3 song search task involving either short (20 songs), medium (75 songs), or long (580 songs) playlists, and a no task condition. Roadway demand varied between baseline, construction, and traffic zones, which were separated by a stoplight. The construction zone increased lateral demand with narrower lane widths. The traffic zone increased longitudinal demand by introducing more vehicles to the roadway. Demand timing referred to when the participant was asked to do the task relative to the demanding section of the roadway; it ranged from no-demand (i.e., no traffic), to just prior to demand (i.e., before entering traffic), to during the demanding sections (i.e., during traffic). The medium and long playlist tasks introduced prior to the demanding zone required drivers to complete the task during the demanding zone. Dependent measures included driving performance (speed, standard deviation of lane position, brake response time), eye movement (mean duration of glance, number of glances), and task performance (task initiation time, task completion time).

Procedure. Upon arrival, participants gave consent and completed the MP3 player usage questionnaire. Next, they drove a five-minute practice route that allowed them to become familiar with the simulator, experience the zone of increased roadway demand, and practice the search tasks while driving. Participants using the aftermarket controller were given extra time to become familiar with the interface. Participants experienced five different drives (each corresponding to an in-vehicle task) and were presented with the same in-vehicle task 16 times during each drive. Before each drive, participants received instructions specific to the drive. Following each drive, participants rated workload using the NASA TLX. At the end of the final drive, participants completed a post-experiment survey.

Capacity Analysis

Wenger and Townsend (2000) outlined a method using task completion times to quantify capacity. This method uses a hazard function to describe task completion and compared integrated hazard functions to identify human information processing limits. This analysis considers the drivers' capacity and task efficiency.

Data for three of the four independent variables were analyzed: interface (two levels), in-vehicle task (four levels, not including the no task condition), and roadway demand (three levels). First, completion times for tasks not finished and those interrupted by the following task were removed. Next, the range covering all task completion times (0-100 seconds) was divided into time bins of 0.1 seconds. The interval of 0.1 seconds was chosen to align with the accuracy of measurements. Each task completion time was assigned to one time bin and the bin totals for each condition were recorded. Then, the following calculations were performed:

Probability Density Function. The empirical probability, *f(t)*, or probability density function (PDF) gives the probability (*P*) that a task completion time (*T*) will be observed in a particular time bin (*t*). The PDF was calculated by dividing the number of bin observations (n_t) by the total number of observations (*n*) in each condition. The sum of the time bins' probabilities from an experimental conditions equals 1.0.

$$f(t) = P(T = t) = \frac{n_t}{n} \tag{1}$$

Cumulative Distribution Function. The empirical cumulative density function, *F(t)*, or CDF, gives the probability (*P*) that the task (*T*) has been completed at or before a particular time (*t*). Accumulating the empirical probabilities from the lowest to highest time bin produced the CDF.

$$F(t) = P(T \le t) = \int_0^t f(t^*)dt^* \tag{2}$$

Survival Function. The survival function, *S(t)*, is the complement to the CDF and gives the probability (*P*) that the driver is still in the process of doing the task (*T*) at a particular time (*t*). It was calculated by subtracting the time bin's CDF value from 1.0.

$$S(t) = P(T > t) = 1 - F(t) \tag{3}$$

Hazard Function. The hazard function, *h(t)*, gives the probability that the driver will finish the task in the next instant, assuming the task has not yet been completed. It was calculated by dividing the PDF by the survival function.

$$h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1-F(t)} \tag{4}$$

Integrated Hazard Function. The integrated hazard function, *H(t)*, gives the probability that the driver will finish the task by a certain time, given that the task has not yet been completed. It is also an indicator of how efficiently the driver can process information, i.e., a high integrated hazard function values indicates a high probability of task completion and a high efficiency of task completion. It was calculated by taking the reciprocal of the natural logarithm of the survival function.

$$H(t) = \int_0^t h(t^*)dt^* = -\ln\,[S(t)] \tag{5}$$

Capacity Coefficient. The capacity coefficient, *C(t)*, is a ratio of integrated hazard functions for each condition.

$$C(t) = \frac{H_1(t)}{H_2(t)} \tag{6}$$

If the ratio is greater than 1.0, the numerator condition has a greater integrated hazard function value and reflects a condition where the driver processes information more efficiently compared to the condition in the denominator. For example, with the radio task in the numerator and the short playlist task in the denominator, having a coefficient greater than 1.0 indicates that the driver is more likely to finish the radio task than the short playlist task by that time, given that the driver has not yet finished. It also shows that the radio task allows greater task efficiency than the short playlist task.

RESULTS

A total of 3,200 tasks (radio tuning and MP3 song selection) were administered. Ninety-seven tasks were removed because they were not completed, leaving 3103 tasks. The data were reduced using MatLab 9.0 and the capacity analysis described above was completed with an emphasis on the integrated hazard function and the capacity coefficient.

Integrated Hazard Function

The integrated hazard function for the MP3 player consistently exceeded the aftermarket controller, indicating that drivers were always more likely to complete the task using the MP3 player versus the aftermarket controller (Figure 1). During the first 20 seconds of the task completion time window, both integrated hazard functions had a similar shape, but the aftermarket controller diverged over time, particularly between 20 and 80 seconds. This indicated that the more time the task took to complete, the less likely drivers were to complete the task with the aftermarket controller compared to the MP3 player.

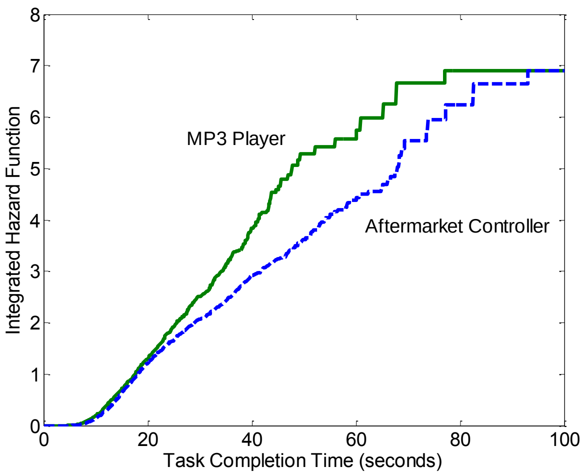

Figure 1. Integrated hazard functions for each interface (MP3 player and aftermarket controller).

Roadway type influenced the integrated hazard functions, with the baseline driving condition always having a higher probability of task completion (Figure 2). The traffic condition resulted in a greater integrated hazard function value compared to the construction condition. All three integrated hazard functions had a different shape. Specifically, the construction integrated hazard function diverged from the other two functions between 40 and 70 seconds. Similar to Figure 1, this indicated that the longer the task went on, the less likely drivers were to complete the task in the construction zone compared to baseline or traffic zones.

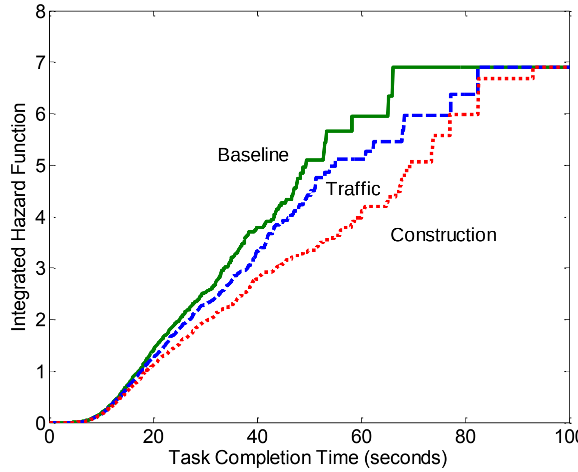

Figure 2. Integrated hazard functions for each roadway condition (baseline, traffic, and construction).

The effect of task type on the integrated hazard function was more complex than interface or roadway. Figure 3 shows that the radio, short, and medium playlist tasks produced similar probabilities for drivers to complete the task by a certain time. The long playlist task consistently showed a lower probability of task completion relative to the other tasks.

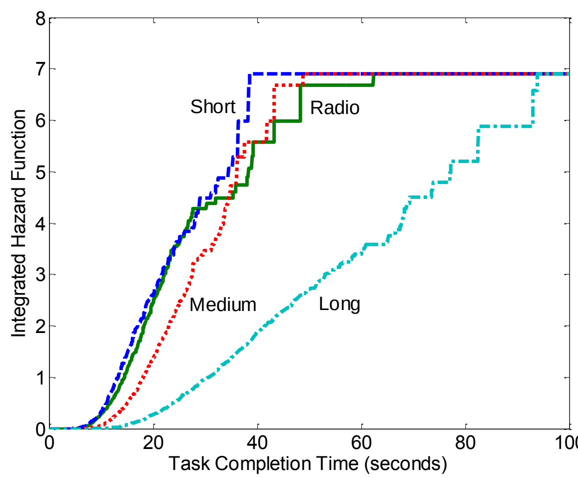

Figure 3. Integrated hazard functions for each task condition (radio, short, medium, and long playlist).

Capacity Coefficient

The capacity coefficient is the ratio of two integrated hazard functions. Because the MP3 player had a consistently higher integrated hazard function value (Figure 1) and the mean task completion time was lower (MP3 player: 17.2 seconds; Aftermarket: 19.0 seconds), the MP3 player condition was placed in the numerator of the capacity coefficient (Figure 4). The initial spike near 5 seconds reflects a minimum task completion time of 5.5 seconds for the aftermarket controller, compared to 4.8 seconds for the MP3 player. The capacity coefficient did not fall below 1.0, indicating that the MP3 player always supported greater task efficiency than the aftermarket controller.

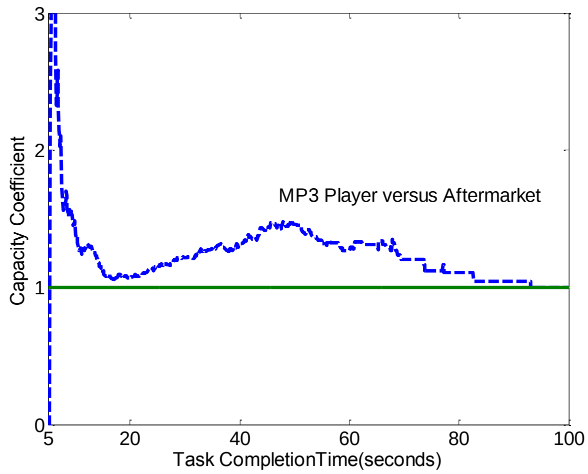

Figure 4. Capacity coefficient for the MP3 player vs. Aftermarket controller.

The interface capacity coefficient was not constant. Near 17 seconds, the MP3 player and the aftermarket controller supported similar task efficiency. However, near 50 seconds, the difference between interfaces reached a maximum. Near the end of the task completion window, the ratio leveled out and approached 1.0 as almost all tasks were completed. These results indicated that the MP3 player supported efficient task completion for tasks lasting less than 17 seconds and those lasting between 40 and 70 seconds. This analysis reveals what elements of task completion time distribution contribute to the lower mean task completion time for the MP3 player.

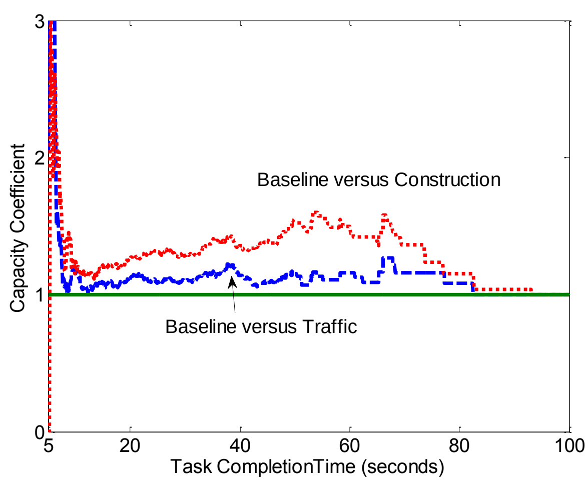

Figure 5. Capacity coefficient for baseline vs. traffic and baseline vs. construction conditions.

The baseline driving condition was placed in the numerator when computing the roadway capacity coefficients because its mean task completion time was lowest (Baseline: 17.1 seconds; Traffic: 18.0 seconds; Construction: 19.7 seconds). The baseline condition supported greater task efficiency than both the traffic and construction condition (Figure 5). Both ratios approached 1.0 between 9 and 12 seconds indicating that all three conditions supported similar task efficiency for tasks being completed at that time. However, the capacity coefficient for the baseline versus construction ratio was not as constant as the traffic versus construction ratio. Although both curves had a similar shape, the baseline versus construction curve reached a maximum between 40 and 70 seconds as it diverged from the baseline versus traffic curve. After 80 seconds, both curves converged towards 1.0, indicating similar task efficiency for tasks that take a long time to complete. These results indicated that the baseline condition allowed drivers to be more efficient than the construction condition for tasks less than 9 seconds and those between 40 and 70 seconds. This may reflect the ability to devote greater attention to tasks in low demand situations that is not possible in high demand situations, and that this greater attention allows tasks to be completed very quickly.

Although the short playlist task resulted in the lowest mean task completion time (Radio: 13.8 seconds; Short: 12.9 seconds; Medium: 17.3 seconds; Long: 29.4 seconds), the radio task was placed in the numerator of the capacity coefficient because radio tuning is often used as benchmark to compare tasks (Angell et al., 2006). The short playlist task resulted in greater task efficiency compared to radio tuning for tasks less than 20 seconds and tasks longer than 30 seconds (Figure 6). At 35 seconds, the medium playlist task also surpassed the radio task, indicating that it supported greater task efficiency than radio tuning, if the radio tuning task was not completed quickly. The long playlist task had the highest capacity coefficient indicating that it never supported greater task efficiency than radio tuning.

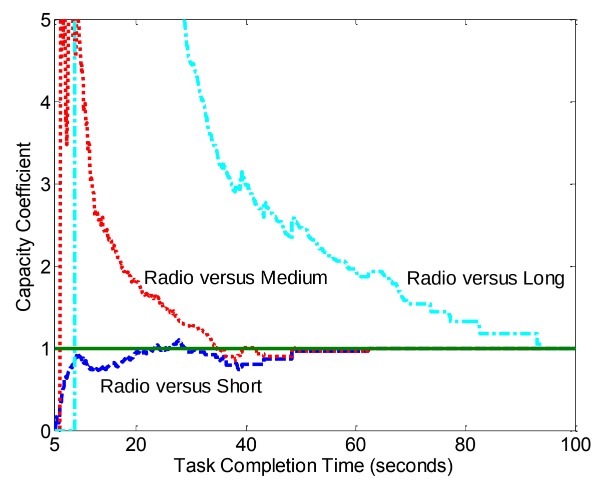

Figure 6. Capacity coefficient for radio vs. short playlist, radio vs. medium playlist, and radio vs. long playlist tasks.

The shape of the capacity coefficient curves in Figure 6 was considerably different from Figure 4 or 5. The radio versus short capacity coefficient hovered near 1.0 throughout the entire window as both tasks had nearly the same probability of completion. Both the medium and long playlist capacity coefficients gradually converged to 1.0 because both tasks did not have a greater probability of task completion than the radio task until 35 and 90 seconds, respectively.

DISCUSSION

The goal of this study was to determine whether the capacity analysis described by Wenger & Townsend (2000) would be valuable in understanding driver behavior. Task completion times were examined using capacity analysis to assess the effect of interface, in-vehicle task, and roadway on drivers' task efficiency. Such an analysis considers the full distribution of task times and not just a measure of central tendency, such as mean task completion time.

Radio tuning and short playlist tasks were similar in that they resulted in nearly the same probability of completion, given the task had not yet been completed. In comparison, the aftermarket device (versus the MP3 player), the long playlist task (versus radio tuning), and the construction roadway condition (versus the baseline roadway condition) led to poorer task efficiency since they had smaller integrated hazard function values. For all three variables analyzed, the capacity coefficient was not constant. For example, the interface capacity coefficient was highest for completion times less than 17 seconds and those between 40 and 70 seconds. This indicates that the MP3 player allowed drivers to be particularly efficient in performing tasks compared to the aftermarket device during those times. Unlike aggregate measures such as the mean, capacity analysis reveals how the probability of task completion changes over time and provides a precise and diagnostic measure to assess devices.

Conclusion

Capacity analysis reveals how interface, in-vehicle task, and roadway affect the probability of task completion over time. Across conditions, tasks with higher probabilities of completion suggest activities that are efficiently timeshared with driving. It also indicates what parts of the task completion time distribution contribute to differences in mean task completion time. Capacity analysis goes beyond commonly used metrics and gives insight into the temporal effect of secondary tasks on the driver's ability to handle demanding situations. Analyzing the temporal distribution of task completion complements previous research considering the distribution of eye glances (Horrey and Wickens, 2007), and demonstrates the benefit of analyzing other properties of distributions beyond measures of central tendency.

Future Work

Although the results from this study are intriguing and give insight into how device design affects efficient task completion, capacity analysis also has limitations. The results from the capacity analysis should not be used as a standalone metric in measuring driver behavior; they must be combined with other commonly used metrics. Applied to secondary task completion, capacity analysis considers how efficiently drivers perform a secondary task and does not consider consequences of this activity on driving. A task with a high capacity coefficient could be one that supports efficient time sharing, or it could be one that compels attention and distracts attention from the road. To address this, capacity coefficients need to be paired with measures of driving performance, such as standard deviation of lane position. Such an analysis would differentiate devices that support efficient time sharing and those that distract drivers.

Interpreting the capacity coefficient as an indicator of efficiency depends on the assumption that the tasks being compared involve a similar amount of work. In this study, the long and short playlist selection tasks violate this assumption as the long playlist task required drivers to scroll through more songs than the short playlist task. Future analysis might normalize across conditions by actions required by the task. It was also assumed that greater capacity implied more efficient task performance, such as when a task becomes easier as the driver nears completion. However, this could also reflect a change in task engagement, where the driver neglects the road to complete the task. Future research needs to refine the capacity measure to reflect the dynamics of attention to driving and secondary activities. Of particular interest are eye glance movements. Capacity analysis of such data could reveal the dynamics of eye glance behavior that complement the processes associated with decisions to engage in distracting activities.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers, Bryan Reimer, Ann Hundt, and the Cognitive Systems Laboratory members for their comments and suggestions. The data presented in this paper were collected as part of a project sponsored by General Motors under the guidance of Dr. Linda Angell.

REFERENCES

Angell, L., Auflick, J., Austria, P. A., Kochar, D., Tijerina, L., Biever, W.,et al. (2006). *Driver workload metrics project: Task 2 final report* (Tech. Rep. #DOT-HS-810-635). Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Chisholm, S. L., Caird, J. K., & Lockhart, L. (2008). The effects of practice with MP3 players on driving performance. *Accident Analysis and Prevention, 40*, 704-713.

Crisler, M. C., Brooks, J. O., Ogle, J. H., Guirl, C. D., Alluri, P., & Dixon, K. K. (2008). Effect of wireless communication and entertainment devices on simulated driving performance. *Transportation Research Record, vol. 2069*, (pp. 48-54), Washington DC: Transportation Research Board.

Horrey, W.J. & Wickens, C.D. (2007). In-vehicle glance duration: Distributions, tails and model of crash risk. *Transportation Research Record, vol. 2018*, (pp. 22-28), Washington DC: Transportation Research Board.

Hugenschmidt, C. E., Hayasaka, S., Peiffer, A. M., & Laurienti, P. J. (2010). Applying capacity analyses to psychophysical evaluation of multisensory interactions. *Information Fusion, 11*, 12-20.

Maglio, P. P., Wenger, M. J., & Copeland, A. M. (2008). Evidence for the role of self-priming in epistemic action: Expertise and the effective use of memory. *Acta Psychologica, 127*, 72-88.

McCartt, A. T., Hellinga, L. A., & Bratiman, K. A. (2006). Cell phones and driving: Review of research. *Traffic Injury Prevention, 7*, 89-106.

Patching, G. R. & Quinlan, P. T. (2004). Cross-modal integration of simple auditory and visual events. *Perception & Psychophysics, 66* (1), 131-140.

Regan, M. A., Lee, J. D., & Young, K. L. (2008). Driver Distraction: Theory, Effects and Mitigation. Boca Raton, Florida: CRC Press.

Salvucci, D. D., Markley, D., Zuber, M., & Brumby, D. P. (2007). iPod distraction: Effects of portable music-player use on driver performance. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 243-250). New York, NY: ACM Press.

Wenger, M. J. & Gibson, B. S. (2004). Using hazard functions to assess changes in processing capacity in an attentional cuing paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 30* (4), 708-719.

Wenger, M. J. & Townsend, J. T. (2000). Basic response time tools for studying general processing capacity in attention, perception, and cognition. *The Journal of General Psychology, 127*(1), 67-99.