Masthead Logo

University of Iowa
Iowa Research Online

Driving Assessment Conference

2005 Driving Assessment Conference

Jun 28th, 12:00 AM

Driver Workload Management During Cell Phone Conversations Chip Wood,

Joshua Hurwitz
Motorola Labs, Tempe, AZ

Chip Wood
Motorola Labs, Tempe, AZ

Follow this and additional works at: https://ir.uiowa.edu/drivingassessment

DRIVER WORKLOAD MANAGEMENT DURING CELL PHONE CONVERSATIONS

Chip Wood, Joshua Hurwitz
Motorola Labs
Tempe, Arizona, USA
E-mail: Chip.Wood@motorola.com

Summary: This study tested if intelligently suspending cell phone conversations during demanding driving situations would improve driver's performance and lessen subjective workload. Using a simulator, drivers followed a lead car approximately 2 secs ahead and braking randomly. The drivers engaged in spontaneous "Neutral," "Intense," or "No" cell phone conversations with the experimenter, who sustained all conversations with scripted non-directive utterances. When the lead car started to decelerate, the conversations were suspended or not for the duration of the event. Objective driving performance and subjective "workload" estimates showed significant differences between conversation topic types. Engaging in an "Intense" conversation during deceleration produced (1) smaller decreases in forward velocity, and (2) delays in releasing the accelerator, applying the brake and decelerating. Suspending "Intense" conversations counteracted these effects, producing larger decreases in forward velocity, and speeding up drivers' responses on the vehicle controls. Furthermore, suspending conversations clearly increased drivers' headways and temporal distances with the lead car. The results show that cell phone conversation intensity has significant effects on both performance and workload. The results also show positive effects of suspending cell phone conversations during critical driving situations on driver performance, thus perhaps counteracting the negative effects of the workload imposed by the conversations. Subjective ratings and value judgments for this system indicate generally positive assessments of the system.

INTRODUCTION

Research suggests that cell phone conversations, and other cognitive distracters, can produce specific deficits in driving performance, mainly increasing the time it takes for drivers to respond to unexpected events. For example, research on car-following indicates that, when the lead car suddenly decelerates, drivers performing a cognitive distraction task take longer to release the accelerator pedal(Hurwitz & Wheatley, 2001; Lee, McGehee, Brown, & Reyes, 2002). The focus in this study is on cellular communications. This focus arises from the concern raised by experts on the relationship between in-vehicle cell phone usage and driver safety (Independent Expert Group on Mobile Phones, 1999; Alm, & Nilsson, 1994; Parkes, & Hooijmeijer, 1999). This concern stems from research showing a correlation between accidents and cell phone usage (Redelmeier & Tibshirani, 1997), although other studies suggest that cell phones are just one of many potential distracters confronting drivers that are associated with accidents (Stutts et al., 2001).

The evidence suggests that increases in workload from cell phone conversations arise from three sources:

1. Conversing with someone who is not present in the vehicle
2. Actively processing information from that speaker
3. Producing verbal output in response to that speaker

If the driver is conversing with someone at another location, the remote individual is unaware of the driving situation confronting the driver, and so cannot adjust the conversation accordingly. Thus, to maintain the flow of the conversation, the driver must continue processing information from the other participant, even in the face of increasing complexity in the traffic situation. There is some evidence that in-vehicle participants who are conversing with the driver modulate the pace of the conversation in response to events in the traffic environment (Parks, 1991), thereby easing the driver's workload and even possibly providing warnings.

In order to display the effects of higher workload from a cell phone conversation, it is not sufficient that the driver be simply listening to another speaker. For example, the evidence suggests that driver response times to unexpected events do not increase when the drivers are simply listening to the radio. Rather, they need to be actively engaged in the conversation, processing information, and producing verbal responses (Strayer, Johnston, & Grison, 1999).

Given the active processing requirements of engaging cell phone conversations, and given that the other participant in such a conversation cannot adapt the conversational demands to account for increases in driving demands, this study was designed to discover whether intelligently suspending cell phone conversations during demanding driving situations would improve driver's performance and lessen subjective workload.

EXPERIMENTAL DESIGN

Facilities

The experiments were conducted in the Motorola driving simulator lab, which is an instrumented car in a surround video virtual world with full visual, audio, and force feedback simulation (although no motion or G-force simulation) of various roads, traffic and pedestrian activity. It consists of a fixed-based car (1997 Saturn LS) surrounded by three front video (approximately 150° forward) and 1 rear screen (approximately 50° rear view for center mirror and driver's side mirror). All driver controls such as steering wheel, brake, and accelerator are monitored and affect the motion through the virtual world in real time. Various hydraulics and motors provide realistic force feedback to driver controls to mimic actual driving.

The basic driving simulator software is a commercial product called DriveSafety from GlobalSim. Motorola has extended this standard product with a set of simulated sensors and controls that, at the behavioral level, simulate a rich set of current and potential future onboard sensors. This set consists of radar for locating other traffic, a GPS system for position information, a camera system for lane positioning and lane marking, a mapping data base for road names, directions, locations of points of interest, etc. There is also a complete car status system for determining the state of engine parameters (coolant temp, oil pressure, etc.) and

driving controls (transmission gear selection, steering angle, window and seat belt status, etc.). The simulator setup also has several video cameras, microphones and eye-tracking infrared sensors to record all driver actions during the drive that are synchronized with all the sensor output and simulator tracking variables. The nominal data rate is 60 samples/sec.

Drivers

There were 20 drivers, 10 males and 10 females. They had a mean age of 32.7 years (s.d. 12.3). The minimum was 20 years and the maximum was 65 years. The drivers had been licensed an average of 16.3 years (s.d. 12.5). Thirteen had had their license less than 13 years. Fifteen reported having had at most 1 accident, with 5 reporting 2 to 4 accidents. Twelve reported having had at most 1 moving violation, and 8 reported receiving between 2 and 4 violations. Twelve reported being moderately or very rested and 3 reported being moderately tired.

Driving Task

This study evaluated performance on a car-following task. It involved driving the car around a predominantly straight square 2-lane track. Each leg of the square was 2,800 meters, and drivers drove five legs for a total of 16.8 km, approximately 20 min/trip.

The drivers were instructed to drive around the track, maintaining a speed of between 45 and 65 mph (approx. 72.4 and 104.6 kph). As the subject car rounded the 1st curve to start Leg 1, another vehicle, the "lead car," pulled onto the track in front of it. This car was programmed to travel at a speed that allowed it to maintain a distance of 2 seconds with the subject car. It maintained this distance as long as the subject car was traveling at least 45 mph. Otherwise, it traveled at 45 mph. Drivers were instructed to follow the lead car and not pass it.

Occasionally the lead car would start to decelerate and no longer maintained the 2-sec margin. It decelerated at a rate of 8 m/s^2 until it reached a speed of 40 kph. It then accelerated again until it was traveling at least 72.4 kph, at which point it once again automatically maintained the 2-second distance with the subject car. The time between the start of the deceleration until the start of the acceleration was 7 seconds.

Conversations

Periodically, as the drivers drove around each track, they answered and conversed over a simulated hands-free cell phone call with the experimenter seated in another room. There were two such conversations during the course of each trip: an Intense conversation and a Neutral conversation. There was also a No-conversation control condition during each trip. Each conversation was approximately 3 minutes in duration.

The topics for Intense conversations focused on controversial political and social issues, whereas the topics for Neutral conversations focused on preferences for leisure activities and technology. Prior to starting the drives, each driver rank ordered the topics according to their interests. This rank ordering was used to determine the topics for the conversations used for that particular driver.

Each conversation began with a question (e.g., "What do you think of the death penalty?"). When the driver had finished an answer, the experimenter continued the conversation with a series of scripted non-directive questions derived from the "Eliza" simulation (Wallace and Dunlop).

Trips

Each driver "drove" 2 trips around the simulator tracks. The conversations were suspended on one of the trips and not suspended on the other, with the sequence of the conversations, suspension conditions across the trips, and the sequence of conversation topics within the trips determined according to a Latin Square design.

Conversation Suspension

According to the design, certain conversations were suspended when the lead car started to decelerate. When this happened, the driver was presented with three short tones, and the experimenter stopped talking. When the critical driving event was over, approximately 7 sec, the conversation was reactivated and the experimenter or the driver started talking again.

Procedure

After completing two practices drives, the driver, while seated and ready to drive in the car, was presented with prerecorded instructions for the first experimental task. They then performed the first experimental driving task, and periodically conversed with the experimenter over the simulated hands-free mobile phone inside their vehicle. After this first trip was completed, they filled out the modified NASA TLX form for that trip. If they had been exposed to suspension during that trip, they also filled out the form asking them for their subjective ratings of the suspension. After these forms were completed, they were presented with the pre-recorded instructions for the next trip. After the final trip was completed, the driver was asked to complete a "Driver History and Sleep Inventory" and a "Driver Behavior Inventory." They were then debriefed and paid for their time.

ANALYSIS AND RESULTS

Driving Performance

The analysis focused on the driver's performance during the deceleration events. The time was broken down into 1-second bins, and average values for the performance variables were computed for each bin. This analysis was performed for 5 variables: (1) forward velocity, (2) accelerator pressure, (3) brake pressure, (4) headway distance (at both minimal and 7sec), and (5) temporal headway (at 7sec). To determine statistical significance (unless otherwise indicated, all results were significant with $p < .01$), each variable was analyzed by performing, for each subject, a second-order polynomial regression of that variable over the time from the start of the lead-vehicle deceleration event until 7 seconds after the deceleration had started. This yielded three parameters for each subject: an intercept, a linear component and a quadratic component. The linear and quadratic components were then analyzed in a Repeated Measures Analysis of Variance, with the factors being Suspension (Yes vs. No) and Type (Intense vs. Neutral vs. None).

Forward Velocity. There was a trend toward a larger decrease in velocity under the *Intense-Suspended* than all other conditions. Also, there was a slight delay in the reduction in velocity under the *Intense-Not Suspended* condition. This is what would be expected if the intense conversation imposed more workload on drivers and suspending the conversation helped drivers attend to the driving task.

There was a significant Suspension X Type interaction ($F_{(2,36)}=5.64$, MSError=0.6256, $R^2=0.24$). This interaction was also significant for the quadratic component ($F_{(2,36)}=5.34$, MSError=0.0143, $R^2=0.23$). A planned comparison showed three effects ($t_{(1,19)}=3.29$, MSError=6.7544, $R^2=0.23$): The linear trend was more negative (drivers slowed more rapidly) when: (1) *Suspended* rather than *Not Suspended*, (2) *Suspended* rather than *None*, (3) *None* rather than *Not Suspended*.

Pair-wise comparisons showed that *Intense-Suspended* was more negative than *Intense-Not Suspended* ($t_{(1,19)}=2.82$, MSError=4.9797, $R^2=0.18$). Also *Intense-Suspended* was significantly more negative than *None* ($t_{(1,19)}=2.89$, MSError=5.2163, $R^2=0.19$).

The planned comparison for the quadratic component showed three effects ($t_{(1,19)}=3.12$, MSError=0.1387, $R^2=0.21$): (1) it took longer to initiate the deceleration when *Not Suspended* than when *Suspended*, (2) deceleration was more rapid when *Suspended* than when *None,* and (3) it took longer to decelerate when *Not Suspended* than when *None.*

Accelerator Pressure. The results showed a trend toward a Suspension X Type interaction for the linear trend ($F_{(2,36)}=3.65$, MSError=0.00286, $p<0.05$, $R^2=0.17$) and for the quadratic trend ($F_{(2,36)}=3.54$, MSError=0.00004, $p<0.05$, $R^2=0.16$). A planned comparison of the linear component revealed a non-significant trend in which drivers released the accelerator pedal at a higher rate in *Intense-Suspended* than all other conditions, and released it at a lower rate in the *Intense-Not-Suspended* than all other conditions ($t_{(1,19)}=2.12$, MSError=0.0129, $p<0.05$, $R^2=0.11$). A planned comparison of the quadratic component revealed a non-significant trend in which most of the accelerator release occurred earlier in *Intense-Suspended* than all other conditions, and occurred later in *Intense-Not Suspended* than all other conditions ($t_{(1,19)}=2.51$, MSError=0.0003, $p<0.05$, $R^2=0.15$).

Brake Pressure. A second-order polynomial regression was performed as a function of time during the deceleration event. A positive linear component indicates more pressure over time, negative is less. A positive quadratic indicates that the rate of pressure is increasing over time, negative is decreasing.

Results for the linear component showed a significant Suspension X Type interaction ($F_{(2,36)}=5.10$, MSError=0.00394, $R^2=0.22$). A planned comparison showed three results ($t_{(1,19)}=3.07$, MSError=0.0372, $R^2=0.21$): (1) *Intense-Not Suspended* decreased the rate at which the brake was applied more than *Neutral* or *None*, (2) *Intense* applied the brake at a higher rate than *Neutral* or *None*, and (3) *Intense-Suspended* applied the brake at a higher rate than *Intense-Not Suspended*. A paired comparison also showed this effect of suspending an intense conversation ($t_{(1,19)}=3.36$, MSError=0.0446, $R^2=0.24$).

Analysis of the quadratic trend also revealed a significant Suspension X Type interaction ($F_{(2,36)}=5.55$, MSError=0.00006, $R^2=0.24$). A planned comparison revealed two effects

(t(1,19)=3.32, MSError=0.0007, R^2=0.23): (1) the brake was applied earlier when *Intense-Suspended* than *Intense-Not Suspended*, and (2) the brake was applied earlier when *Intense-Suspended* than *Neutral* or *None*.

Headway. Two Repeated Measures ANOVAs were performed, one on minimum headway and the other on the headway 7 seconds after the deceleration event had begun.

The results for minimum headway were a trend toward a significant Suspension X Type interaction (F(2,36)=4.34, MSError=23.00, $p<.05$, R^2=0.19). A planned comparison was performed to test the hypothesis that there was a larger benefit for *Intense-Suspended* than for *Neutral-Suspended*, and smaller for *Not Suspended* than *None*. The results of this comparison were significant (t(1,19)=2.72, MSError=l70.33, R^2=0.17).

The results for 7 s headway were a trend toward a significant Suspension X Type interaction (F(2,36)=5.21, MSError=28.07, $p<.02$, R^2=0.22). A planned comparison was performed to test the hypothesis that headways at 7 s were larger when *Suspended*, and were smaller when *Not Suspended* than *None*. The results of this test were also significant (t(1,19)=2.87, MSError=231.00, R^2=0.19).

Temporal Distance. Results from the Repeated Measures ANOVA of temporal distance at 7 s showed a trend toward a significant Suspension X Type interaction (F(2,36)=4.41, MSError=0.20, $p<.02$, R^2=0.20). A planned comparison was performed to test the hypothesis that temporal distances were larger when *Suspended*, and were smaller when *Not Suspended* than *None*. The results of this planned comparison were significant (t(1,19)=3.24, MSError=2.14, R^2=0.23).

Subjective Workload

The NASA TLX data were analyzed using a Repeated Measures ANOVA to assess the impact of conversation suspension and conversation type on subjective workload. The hypothesis was that intense conversations would produce higher workloads than neutral ones, but that suspending the conversation would reduce workload.

The factors for this analysis were Suspension (yes vs. no) and Type (intense vs. neutral vs. none). An ANOVA was performed for each of the NASA TLX items: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration.

The results showed a significant effect of Type on mental demand, temporal demand, performance, and frustration. There was also a trend toward an effect of Type on effort and a significant effect of Suspension on performance.

Summarizing all of these effects is that drivers indicated that they perceived lower demand, higher performance, and lower frustration when *Neutral* than *Intense*. They also indicated that performance was better when the conversations were suspended (mean = 6.7, SEM = 0.34) than not suspended (mean = 6.0, SEM = 0.35).

Subjective Ratings

The drivers' ratings of conversation suspension were analyzed using chi-square tests. The null hypothesis was that there was no difference between the number of drivers who gave to each item ratings of "Definitely Agree" or "Mostly Agree" and the number who gave ratings of "Definitely Disagree" and "Mostly Disagree." The results indicate that drivers mostly (~65%) agreed that the system was helpful and helped them pay attention to their driving when they were distracted by the conversations. Most drivers (~60%) disagreed with statements asserting that the system interfered with driving and made it more difficult for them to concentrate on their driving. Furthermore, most indicated that they did not experience risk compensation, since they typically disagreed with the assertion that they could stop paying attention to their driving and rely only on the system to make them a safer driver.

Regarding the drivers' value judgments, most agreed with statements indicating that the system would be useful for drivers and would make them safer. Most also disagreed with the assertion that they would avoid buying a vehicle that came equipped with this system. There was no general agreement that drivers would be more likely to buy a vehicle with the system installed.

CONCLUSIONS

The results of this study provide support for the hypothesis that conversation intensity has significant effects on driver workload. The results also show positive effects of suspending conversations on driver performance, thus perhaps counteracting the negative effects of the workload imposed by the conversations. Finally, the drivers' subjective ratings of descriptions of conversation suspension and their value judgments for this system indicate generally positive assessments of the system.

The evidence for the workload effects of conversations comes from the NASA TLX results. The drivers indicated in their ratings that driving while engaged in an intense conversation was more mentally and temporally demanding, when compared to driving while engaged in a neutral conversation. They also indicated that it required more effort and was more frustrating, and that driving performance suffered when they engaged in intense conversations.

During car-following, the consequences of engaging in a demanding conversation during deceleration events were

- Smaller decreases in forward velocity
- Delays in releasing the accelerator, applying the brake and decelerating

Suspending intense conversations counteracted these effects, producing larger decreases in forward velocity, and speeding up drivers' responses on the vehicle controls. Furthermore, suspending conversations clearly increased drivers' headways and temporal distances with the lead car 7 s after the deceleration events had begun.

REFERENCES

Alm, H., & Nilsson, L. (1994). Changes in driver behavior as a function of handsfree mobile telephones. *Accident Analysis and Prevention, 26,* 441-451.

Hurwitz, J. B., & Wheatley, D. J. (2001). Driver Choice of Headway with Auditory Warnings. Paper presented at the 2001 Meeting of the Human Factors and Ergonomics Society, Minneapolis, MN.

Independent Expert Group on Mobile Phones. (1999). Report—Mobile Phones and Health. Summary and Recommendations Paper presented at the NHTSA Driver Distraction Forum.

Lee, J. D., McGehee, D. V., Brown, T. L., & Reyes, M. L. (2002). Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. *Human Factors, 44*(2), 314-334.

Parkes, A. M., & Hooijmeijer, V. (1999). The Influence of the Use of Mobile Phones on Driver Situation Awareness. The NHTSA Driver Distraction Forum.

Parks, A. M. (1991). Drivers business decision making ability whilst using carphones. In E. 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.

Strayer, D. L., Johnston, W. A., & Grison, S. (1999). Driven to distraction: studies of driving and cellular phone use. *Abstracts of the Psychonomic Society, 4*(16).

Stutts, J. C., Reinfurt, D. W., Staplin, L., & Rodgman, E. A. (2001). *The Role of Driver Distraction in Traffic Crashes* (Study Report). Washington, DC: AAA Foundation for Traffic Safety.

Tijerina, L., Kiger, S. M., Rockwell, T. H., & Tornow, C. (1995). Workload Assessment of In-Cab Text Messaging System and Cellular Phone Use by Heavy Vehicle Drivers on the Road. *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting,* San Diego, CA (pp. 1117-1121).

Wallace, M., & Dunlop, G. Available at http://www.manifestation.com/neurotoys/eliza.php3