A binary logistic regression model of the driver avoidance manoeuvers in two passenger vehicle crashes

Hayder Mohammed Hayder Al-Taweel [a and c], William Young [a], Amir Sobhani [b]
[a] Department of Civil Engineering, Monash University,
[b] The ARRB Group Ltd, Australia
[c] Department of Highway and Transportation Engineering, Al-Mustansiriyah University, Iraq
hayder.al-taweel@monash.edu, bill.young@monash.edu, amir.sobhani@arrb.com.au

Abstract

The reactions of drivers may influences the risk and severity of car crashes. However, while studies have analysed drivers' reactions in crashes, in general, little is known about the factors affecting crash avoidance maneuvers of two passenger vehicle crashes. To increase understanding in this area a statistical model is developed using the National Automotive Sampling System-Crashworthiness Data System (NASS-CDS) between the years 2009 and 2014. More specifically, a Pearson Chi-Square test is performed first to identify the significant variables, and then a Binary Logistic Regression Model is developed to identify the relative importance of the variables.

The results of the model indicate that, elderly drivers or (65 years or older), are less likely to avoid crashes than other groups. Drunk or drugged drivers have a greater negative effect on engaging in crash avoidance maneuvers. Drivers of large size vehicles are more likely to react than those in smaller cars. Driving in adverse surface conditions, non-level profile, and rural roads, increases the likelihood of reacting. Hitting drivers have higher possibility to react than those in the hit vehicle.

In future research, it is encouraged to investigation the relation between the driver's reaction for hit and hitting drivers and crash severity using Binary Logistic Regression Model.

Keywords: Crash avoidance maneuvers, Binary Logistic Regression Model (BLRM).

1. Research Background

The driver has a major role in the pre-crash event to avoid an imminent crash. Many transportation researchers have attempted to understand the factors associated with drivers' reactions before the crash and its link to traffic crashes (Bishop et al. (1985); Strayer et al. (2003); Bélanger et al. (2010); Schaap (2012); Wang et al. (2011); Leung et al. (2012); Stavrinos et al. (2013); and Muttart (2015)).

Literature reviews have shown that there are three main dimensions associated with drivers' reactions. These dimensions are drivers' factors, road and environmental factors and vehicle factors.

- Drivers' related factors can be described in terms of the demographic factors and drivers intoxication by alcohol or drugs while driving. The available research displays that there is a need to understand the manner in which drivers' age and gender have an influence on the type of reaction. For example, the probability of drivers' reacting decreases with age (Kaplan & Prato, 2012). Drivers' intoxication by alcohol or drugs while driving is another consideration. For instance, descriptive statistics and reports show heavy drinking of alcohol or consumed drugs reduces alertness. This affect the central nervous system by which the driver ability to conduct crash avoidance manoeuvres is reduced (Burns et al. (2002), Strayer et al. (2006) and Leung et al. (2012), Sussman et al. (1985), Harb et al. (2009), Kaplan and Prato (2012)).
- In addition, road and environmental factors such as road profile, road alignment, road speed, road type, sight obstructions, and number of lanes, median type, road surface conditions and weather conditions and time of the crash are also affect drivers' reactions. For example, drivers are less likely to take drivers crash avoidance actions under good weather conditions (Yan et al., 2008).
- Moreover, vehicle related factors also affect drivers' reactions, type of vehicles and role of vehicle in crash involved. For example, type of vehicles involved in the crash is significantly associated with the conducted crash avoidance manoeuvers (Dozza, 2013).

Literature reviews also show that many studies do not clearly analyse the role of vehicle whether they are target or hitting. They assume that both drivers have the same reactions while, in reality the kind of reaction for hitting and target drivers may differ.

The research objective of this paper is analyse the factors affecting drivers' reactions. These factors include drivers, road geometric and environmental, and vehicle factors. Then the paper provides a detailed explanation of the analysis approach. The paper closes with a summary of findings and some recommendations as to future analyses of the data utilised in this study.

2. Data Collection

This paper focuses on two-vehicle crash data which are available from the national website of the United States government. The National Highway Traffic Safety Administration (NHTSA) provides sophisticated reports for motor vehicle crashes. This data includes fatality, severe injury and property damage crashes (Radja, 2014). National Automotive Sampling Systems-Crashworthiness Data System (NASS-CDS) is a national program to collect data.

NASS-CDS has field investigation teams usually archive the data from on-site crashes, studying evidence, photographing vehicles, assessing property damage, determining dynamic factors and categorize victim's medical records for injury severity as seen in Table 1.

Table 1 Crash variables used in this study model

Variables	Code
Role of Diver's in Crash	1: Target , and 2: Hitting
Drivers Reactions	1: Drivers take Reactions , and 2: Drivers do not take Reactions
Driver Age	1: Under 25, 2: 25 - 44, 3: 45-64, and 4 :> 65.
Driver Gender	1: Female, and 2: Male.
Driving Behaviour Alcohol/drugs use	1: No, and 2: Yes.
Speed Limit	1: 25 or less mph 2: 26-40 3: 41-55 4:Over 56 mph.
Type of Crashes	1: Head on crash, 2: Rear end crash, 3: Right Angle Crash, 4: Side swipe Crash Same Direction, 5: Side Swipe Crash Opposite Direction.
Vehicle Type	1: Passenger Cars: Sedan, Hatchback and Station Wagon, 2: LSV: Sport Utility Vehicle, Pickups, and Vans.
Time of the crash	1: Daylight, and 2: Night Time
Road Profile	1: Level and 2: Grade
Road Alignment	1: Curve and 2: Straight.
Road Types	1: Urban, and 2: Rural.
Road Condition	1: Good Dry Surface, 2: Adverse ; Wet, Loose, Muddy or Oily and Snow or Ice Surface
Pavement Type	1: Asphalt , and 2: Others; Concrete, Clay, Sand

The total number of eligible data cases is 7776 drivers in CDS data over the period between 2009 and 2014 CDS database. As seen in Table 1, this database is clearly defined the hitting and target vehicles in two vehicle crashes (Radja, 2014). Age and gender of drivers, which taken from the CDS "*Personal File*" were used to categorize the age of the following drivers involved in pre-crash event. They were grouped from 16 to over 65 years old in 20-year intervals. In addition, the " *Police Official Records for Alcohol Present and Other Drug Present"* variable in the CDS "*Personal Files"* were used to categorize alcohol consumption or illegal drugs used of following vehicle drivers' involved pre-crash events. For each driver there is a *zip code* including Alcohol test, Drug test, Test source, and the Results. It also contains some variables referring to circumstances, conditions, and events that may have contributed to the crash. From the CDS Vehicle/Driver *File*, this study examined the variables, *Time of Day, Weather Conditions* and *crash location related to Profile, Alignment, Surface* *condition, and Road Type.*

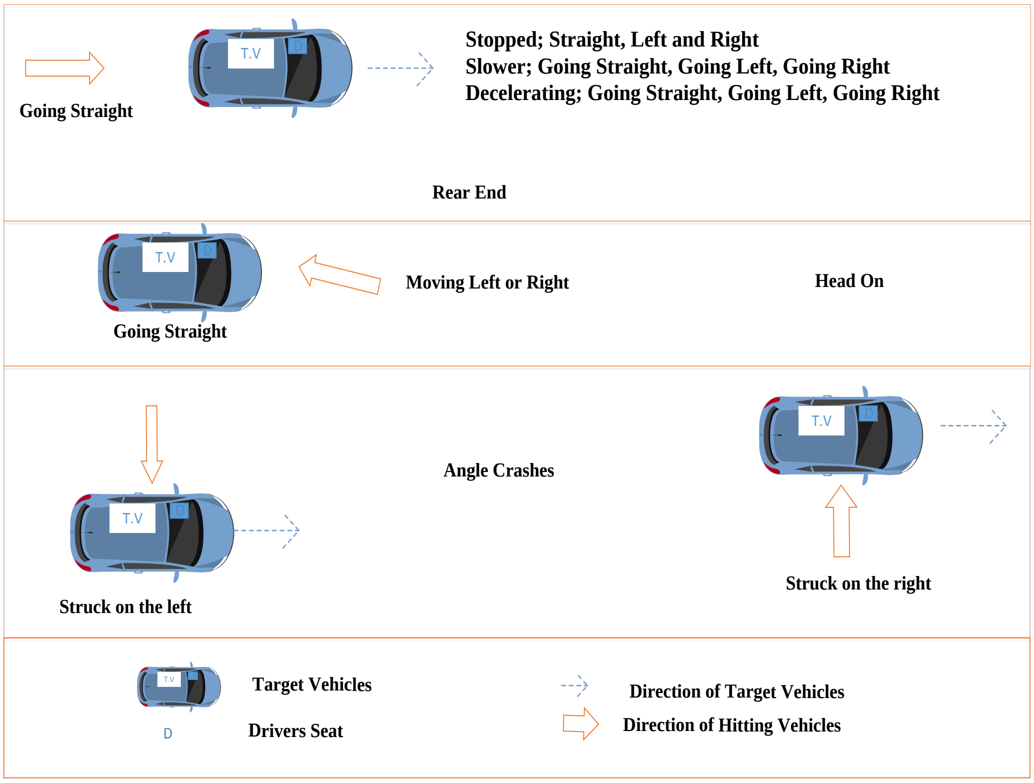

Figure 1 Scenario of rear end, head on, angle crashes adapted from NASS-CDS (Radja, 2014)

Figure 1 shows the scenario of rear end, head on, angle crashes according to NASS-CDS. Rear end crashes, first, present the struck on the rear of target vehicle. The scenario of rear end crashes occur when target vehicle either stopped, slower or decelerating, while hitting going straight. Head on crashes present the struck on the front of the target vehicle. Right angle crashes present the struck on the middle of target vehicles.

3. Data Analysis Method

Previous section outlined the database used in this study. This section explains the statistical methods utilized to analyse the data.

In this study, dependent variable is “Drivers Take Reactions/ Drivers Do not Take Reactions” which describes two possible actions taken by the drivers in response to the unexpected stimulus. In this study three-step analysis are carried out to explore the factors affecting drivers’ reactions.

1) In the first step a Pearson’s correlation test is performed to find out the possible correlation among independent variables each other.
2) In the second step a Chi-Square test is performed to find out the significant independent variables influencing the dependent variable.
3) In the third step a Binary Logistic Regression model is developed in order to explore the relative importance of the significant variables. The significant variables affecting the crash location (Reactions/ No Reactions) are also studied using this model.

3.1 Correlation between Independent Variables Each Other

Pearson’s correlation analysis was first done among various explanatory variables to detect possible correlation among explanatory variables. The results show that these variables are not correlated each other except weather conditions and surface conditions were correlated (Evans, 1996). In the next step the significant variables only are input in the Chi-squared test to detect the possible correlation the correlation between dependent and independent variables.

3.2 Correlation between Dependent and Independent Variables

Pearson chi-square test is carried out using SPSS software to test the correlation between dependent and independent variables. Table 2 summarizes the results of the Pearson chi-square test. The results show that only "pavement type" does not have significant effect on output variable (Field, 2009). All the other variables significantly influence the dependent variable based of Pearson chi-square test with 95% level of confidence. In the next step the significant variables identified using Pearson chi-square test are considered as independent variables of the Binary Logistic Regression model.

Table 2: Results of Pearson Chi-square test

Explanatory Variables	P value (Pearson Chi-Square test)
Driver Age	<0.0001
Driver Gender	0.042
Driving Behaviour Alcohol/Drugs use	0.08
Speed Limit	0.001
Type of Crashes	<0.0001
Vehicle Type	<0.0001
Time	0.03
Road Profile	<0.0001
Road Alignment	<0.0001
Road Type	<0.0001
Road Condition	<0.0001
Pavement Type	0.8
Role of Driver's in Crash	<0.0001

In the next step the significant variables identified using Pearson chi-square test are considered as independent variables of the Binary Logistic Regression model.

3.3 Binary Logistic Regression Model

To determine the possible factors contributing to drivers reactions, the binary logistic regression model is fitted to the set of available independent variables. This study tested the goodness of fit of the binary logit and binary probit models. Binary logit was utilized as it showed better goodness of fit. Binary logistic regression model is a type of Generalized Linear Regression model in the form of Equation 1 model (Field, 2009):

$$P_i = \frac{\mathrm{EXP}(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i})}{1 + \mathrm{EXP}(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i})} \tag{1}$$

When the model is applied to model the driver's reaction, the logit model predicts the probability (π) of the driver taking an action at the crash scene, between 0 and 1 ($0 \leq \pi \leq 1$) for all possible independent variables in Equation 1, where, β_0 is the intercept term in the model, (β = 1, 2, . . . , n) are the regression coefficients for each independent variable. Y is the predicted probability of the event which is coded with 1 (action) rather than with 0 (no action), 1−Y is the predicted probability of no action and X is the set of independent variables. These variables include 'age', 'gender', 'speed', 'surface type', and 'crash type'. These variables can be either discrete, continuous, or a mixed trend. The parameters in the model are estimated using a maximum likelihood approach. The estimated model is evaluated by performing a likelihood ratio test to determine the significance of the covariates in the model. Equation 2 shows the log likelihood function. Because the dependent variable is modelled using a log transformation, logit (π), the interpretation of the estimated coefficient is based on the exponential transformation of the estimated coefficient, which is commonly known as the odds ratio (Equation 3).

$$\mathrm{L}\,(\beta)\mathrm{LL}\,(\beta) = \sum_{i=1}^{n}\{y_i \ln(P_i) + (1-y_i)\ln(1-P_i)\} \tag{2}$$

$$Odds = \frac{P(event)}{P(no\ event)}$$
$$P(no\ event) = 1 - P(event) \tag{3}$$

4. Discussion of Modelling Results

Previous section outlined the method of study. This section explains the discussion of the statistical models utilized to analyse the data (Table 3).

Table 3 Model Estimate.

Dependent Variable	Independent Variable	Level of the Variable	Coefficient	P-Value	Odds Ratio
Drivers take a reaction	Age Groups	<26 Years old	Ref.	-	-
		26–44 Years old	-0.398	< 0.0001	0.672
		45–64 Years old	-0.357	< 0.0001	0.700
		>65 Years old	-0.715	< 0.0001	0.489
	Driving Behaviour Alcohol/Drugs use	No	Ref.	-	-
		Yes	-0.412	0.006	0.662
	Speed Limit in km/hr.	40 or less	Ref.	-	-
		41 – 70	.256	< 0.0001	1.292
		71 – 90	.270	< 0.0001	1.310
		91 and Over	.197	.082	1.218
	Crash Type	Rear end	Ref.	-	-
		Head on	0.398	< 0.0001	1.489
		Angle	-0.126	0.038	0.882
		Side swipe same direction	-0.677	< 0.0001	0.508
		Side swipe opposite direction	0.712	< 0.0001	2.039
	Role of Vehicle in Crash	Target	Ref.	-	-
		Hitting	0.988	< 0.0001	2.685
	Type of Vehicles	Large size vehicles	Ref.	-	-
		Small cars	-0.391	< 0.0001	0.676
	Time of the Crash	Day time	Ref.	-	-
		Night time	-.171	.004	.843
	Road Profile	Level	Ref.	-	-
		Grade	.290	< 0.0001	1.337
	Road Alignment	Straight	Ref.	-	-
		Curve	.216	.004	1.241
	Road Type	Local roads	Ref.	-	-
		Arterial roads	-.322	< 0.0001	.724
	Road Surface Condition	Dry	Ref.	-	-
		Slippery	.221	.001	1.248
The reference category is: No Reaction Number of cases = 7775 Pseudo R-squares=0.09		Constant	-0.709	< 0.0001	0.492

Effect of driver factors:

Driver related factors include demographic information, alcohol and drug use, and speed are significantly associated with the likelihood of performing a reaction before crash.

A 20-year interval is chosen to group driver age. The four driver age groups include, ‘younger than 26 years old’, ‘26–44 years old’, ‘45–64 years old’, and ‘older than 65 years old’. Compared to the reference group (i.e. 25 years old and under), the likelihood of ‘reaction’ is significantly higher in other age groups. Elderly drivers, are less likely to avoid crashes than other groups do. This possibility is the two times highest for age group under 26 compared with other age groups (odds ratio =0.479). Drivers aged between 25 and 64 are 80% more likely to avoid crashes than their elderly counterparts.

Another factor, ‘driver drugs or illegal alcohol consumption’, is reflected in cases where there is use of illegal drugs and alcohol abuse. The blood alcohol content (BAC) of the driver indicating abuse is a measure as less /above 0.08 percent. Results revealed that ‘alcohol and drug’ use decreases the likelihood of conducting reactions. Drunk drivers are 50% less likely to avoid crashes than non-drunk drivers (Odds ratio = 0.662).

In relation to ‘speed limit’, higher speed limits have a higher likelihood of the driver conducting reactions. The estimated parameters illustrates that higher speed limit locations (‘between 26 and 55 mph’) are 30% more likely to perform action than the locations with speed limit of 25 mph or less (odds ratio = 1.3).

Effect of crash characteristics:

Compared to the Rear end crashes (the reference group), the odds ratio shows that ‘Head on’ and ‘Side swipe opposite direction’, are more likely to be associated with ‘reaction’. This is logical, because, in these scenarios, it is easier for the drivers to see each other before the crash.

However, results further show that the lower possibility of ‘reaction is associated with the in ‘Angle’ and ‘Side swipe same direction’ crash type. This is also reasonable as it is difficult to observe the other driver in this type of crash (odds ratio = 0.882 and 0.508, respectively). Accident type and scenario details are shown in the Figure 1.

Effect of vehicle:

Vehicle related factors include vehicle type and role of vehicle in the crash are significantly associated with the likelihood of performing a reaction before crash. As shown in Table 3, ‘vehicle type’ is also significantly associated with drivers’ reactions. Results show that large size vehicles (LSV) drivers are about 30% more likely to take actions than drivers of small cars (odds ratio = 0.676). Role of vehicle is also significantly associated with drivers’ reactions. Results show that hitting drivers are about 2 times more likely to take actions than drivers of target drivers (odds ratio =2.685).

Effect of road related factors:

Road factors such as, surface conditions, profile (grade and level), section type and traffic direction and traffic light, are significantly associated with the likelihood of engaging in reactions (see Table 3).

Compared with driving during the day (the reference group), drivers at night time are less likely to perform any reactions. This is logical, because the drivers with low visibility most likely do not perform action at night. This results are in agreement with the results of (Kaplan & Prato, 2012) study. With respect to ‘Road Profile’, it is found that unleveled roads or up\down grade (slope) significantly reduce the likelihood that the drivers preform action against crashes. The estimated value of the possibility of crashes occurring on roads with a level profile and flat roads is 30% higher on than unlevelled roads.

The effect of 'Road Alignment' is also significant and positive, indicating that drivers at curve alignment are more likely to perform actions than straight roads. The OR statistic tells us drivers are 24 % more likely to perform action. The likelihood of engaging in reactions decreases with appearance of controls functioning traffic light in arterial roads. They are increasing are 30% likelihood of acting positively or to engage in maneuvers than urban roads. When traffic lights are not present on rural roads, the drivers seems to be more cautious.

Slippery pavements conditions, such as wet, snow and ice, are greatly contribute to drivers reactions. The estimated parameters suggested that the likelihood of reactions as a result of these wet pavements conditions is approximately 25% higher than that in dry pavement condition. Slippery pavements are more likely to associate with 'action' rather than 'no action'.

5. Conclusion

Drivers' reactions in the pre-crash situation has been studied for several decades, in general, little is known about drivers' reactions and the role of vehicle. The present study investigates the factors affecting drivers' reactions of two passenger vehicle crashes using real crash data. This data is retrieved from the NASS - CDS crash database for the years between 2009 and 2014. To cut of the ambiguity, only cases where both vehicles described driver's reaction is used.

Pearson Chi-square test is first used to identify the significant of the model. The model parameter estimates shows that 'speed limit, age, surface conditions, profile, traffic flow, and lighting conditions' are the significant independent variables in the model. The results of binary logistic regression are analyzed to estimate the probability of the crash avoidance maneuvers in related to driver, crash, vehicles, road, and environmental factors.

Driver dimensions include the demographic factors influencing the course of driver actions, perceptions, and attitudes which develop the driver responds in pre-crash event. Driver related factors are significantly associated with the likelihood of conducting crash avoidance maneuvers. The likelihood of not taking avoiding crash increases when the age increases, in agreement with the findings of (Yan et al. 2008 & Kaplan and Prato 2012). Elderly drivers or (65 years or older), are less likely to avoid crashes than other groups. Drivers aged between 16 and 24 are two times more likely to avoid crashes than elderly counterparts.

In addition, driving behavior such as drunk and drugged drivers, seems to have a negative impact on the likelihood of perform most crash avoidance maneuvers, in agreement with the findings of (Yan et al. (2008)& Kaplan and Prato (2012)). Therefore, these results suggest that applying effective law and regulations to address alcohol-related crashes for example, lowering blood alcohol concentration (BAC) limits for driving to 0.08 and to 0.05 and promoting educational programs or community-specific models to address the negative consequences of drunk driver's scenarios (Fell et al., 2015).

In comparison "LSVs with cars, drivers of large size vehicles are more likely to take reactions than cars. These results may related to several factors, such as drivers' themselves and confidence while driving a large vehicle, and behavioral differences between nonprofessional and professional drivers. In agreement with results of (Kaplan & Prato, 2012).

An interesting finding, road related factors such as, adverse surface conditions, grade profile, and not availability of traffic light control area or rural roads, were significantly associated with the likelihood of engaging in crash avoidance maneuvers. These results are in agreement with the findings of (Yan et al. 2008 & Kaplan and Prato 2012). The results can be explained by the fact that drivers may be more likely to drive cautiously in those adverse driving surroundings.

This study shows that the kind of reaction for hitting and target drivers is different. For example, the probability of having drivers' reactions for target drivers is very low compares with hitting drivers. This is logical, because target drivers are usually surprised by unexpected struck.

The limitations in this data need to be highlighted. The CDS data provides only a national dataset, not state-level data. Second, driving speed is not included in our study due to the absence data provided. Actual impact speed is certainly give better understanding, however speed limits were used to solve partially this limitation. Thirdly, this study explored the relationship between the probabilities of drivers' Responses/No responses actions and the characteristics of drivers, vehicles, and environments. However, this paper did not distinguish between the detailed behaviour and drivers' reactions, because of insufficient sample size. Fourth, this study does not investigate to the effect of drivers' distractions due to fatigue or using mobile phone while driving. The reason was vast majority of these two variables were almost unknown. Studying drivers' distractions while driving is recommended for future studies.

A further study of the association of drivers' emergency reactions such as braking only, steering, braking and steering, is recommended.

6. Acknowledgement

Authors would like to acknowledge the Higher Committee for Education Development in Iraq, for their financial support to this research project, which is part of master by research thesis. As well as, we would like to acknowledge Richard Amoh-Gyimah PhD candidate, Transport Research Group, Monash University.

7. References

Bélanger, A., Gagnon, S., & Yamin, S. (2010). Capturing the serial nature of older drivers' responses towards challenging events: A simulator study. *Accident Analysis & Prevention, 42*(3), 809-817.

Bishop, H., Madnick, B., Walter, R., & Sussman, E. (1985). Potential for Driver Attention Monitoring System Development.

Burns, P., Parkes, A., Burton, S., Smith, R., & Burch, D. (2002). *How Dangerous is Driving with a Mobile Phone?: Benchmarking the Impairment to Alcohol*: Transport Research Laboratory Berkshire,, United Kingdom.

Dozza, M. (2013). What factors influence drivers' response time for evasive maneuvers in real traffic? *Accident Analysis & Prevention, 58*, 299-308.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*: Brooks/Cole.

Fell, J. C., Thomas, S., Scherer, M., Fisher, D. A., & Romano, E. (2015). Scoring the strengths and weaknesses of underage drinking laws in the United States. *World medical & health policy, 7*(1), 28-58.

Field, A. (2009). *Discovering statistics using SPSS*: Sage publications.

Harb, R., Yan, X., Radwan, E., & Su, X. (2009). Exploring precrash maneuvers using classification trees and random forests. *Accident Analysis & Prevention, 41*(1), 98-107.

Kaplan, S., & Prato, C. G. (2012). Associating crash avoidance maneuvers with driver attributes and accident characteristics: a mixed logit model approach. *Traffic injury prevention, 13*(3), 315-326.

Leung, S., Croft, R. J., Jackson, M. L., Howard, M. E., & Mckenzie, R. J. (2012). A comparison of the effect of mobile phone use and alcohol consumption on driving simulation performance. *Traffic injury prevention, 13*(6), 566-574.

Muttart, J. (2015). Influence of Age, Secondary Tasks and Other Factors on Drivers' Swerving Responses before Crash or Near-Crash Events: SAE Technical Paper.

Radja, G. A. (2014). National Automotive Sampling System – Crashworthiness Data System, 2013 Analytical User's Manual. *Report No.DOT HS 812 066*.

Schaap, T. (2012). Driving Behaviour in Unexpected Situations.

Stavrinos, D., Jones, J. L., Garner, A. A., Griffin, R., Franklin, C. A., Ball, D., . . . Fine, P. R. (2013). Impact of distracted driving on safety and traffic flow. *Accident Analysis & Prevention, 61*, 63-70.

Strayer, D. L., Drews, F. A., & Crouch, D. J. (2006). A comparison of the cell phone driver and the drunk driver. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48*(2), 381-391.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of experimental psychology: Applied, 9*(1), 23.

Sussman, E. D., Bishop, H., Madnick, B., & Walter, R. (1985). Driver inattention and highway safety. *Transportation Research Record, 1047*, 40-48.

Wang, W., Zhang, W., Guo, H., Bubb, H., & Ikeuchi, K. (2011). A safety-based approaching behavioural model with various driving characteristics. *Transportation research part C: emerging technologies, 19*(6), 1202-1214.

Yan, X., Harb, R., & Radwan, E. (2008). Analyses of factors of crash avoidance maneuvers using the general estimates system. *Traffic injury prevention, 9*(2), 173-180.