

NATURALISTIC DRIVING PERFORMANCE DURING SECONDARY TASKS

James Sayer, Joel Devonshire, Carol Flannagan
University of Michigan Transportation Research Institute
Ann Arbor, Michigan, USA
E-mail: jimsayer@umich.edu

Summary: Data from 36 drivers involved in a naturalistic driving study was analyzed to determine the frequency and conditions under which drivers engage in secondary behaviors and to explore the relationship these behaviors might have with driving performance. Researchers coded 1,440 five-second video clips of the drivers' faces for the occurrence of specific secondary behaviors and the duration of glances away from the forward scene. Corresponding performance data from the instrumented vehicles were used to calculate variability of steering angle, mean and variability of lane position, mean and variability of throttle position, and variability of speed. All categories of secondary behavior were associated with significantly higher variability in steering angle. The results for other performance measures were mixed. In summary; driving performance measures vary with differing tasks, with no single driving performance indicator that is obviously predictive of drivers' engagement in secondary tasks.

INTRODUCTION

A field operational test (FOT) was conducted to assess the potential safety impacts, driver acceptance levels, and system maturity of an integrated lane-departure and curve-speed warning system (RDCW). This test was conducted by the University of Michigan Transportation Institute (UMTRI) with its partners Visteon Corporation and AssistWare Technologies (now Cognex Corporation), under a cooperative agreement with the U.S. Department of Transportation. This paper presents the results of analyses performed on the RDCW FOT program (LeBlanc et al., 2006), as well as analyses presented in a separate technical report (Sayer et al., 2005).

Video clips of driving behavior for 36 drivers (out of a pool of 78 drivers who took part in the field test) were analyzed. The mean ages of the drivers were 25.1, 45.6, and 64.2 years old for younger, middle, and older age groups, respectively. All participants were licensed drivers who had been driving for more than two years. Participants drove an RDCW-equipped vehicle as they would their own car for 26 days. A video camera mounted to the inside of the vehicle's A-pillar captured five-second images of the driver's face (at 10 frames/second) at five-minute intervals.

METHODOLOGY

Researchers examined a representative sample of 18,281 video clips from the FOT. The sample was not associated with any RDCW alerts, represented driving at least 25 mph (11.18 m/s), and included drivers with at least 10 qualifying video clips (equivalent to at least 50 minutes of driving) per week. A data acquisition system built specifically for the project collected over 500 channels of data at 10 or 20 Hz (see LeBlanc et al., 2006). To determine whether the RDCW system had an effect on the overall frequency of secondary behaviors from week to week, thus

representing a confounding variable, researchers compared the baseline period, when the system was inactive, to the following three weeks, when the system was active. There were very few differences in the frequency of secondary behaviors. As such, it was deemed appropriate to include data from all four weeks of the drivers' experiences.

Researchers coded the video clips for secondary behaviors using a custom data visualization tool created in Visual Basic. Videos could be played frame-by-frame at various speeds to measure, for example, how often and how long the driver's gaze was not directed toward the forward scene. Dependent measures included variability of steering wheel angle, mean and variability of lane position, variability of speed, and mean and variability of throttle position. Vehicle-based data was recorded at 10 Hz, so each five-second video clip had 50 individual data points for every measured variable. Means of driving performance measures were calculated over the duration of a five-second clip, such that each of the 1,440 clips had associated mean values.

Because time-series data often exhibits autocorrelation (i.e., each observation tends to be highly correlated with immediately preceding observations, violating the assumption of independent observations), the raw observations for each driving performance measure were fit with an autoregressive integrated moving average (ARIMA) model. The random error variance in these models is typically considered noise to eliminate from time-series analyses. In this case, however, these variances were precisely what we wanted to examine. After the best models were fit for each measure, the percent of autocorrelation still present in the data ranged from 8 to 15%.

RESULTS

Drivers were found to engage in secondary behaviors in about one-third of the reviewed clips (486 of 1,440). The most frequently observed secondary behavior was conversation with a passenger, which was present in 220 (or 15.3%) of the clips. Grooming was the second most common secondary behavior (in 6.5% of the clips), and using a hand-held cellular phone was the third most common (in 5.3% of the clips). The frequency of these three behaviors is consistent with the findings of Ervin et al. (2005), and similar to the findings of Stutts et al. (2003) as it relates to the relative frequency of conversation with passengers.

Table 1. Secondary behavior exposure review counts

Observed Behavior	<i>f</i>	%	Multiple Behaviors (<i>f</i>)
No secondary behavior	954	66.2	
Conversation	219	15.3	21
Grooming	96	6.5	26
Cellular phone	76	5.3	10
Eating/Drinking	28	1.9	2
Multiple behaviors	31	2.2	-
Other	36	2.5	5
Total	1,440	100	64

Table 1 shows the final frequency distribution. Because the category of multiple behaviors often included one or more of the categorized behaviors, the rightmost column in Table 1 provides the

frequency with which each individual behavior was observed within multiple behaviors. All drivers had at least four cases of observed secondary behaviors among their exposure clips, and the average per driver was 14 (SD = 6.2). Twenty-four of the 76 observed cellular-phone exposure clips (32%) came from just two drivers.

Linear mixed-effects models were fit on each of the seven driving performance measures. Models were initially fit using a two-level factor of secondary behavior: no secondary behavior versus all types of secondary behaviors combined. This was done to reduce the degrees of freedom and to determine whether secondary behaviors in general had an overall effect on the dependent/outcome measures. The models were then refit using a seven-level factor of secondary behavior (i.e., no secondary behavior and six individual types of behavior) to see if any specific behavior had a unique relationship to the outcome variables.

Unless otherwise specified, all models initially included the factors of age group (three levels), gender (two levels), secondary behaviors (two or seven levels), road type (two levels: limited access vs. all other roads), road condition (two levels: dry vs. wet/snowy), road curvature (two levels: curve vs. no curve), and brake use (two levels: brake application vs. no brake application). A method of “backwards” model selection was used for all analyses in which all main effects were initially included. Each model was then refit multiple times, each time excluding the main effect that was least significant. When only significant main effects remained, the model was again refit to include those main effects and their interaction terms. Finally, the nonsignificant interactions were removed to obtain the final model for each analysis.

The random variance among drivers was included as a parameter within each model. Thus, if the effects of between-subjects variables (age or gender) or within-subjects variables (secondary behaviors) were not significantly greater than random variance among drivers, they would not reach statistical significance in the model.

Glance Duration. Two mixed-effects model analyses were performed, one each on the mean duration of first and second glances away from the forward scene. While the first model found no significant effects, the model for second glances contained a significant main effect of secondary behaviors, $F(6, 479) = 2.52, p < .05$. While the difference between no secondary behaviors and the cellular phone category failed to reach significance when Bonferroni corrections were used for multiple pairwise comparisons, glance durations away from the forward scene were actually at their shortest when drivers used a cellular phone. The longest mean glance duration was associated with other behaviors, and had the greatest influence on the statistical significance of the factor.

Steering Angle Variance. A mixed-effects model was fit on steering angle variance. The main effect of secondary behavior was significant, $F(1, 33.1) = 10.6, p < .01$. Secondary behaviors were associated with greater variance in steering angle. Because the effect of secondary behaviors in the first model was significant, a second mixed-effects model was fit using the seven factors of secondary behavior. There, the effect of secondary behavior just failed to reach significance ($p = .053$). However, each secondary behavior was associated with a higher mean steering angle variance. Figure 1 displays the estimated means to 95% confidence intervals.

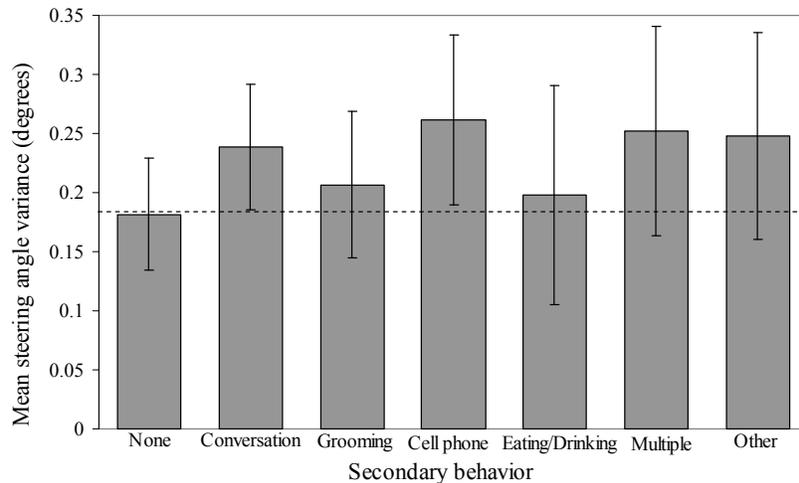


Figure 1. Mean steering angle variance (degrees) for each type of secondary behavior

Cell phone use was associated with the highest mean steering angle variance. Steering angle variance associated with eating/drinking was not as high, although the error was larger (presumably because of the low n in this group). It is plausible that the larger standard error in the last three categories of secondary behavior was a factor in the nonsignificant main effect. An analysis of the standard deviation of steering angle yielded results similar to the ARIMA fit.

There were no significant effects of age group or gender on steering angle variance. However, the main effect of brake use was significant, $F(1, 44.4) = 9.6, p < .01$. Steering angle variance was higher in clips where the driver was braking than when brakes were inactive (estimated mean variance of 0.29° and 0.18° , respectively). There was also a significant interaction between road type and road curvature on steering angle variance, $F(1, 37.4) = 14.1, p < .001$. A large mean steering angle variance was associated with curves on surface (non-limited access) roads. This is not surprising, as surface roads and ramps typically have sharper curves than limited access roads. Both main effects in this interaction were also significant: $F(1, 19.6) = 28.5, p < .0001$ for road curvature and $F(1, 48.9) = 22.3, p < .0001$ for road type. Steering angle variance was generally lower on limited access roads and when the driver was not negotiating a curve.

Mean Lane Position and Lane Position Variance. The distance from lane center (i.e., the absolute value of the lane position) was used to calculate the mean lane position. No distinction was made between being off-center to the left or right. The mixed-effects model for mean distance from lane center showed two significant effects: on curves, the mean distance from lane center was higher than when not on curves, $F(1, 1233) = 18.3, p < .0001$, and mean distance from lane center was also higher on wet or snowy roads than on dry roads, $F(1, 1193) = 8.47, p < .01$. There were no significant effects of age group, gender, or secondary behavior.

Table 2 lists observed frequencies of each secondary behavior in the 1,440 clips. Its rightmost columns represent the observed frequencies of secondary behaviors for a different sample (exposure videos used in the current analyses and exposure clips containing lateral drift warnings excluded from the current analyses). This latter sample represents 854 randomly selected events in which the driver had drifted sufficiently to prompt a lateral drift warning. While the leftmost columns represent driving while relatively centered in the lane, the rightmost columns represent

driving while drifting to the left or right. The percentages of observed secondary behaviors are quite similar in the two samples, suggesting that eliminating the clips that included RDCW alerts was unlikely to significantly alter the outcome of the analyses.

Table 2. Frequencies of observed secondary behaviors for exposures and lateral drift warnings

Secondary Behavior	Exposure Videos		Lateral Drift Warning Videos	
	<i>f</i>	%	<i>f</i>	%
None	954	66.2	548	64.2
Conversation	219	15.3	147	17.2
Grooming	96	6.5	61	7.1
Cellular phone	76	5.3	37	4.3
Eating/Drinking	28	1.9	12	1.4
Multiple	31	2.2	26	3.0
Other	36	2.5	23	2.7
Total:	1,440	100	854	100

Results for lane position variance were mixed. For the ARIMA-fitted measure, no significant differences were seen among any of the independent/predictor variables. However, the more common measure of standard deviation of lane position showed several significant differences. The effect of secondary behaviors on the standard deviation of lane position was significant, $F(6, 326) = 2.2, p < .05$. No direct relationship emerged from examining the means; conversation, cellular phone use, and multiple behaviors were associated with greater variability in lane position, while other secondary behaviors were associated with lower variability.

There was a significant effect of age group on the standard deviation of lane position, $F(2, 40.5) = 4.3, p < .05$. There was higher variability in lane position for the younger age group (0.19 m standard deviation), while the middle and older age groups showed a similar level of variability (.15 m and .16 m, respectively). There was no significant effect of gender.

Mean Throttle Position and Throttle Variance. The mixed-effects model for mean throttle position showed no significant main effect of secondary behavior (for neither the two- nor seven-level factors). There were, however, two expected effects: brake use was associated with lower mean throttle positions, $F(1, 92.7) = 411.2, p < .0001$ and driving on limited access roads was associated with higher mean throttle positions, $F(1, 47.9) = 18.4, p < .0001$. There were no significant effects of age or gender on mean throttle position.

The results for mean variance in throttle position were difficult to interpret. Mixed-effect model analyses of both the two- and seven-level factors of secondary behavior did not show any significant main effects. Nonetheless, it is worthwhile to compare the means of throttle variance by secondary behavior to those of steering angle variance. Many behaviors were associated with a higher throttle variance (similar to the findings for steering angle variance), but few behaviors (noticeably eating/drinking) were associated with lower mean variances.

The percentage of cases in which either throttle or braking behavior was present as a function of secondary behavior type was also examined. While it seems that there is no clear relationship between secondary behavior and throttle variance, examining the data slightly differently raises

an interesting issue. In the majority of cases (in which no secondary behavior was observed), throttle engagement (by any amount and/or duration) occurred in about 90% of the cases, while brake engagement (by any amount and/or duration) occurred in 13% of cases. However, in the respective proportions for eating/drinking, there were noticeably fewer cases of throttle engagement (a 10.7% reduction) and more cases of brake engagement (a 15.6% increase). In other words, drivers were less likely to use the throttle, as opposed to the brake, while eating or drinking. While throttle variance is lower in this category, it does not necessarily indicate lower variance in longitudinal control of the vehicle. Variance in speed may more accurately capture longitudinal control in certain circumstances. No significant effects were found when examining the more conventional measure of the standard deviation of throttle.

Speed Variance. Both mixed-effects models of speed variance (including the two-level and seven-level factors of secondary behavior) showed the same set of significant effects. Results are therefore presented for the seven-level model (Figure 2). The main effect of secondary behavior on the ARIMA-modeled measure of speed variance was significant, $F(6, 1,371) = 3.2, p < .01$. For most behaviors, regardless of braking, speed variance was lower when drivers were engaging in secondary behaviors. While it is difficult to interpret this in light of the higher throttle variance findings, it may suggest that drivers were exercising more longitudinal control (and smoother longitudinal movements) while engaging in secondary behaviors, but needed to make more throttle corrections to achieve this.

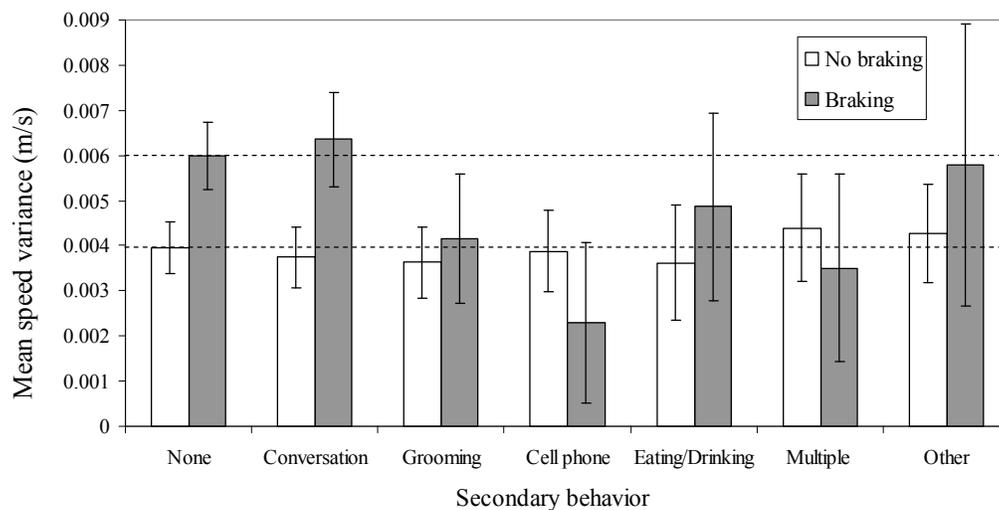


Figure 2. Mean speed variance (m/s), with and without brake application, for secondary behaviors

There was also a significant interaction effect between brake use and secondary behaviors, $F(1, 1,354) = 3.7, p < .01$. Cases not associated with any braking were roughly equivalent in speed variance across secondary behavior types. However, when drivers braked, their speed variance changed considerably, depending on what type of secondary behavior they were engaged in. That is, drivers' speeds tended to change more smoothly when they engaged in certain types of behaviors, particularly when using the cellular phone. In fact, there was actually less variance in speed when drivers were using the cellular phone relative to any other behavior. This is a complicated interaction, perhaps made more difficult to interpret due to the modified ARIMA-modeled measure of speed variance. Therefore, the same model was fit to the pre-ARIMA

measure of speed variability (i.e., the standard deviation of speed over the five-second duration). The interaction was still significant, with the only major difference in the eating/drinking category in which speed variability when braking was higher than braking without secondary behaviors. Otherwise the relationship was quite similar.

CONCLUSIONS

In general, secondary behaviors are neither equal in frequency of occurrence nor in their effect on driving performance. The present study showed relatively little effect of secondary behaviors on basic driving performance measures. Steering angle variance seemed to be most affected by secondary tasks, with cellular phone use, eating and drinking, and conversation all associated with higher steering angle variance. However, these behaviors were associated with few other differences in driving performance measures. Outside of driving performance, some differences in glance behavior were detected. In particular, cellular phone use was associated with fewer and shorter glances, while eating/drinking was associated with more, but shorter, glances.

The present study suggests that secondary behaviors have limited effects on continuous driving performance measures in naturalistic driving conditions. Perhaps more importantly, this study demonstrates the importance of conducting such a naturalistic study, as controlled studies cannot always account for the effects of driver choice and perceived risk. In this vein, our ongoing investigation of reaction time, response to critical events, and the specifics of eye glance behavior in naturalistic driving will provide an important complement to the results in this report.

ACKNOWLEDGMENTS

The authors thank colleagues at Visteon and AssistWare/Cognex Corporation. A portion of this work was conducted under agreement DTFH61-01-X-00053 with the U.S. DOT, with the balance of the effort supported by the Industry Affiliation Program for Human Factors in Transportation Safety (<http://www.umich.edu/~industry/index.html>).

REFERENCES

- Ervin, R., Sayer, J., LeBlanc, D., Bogard, S., Mefford, M., Hagan, M., Bareket, Z., & Winkler, C. (2005). Automotive collision avoidance system (ACAS) field operational test methodology and results (U.S. DOT HS 809 901). Washington, DC: Department of Transportation.
- LeBlanc, D., Sayer, J., Winkler, C., Bogard, S., Devonshire, J., Mefford, M., Hagan, M., Bareket, Z., Goodsell, R., & Gordon, T. (2006). Road departure crash warning system (RDCW) field operational test final report (Report No. 2006-9-1). Ann Arbor: The University of Michigan Transportation Research Institute.
- Sayer, J.R., Devonshire, J., & Flannagan, C.A. (2005). Effects of secondary tasks on driving performance (Report No. UMTRI-2005-29). Ann Arbor: The University of Michigan Transportation Research Institute.
- Stutts, J.C., Gish, K., Feaganes, J., Mercadante, M., Rodgman, E.A., Staplin, L., Hamlett, C., Meadows, T., & Reinfurt, D. (2003). *Distractions in everyday driving*. Washington, DC: AAA Foundation for Traffic Safety.