

ARE REAR-END CRASHES CAUSED MAINLY BY AN INTERACTION BETWEEN GLANCE DURATION AND CLOSURE RATE?

Richard A. Young
Driving Safety Consulting, LLC, Troy, MI, USA
Email: richardyoung9@gmail.com

Summary: Victor and colleagues recently analyzed rear-end crash and near-crash data from the Strategic Highway Research Program Phase 2 naturalistic driving study. They measured the last off-path glance duration just before a crash, and the *closure rate* (the change in looming rate during the last glance). They concluded that the predominant cause of rear-end crashes was a “mismatch”— a short glance with a high closure rate or a long glance with a low closure rate. The current study independently tested this “mismatch” hypothesis using two epidemiological methods to estimate odds ratios for combined crashes and near-crashes (*events*). First, the glance and closure rates were stratified, and compared to a “just following,” no-crash baseline (short glances <0.5 s and closure rates near 0). Second, a logistic regression analyzed the data on continuous scales. Both tests confirmed a strong interaction between glance duration and closure rate. At closure rates <0.056 s⁻², glances *decreased* event risk relative to baseline, proportional to glance duration. At closure rates >0.056 s⁻², glances *increased* event risk, proportional to glance duration. However, a major data limitation that potentially upwardly biased the interaction OR estimate is an inherent dependency of closure rate on glance duration, simply because of the way closure rate was defined. The mismatch hypothesis for rear-end events must be tested with other rear-end event datasets not subject to this limitation before being considered fully validated.

INTRODUCTION

Victor and colleagues (2015) (hereafter Victor et al.) analyzed rear-end crashes and near-crashes (hereafter, *events*) from the Strategic Highway Research Program Phase 2 (SHRP 2) naturalistic driving study. They used two different methods to estimate the contribution of three factors to crashes: (1) the *last glance duration* before a crash; (2) the *closure rate* between the lead and following vehicle; and (3) the *interaction* between (1) and (2). Using logistic regression with matched baseline controls, they concluded that the interaction was dominant, and risk factors (1) and (2) had minimal effect on crash risk. Using a “mechanism” method not based on risk, they concluded that the interaction effect or “mismatch” hypothesis was the main causal mechanism for rear-end crashes, but implied that glances by themselves also caused rear-end crashes.

Victor et al. analyzed glances in the 12 s immediately preceding either a crash, or the minimum time-to-collision for a near-crash. They defined *last glance* as “the last glance away from the forward path initiated before the reference point (the crash point for crashes; minimum TTC for near crashes; a random point for the matched baselines)” (p. 68). They defined *closure rate* as *inverse time-to-collision change rate* (p. 10), and calculated it from speed, distance, and manual estimates of the optical angle θ , or width, of the lead vehicle. *Time-to-collision* (TTC) in seconds was optically approximated as θ in degrees divided by its time derivative $\dot{\theta}$ in deg/s (Lee, 1976,

Eq. 6). The inverse of TTC (*invTTC*) thus has units of s^{-1} , a rate. Because the closure rate is the *change* in *invTTC* over time (the rate of a rate), it has units of s^{-2} . Victor et al. measured closure rate by the slope of a line fitted to the *invTTC* curve during the last off-path glance (their Figure 7.6). Thus, *closure rate* is the change rate of the looming rate during the last off-path glance.

Interaction Effect of Closure Rate and Glance Duration: “Mismatch” Hypothesis

Victor et al. (their Chapter 7) investigated in detail to what extent the joint occurrence of an off-path glance and closure rate (i.e., an *interaction effect*) “produces” (i.e., causes) crashes. An *interaction effect* means by definition that both factors are required, or the crash would not have occurred. They conclude that, “The results clearly demonstrate that inopportune glances of normal duration with the wrong timing relative to high lead-vehicle closure rates often produce rear-end crashes” (p. 106) and further, “Thus, the key mechanism behind these types of rear end crashes... can be understood as a ‘perfect mismatch’ between last glance duration and looming change rate” (p. 7). This interaction hypothesis between demands of the roadway and a competing activity was advanced in a general fashion by Lee et al. (2009, p. 35), and extended by Engström et al. (2013, Section 4.3.1) into a “mismatch” concept of driver “inattention.” The mismatch hypothesis predicts that a glance of any duration (including a short glance) will cause a rear-end crash if a sufficiently high positive closure rate occurs during the glance. The current objective is to independently test the “mismatch” hypothesis with the Victor et al. rear-end data.

METHODS

Two epidemiological analysis methods with baseline controls were employed. A *stratification* of the data estimated the individual effects of glance duration and closure rate on events, adjusting each for the effect of the other. A *logistic regression* analyzed glance duration, closure rate and their interaction, to calculate the relative risk (estimated by the Odds Ratio) of these three risk factors individually and conjointly, adjusting each risk factor for the influence of the other two.

Figure 1 (next page) is redrawn from Victor et al. (their Figure 7.8), which plots the data points they used to examine the interaction between last glance duration and closure rate. Figure 1 has 310 visible symbols, indicating 36 crashes, 147 near-crashes, and 127 baselines. The *x*-axis is the last glance duration, plotted from 0 to 7 s. Victor et al. labelled the *y*-axis “inverse TTC change rate” (same as “closure rate”) during the off-path glance. A zero value indicates the lead vehicle has a constant time and distance headway. Positive values (i.e., higher on the *y*-axis) indicate the lead vehicle is approaching (i.e., “looming”) at an increasing rate. Negative values indicate the lead vehicle is receding at an increasing rate.

Stratification is the standard epidemiological method that is recommended before regression analysis, to allow a better understanding and interpretation of the relationships between variables with multiple exposure levels (Rothman, 2012, Chapter 10). The stratum cut points in Figure 1 were selected to avoid a count of 0 for an event or baseline in any stratum, and to provide at least 3 strata for each variable. Stratum cut points of 0.5, 1.0, and 1.5 s were chosen for last glance duration, and 0.05 and 0.10 for closure rate (see labels in *italics* on top and right of Figure 1). Other multi-level stratifications (not shown) gave rise to similar results. Counts of data points

were made in each of the 12 strata in the Figure 1 grid. In the inset table, the first number in each pair is the baseline count and the second is the event count (combined crashes and near-crashes). Victor et al. did not plot the closure rates for “no glances” (i.e., 0 s glance

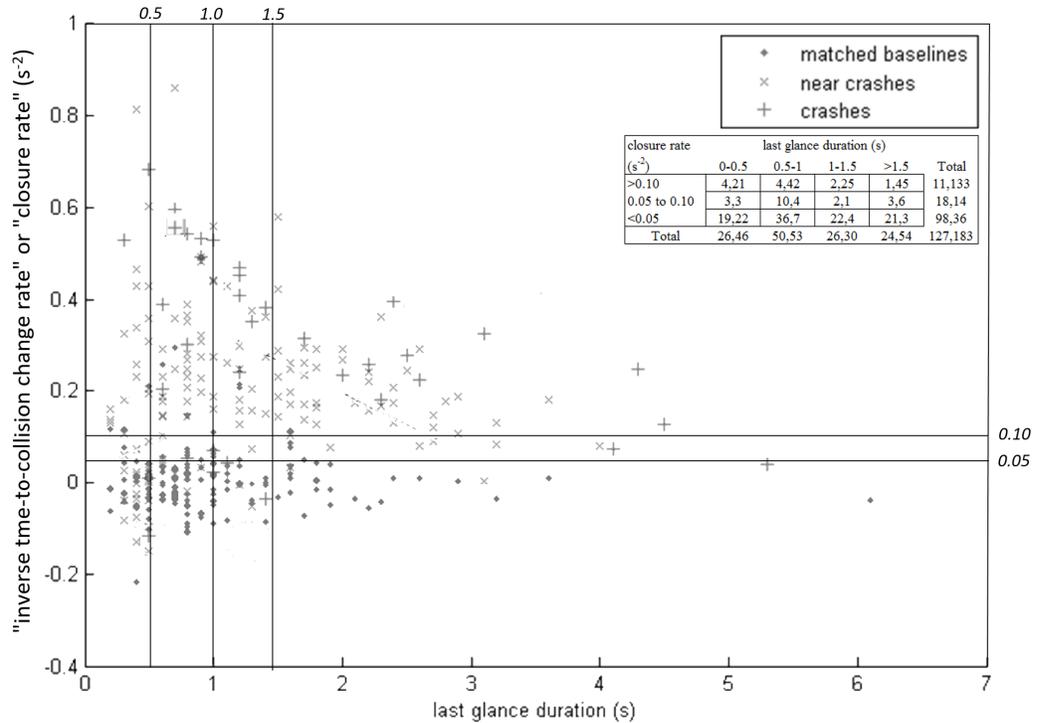


Figure 1. Closure rate vs. last glance duration for baselines, near-crashes, and crashes

duration) because closure rate was defined only for glances. Therefore, the stratification method needed a suitable baseline control reference, which was chosen as the stratum in the lower left corner of Figure 1 (last glance duration 0 to 0.5 s, closure rate <0.05 s⁻²). This stratum is plausibly close to normal driver following behavior, with small variations in headway and short off-path glances. The paired counts in every stratum in the inset Table were compared to this control reference pair count to estimate odds ratios. Standard epidemiological methods (Rothman, 2012, pp. 206-207; Breslow and Day, 1980, Table 2.8) then adjusted the odds ratios for the effect of each risk factor on the other risk factor.

Logistic regression analysis acts on all the information in the continuous variables which improves statistical power (Royston et al., 2006). The original data points were requested from the lead author T. Victor but were no longer available (personal communication), so the data points were digitized from Figure 1. Stata13 was used for the analyses and confirmed using Minitab 17. All 310 data points were entered into the analyses. The logistic regression model was a linear combination of closure rate, last glance duration, and their interaction.

RESULTS

Table 1 gives the odds ratio (OR) estimates (both crude and adjusted) for the stratified analysis using the data in the inset table in Figure 1. The right-most column shows that increasing closure rate from <0.05 to 0.05-0.10 s⁻² increased its OR estimate (adjusted for glance duration) from 1 to 2.61, 95% Confidence Interval (CI) of 1.14

Table 1. Odds ratios for indicated strata

closure rate (s ⁻²)	last glance duration (s)				adjusted for lgd
	<0.5	0.5 to 1	1 to 1.5	>1.5	
>0.10	4.53	9.07	10.80	38.86	25.47
0.05 to 0.10	0.86	0.35	0.43	1.73	2.61
<0.05	1	0.17	0.16	0.12	1
adjusted for closure rate	1	0.38	0.39	0.58	

to 5.99. Closure rates $>0.10 \text{ s}^{-2}$ increased its adjusted OR estimate to 25.47 (CI 10.98-59.10). The bottom row shows that after adjustment for closure rate, glance durations $>0.5 \text{ s}$ decreased OR estimates (a protective effect) relative to control (0 to 0.5 s), with adjusted OR estimates of 0.38 (CI 0.18-0.77), 0.39 (CI 0.16-0.93), and 0.58 (CI 0.27-1.24). Note that none of these OR estimates are yet adjusted for possible interaction effects of glance duration and closure rate.

To illustrate the interaction, note that at minimum closure rates $<0.05 \text{ s}^{-2}$ (next to bottom row in Table 1), glance durations $>0.5 \text{ s}$ decreased OR estimates (a protective effect), relative to control (closure rate $<0.05 \text{ s}^{-2}$ and glance duration $<0.5 \text{ s}$). However, at closure rates $>0.10 \text{ s}^{-2}$ (top row), glances of all durations increased OR estimates relative to control, to a maximum OR estimate of 38.86 (CI 3-450) for conjoint exposure to $>0.10 \text{ s}^{-2}$ closure rate and glances $>1.5 \text{ s}$. The fact that both the effect size and direction (protective or causal) of glance duration are dependent upon closure rate indicates a strong interaction effect. A standard epidemiologic test confirms this interaction. Consider the OR estimates in the upper left (4.53), and the lower right (0.12) of the matrix in Table 1. With no interaction, the sum of these two OR estimates should be 4.66, but the observed estimate equals 38.86, indicating a strong supra-additive interaction effect. On a multiplicative scale (the usual scale for an OR), the product of these two OR estimates should be 0.56, but it is 38.86, indicating a strong supra-multiplicative interaction. Formal epidemiological methods can test if this interaction is a “causal biologic” one (not a mere statistical correlation) (Rothman, 2012, p. 208). However, Table 1 contains both causal and preventive estimates for the same factor (depending upon level), so a formal biologic interaction assessment is complicated (Rothman, 2012, pp. 205-209), as is adjustment of all ORs in Table 1 for that interaction effect.

Logistic regression analysis is a simpler method to assess and adjust for interaction effects in this situation. Derived from all 310 digitized data points in Fig. 1, Eq. 1 gives the logistic regression solution for predicting the OR of an event as a function of closure rate (cr), last glance duration (lgd), and their interaction (cr*lgd), with each factor adjusted for the two other factors.

$$\text{Predicted OR} = \exp(-0.404\text{cr} - 0.903\text{lgd} + 16.32\text{cr}*\text{lgd} - 0.097) \quad (1)$$

Fig. 2 plots some key values for Eq. 1. The regression used a baseline (OR = 1) of 0 last glance duration and 0.056 s^{-2} closure rate. Glance duration has no event effect at that “baseline” closure rate of 0.056 s^{-2} (thick horizontal line). At lower closure rates, increased glance duration reduces the OR (a protective effect, dotted lines). At higher closure rates, the reverse occurs (a causative effect, thin solid lines). Thus, both the effect magnitude and direction (causal or protective) of glance duration depend upon closure rate, consistent with the Table 1

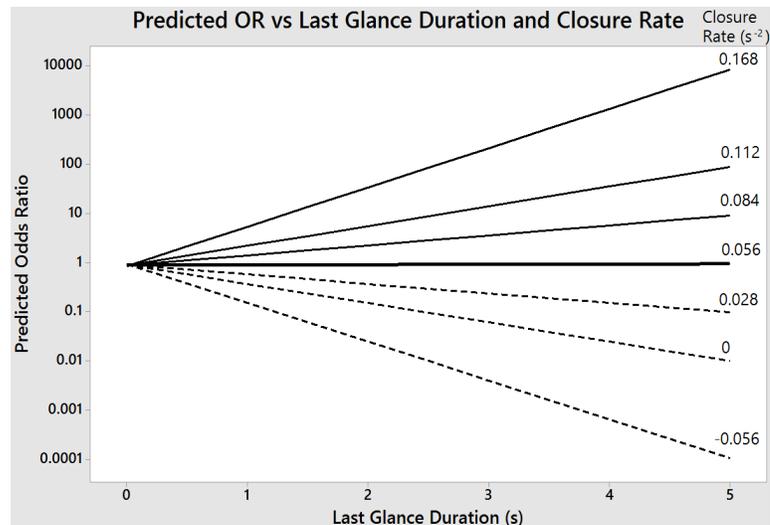


Figure 2. Predicted OR of an event from the logistic regression model of closure rate, last glance duration, and their interaction

stratification. Plus, last glance duration (adjusted for closure rate and interaction effects) had an OR estimate of 0.40 (CI 0.21-0.76), consistent with the protective adjusted OR estimates in Table 1, bottom row. Closure rate (adjusted for last glance duration and interaction effects) had a negligible OR estimate of 0.67 (CI 0.002-182). This estimate is lower than the adjusted OR estimates for closure rate in Table 1, right column, because the regression method adjusted not just for glance duration, but also for interaction effects. Finally, the interaction had an enormous OR estimate of 12.3 million (CI 6,539-23.2 million).

DISCUSSION AND CONCLUSIONS

The current study analyzed Victor et al.'s (2015) SHRP 2 data on rear-end crashes, near-crashes and baselines (Figure 1), with two independent tests of their results and conclusions about the crash/near-crash event effects of closure rate, off-path glance duration, and their interaction. Both tests confirmed a strong interaction effect, supporting Victor et al.'s hypothesis that an interaction effect, or "mismatch" between off-path glance duration and closure rate during that glance, is the major cause of rear-end crashes in the SHRP 2 data they examined.

The current results also confirm Victor et al.'s (Chapter 7) conclusions about individual glance and closure rate effects based on their logistic regression modelling. They found (p. 72) that the "individual effect of the glance duration was much weaker" than the interaction effect, and that the "additional contributions" of the individual closure rate and individual last glance duration factors were both "marginal" compared to the interaction effect. The current analysis confirms that the individual effect of closure rate was negligible after adjusting for the last glance duration and interaction effects. However, the current results extended their "weak" or "marginal" glance effect conclusion based on their regression models, by finding that the individual effect of last glance duration was actually protective (i.e., OR estimate < 1), after controlling for closure rate and interaction effects, particularly at closure rates < 0.056 s⁻².

This result is consistent with a protective glance effect found in other studies. Klauer et al. (2006) found in the 100-Car study that short in-vehicle glances (e.g., to speedometer or rear-view mirror) decrease risk. Angell et al. (2006, p. 3-58) found in an experimental study that after a lead vehicle brake light was activated and responded to by the driver in a following vehicle, glance durations increased to situation awareness locations (e.g., mirrors), and decreased for all other locations. Tijerina (2004) likewise found that many drivers before an imminent rear-end crash looked to the left, right, or center rearview mirror during lead vehicle deceleration. These off-path glances would improve situational awareness, with a plausible mitigating effect on crash likelihood or severity. For example, a safety benefit would accrue to drivers if their goal is to look for a safe pathway to steer toward to avoid the rear-end crash, or to the rear-view mirror to see if there is another vehicle that might crash into their vehicle if they engaged in high-G braking, thereby encouraging emergency steering rather than braking. Consistent with this hypothesis, Markkula et al. (2016, p. 211) recently found in 17 SHRP 2 rear-end crashes, that the last glance before the crash occurred *after* the onset of a driver's physical reaction or defensive deceleration in response to looming cues. This result suggests that these glances were in *response* to the critical situation; e.g., to look for escape routes, or to interior locations that could be used to brace for impact. However, in another dataset, Eiríksdóttir (2016) excluded such glances initiated after significant looming occurred, and still found an interaction effect.

Victor et al. (Chapter 7) also investigated, without using risk analysis methods, what they claim are direct causal *mechanisms* of rear-end crashes, based on timing of off-path glances relative to “situation kinematics” (i.e., closure rate). Based on this mechanism analysis, Victor et al. (p. 60) concluded that long glances ≥ 2 s had a “strong influence” on rear-end crashes (meaning that they increase crashes) They further concluded about short glances (p. 85, emphasis added) that, “...the majority of Category 1 and Category 2 crashes are produced by relatively short glances (<2 s).” And furthermore, “Even in the hypothetical situation where all glances longer than two seconds are eliminated, the majority of crashes directly caused by off-path glances would remain.” For such reasons, they recommended (p. 106) that, “HMI design should thus also minimize occurrence of shorter glances.”

These statements by Victor et al. about the causal effects of glances from their non-risk “mechanism” analysis methods appear to be inconsistent with their (and the current study) conclusions based on risk methods. It also logically follows from the current risk analysis of the interaction effect that any claim of a rear-end crash being “directly caused” by a glance *per se* is at best incomplete, because the effect of glance duration on rear-end events can only be determined given a specific closure rate (or vice versa). For example, Figure 2 predicts that glances increase rear-end event risk only for closure rates >0.056 s⁻². For closure rates <0.056 s⁻², Figure 2 predicts a *protective* effect of glances of all durations, when compared to normal baseline car-following with eyes-on-threat at 0.056 s⁻² closure rate. Some may believe that these relative risk results do not contradict the “mechanism” analysis, because they believe that relative risks are not “causal,” but only non-casual associations (like correlations). These false beliefs appear to be based upon a misunderstanding of epidemiological concepts of cause (see Rothman, 2012, Chapter 3). Indeed, “A central objective of epidemiologic research is to study the causes of disease” (Rothman, 2012, p. 57). To measure a causal effect in epidemiology, one typically contrasts “the experience of exposed people with what would have happened in the absence of exposure” (Rothman, 2012, p. 57), exactly as was done in the current regression and stratification analyses, using the causal OR (Rothman et al., 2008, p. 53). If there were a direct causal effect of last glance duration by itself, then the adjusted OR estimate of the glance duration must be greater than 1, and it was actually less than 1. Furthermore, the current results show that their recommendation (p. 106) that, “HMI design should minimize occurrence of shorter glances” would actually increase rear-end crashes, particularly at closure rates <0.056 s⁻².

There was a large discrepancy in the interaction effect size between the two epidemiological analysis methods (39 vs. 12.3 million). This discrepancy may partly arise because logistic regression takes all information in the continuous data into account, and is thus more powerful than stratification (Royston et al., 2006). Also, the baseline of the regression was 0 s last glance duration and 0.056 s⁻² closure rate. The baseline of the stratification included glance durations up to 0.5 s and closure rates below 0.05 s⁻², which may have also contributed to its lower effect size.

A major inherent limitation of the Victor et al. data for estimating an interaction effect is that their method of measuring the closure rate created a direct dependency of change rate on last glance duration. Specifically, their Figure 7.6 shows that the closure rate was estimated by the slope of a regression line (i.e., the change in looming divided by the glance duration). Long glance durations may thus inherently lower the closure rate, simply because of the way closure rate was measured, potentially upwardly biasing the interaction effect in this study

and theirs. However, Eiríksdóttir (2016), analyzed a different rear-end crash dataset with a different closure metric, invTau difference, which is not subject to this dependency, and still found an interaction effect. Future studies should measure closure rates independently of glance duration to avoid this bias. It is predicted that the interaction would still exist, but it would have a smaller effect size after eliminating this major potential bias.

A minor limitation of the regression analysis is that the original Victor et al. data were not available, only the digitized data from their published figure. The effect sizes are sufficiently strong, however, that any small errors introduced by digitization are unlikely to change the conclusions. Also, the stratification method used only counts of the observed, non-digitized data points, and similar results were obtained, just with smaller effect sizes. Another minor limitation is that Victor et al. defined closure rate only during the last off-path glance, so closure rate was undefined with no glances and could not be plotted in Figure 1. However, the regression successfully extrapolated the data to a glance duration of zero for its baseline (see Figure 2).

In conclusion, two independent epidemiological tests of the Victor et al. data confirmed their “mismatch” hypothesis of an interaction between the closure rate and the last off-path glance before a rear-end event. Off-path glances increased event risk estimates for closure rates $>0.056 s^{-2}$; below that, glances reduced risk estimates. However, a major data limitation was an inherent dependency of closure rate on glance duration because of the way closure rate was defined. Hence, the mismatch hypothesis for rear-end events must be tested with other rear-end event datasets not subject to this limitation before being considered fully validated.

ACKNOWLEDGMENTS

I thank the anonymous reviewers and especially Ja Young Lee for helpful comments on earlier drafts, and Amanda Zeidan for digitizing the data points in Figure 1.

REFERENCES

- Angell, L., Auflick, J., Austria, P., Kochhar, D., Tijerina, L., Biever, W., Diptiman, T., Hogsett, J., & Kiger, S. (2006). *Driver workload metrics task 2 final report* (DOT HS 810 635), NHTSA.
- Breslow, N. E., & Day, N. E. (1980). Statistical methods in cancer research. Volume I - The analysis of case-control studies. *IARC Scientific Publications* (32), 5-338.
- Engström, J., Monk, C., Hanowski, R., Horrey, W., Lee, J.,..., & Tuukkanen, M. (2013). A conceptual framework and taxonomy for understanding and categorizing driver inattention. *Brussels, Belgium: European Commission*.
- Eiríksdóttir, H. H. (2016). Quantitative analysis of rear-end crash causation mechanisms based on naturalistic crash data. (Masters), Chalmers University of Technology, Göteborg, Sweden.
- Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). The impact of driver inattention on near-crash/crash risk: An analysis using the 100-Car naturalistic driving study data. Washington, DC: NHTSA.
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, 5(4), 437-459.

- Lee, J. D., Young, K. L., & Regan, M. A. (2009). Defining driver distraction. In M. A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver Distraction: Theory, Effects and Mitigation* (pp. 31-40). Boca Raton, Florida: CRC Press.
- Markkula, G., Engström, J., Lodin, J., Bärghman, J., & Victor, T. (2016). A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. *Accid Anal Prev*, 95, Part A, 209-226.
- Rothman, K., Greenland, S., & Lash, T. (2008). *Modern Epidemiology* (3rd ed.). Phila., PA: Lippincott Williams & Wilkins.
- Rothman, K. J. (2012). *Epidemiology: An Introduction* (2nd Ed.). NY, NY: Oxford Univ. Press.
- Royston, P., Altman, D. G., & Sauerbrei, W. (2006). Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med*, 25(1), 127-141. doi:10.1002/sim.2331.
- Tijerina, L., Barickman, F., & Mazzae, E. (2004). Driver eye glance behavior during car following. NHTSA, Wash., D.C., USA.
- Victor, T., Dozza, M., Bärghman, J., Boda, C.-N., Engström, J., Flannagan, C., Lee, J. D., & Markkula, G. (2015). Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk. Report S2-S08A-RW-1, TRB.