ADJUSTED CRASH ODDS RATIO ESTIMATES OF DRIVER BEHAVIOR ERRORS: A RE-ANALYSIS OF THE SHRP 2 NATURALISTIC DRIVING STUDY DATA

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

Summary: Dingus and colleagues recently estimated crash odds ratios (ORs) for "driver behavior errors" (hereafter, "Behaviors") in the Strategic Highway Research Program Phase 2 naturalistic driving study. *Behaviors* are illegal, improper, aggressive, and/or reckless driving maneuvers. For example, the Dingus study OR estimate for "Speeding over limit and too fast for conditions," (hereafter, "Speeding"), was 12.8, with a 95% confidence interval (CI) from 10.1 to 16.2. The current study identified four issues in the Dingus study. First, heterogeneous Behaviors were pooled; e.g., "Exceeded speed limit," and "Exceeded safe speed but not speed limit" were apparently improperly pooled to form Speeding. Second, exposed drivers often had other Behaviors in the same time window, but unexposed drivers had none, a selection bias that inflated Behavior ORs by 30%. Third, impairments were not filtered out. Fourth, secondary tasks were not filtered out, creating a confounding bias that deflated Behavior OR estimates by 50%. To correct these issues, the current study stratified the heterogeneous categories, then filtered out other Behaviors, impairments, and secondary tasks. "Pure Behavior" (no other Behaviors, secondary tasks, or impairments) was thus compared to "Pure Driving" (no Behaviors, secondary tasks, or impairments). The Pure OR estimate for "Exceeded speed limit" was 5.4 (CI 2.7-10.1), and for "Exceeded safe speed but not speed limit" was 71.5 (CI 36.0-136.2), both substantially different than the Dingus study Speeding estimate. All Behavior OR estimates in the Dingus study should be similarly corrected and adjusted to improve their validity.

INTRODUCTION

Driver Behavior Errors in the 100-Car Naturalistic Driving Study

Driver behavior errors (hereafter, *Behaviors*) refer to activities intentionally undertaken by a driver that are improper, illegal, aggressive, and/or reckless driving maneuvers. Behaviors are a separate classification from secondary tasks or driver impairments. Young (2015, Appendix G) estimated the odds ratios (OR) of Behaviors in the 100-Car naturalistic driving study (NDS) database (VTTI, 2013), and found that almost all Behaviors substantially increase crash/near-crash risk, with much higher OR estimates than secondary tasks. For example, "Sudden or improper stopping on roadway" had an OR estimate of 684, with a 95% confidence interval (CI) of 284 to 2149. Young further noted (Section 7) that such Behaviors often occurred concurrent with secondary tasks, and suggested that such Behaviors "may provide better causal explanations of the crash or near-crash than the secondary tasks that occurred in the same 6-second case window" (p. 17). However, Young did not analyze possible interactions between Behavior OR estimates and secondary tasks or impairments in the 100-Car database.

Behaviors in the SHRP 2 NDS Data as Analyzed by Dingus et al. (2016)
Dingus et al. (2016) (hereafter, the *Dingus study*) estimated the crash odds ratios (ORs) for Behaviors (as well as secondary tasks and impairments) in an early version of the Strategic Highway Research Program Phase 2 (SHRP 2) database (TRB, 2013). The Dingus study (its Figure 2) tabulated OR estimates, 95% confidence intervals (CI), and baseline prevalence for Behaviors. It classified the observed single Behaviors into larger categories: "*driver performance error*, including a variety of vehicle operation and maneuver errors (e.g., failing to yield properly to other traffic, making an improper turn)"; and "*momentary driver judgment error*, including such factors as aggressive driving and speeding" (Dingus study, p. 2636, and their Figure 2). The Dingus study found that the OR estimates for individual Behaviors were substantially higher than for secondary tasks, confirming the finding by Young (2015) based on the 100-Car NDS. However, the Dingus study states, "Distraction and impairment were not filtered in the driving error evaluations" (p. 2637). (Note: The Dingus study used the term *Distraction* as a synonym for "secondary task.") However, Young (2017a) found a strong biologic interaction between Behaviors and secondary tasks in the SHRP 2 database. The Dingus study may therefore have confounded its Behavior OR estimates by not filtering out secondary tasks.

Objective and Benefits of Current Study
The objective of the current study is to improve the validity of the Dingus study estimates of the real-world relative crash risk of Behaviors. The benefit of valid estimates is that society, government, and industry can have them available to inform decisions regarding allocation of resources to reduce traffic crashes, injuries and death. For example, if valid OR estimates (and population attributable risk) for Behaviors are higher than for secondary tasks and impairments, then private and public funding would have a greater safety benefit (i.e., larger crash reductions) if devoted to reducing Behaviors, rather than secondary tasks or impairments.

METHODS AND PROCEDURE

A preliminary study investigated overall count differences between the SHRP 2 database versions. A three-step procedure was then followed for the main study: 1) Replicate the Dingus study Behavior parameter estimates using its analysis methods; 2) Identify issues (if any) in those analysis methods; and 3) Apply corrections and adjustments in the analysis methods as needed to improve the validity of the Dingus study Behavior OR estimates.

Preliminary Study: Database Versions, Overall Counts, Caveats, and Time Windows
The Dingus study used version 1.0 of the SHRP 2 database, and only the revised version 2.1.1 was available at the time of the current study. Hence, there are some slight differences in overall counts of cases, controls, secondary tasks, behaviors, and impairments between the databases (Table 1, next page). These differences are immaterial to the main results of the current paper. For example, the Dingus study reported 905 "injurious and property damage" crashes. It did not state that it pooled crash severity levels I-III (Severe Crashes, Police-Reported Crashes, and Minor Crashes), but that was assumed. The current database has 834 such cases, an 8% decrease.

Table 1. Differences in total counts between Dingus study and current study

		Total		Secondary Task				Behavior Error				Impairment			
DB	Study	cases	controls	cases	%	controls	%	cases	%	controls	%	cases	%	controls	%
1.0	Dingus	905	19,732	618	68.3%	10,247	51.9%	667	73.7%	1,782	9.0%	154	6.4%	379	1.9%
2.1.1	Current	834	19,998	574	68.8%	10,368	51.8%	423	50.7%	1,622	8.1%	58	7.0%	381	1.9%
	difference	-71	266	-44	0.5%	121	-0.1%	-244	-23.0%	-160	-0.9%	-96	0.6%	2	0.0%

Caveats about Case Behaviors. There were 3 "slots" (Driver Behavior 1, 2, and 3) that could be filled with 1-3 Behaviors, if any, for each case. However, "Distracted" and "Drowsy, sleepy, asleep, fatigued" were tabulated as Behaviors, but they represent instead secondary tasks and impairments, and must not be counted as Behaviors. If "Distracted" and "Drowsy…" were both present in a given case record, only 1 slot remained for recording Behaviors. This potential bias was resolved in the final analysis by using only the single Behavior of interest in the exposed group, and no Behaviors in the unexposed group.

Total Controls. The Dingus study reported 19,732 records in the "balanced-sample" baseline (i.e., driving with no safety-related event). These baselines had been randomly pre-selected by VTTI analysts from each driver's videos, with the number of records for each driver being proportional to that particular driver's total driving time over 5 mph (hence the term "balanced"). The SHRP 2 database v2.1.1 has 19,998 records, a 1% increase compared to the Dingus study.

Caveats about Control Behaviors. As with cases, there were 3 Behavior "slots" in the control database. Secondary tasks were not tabulated as "Distractions" in the control Behavior variables, as they were in the case variables, but "Drowsy" was. That left only 2 slots available in the control database for recording Behaviors if a slot was occupied by Drowsy. Hence, there was potentially greater opportunity for Behaviors in controls than cases. This potential bias was resolved in the final analysis by using only the single Behavior of interest in the exposed group, and no Behaviors in the unexposed group.

Time Windows for Secondary Tasks. The Dingus study (p. 2637) states that the exposure for crash cases was extracted from "short time windows (6 s …) of video surrounding the onset of crashes." The term "onset of crashes" refers to the "Precipitating Event of crashes." The *Precipitating Event* is what began the "conflict" that was followed by the crash. For cases, the secondary task exposures were coded starting 5 s before the Precipitating Event time (*Event Start*) through the end of the conflict (*Event End*). For controls, there is no Precipitating Event, so an arbitrary "Event Start" time was defined and from there, secondary tasks were coded 5 s before through 1 s after that.

Time Windows for Behaviors. The Dingus study (p. 2637) states that the case Behavior exposure was extracted from a 20 s time window of video surrounding the onset of crashes, or 20 s prior to the "Event Start" time. The SHRP 2 Behavior dictionary defines them as, "those that either occurred within seconds prior to the Precipitating Event or those resulting from the context of the driving environment) that include what the driver did to cause or contribute to the crash or near-crash. Behaviors may be apparent at times other than the time of the Precipitating Event, such as aggressive driving at an earlier moment which led to retaliatory behavior later." For example, the driver may have been weaving in and out of traffic or "acting out" towards another driver for some time prior to the Precipitating Event, and this aggressive behavior could still have an effect on driver performance and the actions of other vehicles near the time of the Precipitating Event. For baseline controls, Behaviors were coded in an arbitrary 20 s time window.

OR Estimates and Confidence Intervals. The relative crash risk in a case-control analysis of naturalistic data can be estimated by the *Odds Ratio* (OR), because crash events are rare. For any

given Behavior, the OR estimate is the ratio of the odds of being exposed to the Behavior among those who crashed, to the odds of being exposed to the Behavior among those who did not crash (Rothman, 2012, p. 57). The 95% confidence interval (CI) of the OR was estimated here using exact methods with Stata 13 (https://www.stata.com/). Note that the CIs in this paper are used solely as a measure of effect size, and are not, and should not, be used as a measure of "statistical significance" (Greenland et al., 2016; Rothman, 2016).

Step 1 – Replicate Dingus Study Parameter Estimates for Behaviors
The Dingus study (Table 2) lists four parameters for each Behavior: the OR estimate, its lower and upper 95% confidence limits, and the % baseline exposures. Replicating these parameters for a given Behavior requires tabulating the counts of the exposed and unexposed cases and controls for that Behavior. However, such counts were not provided in the Dingus study. Therefore, the current analysis solved for the counts by querying the SHRP 2 database v2.1.1 in different ways until the four Dingus study parameters were closely replicated. Slight differences in the final replicated parameter estimates were expected because of the different database versions.
Steps 2 and 3 – Identify the Dingus Study Analysis Issues and Correct or Adjust for Them
Four issues were readily identified during Step 2 and are detailed in the Results. To correct or adjust for these issues, the current study used the standard epidemiological method to estimate exposure ORs, by comparing the odds of being exposed to the Behavior among those who crashed, to the odds of being exposed to the Behavior among those who did not crash (Rothman, 2012, p. 57; see also Young, 2013a,b; 2014a). Specifically, the exposed odds were compared to the unexposed odds of what is here termed *pure driving* (no impairments, no errors, and no secondary tasks). This method differs from the Dingus study, which did not filter out impairments or secondary tasks from exposed or unexposed groups when analyzing Behaviors.

RESULTS

The Dingus study Behavior of "Speeding (over limit and too fast for conditions)" (hereafter, *Speeding*) is used to illustrate the 3 procedure steps: replication, identifying issues, and correcting or adjusting for these issues. The results for other Behaviors were similar, but are not shown for space reasons.
Replication. The Dingus study Speeding OR estimate was 12.8 (CI 10.1-16.2), with a Speeding baseline percentage of 2.77%. Unfortunately, the Dingus study Behavior of "Speeding (over limit and too fast for conditions)" is not a coded Behavior in the SHRP 2 database v2.1.1. It is possible that this category may have existed in the early version of the database used in the Dingus study. More likely, the Dingus study combined several actually observed speeding-related Behaviors in the database to form its Speeding category, although its methods do not state that it did this pooling (many other Behaviors in the Dingus study Table 2 were likewise pooled). Table 2A gives the results for the replication with this pooling assumption.

Table 2. A. Replicate Dingus study Speeding OR estimate. B. Stratify Speeding. C. Remove selection bias. D. Remove impairment bias. E. Remove secondary task bias

A. Replicate

	Speeding1*	No Behaviors
Crashes I-III	132[a]	411[b]
Balanced-sample Baseline	554[c]	18,376[d]
OR exact (95% CI, % Baseline)	10.7 (8.5-13.2) 2.77%	
p-value for homogeneity	1.1E-33	
Dingus OR (95% CI, % Baseline)	12.8 (10.1-16.2) 2.77%	

B. Stratify

	Speeding2**	No Behaviors	Speeding3***	No Behaviors
Crashes I-III	40[e]	411[b]	92[g]	411[b]
Balanced-sample Baseline	487[f]	18,376[d]	67[h]	18,376[d]
OR exact (95% CI)	3.7 (2.6-5.2)		61.4 (43.6-86.6)	

C. Remove selection bias

	Speeding2	No Behaviors	Speeding3	No Behaviors
Crashes I-III	26[i]	411[b]	54[k]	411[b]
Balanced-sample Baseline	449[j]	18,376[d]	58[l]	18,376[d]
OR exact (95% CI)	2.6 (1.7-3.9)		41.6 (27.8-62.2)	

D. Remove impairment bias

	Speeding2	No Behaviors	Speeding3	No Behaviors
Crashes I-III	22[m]	376[n]	53[q]	376[n]
Balanced-sample Baseline	445[o]	18,025[p]	58[r]	18,025[p]
OR exact (95% CI)	2.4 (1.5-3.7)		43.8 (29.2-65.6)	

E. Remove secondary task bias

	Speeding2	No Behaviors	Speeding3	No Behaviors
Crashes I-III	12[s]	94[t]	21[w]	94[t]
Balanced-sample Baseline	204[u]	8,642[v]	27[x]	8,642[v]
OR exact (95% CI)	5.4 (2.7-10.1)		71.5 (36.9-136.2)	

Key:
*Speeding1 = "Speeding (over limit and too fast for conditions)" (Dingus)
**Speeding2 = "Exceeded speed limit" (SHRP 2 database)
***Speeding3 = "Exceeded safe speed but not speed limit" (SHRP 2 database)

Notes	Speeding Exposure	Count	% Add. Behav.	% Scndry. Tasks	% Impair.
a	Speeding1	132	33%	44%	6%
b	none	411	0%	74%	9%
c	Speeding1	554	7%	47%	2%
d	none	18376	0%	52%	2%
e	Speeding2	40	25%	45%	10%
f	Speeding2	487	6%	54%	2%
g	Speeding3	92	37%	61%	4%
h	Speeding3	67	13%	51%	0%
i	Speeding2	26	0%	38%	4%
j	Speeding2	449	0%	54%	1%
k	Speeding3	54	0%	61%	2%
l	Speeding3	58	0%	53%	0%
m	Speeding2	22	0%	45%	0%
n	none	376	0%	25%	0%
o	Speeding2	445	0%	54%	0%
p	none	18025	0%	48%	0%
q	Speeding3	53	0%	60%	0%
r	Speeding3	58	0%	53%	0%
s	Speeding2	12	0%	0%	0%
t	none	94	0%	0%	0%
u	Speeding2	204	0%	0%	0%
v	none	8642	0%	0%	0%
w	Speeding3	21	0%	0%	0%
x	Speeding3	27	0%	0%	0%

In Table 2A, the "Speeding1" exposed counts are the sum of two speeding Behaviors in the v2.1.1 database. These are "Exceeded speed limit" (Speeding2 in Table 2B) and "Exceeded safe speed but not speed limit" (Speeding3 in Table 2B). The Table 2A OR estimate of 10.7 (CI 8.5-13.2; % baseline 2.77%) reasonably replicated the Dingus study Speeding estimate. The baseline prevalence of 2.77% is the total count of Speeding1 in the exposed group (554) divided by the total baseline records (19,998), and exactly matched the Dingus study prevalence of 2.77%.
Issue 1: Heterogeneity. Table 2B shows the strata Speeding2 and Speeding3 whose exposed counts (left columns) sum to form the exposed column in Table 2A. Table 2B shows that the stratum-specific OR estimates (Speeding2 and Speeding3) are 3.7 and 61.4. These strata are strongly heterogeneous with p = 1.1E-33, meaning that they represent fundamentally different speeding risk factors. Because the data do not conform to the assumption that the effect is constant across strata, pooling is not applicable (Rothman, 2012, p. 178). Thus, in Table 2A, both the Dingus study pooled OR estimate (12.8) and its replication (10.7) are invalid.
Issue 2: Selection Bias. The Dingus study selected records for its exposed group that had more Behaviors than just Speeding: 33% of the cases (note a) and 7% of the controls (note c) had one or two additional Behaviors besides Speeding. However, the Dingus study selected records for its unexposed group with 0% Behaviors for cases (note b) and controls (note d). Such a differential selection criterion is termed *selection bias* in epidemiology (Rothman, 2012, pp. 126-133). The exposed cells in Tables 2A and 2B effectively pooled Speeding Alone, Speeding plus 1 additional Behavior, and Speeding plus 2 additional Behaviors. Selection bias was removed by filtering out all records with additional Behaviors from the exposed column in Table 2B, leaving Speeding by itself (Table 2C). Removing selection bias caused the speeding OR estimates to each decline about 30% (Speeding2 from 3.7 to 2.6 and Speeding3 from 61.4 to 41.6).

Issue 3: Confounding by Impairments. Impairments were present in all cells in Tables 2A-C, which could also potentially bias the Speeding OR estimates. Impairment bias was removed from Table 2C by selecting only records with 0 impairments (Table 2D). Removing selection bias had little effect on the speeding OR estimates (Speeding2 was 2.6 vs. 2.4, and Speeding3 was 43.8 vs. 41.6), probably because of the low prevalence of impairments. Notes b, d, i, j, k, and l indicate that impairments in the Table 2C cells ranged from 0% (note l) to only 9% (note b).

Issue 4: Confounding by Secondary Tasks. Secondary tasks were prevalent in a high percentage of all cells in Tables 2A-D. Their presence biases the speeding OR estimates, because of the strong interaction between secondary tasks and Behaviors (Young, 2017a, Appendices B and C). This bias was removed from Table 2D by selecting exposed records with no secondary task ("Pure Speeding"), and likewise for unexposed records ("Pure Driving"). Table 2E shows the unbiased "Pure OR" estimate for Speeding2 about doubled (2.4 to 5.4), and likewise for Speeding3 (43.8 to 71.5). The final "Pure OR" estimates of Speeding2 and Speeding3 (Table 2E) were thus increased by about 50% by removing records with secondary tasks concurrent with Behaviors from Table 2D. The converse is that the presence of secondary tasks (Table 2D) *reduced* the speeding OR estimates by about 50% compared to their non-presence (Table 2E).

DISCUSSION

The current study replicated the Dingus study Behavior parameters, and then identified four issues in their analysis. First, the Dingus study incorrectly pooled across strongly heterogeneous Behaviors. Second, it selected exposed drivers who often had more than one Behavior in the same time window, but selected unexposed drivers with none, a selection bias that overestimated Behavior ORs. Third, it did not filter out drivers with impairments, with little effect. Fourth, it did not filter out drivers with secondary tasks, which biased its Behavior OR estimates too low.

Speeding Behavior illustrated these issues and their corrections. To avoid heterogeneity, the current study analyzed individual speeding strata, "Exceeded speed limit" and "Exceeded safe speed but not speed limit." To avoid the identified biases, it selected drivers exposed to only a single speeding Behavior, with no secondary tasks or impairments ("Pure Speeding"), and compared these drivers to unexposed drivers with no Behaviors, secondary tasks, or impairments ("Pure Driving"). The final adjusted OR estimates for the "Pure" speeding strata were 5.4 and 71.5, each substantially different than the Dingus study Speeding estimate of 12.8. The presence of secondary tasks reduced both "Pure" speeding OR estimates by about 50%. Hence, speeding is an independent driver action that substantially increases crash risk, not a side effect of secondary tasks. Secondary tasks thus help mitigate Behavior risk (Young, 2017b), likely because of driver self-regulation during secondary task performance (Young, 2014b, 2017b).

Limitations

Neither the current study nor the Dingus study matched its baseline control records to crash case records by demographic, time-of-day, or closeness-to-junction variables, which tend to upwardly-bias secondary task OR estimates (Young, 2015). Young (2015) adjusted for such confounding in the 100-Car study dataset by using the matched baseline data from Klauer et al. (2010), and all secondary task OR estimates declined. Bias from such confounding variables, if not adjusted for, may also occur for Behavior OR estimates. Hence, the revised speeding OR estimates here may change if matched baselines could have been used, but these were not available in the web-based SHRP 2 database v2.1.1 (TRB, 2013) at the time of the current study.

CONCLUSION

The Dingus study speeding OR estimate combined heterogeneous speeding categories. After stratification to remove heterogeneity, the speeding ORs were substantially overestimated because of selection bias: Additional Behaviors were present in the speeding-exposed but not unexposed group. Also, the presence of secondary tasks underestimated the speeding ORs. It is recommended that all Behavior OR estimates in the Dingus study be adjusted for these issues.

REFERENCES

Dingus, T. A., Guo, F., Lee, S.,…, & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proc. Natl. Acad. Sci. U.S.A.*, 113(10), 2636–2641. http://www.pnas.org/content/113/10/2636.abstract.

Greenland, S., Senn, S.J., Rothman, K.J., Carlin, J.B. et al. (2016). Statistical tests, p values, confidence intervals, and power: A guide to misinterpretations, *Eur. J. Epidemiol.* 31(4):337-350. http://dx.doi.org/10.1007/s10654-016-0149-3.

Klauer, S.G., Guo, F., Sudweeks, J. and Dingus, T.A., (2010). An analysis of driver inattention using a case-crossover approach on 100-Car data: Final report, DOT HS 811 334, U.S. DOT, Washington, D.C.

Rothman, K. J. (2012). *Epidemiology: An Introduction* (2nd ed.). New York, NY: Oxford University Press.

Rothman, K.J. (2016). Disengaging from statistical significance, [In Eng], *Eur J Epidemiol.,* 31(5):443-444.

TRB (2013). Transportation Research Board of the National Academy of Sciences. The 2nd Strategic Highway Research Program Naturalistic Driving Study Dataset. SHRP 2 NDS InSight Data Dissemination web site: https://insight.shrp2nds.us.

VTTI (2013). VTTI data warehouse: 100-Car data. http://forums.vtti.vt.edu/index.php?/files/category/2-vtti-data-sets/.

Young, R. A. (2013a). Naturalistic studies of driver distraction: effects of analysis methods on odds ratios and population attributable risk. *7th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Bolton Landing, NY. http://drivingassessment.uiowa.edu/sites/default/files/DA2013/Papers/077_Young_0.pdf.

Young, R. A. (2013b). Cell phone conversation and automobile crashes: Relative risk is near 1, not 4. Paper from the *Third International Conference on Driver Distraction and Inattention*, Gothenburg, Sweden. http://document.chalmers.se/download?docid=cfd54630-edad-4476-b145-bd46fc08d9b7.

Young, R. A. (2014a). An unbiased estimate of the relative crash risk of cell phone conversation while driving an automobile. *SAE Int. J. Trans. Safety, 2*(1), 46-66. http://papers.sae.org/2014-01-0446/.

Young, R. (2014b). Self-regulation minimizes crash risk from attentional effects of cognitive load during auditory-vocal tasks. *SAE Int. J. Trans. Safety, 2*(1), 67-85. http://papers.sae.org/2014-01-0448/.

Young, R. A. (2015). Revised odds ratio estimates of secondary tasks: A re-analysis of the 100-Car naturalistic driving study data. SAE Technical Paper 2015-01-1387. http://papers.sae.org/2015-01-1387/.

Young, R. A. (2017a). Removing biases from crash odds ratio estimates of secondary tasks: A new analysis of the SHRP 2 naturalistic driving study data. Paper #2017-01-1380 from SAE, Detroit, MI USA. http://papers.sae.org/2017-01-1380/.

Young, R. A. (2017b). Predicting relative crash risk from the attentional effects of the cognitive demand of visual-manual secondary tasks. Paper #2017-01-1384 from SAE, Detroit, MI USA. http://papers.sae.org/2017-01-1384/.