Louisiana Transportation Research Center

Final Report 580

Exploring Naturalistic Driving Data for Distracted Driving Measures

by

Sherif Ishak, Ph.D., P.E.
Osama A. Osman, Ph.D.
Julius Codjoe, Ph.D.
Syndney Jenkins
Sogand Karbalaieali
Matthew Theriot
Peter Bakhit

Louisiana State University

4101 Gourrier Avenue | Baton Rouge, Louisiana 70808
(225) 767-9131 | (225) 767-9108 fax | www.ltrc.lsu.edu

1. Report No. **FHWA/LA.17/580**	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle **Exploring Naturalistic Driving Data for Distracted Driving Measures**	5. Report Date October 2017 6. Performing Organization Code LTRC Project Number: 15-1SA State Project Number: DOTLT1000053	
7. Author(s) Sherif Ishak, Osama A. Osman, Julius Codjoe, Syndney Jenkins, Sogand Karbalaieali, Matthew Theriot, Peter Bakhit	8. Performing Organization Report No.	
9. Performing Organization Name and Address Department of Civil and Environmental Engineering Louisiana State University Baton Rouge, LA 70803	10. Work Unit No. 11. Contract or Grant No.	
12. Sponsoring Agency Name and Address Louisiana Department of Transportation and Development P.O. Box 94245 Baton Rouge, LA 70804-9245	13. Type of Report and Period Covered **Final Report** **02/16/15 – 02/15/17** 14. Sponsoring Agency Code	

15. Supplementary Notes
Conducted in Cooperation with the U.S. Department of Transportation, Federal Highway Administration

16. Abstract
The SHRP 2 NDS project was the largest naturalistic driving study ever conducted. The data obtained from the study was released to the research community in 2014 through the project's InSight webpage. The objectives of this research were to (a) explore the content of this large dataset and perform statistical analysis to identify useful performance measures to detect distracted driving behavior, and (b) provide an outline for a crash index model that can be used to quantify the crash risk associated with distracted driving behavior. Time series data on driver GPS speed, lateral and longitudinal acceleration, throttle position, and yaw rate were extracted as five appropriate performance measures available from the NDS that could be used for the purpose of this research. Using this data, the objective was to detect whether a driver was engaged in one of three specific secondary tasks or no secondary task at all using the selected performance measures. The specific secondary tasks included talking or listening on a hand-held phone, texting/dialing on a hand-held phone, and driver interaction with an adjacent passenger. Multiple logistic regression was used to determine the odds of a driver being engaged in one of the secondary tasks given their corresponding driving performance data. The results indicated that while none of the models provided a statistically good fit of the data, the lateral acceleration measure seemed to be a useful indicator of drivers' engagement in talking/listening and texting/dialing on the cell phone. The analysis of distracted driving behavior for by age and gender showed slightly different results. The longitudinal acceleration variable appeared to perform better in predicting talking/listening and texting/dialing for drivers aged 70-89. The lateral acceleration measure, however, performed better in predicting the engagement of younger drivers (16-29) in the same secondary tasks. When considering the gender of drivers, the lateral acceleration performance variable proved to be more effective in predicting texting/dialing and talking/listening for both genders. Still, these results are inconclusive due to the undesirable Hosmer and Lemeshow Test p-values observed in all the models. Thus, the same analysis was performed using neural networks modeling which is recognized for its capability of nonlinear pattern recognition. The neural network analysis showed that the five performance measures can be used as surrogate measures of distracted driving. The developed neural network models also proved to be good tools for detecting drivers' engagement in secondary tasks. A proposed framework of crash index calculation provides an insight into how the crash risk associated with distracted driving behavior can be quantified. Further research is required to identify the required statistical analysis for the crash index calculation as well as provide further details on how such index can be used.

17. Key Words Traffic Safety, Distracted Driving, Naturalistic Driving		18. Distribution Statement **Unrestricted. This document is available through the National Technical Information Service, Springfield, VA 21161.**	
19. Security Classif. (of this report) N/A	20. Security Classif. (of this page) N/A	21. No. of Pages 95	22. Price $124,321

Project Review Committee

Each research project will have an advisory committee appointed by the LTRC Director. The Project Review Committee is responsible for assisting the LTRC Administrator or Manager in the development of acceptable research problem statements, requests for proposals, review of research proposals, oversight of approved research projects, and implementation of findings.

LTRC appreciates the dedication of the following Project Review Committee Members in guiding this research study to fruition.

LTRC Administrator/Manager
Kirk Zeringue
Special Studies Research Administrator

Members
Autumn Goodfellow-Thompson
Dan Magri
April Renard
Cathy Gautreaux
Ken Trull
David Staton
Betsey Tramonte

Directorate Implementation Sponsor
Janice P. Williams
DOTD Chief Engineer

Exploring Naturalistic Driving Data for Distracted Driving Measures

by

Sherif Ishak, Ph.D., P.E.
Osama A. Osman, Ph.D.
Julius Codjoe, Ph.D.
Syndney Jenkins
Sogand Karbalaieali
Matthew Theriot
Peter Bakhit

Department of Civil and Environmental Engineering
3418 Patrick F. Taylor Hall
Louisiana State University
Baton Rouge, LA 70803

LTRC Project No. 15-1SA
State Project No. DOTLT1000053

conducted for

Louisiana Department of Transportation and Development
Louisiana Transportation Research Center

October 2017

ABSTRACT

The Second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS)project was the largest naturalistic driving study ever conducted. The data obtained from the study was released to the research community in 2014 through the project's InSight webpage. The objectives of this research were to (a) explore the content of this large dataset and perform statistical analysis to identify useful performance measures to detect distracted driving behavior, and (b) provide an outline for a crash index model that can be used to quantify the crash risk associated with distracted driving behavior. Time series data on driver GPS speed, lateral and longitudinal acceleration, throttle position, and yaw rate were extracted as five appropriate performance measures available from the NDS that could be used for the purpose of this research. Using this data, the objective was to detect whether a driver was engaged in one of three specific secondary tasks or no secondary task at all using the selected performance measures. The specific secondary tasks included talking or listening on a hand-held phone, texting/dialing on a hand-held phone, and driver interaction with an adjacent passenger. Multiple logistic regression was used to determine the odds of a driver being engaged in one of the secondary tasks given their corresponding driving performance data. The results indicated that, while none of the models provided a statistically good fit of the data, the lateral acceleration measure seemed to be a useful indicator of drivers' engagement in talking/listening and texting/dialing on the cell phone. The analysis of distracted driving behavior for by age and gender showed slightly different results. The longitudinal acceleration variable appeared to perform better in predicting talking/listening and texting/dialing for drivers aged 70-89. The lateral acceleration measure, however, performed better in predicting the engagement of younger drivers (16-29) in the same secondary tasks. When considering the gender of drivers, the lateral acceleration performance variable proved to be more effective in predicting texting/dialing and talking/listening for both genders. Still, these results are inconclusive due to the undesirable Hosmer and Lemeshow Test p-values observed in all the models. Thus, the same analysis was performed using neural networks modeling, which is recognized for its capability of nonlinear pattern recognition. The neural network analysis showed that the five performance measures can be used as surrogate measures of distracted driving. The developed neural network models also proved to be good tools for detecting drivers' engagement in secondary tasks. A proposed framework of crash index calculation provides an insight into how the crash risk associated with distracted driving behavior can be quantified. Further research is required to identify the required statistical analysis for the crash index calculation as well as provide further details on how such index can be used.

ACKNOWLEDGMENTS

This project was completed with support from the Louisiana Department of Transportation and Development (DOTD) and the Louisiana Transportation Research Center (LTRC). The research team also gratefully acknowledges the assistance received from the Project Review Committee (PRC) members for their valuable feedback and all other DOTD personnel involved during the course of this project.

IMPLEMENTATION STATEMENT

Distracted driving has long been acknowledged as one of the main contributors to crashes in the US. Distracted driving has captured the attention of many researchers and transportation officials due to its significant impact on traffic safety. A recent study funded by LTRC and University Transportation Center (UTC), "Distracted Driving and Associated Crash Risks," concluded that texting and talking to passengers while driving impaired driving performance but failed to find any significant effects for cell phone conversation. The study was however unable to make any statistical findings on the driving performance based on demographics and road facility type because of the limited sample utilized. The Second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS) collected large amounts of data on people's driving behavior in six states across the US. This data offers ample opportunity to utilize a bigger sample size that will allow statistical conclusions to be drawn on various strata including gender, road facility type, age, and time of day. This report presents findings of a comprehensive exploration study on the SHRP 2 NDS data to identify appropriate performance measures that can be used as surrogate measures for distracted driving behavior, and outline a methodology of developing a crash index. The findings of this report provide an insight on the usefulness of the SHRP2 NDS data for distracted driving studies to the officials of DOTD and other interested transportation officials within Louisiana. Based on the reported findings of this study, some performance measures were identified as surrogates to detect distracted driving behavior. However, these findings were inconclusive as the powers of the performed statistical tests were very low. This performance can be explained by the nonlinearity in driving behavior which needs more advanced analysis tools. Thus, artificial intelligence was implemented and proved to have high accuracy in detecting drivers' engagement in secondary tasks. Moreover, the artificial intelligence tool proved that the five measures used in the analysis can be used as surrogate measures for distracted driving behavior.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

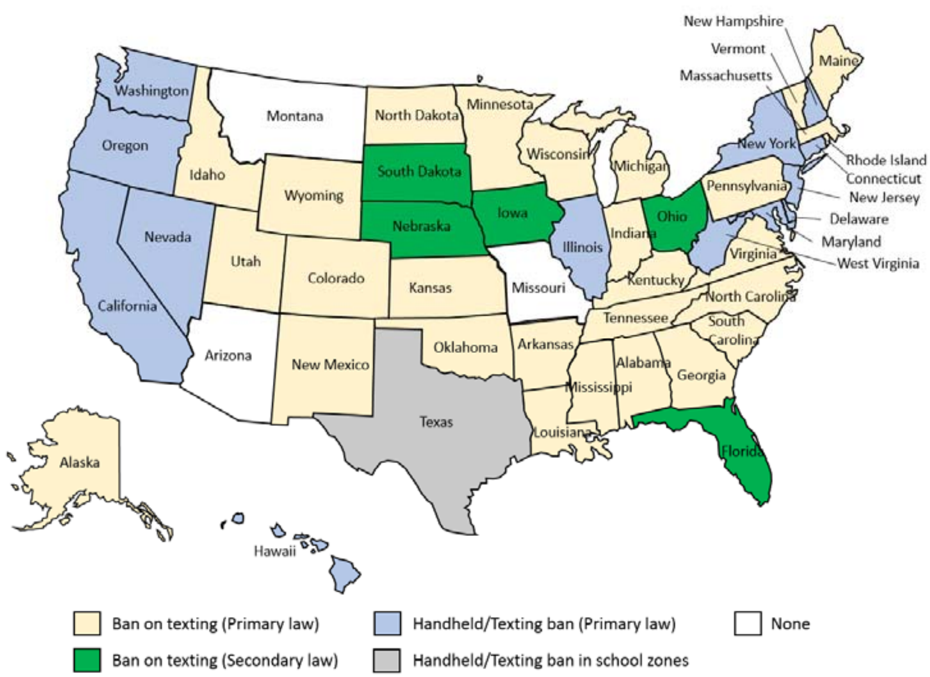
Distracted driving is a dangerous epidemic that continues to cause deaths and injuries in related crashes throughout the U.S. According to the National Highway Traffic Safety Administration, 3,328 people (including 540 non-occupants) were killed and an estimated additional 421,000 were injured in 2012 from distraction-affected crashes *[1]*. In Louisiana, a reported 675 people were killed in 2011 from motor vehicle crashes, and it is estimated that 10% (national estimate from NHTSA) of these were a result of distracted driving. Causes of distracted driving involve activities that divert the driver's attention from the driving task and may include eating, adjusting the radio or climate controls, talking to passengers, cell phone use and texting, as well as many other external distractions. Such distractions are likely to affect the driving performance and consequently elevate the crash risk of drivers.

To minimize the effect of distracted driving on safety, proactive laws have been established banning secondary tasks while driving, specifically the use of cell phones as a main reason for distraction. These laws vary from state to state and can be established as either primary or secondary laws. When a law is established with primary enforcement, officers are permitted to ticket the driver for this offense without the driver disobeying any additional laws. On the other hand, for an officer to enforce a secondary law, a primary law must have been violated first. The different primary and secondary laws issued in the all the united states are shown in Figure 1. As shown in Figure 1(a), most of the states have issued a ban on cellphone texting for all drivers as a primary law. Although several states have not included texting bans for all drivers, other precautionary measures were taken for novice drivers, as shown in Figure 1(b). The majority of states have chosen to ban cellphone use entirely from novice drivers, with the assumption that they are more prone to cellphone related incidents.

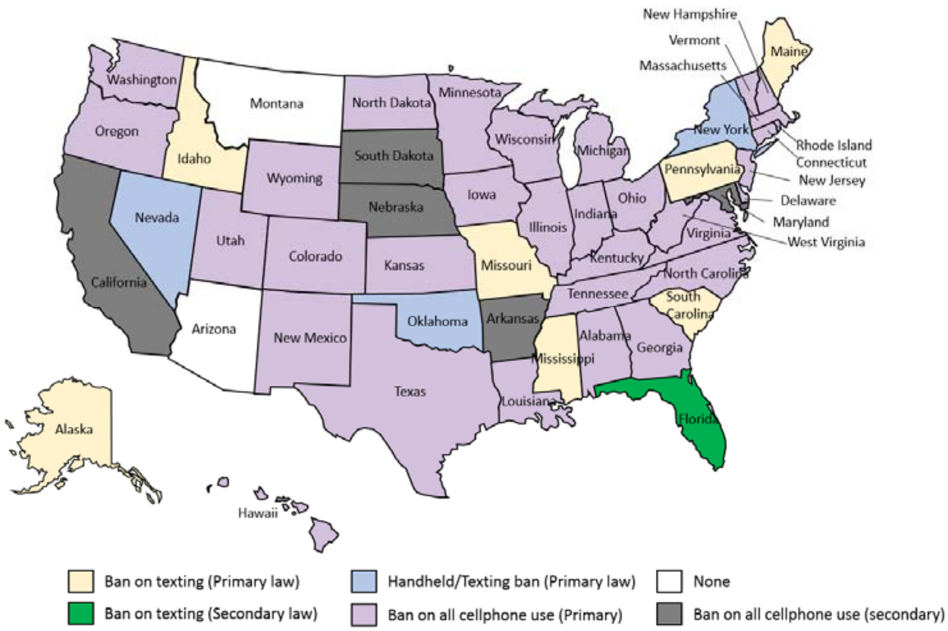
Given the effect of cellphone use while driving on safety and the significance of any related incidents that might take place for bus drivers, several states have issued regulations for cellphone use specifically for bus drivers. While there is a discrepancy between states on the best way to regulate bus drivers' use of cellphones, all but two states have banned the use of cellphones for such a category of drivers, as seen in Figure 1(c).

The cellphone-use-while-driving regulations are meant to reduce the effect of distraction on safety. Enforcement of these regulations leads many drivers to avoid being ticketed, and hence accidents related to cellphone use while driving are minimized and many lives are saved. Most of these regulations, if not all, are based on results from research studies performed in collaboration between universities, research institutes, and government officials. In Louisiana, a recent study funded by LTRC and UTC, "Distracted Driving and

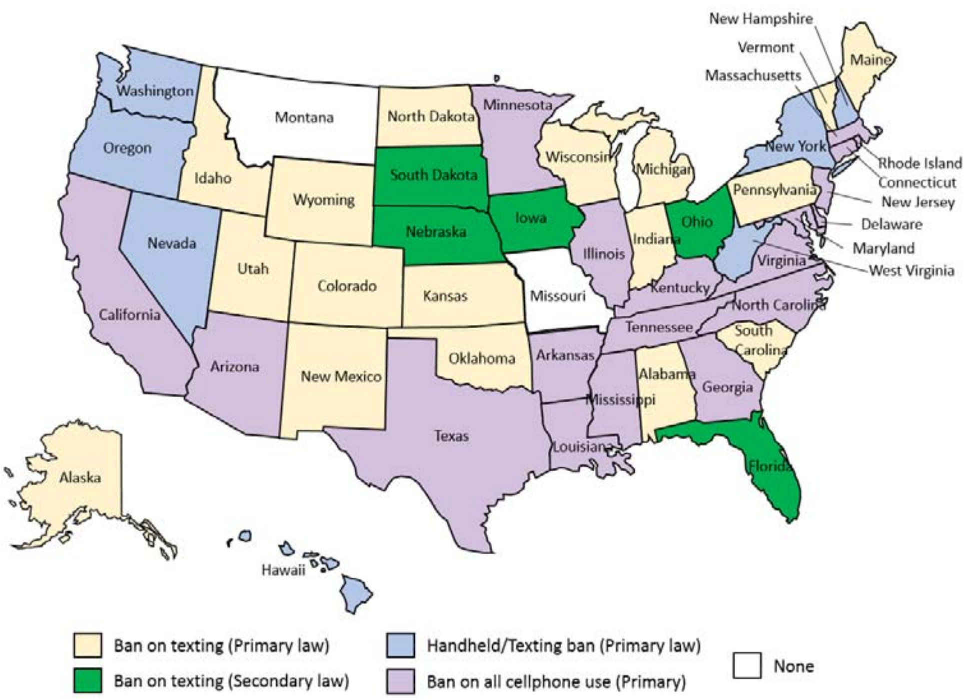
Associated Crash Risks," concluded that texting and talking to passengers while driving impaired driving performance but failed to find any significant effects for cellphone conversation. The study was, however, unable to make any statistical findings on the driving performance based on demographics and road facility type because of the limited sample utilized. With the recent availability of data from the Strategic Highway Research Program (SHRP 2) Naturalistic Driving Studies (NDS), there may be ample opportunity to utilize a bigger sample size in a further study that will allow statistical conclusions to be drawn on various strata including gender, road facility type, age, and time of day. NDS offers the ability to observe drivers in their own vehicles, driving their typical commutes, and exhibiting their normal driving behavior *[2]*. This aspect, that is unique to NDS, more accurately reflects actual driving behavior when compared to driver simulator studies that use a simulation vehicle and ask the driver to maneuver through a simulated environment. However, the SHRP 2 data is relatively new, and it is not clear whether the data needs for the further study can be met solely from what is available. Therefore, this study aims to perform a comprehensive exploration of the SHRP 2 NDS data with the view of identifying if it can provide the data required for an enhanced study on the crash risks of distracted driving. This study also includes an outline for the development of a Crash Risk Index to evaluate potential risk associated with drivers based on their socioeconomic characteristics and secondary task involvement.



(a) All drivers



(b) Novice drivers



(c) Bus drivers

Figure 1
Cellphone laws across the United States

Literature Review of Distracted Driving

Distracted driving continues to be a risky behavior that poses a danger to drivers, vehicle occupants, and non-occupants such as pedestrians and cyclists. Causes of distraction range from external sources (outside object, crash incident, scenery, advertisements, finding direction, etc.) to internal sources (in car moving object, reading or writing, eating or drinking, grooming, etc.). It was not until the past decade, however, that distracted driving came to the forefront of public awareness, stemming in large part from the rapid increase in cell phone ownership and the explosion in portable and in-vehicle devices that have become available. These devices allow drivers to engage in activities that were previously inconceivable (e.g., browsing the Internet) and have the capacity to absorb drivers' attention to a whole new degree. Nationwide, this has increased the crash risk of drivers and in the year 2012, resulted in increased number of fatal crashes (10%), injury crashes (18%), and motor vehicle traffic crashes (16%) *[1]*. It has become one of the focuses of state departments of transportation to reduce the occurrence of distracted driving and raise awareness of its dangers.

Distracted driving has captured the attention of many researchers and transportation officials due to its significant impact on traffic safety. Several studies showed that distracted driving is likely to increase the reaction time of drivers and their response time *[3]*. When analyzing the impact of specific secondary tasks, studies have shown that: (a) talking on a handheld cellphone impairs the drivers' ability to maintain their speed and position on the road *[4]*;and (b) texting increases braking reaction times and increases lane-position variability with no change in speed *[5]*. In another study by Klauder et al., the researchers investigated the crash risk associated with performing secondary tasks *[6]*. The results indicated that crash risk significantly increased for novice drivers when they were dialing a cellphone, texting, reaching for objects, looking at roadside objects, and eating. On the other hand, for experienced drivers, the crash risk increased significantly only when drivers were dialing cellphones.

According to Elander et al., unsafe driving behavior is a type of driving style that is developed over time. This unsafe driving behavior becomes a habit that differs from one driver to another according to some socioeconomic characteristics *[7]*. Based on a detailed survey of 834 licensed drivers, Poysti et al. concluded that younger and male drivers tend to use phones more often compared to older and female drivers *[8]*. The survey also showed that driving for longer distances increases the likelihood of cellphone use. More so, people tend to use cellphones more often when they perceive themselves as skilled drivers. Based on a survey conducted by Strayer et al., most drivers may not be aware of their impaired driving behavior while engaged in distracted driving *[9]*.

Driving simulator studies and naturalistic driving studies are two ways that distracted driving can be investigated. Experiments in driving simulators are easier to control and data collection is relatively easier and non-invasive since vehicles are designed with the data acquisition component in mind from the onset. They provide an inexpensive alternative to conventional experiment and sometimes impossible (unethical or safety implications) field tests that cannot be achieved in real life situations *[10]*. Nevertheless, the controlled settings and environments provide a lesser degree of realism compared to NDS. The NDS data include observations of drivers in their own vehicles while driving their normal commutes. To collect these observations, the vehicles were equipped with sensors and other data collection gadgets, which are usually add-ons to the in-vehicle systems a vehicle will normally be equipped with. While NDS will produce more realistic scenarios, and thereby more valuable data to study driver behavior and performance, the collection of data could be problematic and they are very expensive. The first large-scale NDS conducted was the 100-Car Naturalistic Driving Study which involved 241 drivers over an 18-month period resulting in about 3 million vehicle miles that yielded 42,300 data hours, 82 crashes, 761 near-crashes, and 8,295 critical incidents *[11]*. Due to NDS being a behavioral-based observational method of analysis, there are many ways this data can be used to study driver behavior and risk analysis. Some of the studies that have been conducted using the 100-car NDS include validation of near-crashes as crash surrogates, assessing safety critical braking events, prediction of high-risk drivers based on demographic, personality, driving characteristic data, modeling of driver car-following behavior and examining driver inattention *16]*.

The SHRP 2 NDS is the second large-scale and the largest NDS conducted with 3,147 drivers using all light vehicle types over a 3-year period in 6 sites across the nation: Bloomington, Indiana; Central Pennsylvania; Tampa Bay, Florida; Buffalo, New York; Durham, North Carolina; and Seattle, Washington. This study, amounting to over 35 million vehicle miles, is on a scale of 40 times larger than that of the 100-car NDS and specifically recruited drivers at different geographical locations to accommodate variations in weather, geographical features, and rural, suburban, and urban land use. The data collection package includes roadway information database (RID) which provides information on lane departures, intersection crashes, and roadway characteristics such as grade, curvature, and posted speed limits. The detailed nature of the data will allow analyses on the effect of road design characteristics or weather condition on the interaction between the driver and vehicle; driving style comparisons for specific road user groups; prevalence of mobile phone or other in-car information devices and the relationship with particular behavior patterns; the effect of particular interventions; effect of passengers on distraction; and exploration of the interaction between motorized vehicles and vulnerable road *[2]*. While the 100-car NDS data is already 10 years old, the SHRP 2 NDS data has just been released and can remain useable for the next 20 years or more. Very few publications have been released on this relatively new data, providing guidance on how to use the large dataset and also documenting the effort of the data collection process *[2]*.

100-Car NDS Studies

Although it is not the most extensive NDS data set available, the 100-car NDS study provides an insight on several safety concerns which has been available for several years and has been investigated extensively. For instance, Montgomery et al. analyzed the impact that a driver's age and gender has on their ability to break in normal driving situations *[17]*. For their experiment, near-crash and crash data was excluded from the data set. The overall goal of the study was to determine if forward collision warnings (FCW) should be designed to tailor alert timings to the target demographic of a vehicle. Therefore, the authors analyzed time to collision (TTC) data from the 100-car NDS dataset. The results determined that males TTC at braking was 1.3 seconds lower on average than women's TTC. The results also showed that participants aged over 30 had a TTC at braking of 1.7 seconds higher than participants aged under 30 years. With such a significant difference in TTC for both age and gender it was determined FCWs should be designed based on the demographic of their particular vehicle to maximize the effectiveness of this warning system *[17]*.

Another study by Bagdadi analyzed the NDS data using a new method based on critical jerk to determine when critical braking events have occurred *[13]*. The author compared his new method to another method commonly used to analyze longitudinal acceleration measures. The study investigated only the NDS data where evasive braking action was taken before near-crash events. To measure the braking, Bagdadi analyzed the jerk rate, which is the rate of change of acceleration by 1.0 g/s as the threshold for critical jerks. While the longitudinal acceleration method produced a success rate of 54.2% with a threshold of 0.6g as seen in Figure 2, the new method provided a success rate of 86%. Bagdadi also performed the test using critical jerk thresholds of 0.8g/s and 1.2g/s to compare. The analysis results showed that the success rate increased by 9% and decreased by 9% for 0.8g/s and 1.2g/s, respectively. A similar procedure was done for acceleration as shown in Figure 2. The study results showed the proposed method outperformed the longitudinal acceleration method by 1.6 times. However, the proposed method was not able to determine the false rate of near-crash events, which can be easily performed with the longitudinal acceleration method.

Jonasson analyzed the available near-crash identification method used for the 100-car study *[12]*. In this method, near-crash selection occurs in a two-step process. The first step uses kinematic triggers for automated identification of potential candidate events. Next, the visual recordings within the time windows of the events must be reviewed to select the near-crash events based on specified criteria. Viewing the recordings to make the selections in this method allows for a subjective decision in determining near-crash events. Jonasson noted two situations where there seems to be selection bias in the 100-car study. The 100-car study showed that 34% of crashes involved no reaction from the driver, but only 5% of near-crashes involved no reaction because these events were not captured by the kinematic triggers for near-crashes. This was likely because these events were not captured by the kinematic triggers for near-crashes. Another instance of bias is with rear-end striking at speeds under 25 km/h. The data showed to drive slower than 25 km/h is 48 times more dangerous which seems highly unlikely.

Table 7
Sensitivity analysis by testing alternative threshold values for the critical jerk method.

Threshold values	Detected (total)	Success rate (%)	Change in success rate
Critical jerk = 1.0 g/s	548 (637)	86	–
Critical jerk = 0.8 g/s	595 (637)	93	9% increase
Critical jerk = 1.2 g/s	496 (637)	78	9% decrease

Table 8
Sensitivity analysis by testing alternative threshold values for the longitudinal acceleration method.

Threshold values	Detected (total)	Success rate (%)	Change in success rate
Acceleration = −0.60 g	345 (637)	54	–
Acceleration = −0.72 g	214 (637)	34	38% decrease
Acceleration = −0.48 g	487 (637)	76	41% increase

Figure 2
Critical jerk vs. longitudinal acceleration analysis *[13]*

To overcome these limitations, Jonasson applied two methods based on extreme value statistics to validate near-crash events differently *[12]*. The first method used near-crashes to predict crash frequency in the 100-car study data. This was performed by fitting a generalized extreme function (GEV) distribution to the observed maxima -TTC in all near-crashes. Then, if a crash occurs when the TTC value crosses 0, an estimate of this probability was computed using the fitted GEV and compared with the observed crash frequency. The second method involved multivariate near-crash modeling which was performed by finding continuous variables that could contribute to causing crashes. Then, this was fitted to a multivariate GEV to max (–TTC), and the data were compared to the distribution of the same variables in the crashes. The results of this study showed a discrepancy between the distribution of maximum speeds for crashes and maximum speeds for near-crashes, portrayed in Figure 3. He confirmed that there was considerable bias in the selection of near-crashes as shown below.

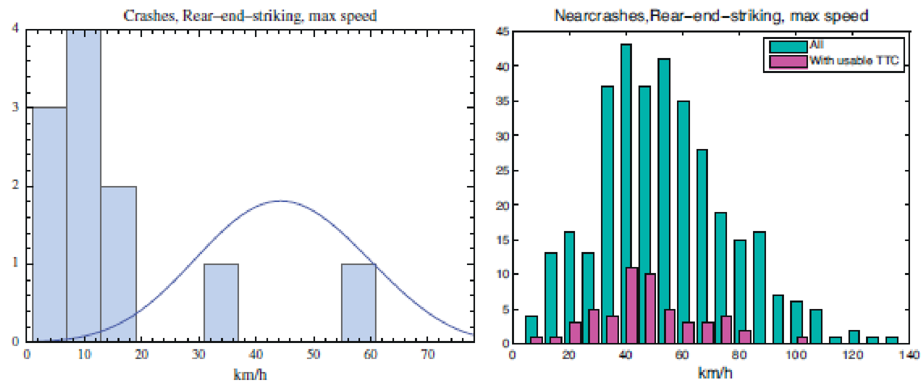


Figure 3
Crash data and near-crash selection vs. speed *[12]*

Klauer et al. studied the impact of driver inattention on near-crash and crash risk using the 100-car data *[18]*. In this study, distracted driving (driver inattention) data were obtained from baseline events and compared to those obtained from combined crash and near-crash events. Based on eye glance data, several driver inattention instances were reported including engagement in secondary tasks, drowsiness, driving related inattention to forward roadway, and non-specific eye glances away from the forward roadway. The study showed that drowsiness increased near-crash/crash risk by four to six times and engagement in secondary tasks increased risk by two times compared to normal driving. On the other hand, driving-related inattention to the forward roadway increased safety by almost two times. This increase in safety was expected as driving related inattention included actions such as checking rearview mirrors, meaning that drivers were more alert. The study also showed that drowsiness contributed to 22% of all the near-crash /crashes and occurred much more frequently during free flow situations. For the baseline data, secondary tasks occurred during 54% of the datasets, driving related inattention occurred during 44%, drowsiness occurred during 4%, and non-specific eye glances occurred during 2%. The analysis showed that eye glances of fewer than 2 seconds were useful for the drivers, whereas those that lasted over 2 seconds were considered to impact drivers' safety significantly.

SHRP2 NDS Studies

The literature is being enriched with studies using SHRP 2 NDS data. Example studies include the Iowa State University Center for Transportation Research and Education (CTRE): Lane departures on rural two-lane curves *[19]*; MRIGlobal: Offset left-turn lanes *[20]*; University of Minnesota Center for Transportation Studies (CTS): Rear-end crashes on congested freeways *[21]* ; and SAFER Vehicle and Traffic Safety Centre at Chalmers University, Sweden: Driver inattention and crash risk *[22]*. These studies were only able to use limited data from the SHRP2 study, since they began before the data collection process was complete.

Researchers at Chalmers University of Technology in Sweden performed the first study incorporating the SHRP 2 and RID data *[22]*. Their study analyzed the effects of driver distractions using the SHRP 2 data. The primary goal of the study was to develop inattention-risk relationships that determine the relationship between driver inattention and crash risk in lead-vehicle pre-crash scenarios. These relationships help determine which glances are most dangerous for drivers.

The dataset used for this study included 46 rear end crashes, 211 near-crashes, 257 matched baseline events, and 260 random baseline events. Matched baseline events allow the researchers to compare glance data by matching factors such as driver, trip, traffic flow, speed, and weather to the near-crash /crash events. Over 50 distracting activities were examined, but many of these distractions did not occur frequently enough to have statistical significance.

The analysis confirmed some of the findings from previous studies. It confirmed that distracting activities occurred more frequently in near-crash events, visually demanding tasks involved more risk, and texting had the highest odds ratio, meaning it leads to a significant risk. The danger of glances was quantified using a three metric model including inopportune glance, mean glance duration, and the driver's uncertainty of the driving scenario. Figure 4 shows that crash risk increases, the longer a driver's eyes are off path. The results also found that lead vehicle crashes are caused by a combination of glance duration and closure rate. The researchers note that their results suggest the need for FCW, autonomous cruise control, and autonomous emergency braking *[22]*.

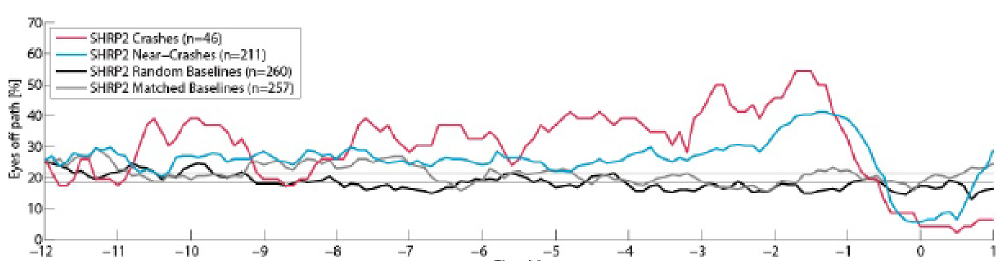


Figure 4
Driver glance duration's impact on crashes *[22]*

The second project assigned to the SHRP 2 data was MRIGlobal's study, which uses NDS and RID data to provide guidance for safety countermeasures to offset left turn lanes *[20]*. Gap acceptance behavior was a contributing analysis factor to this study. The main goal of their research was to evaluate left turning gap acceptance by an extensive sample of drivers at different intersections that incorporate left turn lane offsets.

Left turns at intersections can have a negative offset, positive offset, or no offset. The study analyzes situations where the drivers' view was both obstructed by oncoming left turn vehicles and not obstructed. The data set included 6,500 intersections, 44 signalized intersection left-turn offset pairs, and 14 two way stop controlled intersection left turn offset pairs. The research team analyzed video footage when NDS drivers made left turning maneuvers at these intersections and collected data including weather conditions, signal indications, presence of other vehicles, and the start and end time of each gap rejected or accepted by the driver.

The analysis used a logistic regression to predict the critical gap from left turning vehicles in each offset category. The results determined that as the offset became more negative, the critical gap length increased. Critical gaps were also 2 seconds longer when the sight was restricted from an oncoming left turning vehicle, but this result is not considered a statistically significant amount. It was also determined that intersections designed to allow vehicles' view to be blocked from oncoming left turn vehicles decreased the operation efficiency of the intersection. Since there was no crash data from these intersections, data was too limited to determine crash related safety *[20]*.

The University of Minnesota's study was not completed, so only preliminary analysis is available *[21]*. The primary goal of the study was to determine how drivers behave when encountering a freeway stopping wave. This information can be used to reduce congestion on urban roadways.

The NDS data includes 250 freeway trips containing break-to-stop events. From the NDS data, researchers can obtain braking deceleration data, along with following vehicle reaction time and following distance. With this information, it is possible to gain more insight on drivers' behavior on congested freeways *[21]*.

Iowa State University's research team performed a study to analyze roadway departures on rural two lane curves *[19]*. The purpose of the study was to use the NDS and RID data to determine how driving behavior, roadway factors, and environmental factors relate to these departures. Only paved roadways over one mile out of the urban area with speeds posted 40-60 mph were included in the study.

The research helped to determine what defines a curve's area of influence, normal behavior on a curve, and relationships between driver distractions and risk of roadway departure. To define curves' area of influence they had to determine where drivers begin to react to the curve. By using time series data, regression models were able to determine that drivers began reacting 538-591 feet upstream of the point of curvature. This information is useful for signage and other traffic control measures.

Time series models were also used to evaluate lane position and speed of the vehicles. The results showed that drivers tended to maintain their upstream position during the curve and that distractions caused them to shift in the lane. If they were on the inside and encountered a distraction, they tended to shift 0.46 feet towards the right at the next point in the curve. This shows the need for rumble strips or paved shoulders as a counter safety measure. Younger drivers were found to speed into curves more than older drivers by 0.5 mph per every 10 years.

In addition to these models, four multivariate logistic regression models were used to evaluate how environmental factors affect roadway departure. The results showed that right side lane departure is 6.8 times more likely on the inside of a curve. The presence of a guardrail decreased inside departures by 66%. Also, males were found to have outside lane departures four times as often as females *[19]*.

A more recent study by Dingus et al. used the SHRP 2 data to evaluate driver crash risk factors and prevalence *[23]*. This research provided important insight, as drivers tend to become distracted when they are involved in secondary tasks such as texting, interaction with a passenger, talking on a handheld cell phone, eating, and adjusting the radio among others. Their research team conducted analyses on crashes and controls for impairment, performance error, judgment error, and distraction. Through their findings it was determined that drivers tend to be engaged with at least one secondary activity during 51.93% of the time while driving, which raises the crash risk to at least 2 times higher than it is during normal driving.

OBJECTIVES

The main focus of this exploratory study was to compile a technical summary of the limitations and capabilities of the SHRP 2 NDS data for an enhanced research on distracted driving that will provide valid statistical inferences to be applied to Louisiana drivers based on gender, age, and road facility type.

SCOPE

This study focused on exploring the naturalistic driving data collected under the SHRP2 Naturalistic Driving Study (NDS) at Virginia Tech Transportation Institute (VTTI).

DATA EXPLORATION: SHRP 2 NATURALISTIC DRIVING STUDY

The SHRP 2 program was created to address three national transportation challenges: improving highway safety, reducing congestion, and improving methods for renewing roads and bridges. The Naturalistic Driving Study (NDS) was developed to target the safety component of the program. The goal of the SHRP 2 NDS was to "improve traffic safety by obtaining objective information on driver behavior and driver interaction with the vehicle and the roadway" *[2]*. What do drivers actually do in their vehicles? What were they doing immediately before they crashed? These are examples of the type of research questions this study aimed to answer. The SHRP 2 NDS was 40 times larger than the 100-Car NDS Study, and was the first of its kind to obtain data from all over the nation. In total the study included 3,147 drivers, about 50 million miles of driving, and 3 years' worth of data from 6 data collection sites.

Data Description

To collect the NDS data, each vehicle was equipped with a data acquisition system (DAS) developed by the Virginia Tech Transportation Institute. The DAS includes forward radar, accelerometers, vehicle network information, Geographic Positioning System (GPS), on-board computer vision lane tracking, data storage capability and four video cameras, including one forward-facing, color and wide-angle view *[2]*. The DAS continuously recorded data while the participant's vehicle was in operation. A depiction of the equipment installed in each vehicle is shown in Figure 5.

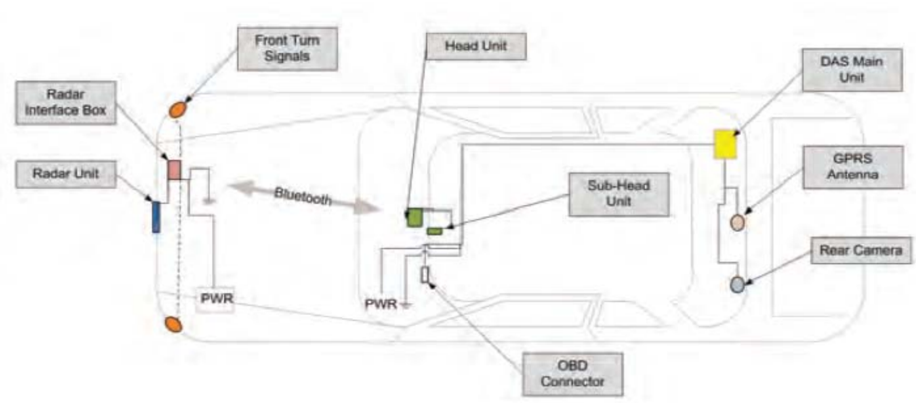


Figure 5
Data Acquisition System installed in participant's vehicles

The SHRP 2 NDS used all light vehicle types over a three-year period in 6 sites across the nation were specifically recruited across these six different geographical locations to accommodate variations in weather, geographical features, and rural, suburban, and urban land use. In the next sections, the method in which SHRP 2 officials distributed the data obtained from the NDS is discussed. Much of the data can be viewed on the SHRP 2 NDS Insight website. In order to gain access to the site, researchers must register as either a “guest” or under “qualified researcher” status. To obtain qualified researcher status, one must present acceptable proof of completion of Institutional Review Board (IRB) training for dealing with Personal Identifiable Information. As a qualified researcher, more of the dataset is viewable online; however, even under this recognition, the data presented cannot be downloaded or exported directly from the webpage. Researchers must complete a Data Sharing Agreement with SHRP 2 officials in order to receive the desired datasets in a usable form.

NDS Data on InSight Website

The website divides the database into the following five categories: Vehicles, Drivers, Trips, Events, and Query Builder, as shown in Figure 6. Within each category there is a description of the data available and an “Info” tab that when accessed provides background, conversions, coordinates, version history and an overview of all variables comprised within the dataset. Figure 7 shows a portion of one of these Info tabs.

The Vehicles category contains summary information on the vehicles that were driven throughout the study. Graphs are used to display data on vehicles by classification, model year, beginning mileage, amount of data collected, timing of equipment installation and number of vehicles actively collecting data per month. Example information is shown in Figure 8 to Figure 13. In addition to these graphs, a Vehicle Detail Table that provides detailed data on each vehicle used in the study.

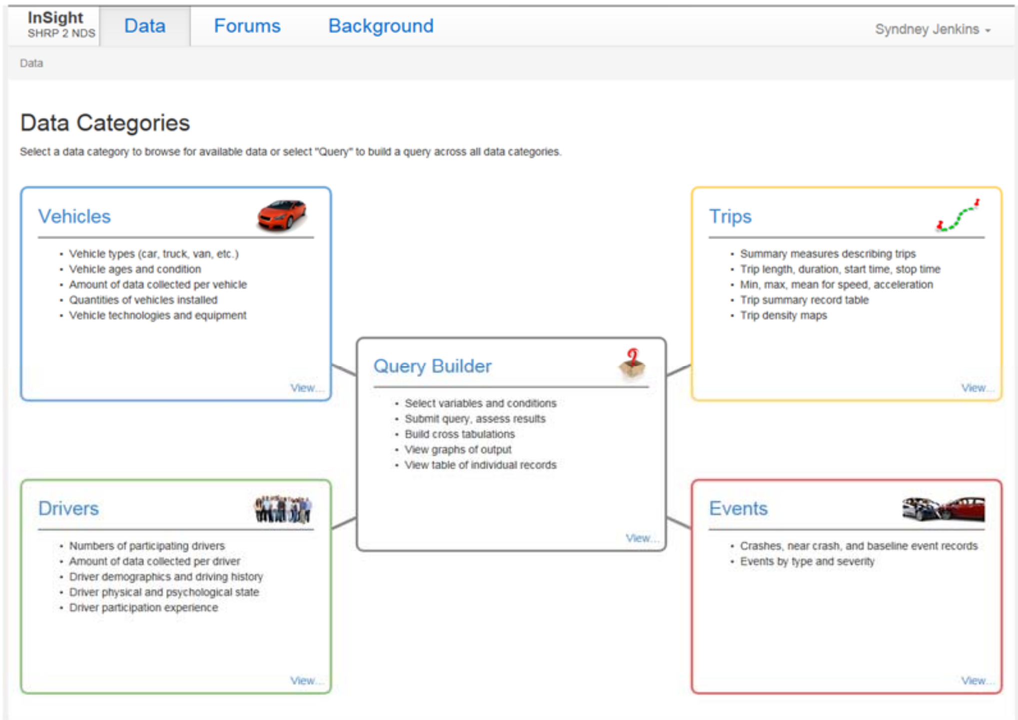


Figure 6
Data available on InSight webpage ***[2]***

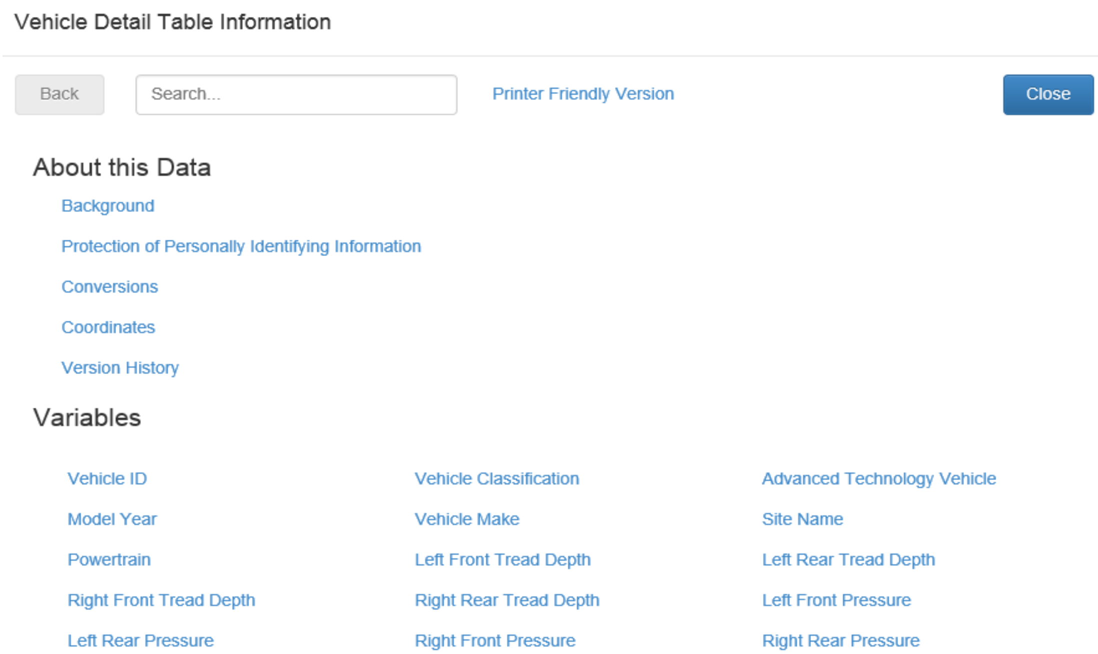


Figure 7
Portion of Vehicle Category overview on InSight webpage ***[2]***

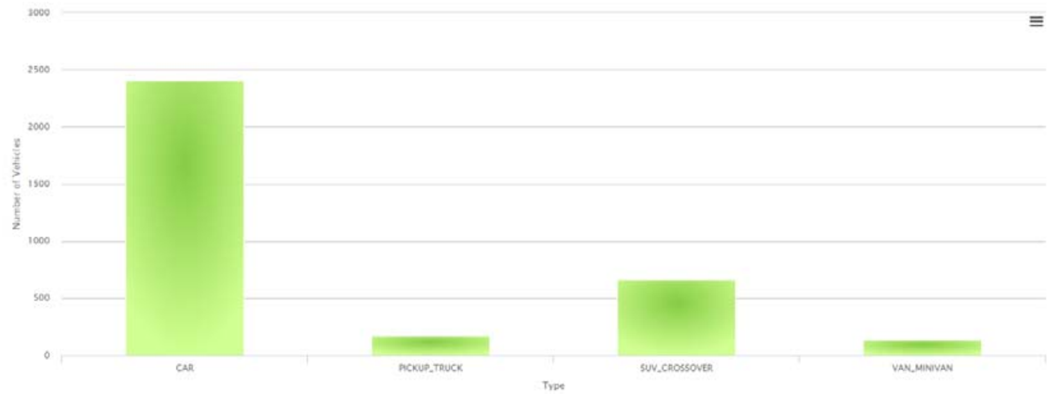


Figure 8
Number of participating vehicles in the study per type *[2]*

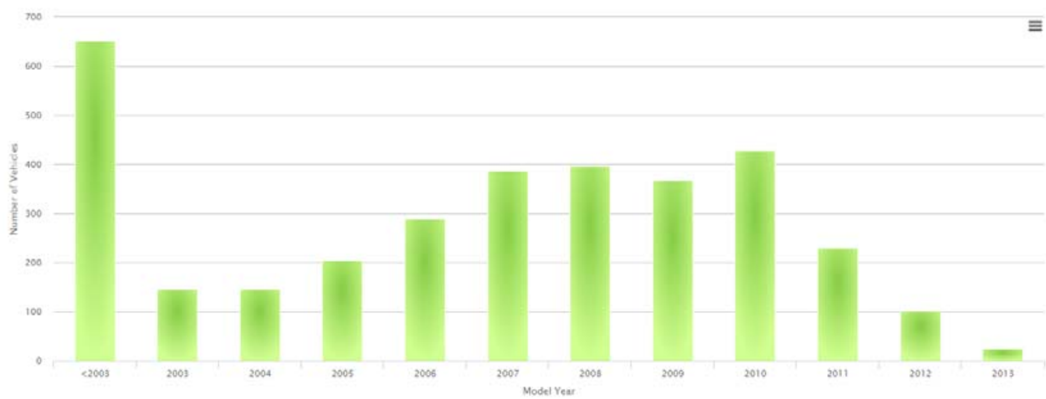


Figure 9
Number of participating vehicles in the study per model year *[2]*

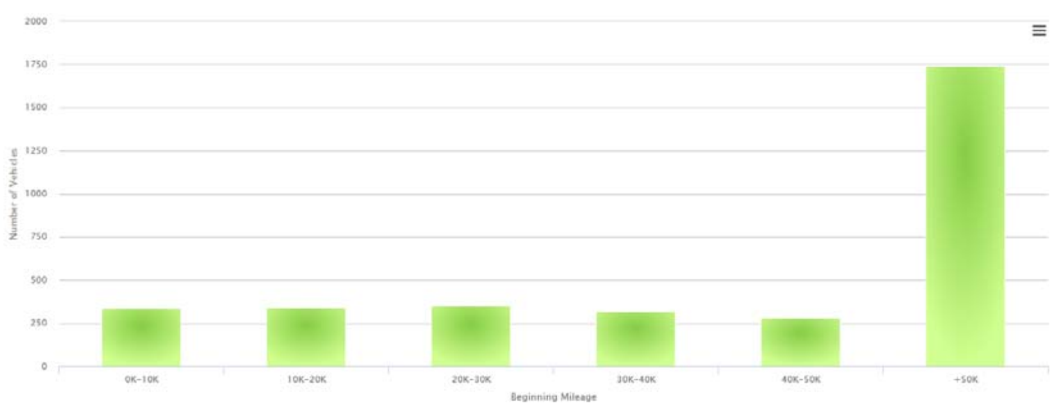


Figure 10
Number of participating vehicles in the study categorized by beginning mileage *[2]*

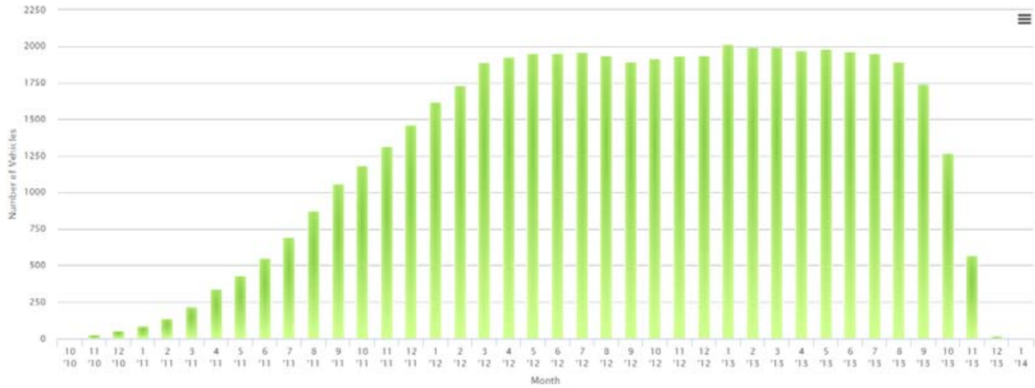


Figure 11
Number of participating vehicles in the study categorized by date of participation *[2]*

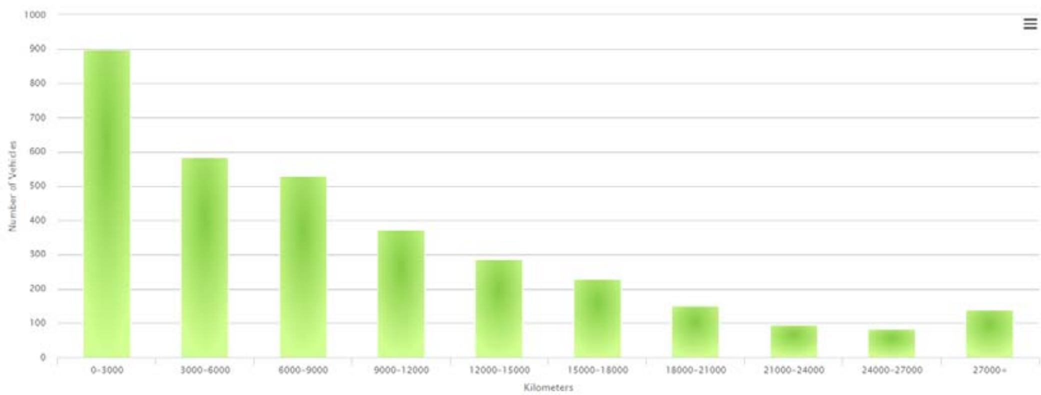


Figure 12
Number of participating vehicles in the study categorized by the travelled kilometers in the study *[2]*

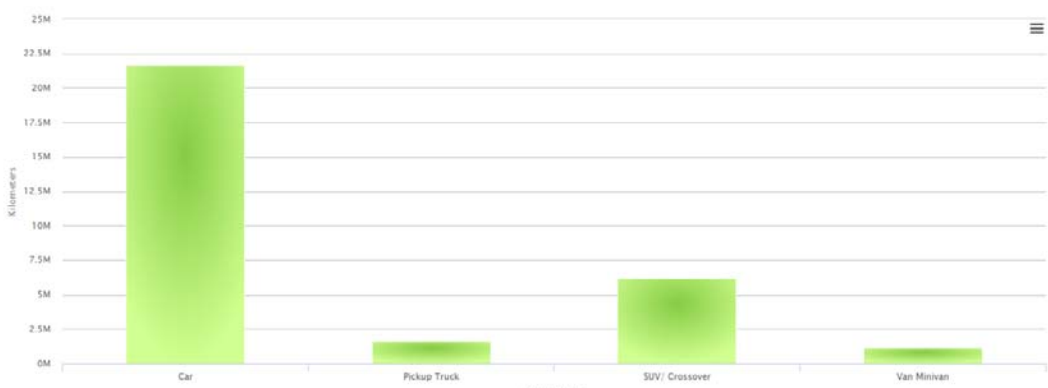


Figure 13
Travelled distance in the study categorized by vehicle type *[2]*

The Drivers category houses data on the numbers of participating drivers, amount of data collected per driver (example shown in Figure 14 and Figure 15), driver demographic and driving history (example demographics in Figure 16 and Figure 17), driver physical and psychological state, and driver participation experience. The drivers were given physical strength tests that include hand strength measurements through a hand dynamometer, and raw walk time test that measured the time it took participants to complete a 10 feet walk each way. To measure driver's psychological condition, they were given Barkley's ADHD Screening Test, a Risk Perception Questionnaire, Risk Taking Questionnaire, Sensation Seeking Scale Survey, and a Driver Behavior Questionnaire.

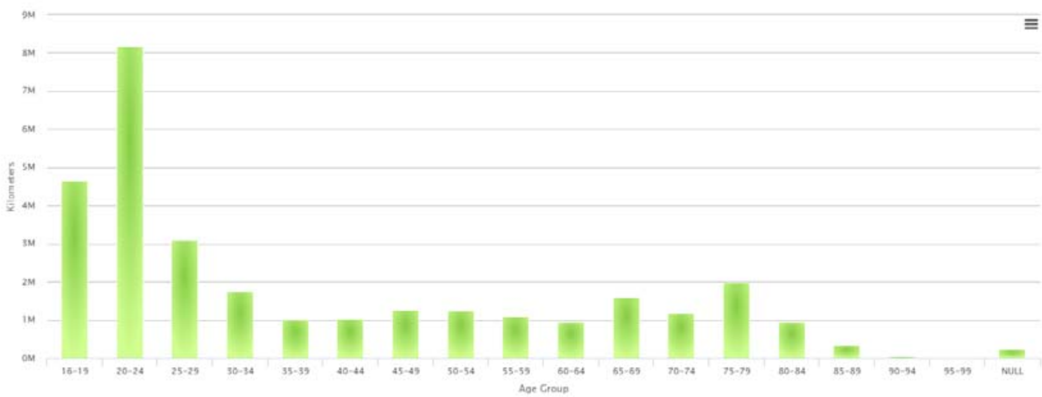


Figure 14
Trips travelled by each age group *[2]*

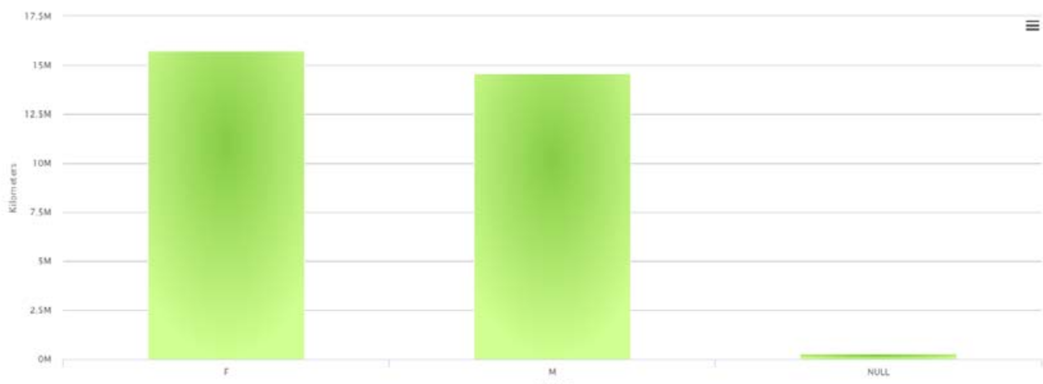


Figure 15
Trips travelled by each gender *[2]*

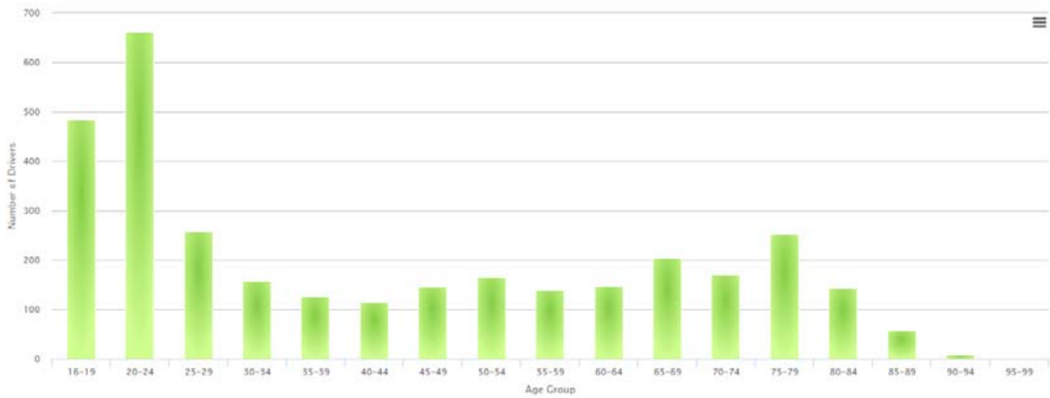


Figure 16
Number of participating drivers in the study categorized by age *[2]*

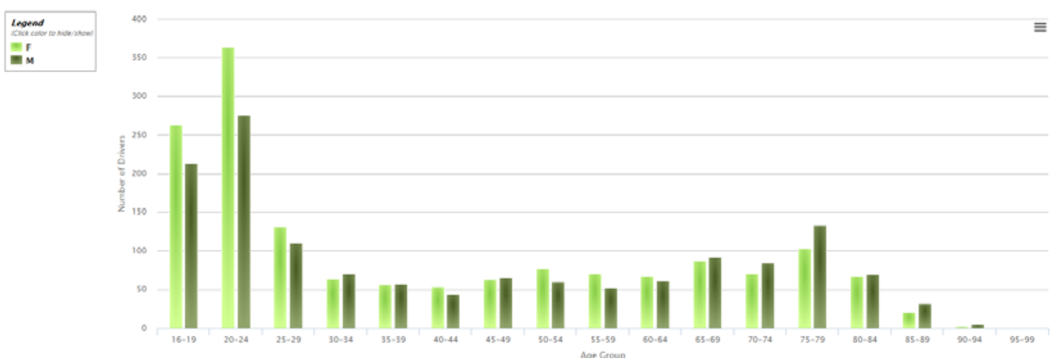


Figure 17
Number of participating drivers in the study categorized by age and gender *[2]*

A summary of the distribution of drivers sampled in the NDS study grouped by gender and age is provided in Table 1 and Table 2. The sample size consisted of 52% women to the remaining 48% of men. Driver ages were combined into unique groups ranging from 1-16. Table 3 defines the ages that make up each age group. As shown in Table 1 and Table 2 there was not an equal distribution of drivers per age group. The sample consisted of more drivers in age groups 1 and 2 than that of the remaining groups. While the Vehicles and Drivers categories contain useful background information on the overall study, the Trip Data and Events categories were most relevant to this research.

The Trip Data category contains summary measures describing trips, trip length, duration, start and stop time, summary statistics for speed and acceleration, trip summary record table and trip density maps. This section also details maximum deceleration and speed by vehicle classification, gender, age group, and data collection site. More specifically, the Trip Summary Table contains a plethora of point data, or data measured at one point in time.

Examples of this are the trip duration, maximum, minimum and mean speeds which are all contained within the Trip Summary Table. Time series trip data was also recorded throughout the NDS. However, time series data is not displayed on the website, only the variables on which time series data were collected are shown online. Researchers must contact SHRP 2 personnel in order to receive instruction on how to acquire this data. This action was completed in order to get data that was required to conduct this research.

Table 1
Summary of female drivers sampled in NDS

	Age Group	State						Total	% of Total
		FL	IN	NY	NC	PA	WA		
Female Drivers Sampled	1	57	22	56	44	15	67	261	17%
	2	98	31	95	48	19	74	365	23%
	3	28	8	36	19	12	28	131	8%
	4	14	3	19	10	3	15	64	4%
	5	17	6	13	9	3	8	56	4%
	6	10	7	10	12	3	11	53	3%
	7	10	7	15	10	7	14	63	4%
	8	15	9	20	11	10	12	77	5%
	9	14	5	13	11	13	14	70	5%
	10	14	7	21	9	5	11	67	4%
	11	21	7	18	16	6	18	86	6%
	12	14	6	23	7	7	12	69	4%
	13	20	7	31	13	6	26	103	7%
	14	14	6	16	12	3	16	67	4%
	15	3	3	1	3	1	10	21	1%
	16	1	0	0	0	0	1	2	0%
	Total	350	134	387	234	113	337	1555	100%

The Events category provides records of baseline drives, crashes, and near-crash event records by event type and severity. The Event Detail Table contains information that may or may not have contributed to a crash or near-crash event such as lighting, road grade, alignment, weather, and surface condition. A Post Crash Interview was conducted after an incident occurred. There, drivers detailed specific information regarding passengers in-vehicle, description of the crash itself and of surrounding conditions that may or may not have contributed to the collision.

Table 2
Summary of male drivers sampled in NDS

	Age Group	State						Total	% of Total
		FL	IN	NY	NC	PA	WA		
Male Drivers Sample	1	54	21	36	41	8	53	213	15%
	2	83	25	60	24	32	53	277	19%
	3	18	12	24	25	9	23	111	8%
	4	15	2	14	15	11	13	70	5%
	5	6	3	16	14	5	13	57	4%
	6	14	3	8	8	3	8	44	3%
	7	12	4	16	15	3	16	66	5%
	8	7	3	21	11	6	12	60	4%
	9	13	3	13	14	1	8	52	4%
	10	14	8	14	8	5	12	61	4%
	11	22	10	24	8	5	22	91	6%
	12	16	7	19	24	6	11	83	6%
	13	24	12	28	30	8	30	132	9%
	14	13	6	13	13	3	21	69	5%
	15	4	3	6	4	2	13	32	2%
	16	1	0	1	1	0	2	5	0%
	Total	316	122	313	255	107	310	1423	100%

Table 3
Description of age categories

Age	Age Group
16-19	1
20-24	2
24-29	3
30-34	4
35-39	5
40-44	6
45-49	7
50-54	8
55-59	9
60-64	10
65-69	11
70-74	12
75-79	13
80-84	14
85-89	15
90-94	16

Finally, the last section of the website database is the Query Builder. Here site users can select variables or conditions of interest to create a query. Results can display graph output and cross tabulations or a table of individual records. The complete list of variables available for all categories in the NDS dataset can be found in Appendix A.

Events Category Variable Options

Due to the nature of naturalistic data, video cameras, and video reductionists that manually review the film and draw conclusions, were used frequently to collect and categorize data. Therefore, it is important to describe how each variable in the Event category used in this study was explicitly defined in the NDS. A Crash was here defined as “any contact the subject vehicle has with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated” *[2]*. Any non-premeditated roadway departures where at least one tire left the travel surface are also categorized as a crash.

Near-crashes tend to be more ambiguous and require more attention before an accurate categorization can be made. A near-crash equals “any circumstance that requires a rapid evasive maneuvers by the subject vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash” *[2]*. Also, a near-crash meets the following criteria: not a crash, not pre-meditated, evasion required, and rapid evasive maneuver required.

Crash relevant was described as a situation “that requires an evasive maneuver on the part of the subject vehicle or any other vehicle, pedestrian, cyclist, or animal that is less urgent than a rapid evasive maneuver, but greater urgency than normally required to avoid a crash.” Non-conflict was defined as an incident that is within the bounds of “normal” driving behaviors and scenarios that is accurately represented by the time series data that created a flag. Non-subject conflict was referred as any incident that was captured on video that did not involve the subject driver.

Baseline drives were defined as those did not result in the pre-defined Crash, near-crash, Crash Relevant, Non-Conflict or Non-Subject Conflict and are represented of “regular” driving. Only data from baseline drives were used to create the prediction models described in this paper. This is because in order to analyze the effect of distraction on the driver, the researcher wanted to target drives both with and without a secondary task that did not result in any sort of crash or conflict.

Data Acquisition

The data used in this research was obtained through data user license agreements No. SHRP2-DUL-A-16-178 and SHRP2-DSA-15-62 from VTTI. The acquired NDS data included Event Detailed Tables, Participants Demographics, and Time Series Data for several performance measures. In addition to these categories, additional information was obtained to link the driver to their trip and event information. This linkage was important because it enabled comparisons of driver performance measures based on driver gender and age. A summary of the obtained data is shown in Table 4.

Table 4
NDS data used in this study

Data Category	Variable Used	Variable Definition	Variable Options Used
Events	Event ID	Identification number of Event	
	Event Severity 1	Describes outcome of event type	Baseline, Crash, Near-Crash
	Secondary Task 1	Driver engagement in any activity other than driving, observed on video by data reductionist	• No Secondary Tasks • Passenger in Adjacent Seat Interaction • Cell phone: Talking/Listening hand-held • Cell phone: Texting* • Cell phone: Dialing hand-held* • Dancing • Eating • Grooming • ….
Trip Time Series	GPS Speed	Vehicle speed from GPS	
	Longitudinal Acceleration	Vehicle acceleration in the X-axis direction versus time	
	Lateral Acceleration	Vehicle acceleration in the Y-axis direction versus time	
	Yaw Rate (Z Axis)	Vehicle angular velocity around the vertical axis	
	Throttle Position (Pedal Accelerator)	Position of the accelerator pedal collected from the vehicle network and normalized using manufacturer specifications	
Participants Demographics	Participant ID	Identification number of Driver	
	Participant State of Origin	State in which Driver resides	
	Participant Age Group	Age Range of Driver	
	Participant Gender	Sex of Driver	
	Household Income	Income Level of Driver's Household	

*These two variables were combined into one category in the analysis

The driving performance measures of GPS speed, lateral and longitudinal acceleration, throttle position and yaw rate, (reflected in variables used in Trip Time Series Category) were selected because literature revealed they were most frequently used in driver behavior research *[24]*. Figure 18 displays a graphical depiction of the coordinate system used to define the lateral and longitudinal directions as well as the yaw axis *[2]*.

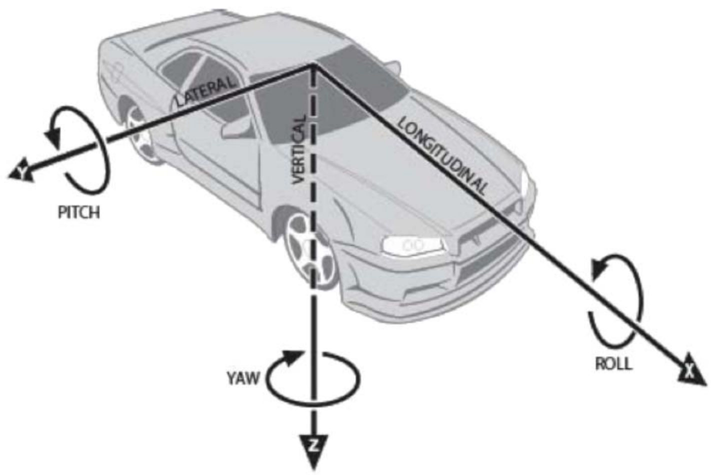


Figure 18
SAEJ760 Coordinate System used in data collection

The data categories displayed in Table 4 were described in the previous section. The Event Severity 1 variable described the outcome of the event, denoted as either Baseline, Crash, Near-crash, Crash Relevant, Non-Conflict or Non-Subject Conflict. There was also an Event Severity 2 designated, which was used when an additional event severity option described the corresponding event. However, only Event Severity 1 was used in this research. Secondary Task 1 described the observable driver engagement in one of many listed secondary tasks. There are also Secondary Task 2 and Secondary Task 3 variables defined that were used when the driver was engaged in two or three tasks respectively. However, only Secondary Task 1 was used in this study. Appendix B contains the entire listing of the available secondary tasks.

METHODOLOGY

The main focus of this exploratory study is to compile a technical summary of the limitations and capabilities of the SHRP 2 NDS data for an enhanced research on distracted driving that will provide valid statistical inferences to be applied to Louisiana drivers based on gender, age, and road facility type. More specifically, this research aims to thoroughly explore the SHRP 2 NDS database in order to (a) identify appropriate performance measures that can be used as surrogate measures of distraction, and (b) outline a methodology of developing a crash index.

The methodology to achieve the research objective included performing a comprehensive review of the NDS data to identify the appropriate sample that can potentially represent Louisiana drivers, reviewing the available performance variables in the SHRP 2 NDS, and conducting a statistical assessment on each variable's appropriateness as a surrogate measure to quantify distractions. For the surrogate measure selection, statistical analysis and artificial intelligence were utilized. For each type of modeling, the data had to go through several steps of data cleaning and reduction. Finally, researchers explored the NDS data to develop an outline for a crash index.

Creation of Appropriate Sample

Within the NDS dataset sample, drivers were extracted from the following six states: Florida, Indiana, New York, North Carolina, Pennsylvania, and Washington. The Louisiana Transportation Research Center (LTRC) took interest in the NDS dataset and its potential to be used in future research regarding Louisiana roads. In order for LTRC to use the NDS data for future research efforts that are of particular interest to their Louisiana constituents, it was important to select a sample from within the dataset that could be statistically representative of Louisiana drivers. In order to obtain this representative sample, information on Louisiana drivers was statistically compared to that of the six states in the NDS study using a Chi-square procedure.

The Chi-square method was developed in 1900 by Karl Pearson, and used as a goodness-of-fit test on non-normal distributions *[25]*. Chi-square tests if frequencies of an occurrence measured for a particular category are distributed as expected given only random chance influenced the outcome *[25]*. Therefore, the null hypothesis for a Chi-square test would be that the frequencies observed are statistically equal to the frequencies expected or those observed frequencies do not significantly diverge from what was expected. In performing, the Chi-square test it is often difficult to establish what is expected. In this application of the Chi-square, the expected frequencies were equal to selected Louisiana driver demographics.

The Federal Highway Administration's 2012 Highway Statistics data were sourced in order to extract the percentage of licensed drivers in the state of Louisiana as of January 2012 *[27]*. This data along with corresponding information in the NDS data was used in the Chi-square analysis.

Chi-square Procedure

In order to prepare the NDS data for Chi-square analysis, the first step was to record the percent of drivers studied in the NDS broken down by state origin, age group and gender of the driver. This was also done using the Louisiana driver data. Driver ages were divided into 15 age groups using the same age groups defined in the NDS as shown in Table 1,Table 2, and Table 3. It should be noted that there was a discrepancy in the age labeling between the NDS data and the FHWA Highway statistics on Louisiana drivers. Louisiana elderly drivers were simply categorized as aged 85 and over, while in the NDS they provided a more detailed breakdown of the elderly (ages 85-89 and 90-94). To account for this difference in the analysis, all NDS drivers aged 85 or older were combined into one category (Category 20). Gender was coded dichotomously, where the value 1 represents males and 2 represents females. A new variable titled "delta frequency" was created to aid in the analysis. Delta frequency equaled the absolute value of the difference in percentage between licensed Louisiana drivers and drivers in each of the states represented in the NDS. Table 5 displays an example of the organized data used in the analysis, all frequency data represents the percentage for each category. Here delta frequency equaled the absolute value of the difference between percentage of Louisiana drivers and percentage of Florida drivers.

Table 5
Data used for Chi-Square test of Louisiana drivers vs. Florida drivers

Age Group	Gender	LA Frequency	FL Frequency	Delta Frequency
1	1	2.55	8.12	5.57
1	2	2.45	8.56	6.11
2	1	4.41	12.46	8.05
2	2	4.59	14.71	10.12
3	1	4.32	2.7	1.62
3	2	4.68	4.2	0.48
4	1	4.32	2.25	2.07
4	2	4.68	2.1	2.58
5	1	3.84	0.91	2.93
5	2	4.16	2.55	1.61
6	1	3.84	2.1	1.74
6	2	4.16	1.5	2.66
7	1	4.32	1.8	2.52
7	2	4.68	1.5	3.18
8	1	4.8	1.05	3.75
8	2	5.2	2.25	2.95
9	1	4.32	1.95	2.37
9	2	4.68	2.1	2.58
10	1	3.84	2.1	1.74
10	2	4.16	2.1	2.06
11	1	2.88	3.3	0.42
11	2	3.12	3.15	0.03
12	1	1.88	2.4	0.52
12	2	2.12	2.1	0.02
13	1	1.38	3.6	2.22
13	2	1.62	3.04	1.42
14	1	0.9	1.95	1.05
14	2	1.1	2.1	1
20	1	0.45	0.75	0.3
20	2	0.55	0.6	0.05

SAS Enterprise Guide 6.1 software was employed to run the Chi-square test for all delta frequency values, representing the difference in percentage of drivers in Louisiana against all 6 states individually. For each test the null hypothesis equaled cell values are identical and equal to 0 (% of drivers in Louisiana - % of drivers in state examined = 0). The alternative hypothesis equaled cell values are not identical and not equal to zero. Table 6 displays the results of each Chi-square test.

Table 6
Results of each Chi-square Test

State	Chi-square Value	P-value
LA vs. FL	2.149	0.9999
LA vs. IN	4.5377	0.9913
LA vs. NC	7.7674	0.9011
LA vs. NY	2.0004	0.9999
LA vs. PA	11.8521	0.6182
LA vs. WA	2.2588	0.9998

For the purpose of this test, a higher p-value was desired in order to fail to reject the null hypothesis. That would mean it could not be stated with statistical certainty that the drivers in each state used in the NDS and the Louisiana drivers were not identical. A higher p-value provides a corresponding small Chi-square value. Therefore, a smaller Chi-square value was also desirable because as the Chi-square value decreases, the drivers would become more similar. As shown in Table 6, New York and Florida had the largest p-values (0.9999) and their Chi-square values were also very close with values of 2.0004 and 2.149, respectively. Since the Chi Square values were only minimally different, another criterion, the geographical factor, was added into the test in order to finalize which data would be selected as the appropriate representative sample. Since Florida and Louisiana are closer geographically, Florida was chosen as the sample that would be most representative of Louisiana drivers. A more inconspicuous factor that contributed to Florida's selection is the logic that a state like New York has a different social fabric, where driving characteristics are innately different than that of southern states such as Louisiana or Florida. Due to those reasons, Florida data was selected as the representative sample.

Data Reduction and Preparation (Statistical Analysis)

To identify the surrogate measures of distracted driving, the available performance variables in the SHRP 2 NDS were reviewed and a statistical assessment on each variable's appropriateness as a surrogate measure to quantify distractions was conducted. Before doing so, the data was grouped, edited, and reduced to make the statistical analysis process easier.

Group Division, Data Aggregation and Editing

In order to perform the desired statistical analysis, the data were divided into groups based on the secondary tasks in which the drivers were engaged. After grouping, the data were aggregated and edited as further preparation for the eventual statistical analysis.

Group Division. In this research, the NDS time series data was divided into four groups: Group 0, Group 1, Group 2, and Group 3. The secondary tasks that were analyzed in this research were: No Secondary Task, Passenger in Adjacent Seat Interaction, Cell phone: Talking/Listening hand-held, Cell phone: Texting, and Cell phone: Dialing hand-held. From these five tasks, four groups were created for analysis. The control group (designated as Group 0) contained event data when the driver was engaged in no secondary task. Group 1 consisted of event data for Cell phone: Talking/Listening hand-held. Group 2 combined the data for Cell phone: Dialing hand-held and Cell phone: Texting. These two tasks were combined into one group because these tasks are very similar in nature and putting them together allowed for a larger sample size in Group 2. Finally, Group 3 contained event data for Passenger in Adjacent Seat Interaction.

Data Reduction and Cleaning. Proper data editing before applying data as input into analyses can aid in the assurance that the results obtained are accurate. The data editing process included checking the time series data entries for the selected five performance measures to ensure their values were within an acceptable range and logically reasonable as well as identifying outliers or missing data. Since the used data were time series, the first step taken in the data editing process dealt with aggregating the time intervals.

Data on the five time-series variables were collected over a 20-second time interval for each driver. Within the twenty-second time interval, the data were broken down into 0.1-second intervals. For example, the data for the GPS speed variable were represented by 200 data points displayed in 0.1-seconds increments to account for the twenty seconds of data collected. In order to reduce the data size, the time series data were aggregated into 1-second increments instead of the original interval of 0.1 seconds, using the time series procedure in SAS statistical software. The 200 data points for the time series variables were averaged to the point where it became organized into 20 data points representing each of the 20 seconds worth of data recorded. The code for the procedure used is displayed in Figure 19.

```
proc timeseries data=baseline_gps_speed
            out=baseline_gps_speed_timeseries;
         id time interval=seconds accumulate=average;
         by event_id;
   var value;
run;
```

Figure 19
Example SAS code used to aggregate time series data

After the data was aggregated into 1-second intervals, the next step in the data editing process was to ensure the values were within an acceptable range. The upper and lower data ranges of each time series variable were defined in the Trip Data category on the InSight webpage. Other useful information on each variable was displayed as well such as variable units, accuracy and sign convention as seen in Figure 20.

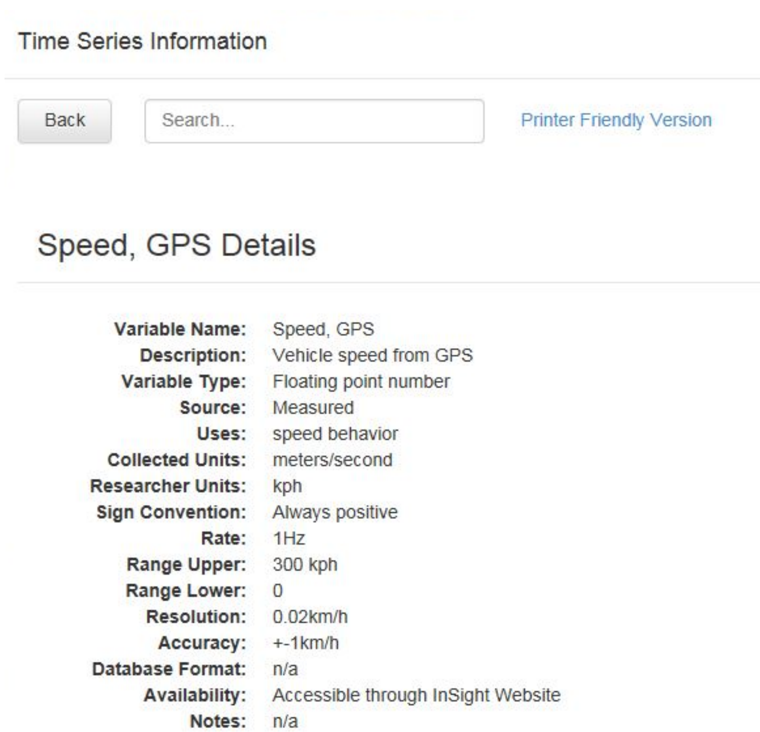


Figure 20
GPS speed details displayed on InSight webpage

All values outside of the predefined range limits were removed from the dataset for each of the five time-series variables studied. Next, any entry that contained missing information was also removed. Potential outliers were inspected using the distribution analysis task in SAS Enterprise Guide statistical software and removed once identified. A summary of the amount and type of data removed can be found in Appendix C.

Test of Normality. The next phase of data analysis involved conducting tests for normality on each of the performance measures. The result affected the statistical analysis to identify the distracted driving surrogate measures. The Kolmogorov-Smirnov test for normality was used because it is recommended when data entries exceed 2,000 and each variable of interest fits this criterion. For a level of significance value set at 0.05, all of the tests resulted in statistically significant outcomes. Therefore, under the null hypothesis that the data was distributed normally, this hypothesis was rejected in each test. The p-values were identical regardless of the variable type and almost all of the normality tests resulted in a p-value equal to <0.01. Only the Group 2 tests resulted in a different p-value (<0.0001). Due to these findings, it could be concluded that the datasets contained non-normal distributions.

Independence of Groups

In the data grouping and editing phase, multiple secondary tasks were found to be associated with each participant. In order to perform many statistical tests, a major assumption is that each group of data is independent from the other. This would mean that the same participant could not be placed in more than one group. For example, even though a driver may have been texting (Group 2 task) and interacting with a passenger in the adjacent seat (Group 3 task) during one of their trips, the associated performance measures for both of the groups cannot be used simultaneously; the driver must either be placed in Group 2 or Group 3 exclusively. This issue had to be addressed before proceeding to identify the distracted driving surrogate measures and ensure group independence.

A large amount of effort was required to ensure each group was created independently of the other. Due to the amount of data, it was imperative that each group of data be sorted first by Participant ID and then group Number to determine which participants originally had data that would place the participant in two, three or even all four of the groups. Of the 2183 unique Event IDs used in this research, there were 746 instances where a participant belonged to more than one group.

After the participants belonging to multiple groups were extracted, a method of randomization was enacted to help ensure group independence. In Excel, each field was given a random ID according to the Random ID function embed in the Excel program. Then, those random IDs assigned were randomly ordered in the spreadsheet. From here, each field associated with the random IDs were numbered in order (#1-746). The spreadsheet was again sorted first by Participant ID and then by group Number.

Group Assignment Rules. Groups were assigned after following two steps of rules which are reflected in Table 7. The first step involved only the groups where drivers engaged in some secondary task (Group 1, 2, or 3). In this step of group assignment, participants were randomly assigned if previously in Group 1 and 2 or in Group 1 and 3. However, Group 2, where the driver was engaged in texting/dialing a hand-held cell phone, had a significantly lower amount of participants than that of those in Group 3. Therefore, stratification was utilized in the special case when a driver was in Group 2 and 3. If that driver was in both Group 2 and 3, they were automatically kept in Group 2 in order to increase the sample size within Group 2.

The second step of group assignment included only participants who had instances where they engaged in Group 0. In this case, if a driver was in Group 0 in addition to the new group they were just assigned according to step one, then that driver automatically stayed in their group assigned in step one. This rule was created because there was already a large sample size for participants that were in Group 0 exclusively before any group assignment was necessary. A summary of the final sample sizes for each group after the group assignment process is displayed in Table 8. An important note, the "# of Samples before group Assignment" column represents the number of samples per group right before the rules were enacted. These values are much lower than that of the original dataset because group assignment was performed after the data editing procedures of removing outliers and data values outside of the acceptable range were conducted.

The five-time series attributes were utilized for the calling (Talking/Listening), texting (Texting/Dialing), and adjacent passenger interaction events. Each event contains time series records for the five performance attributes over a period of nearly one minute with a resolution of 0.1 seconds. Each event also included the starting and ending times of each secondary task that lasted around 6 seconds.

To prepare the time series data for the ANN model, each observation was coded as "1" if associated with a secondary task (i.e., from the beginning to the end of the secondary task) and "0" otherwise (i.e., before the beginning or after the end of the secondary task). By the end of the preparation phase, each secondary task had the total number of around 10,000 time-series coded records.

Table 7
Rules Created for Group Assignment

Step	Scenario	Rule
1	Participant in Group 1 & 2	If # assigned to Event ID where Group=1 is odd, keep that participant in Group 1 and eliminate all other group information for that participant
		If # assigned to the Event ID where Group=1 is even, keep that participant in Group 2 and eliminate all other group information for that participant
	Participant in Group 1 & 3	If # assigned to Event ID where Group=1 is odd, keep that participant in Group 1 and eliminate all other group information for that participant
		If # assigned to the Event ID where Group=1 is even, keep that participant in Group 3 and eliminate all other group information for that participant
	Participant in Group 2 & 3	Automatically keep that participant in Group 2
	Participant in Group 1, 2, & 3	If # assigned to Event ID where Group=2 is odd, automatically eliminate all group information associated with Group 3, then refer to rules for when Participant is in Group 1 & 2
		If # assigned to the Event ID where Group=2 is even, automatically eliminate all group information associated with Group 1, then refer to rules for when Participant is in Group 2 & 3
2	Participant in Group 0 & any other Group	Automatically eliminate all group data associated with Group 0 and keep Participant in the already assigned Group
All data remaining after both Steps is automatically left for Group 0		

Table 8
Final Sample Size Count by Group

Group	# of Samples Before Group Assignment	# of Final Sample Size
0	1501	1127
1	162	102
2	78	69
3	442	299

Data Reduction and Preparation (Artificial Intelligence)

Data Cleaning and Reduction

The next step was to clean and mine the data for the required information to analyze the pattern in the driving behavior associated with each secondary task. As shown in

Figure 21, several steps are followed to clean and mine the data. The acquired data had several missing observations, causing a problem in pattern recognition which is the main step in developing the ANN model. Thus, overcoming the missing data problem was the first procedure in this data cleaning and mining stage.

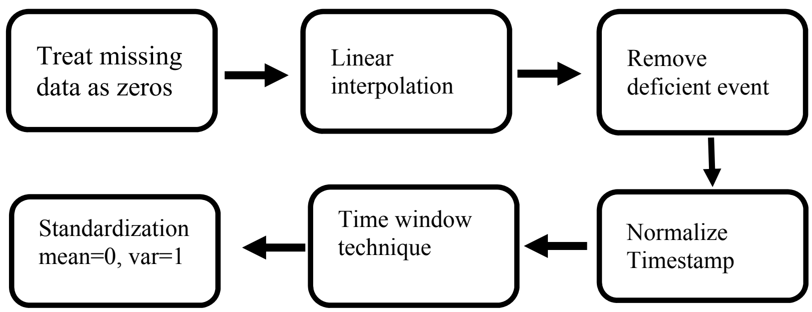


Figure 21
Data cleaning and mining

Each missing record was replaced with an interpolated value based on the preceding and following observations. In order to differentiate between a missing-value "zero" and an actual zero-observation in the data, interpolation was only conducted in three cases: (a) the speed is less than 6.5 mph; (b) pedal position is less than 5%; (c) lateral acceleration, longitudinal acceleration, or yaw rate are transitioning from negative to positive values or vice versa. It was noted that some performance attributes did not have any values throughout some events (e.g., longitudinal acceleration values in Event ID 1200034 were all zeros); these events were considered deficient and were removed completely from the analysis. The next step was to normalize the events based on the associated timestamp. Each event has different length of time duration around one minute. In order to maintain the same duration, a total of 59 seconds interval that was the duration that all events satisfy, was set for all events.

Distracted driving is a continuous action a driver takes over a specific time period. Having said that, to understand the pattern in driver behavior while performing a secondary task or the primary task of driving, it is important to analyze the data over time and not as independent observations. Thus, a moving time window technique is used to recognize the driving pattern along each time-series performance attribute. The moving time window technique is an approach used to capture changes in an entire set of features over time. The time window size is defined as the number of time steps over which the change in driving behavior is to be detected. In this study, a time window of one second (10 time steps) is used. The small time step of one second was selected to detect any minimal change in the driving pattern. For each time window, the standard deviation, as an attribute of the changes in driving patterns, was calculated for the observations within the time window for every performance attribute. The time window technique operates by moving the time window one-time step (0.1 second) at a time, calculating the standard deviation for the observations within the time window. For the likelihood coding (0/1), the average values within each time window were obtained. At the end, a new dataset of standard deviations for the five performance attributes along with average likelihoods was obtained with a resolution of “1” second.

When drivers are distracted, the selected five performance attributes change significantly according to the literature. However, this change is not the same across all performance attributes. Some attributes may show more significant changes compared to the others. This may confuse the ANN model while analyzing the pattern of each attribute. In other words, the ANN may treat some attributes as more important than the others, which could result in a biased detection accuracy. Therefore, the resulting dataset of standard deviations from the previous step was standardized to have a mean value of “0” and a variance of “1” for each performance attribute. This was done based on the formula in equation (1). The resulting dataset from this step will allow the ANN model to treat all performance attributes equally.

$$x_{new} = \frac{x - \mu}{\sigma} \tag{1}$$

Where, x_{new} is the standardized value, x is the original variable value, μ is the mean of the standard deviations across each performance attribute, and σ is the standard deviation of each performance-attribute’s standard deviations. With this step, the time-series data for the different performance attributes along with the likelihoods of performing secondary tasks were ready to train the ANN models.

SELECTION OF DISTRACTED DRIVING SURROGATE MEASURES

The time series variables were analyzed to identify the surrogate measures that can be used to detect distracted driving behavior. To achieve this, a statistical model was developed to detect distracted driving behavior using the five time-series performance measures: GPS Speed, Lateral Acceleration, Longitudinal Acceleration, Throttle Position, and Yaw Rate. In other words, the five driving performance measures were used to detect the secondary task in which a driver was engaged. Selecting the appropriate statistical method was an important task that had to be completed in order to achieve the detection goal. In the analysis process, the statistical model will identify the performance measures that can be used to detect the type of secondary task a driver was engaged with an acceptable accuracy.

Detection Model Selection

Discriminant analysis and logistic regression were two statistical methods considered for use in the development of the detection model. Discriminant analysis can be used to classify an observation into one of several populations, while logistic regression relates qualitative variables to other variables through a logistic cdf function *[26]*. Both methods have the ability of accomplishing a similar goal, but depending on the normality of the data, one method is generally recommended over the other. For data of non-normal distribution, logistic regression is recommended because of its use of Maximum Likelihood Estimators (MLE). Although discriminant analysis and logistic regression will likely yield similar results in most cases, MLE used in the latter method were proven to outperform classical linear discriminant analysis under non-normal data conditions *[26]*. As discussed in the previous chapter, according to the results of the tests for normality, all data used in this research were deemed non-normally distributed. Therefore, logistic regression was chosen over discriminant analysis as the tool used to develop the prediction model.

Multiple Logistic Regression Analysis

Logistic regression is frequently used in research to predict the probability that a particular outcome will occur. The outcome can either be a continuous-level variable or a dichotomous (binary) variable *[25]*. However, the outcomes are usually classified in a binary nature in Logistic regression. In this case, the dependent variable is dichotomous and is coded as “1” if the event did occur and “0” if the event did not occur. During the analysis, the logistic function estimates the probability that the specified event will occur as a function of unit change in the independent variable(s) *[27]*. The logistic function used to calculate the expected probability that Y=1 for a given value is shown in equation (2).

$$\hat{p} = \frac{\exp(B_0 + B_1 X)}{1 + \exp(B_0 + B_1 X)} = \frac{e^{B_0 + B_1 X}}{1 + e^{B_0 + B_1 X}} \qquad (2)$$

In the literature, logistic regression has been described as "conceptually analogous" to linear regression. This similarity is because a single dependent variable is predicted from either a single predictor (simple logistic regression) or multiple predictors (multiple logistic regression) *[25]*. In the logistic function displayed in equation 2, $B_0 + B_1X$ is directly pulled from the equation for the regression line *[28]*. The intent of the analysis was to use all five independent variables (GPS speed, lateral acceleration, longitudinal acceleration, throttle position and yaw rate) to detect whether the driver was or was not engaged in a secondary task. Since five independent variables were considered, multiple logistic regression (MLR) was used instead of simple logistic regression.

Three separate MLR tests were completed to compare the overall statistical output between the control and the three individual cellphone bases secondary tasks of concern. The control group for each test was equal to the NDS Florida driver "events" where the driver was not engaged in any secondary task. As stated earlier, only NDS events with an event severity defined as "Baseline" were used in the analysis. This is because the surrogate measures identification was focused exclusively on driver behavior and not crash risk, and the baseline event severity described drives that did not result in a crash or near-crash scenario. Table 9 describes each of the tests.

Table 9
Description of Multiple Logistic Regression Tests

MLR Test	Description
Control vs. Group 1	Engaged in No Secondary Task vs. Talking/Listening on Cell Phone (hand-held)
Control vs. Group 2	Engaged in No Secondary Task vs. Texting/Dialing on Cell Phone (hand-held)
Control vs. Group 3	Engaged in No Secondary Task vs. Adjacent Passenger Interaction

In order to accurately interpret the results of the MLR, the binary predictor variables used must be coded in a very specific manner. Typically, in MLR, the group that is to be used as the focal or reference group is coded as "0" and the other outcome is coded as "1". The focal group in each of the comparison tests was the individual secondary task in which the driver was engaged. Therefore, for each comparison test the variable that described No Secondary Task was coded as "1" and the specified secondary task was coded "0".

Neural Network Modeling

Since driving behavior detection is a nonlinear pattern recognition problem, artificial intelligence, and more specifically neural network modeling, was also deployed for secondary task detection and hence distracted driving surrogate measures selection. Artificial neural network is a modeling approach inspired by how the human brain works. It is an adaptive technique that has been used in several detection and pattern recognition studies. This modeling technique has been recognized for its ability to detect patterns in datasets and find the best non-linear function to fit these data. Each ANN model is defined based on its topology and updating rules. The topology indicates the arrangement of neurons and the way they are connected, while the updating rules are chosen based on ANN architecture and the data that are dealt with.

In this study a supervised feed-forward network with backward propagation (FFBP) was used to develop the detection models. FFBP architecture is well-known for its ability in solving pattern recognition problems. A sigmoid function was used as an internal transfer function. Three hidden layers were selected because of the large size of the data (10,000 observations for each secondary tasks) such that a reasonable number of neurons can be selected in each layer. The Levenberg-Marquardt algorithm was selected for the performance (optimization) function. The model output defines whether a secondary task was associated with the driving behavior or not. Therefore, a binary outcome of 0 or 1 was used, where 1 indicates association with a secondary task and 0 otherwise.

Three ANN models were developed individually for each type of secondary task. The input layer included five neurons to represent the selected five driving performance attributes. After a preliminary analysis to improve the model accuracy, the number of neurons in the hidden layers was selected as 9, 13, and 7 for the first, second, and third hidden layer, respectively. To develop the three models, the dataset for each secondary task was randomly divided into 70% for training, 15% for validation, and 15% for testing. The ANN model structure is depicted in Figure 22.

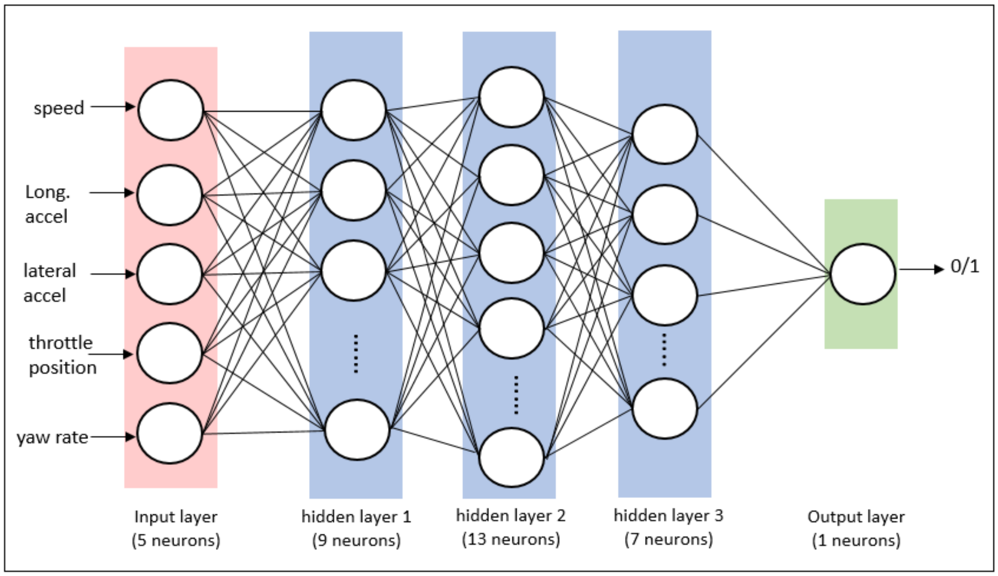


Figure 22
ANN training model structure for cellphone calling

CRASH INDEX OUTLINE

The literature shows the significance of driver's engagement in secondary tasks in determining roadway safety. Past research also shows that socioeconomic attributes have a significant impact on the crash risk and the likelihood of performing secondary tasks. To the research team's knowledge, there is no previous published work on the quantification of the effect of the socioeconomic characteristics on drivers' involvement in secondary tasks and crash risk. Thus, this research uses the SHRP 2 NDS data to develop an approach (grading system) to quantify the crash risk associated with the driving behavior based on drivers' socioeconomic characteristics. Based on this grading system, a Crash Index measure is developed.

Crash Index Development

The proposed Crash Index, also called a Crash Risk Index, could be developed in three main steps: (a) extract data for all secondary tasks and socioeconomic attributes from the NDS database; (b) select the records with high crash risk and significant effect on drivers' involvement in secondary tasks; and (c) develop a grading system and a Crash Risk Index for quantification of crash risk.

Extracting Secondary Tasks and Socioeconomic Attributes

The NDS data provides detailed demographic and history questionnaires for each participant of the study. The demographic questionnaires provide information on the participant's personal background, while the history questionnaires provide information on their driving record. The questionnaires can be used to extract several socioeconomic attributes including Age, Gender, Marital Status, Work Status, Average Annual Miles Travelled, Years (has been) Driving, Annual Household Income, Education Level, Vehicle Classification, and the State (participants' location).

Selection of Secondary Tasks and Socioeconomic Attributes

Selection of Secondary Tasks. The Crash Risk Index quantifies the crash risk associated with drivers' socioeconomic characteristics based on their tendency to perform secondary tasks. As such, only secondary tasks with significantly high crash risk should be first identified. Through statistical analysis, the likelihood of crash when a specific secondary task is performed is determined. Based on that, secondary tasks with higher crash likelihood are identified as high crash risk secondary tasks.

Selection of Socioeconomic Attributes. For each secondary task with high crash risk, the socioeconomic attributes with significant association to drivers' involvement in that task should be determined. Statistical association relationships can be determined using statistical analysis techniques such as MLR. Socioeconomic attributes that significantly increase the likelihood of drivers' engagement in high crash risk secondary tasks should then be selected.

Grading System and Crash Risk Index

This section describes a proposed grading system to quantify the crash risk associated with drivers based on their socioeconomic characteristics. This could be accomplished in three consecutive steps performed at three different levels as shown in Figure 23. First, the crash risk associated with each secondary task should be quantified. Then, the effect of the different socioeconomic attributes on the likelihood of performing that secondary task should be quantified. Finally, the effect of each socioeconomic attribute should be broken down by the different categories within that attribute. This is to determine the relative effect of each category for a socioeconomic attribute on the likelihood of performing the secondary task in step 1. The effects in the three steps can be labeled as (a) Crash Risk Coefficient for the first step, (b) Significance Level Coefficient for the second step, and (c) Category Contribution Coefficient for the third step.

The Crash Risk coefficient (R_i) is defined based on the odds ratios for each task (i), the Significance Level Coefficient (a_{ij}) quantifies the effect a socioeconomic attribute (j) on the likelihood of drivers' involvement in secondary task (i), and the Category Contribution coefficient (b_{ijk}) will be measured for the category (k) within each Significant Attribute (j). These three coefficients could represent the grading system for the crash risk associated with performing the different secondary tasks based on drivers' socioeconomic characteristics. The three coefficients can be used to quantify the crash risk associated with drivers' tendency to conduct secondary tasks based on their socioeconomic characteristics. This crash risk can be measured as a Crash Risk Index (CRI) that is calculated using equation (3).

$$\boldsymbol{CRI_i = R_i * \sum_{j \in s(i)} (a_{ij} * b_{ijk})} \qquad (3)$$

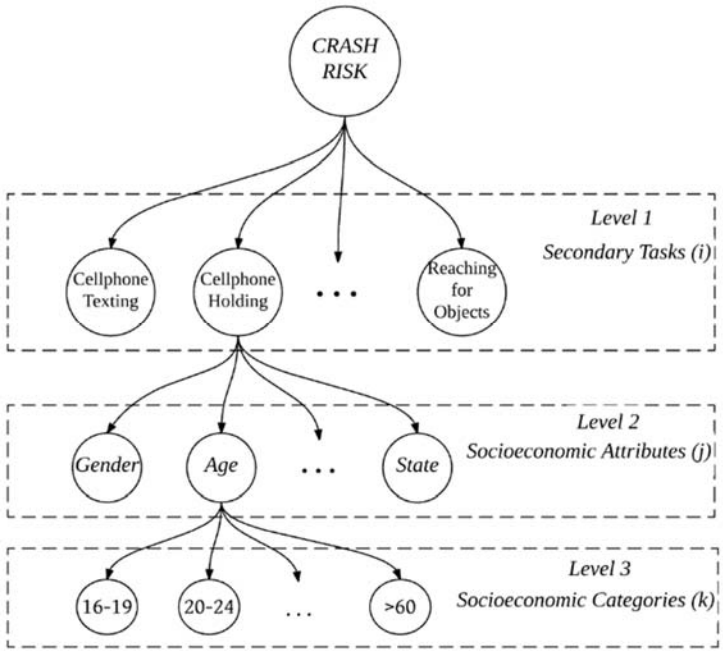


Figure 23
Crash risk quantification tree

Where, *s(i)* is the set of attribute (*j*) with significant influence on the likelihood of performing secondary task (i) and CRI_i is the CRI value associated with the high-crash-risk secondary task (*i*). CRI is calculated for each secondary task individually. To calculate the overall CRI value associated with all high-crash-risk secondary tasks, equation (4) is used.

$$CRI_o = \sum CRI_i \tag{4}$$

DISCUSSION AND RESULTS

Selection of Distracted Driving Surrogate Measures

SAS Enterprise Guide 6.1 was called upon again to run the MLR analysis. In SAS Enterprise Guide, the output generated in the MLR consisted of a Chi-square value and the corresponding p-value that described how well the model performed in predicting the outcome of the event or focal group. The Chi-square procedure used to test the statistical significance of the logistic regression model is similar to the analysis of variance procedure that is used linear regression. In logistic regression, a true R-square value cannot be computed but SAS has the ability to estimate "pseudo" R-square values *[25]*. These values are interpreted the same way as the actual R-square from linear regression. Logistic regression also estimates the amount of dependent variable variance accounted for by the model.

The Hosmer and Lemeshow test is used in MLR to assess whether the predicted probabilities match the observed probabilities using a Chi-square statistic. If the p-values for this test are significant, this means the model predictions are not in accordance with those observed. If the converse is true, this is an indication that the model predictions and actual observations are about the same and the model provided a good fit of the data.

All of the output is useful; however, the outputs of most interest in the MLR are the p-values for the Hosmer and Lemeshow Test, the estimates for the Analysis of MLE, and the Odds Ratio Estimate. The estimates computed in the Maximum Likelihood Estimation (MLE) are the coefficients used to create a regression line. The regression lines for each of the three tests represent the prediction model that can help to identify the surrogate measures of distracted driving. The odds ratio is associated with each predictor describes the odds of a case being coded as "0" on the dependent variable. It indicates the amount of change expected in the odds when there is a 1-unit change in the predictor variable.

Results of Overall MLR Tests

In each of the three tests against the control, the R-square values were very low, all of which below 0.02. This signifies very little of the dependent variable variance that could be accounted for by the models. Estimates generated for the analysis of MLE have a p-value for each estimate respectively. These p-values reveal whether the MLE estimates are of statistical significance, however, since the R-square values revealed the models were weak to begin, the p-values provided little further insight.

The Control vs. Group 1 test compared no secondary tasks to drivers engaged in talking/listening on a hand-held cell phone. The results indicated the model predictions do not statistically match the observed data and therefore the model did not fit the data well. The p-value for the Hosmer and Lemeshow Goodness-of-Fit Test equaled <0.0001 and was used to draw this conclusion. The odds ratios for three out of the five variables were around 1.0; values of 1.008, 1.003, and 0.978 for GPS speed, throttle position and yaw rate respectively. In the case of the variable GPS speed, this means for a one-unit increase in driver's GPS speed, the odds of that driver being classified in Group 1 are increased by 1.008 times. Since the odds ratio for GPS speed, throttle position and yaw rate are all so close to 1.0, one can conclude there is no accuracy in the prediction for these values. The acceleration variables had odds ratios that were not as close to 1, as longitudinal acceleration's odds ratio was equal to 0.393. More interestingly, the odds ratio for lateral acceleration was 17.013. Thus, the drivers lateral acceleration seemed to have a major impact on whether they would be classified in Group 1 (talking/listening on the phone). The regression line formed using the MLE estimates is shown in equation (5). This equation detects whether a driver was categorized in talking/listening on a hand-held cell phone, where x_1=GPS speed, x_2= lateral acceleration, x_3= longitudinal acceleration, x_4= throttle position and x_5=yaw rate.

$$\boldsymbol{y = -3.2321 + 0.00774x_1 + 2.834x_2 - 0.933x_3 + 0.00329x_4 - 0.0218x_5} \tag{5}$$

The Control vs. Group 2 test compared no secondary tasks to texting/dialing on a hand-held cell phone. This test produced slightly different results from the previous test but followed a similar trend. The Hosmer and Lemeshow p-value result was again significant at 0.0003. This signifies that the model was not a good fit for this data, as a significant value for the Hosmer and Lemeshow test means that the predictions made in the model are different from those observed. The GPS speed, throttle position and yaw rate odds ratios had values of 1.0, 1.021, and 0.975 respectively. This again meant that the model had little accuracy in regards to prediction using these variables. Longitudinal acceleration's odds ratio was similar to the previous test as well at 0.051, and lateral acceleration's odds ratio was quite high again 21.556. Therefore, with a one-unit increase in the lateral acceleration, the driver was 21 times more likely to be categorized as texting/dialing on his or her cell phone. The regression line for this test is reflected in equation (6) where the variables are the same as described previously.

$$\boldsymbol{y = -3.1999 + 0.00037x_1 + 3.0711x_2 - 2.978x_3 + 0.0208x_4 - 0.0254x_5} \tag{6}$$

The Control vs. Group 3 test analyzed no secondary tasks to some form of driver interaction with an adjacent passenger. The Hosmer and Lemeshow test p-value was again <0.0001.

The GPS speed, throttle position and yaw rate odds ratios were all around 1.0 as well and longitudinal acceleration was similar to the previous tests at 0.227. However, the lateral acceleration odds ratio trended differently this time around at a value of 0.004, significantly lower than the Control vs. Group 1 and Group 2 tests. In summary, there is no accuracy in the detection of the different types of secondary tasks using the GPS speed, throttle position and yaw rate. Only slightly better detection power can be achieved with the longitudinal and lateral accelerations. The last regression equation is defined in equation (7).

$$\boldsymbol{y = -1.6721 + 0.00206x_1 - 5.4096x_2 - 1.4833x_3 + 0.0113x_4 - 0.0647x_5} \quad (7)$$

Figure 24 displays a summary of the odds ratio values calculated for each test. The horizontal axis distinguishes each test where 1 = Test 1 (Control vs. Group 1) and so forth. It is clear from the figure that the lateral acceleration is the only variable that can be identified as a surrogate measure for distracted driving behavior resulting from secondary tasks in Group 1 and Group 2. A 1-unit increase in drivers' lateral acceleration increased their chances of being categorized in both Group 1 and Group 2. In practical terms, if a driver increased his/her lateral acceleration slightly, that driver was 17 times more likely to be talking/listening on the phone and 21 times more likely to be texting. Very little other information could be ascertained from the three overall MLR tests. Also, although the lateral acceleration variable proved to have predictive power within the tests against Groups 1 and 2, the overall test itself for all three groups was proven very weak and unreliable. Thus, the claims just made regarding the lateral acceleration variable should not be taken as actual fact due to the fact that the data was not a good fit for the models it developed.

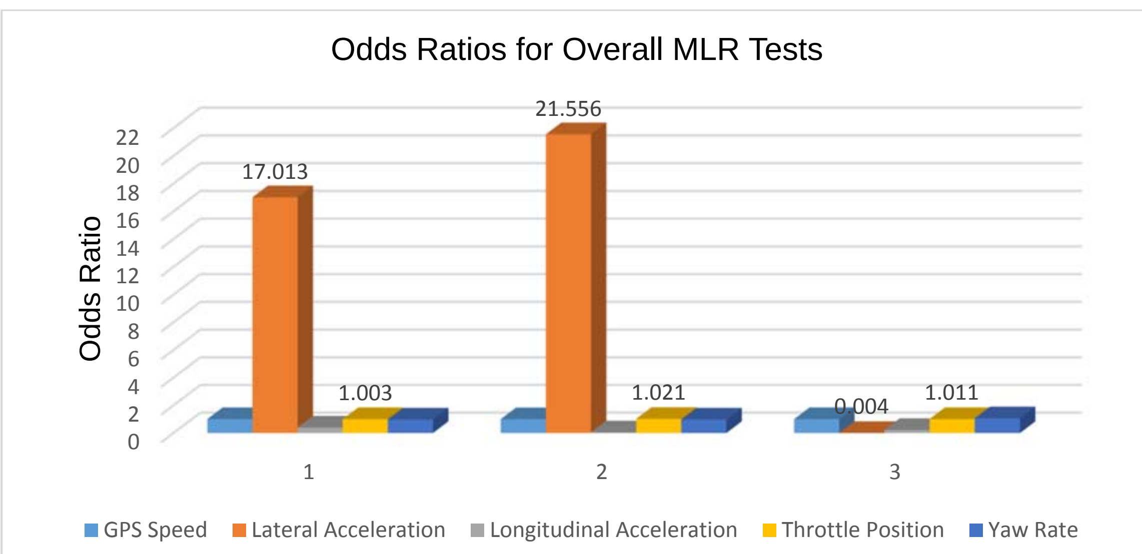


Figure 24
Summary of Odds Ratio results for overall multiple logistic

Regression Tests. Since all three models were proven to have a weak predictive power, there was no need to validate the equations using regression analysis. It is recommended that further study be conducted to produce stronger models and validation be tested on those. Although the three initial multiple logistic regression tests had extremely low pseudo-R values and Hosmer and Lemeshow p-values, more tests were conducted in order to discover if any trends were apparent based on the drivers' age and gender. The data was partitioned by driver age and driver gender in order to run the additional MLR tests.

Results of MLR based on Driver Age. Instead of using the original 16 driver age groups, five groups were created to provide a larger sample size within each group for the multiple logistic regression test: ages 16-29, 30-49, 50-69, 70-89, and 90 and older. A MLR test was run for the Control group vs. Group 1, Control group vs. Group 2, and Control group vs. Group 3 for each of the five age groups excluding the oldest drivers in the 90 and older group as there were no instances of drivers engaged in any secondary tasks within this age group. This was also true for the Group 2 task (texting/dialing) with regard to age groups 50-69 and 70-89. Cases where the MLR test could not be run due to no instances of the appropriate secondary task have been left blank in the summary tables to reflect this fact.

The pseudo-R-square and Hosmer and Lemeshow p-values resulting from these tests were similar to those found in the original round of MLR tests. All R-square were less than 0.3 which means very little of the variance could be accounted for by the dependent variables. The p-values for the Hosmer and Lemeshow test were significant for most of the tests meaning that the predicted estimates did not statistically match those observed. However, for drivers' ages 30-49, the Hosmer and Lemeshow p-values were not significant for Group 2 and Group 3 tests at 0.0889 and 0.0759 respectively. Due to this result, the odds ratios for these age groups within the Group 2 and Group 3 tests can be deemed more credible. However, the p-values were still very close to the alpha specified (0.05), therefore these values barely made the cut of being considered not statistically significant. The Hosmer and Lemeshow p-values for each of the tests previously described are displayed in Table 10.

Table 10
Hosmer and Lemeshow p-values for MLR tests partitioned by age

Driver Age	Control vs. Group 1 Test	Control vs. Group 2 Test	Control vs. Group 3 Test
16-29	0.0008	<0.0001	<0.0001
30-49	<0.0001	0.0889	0.0759
50-69	<0.0001	-	<0.0001
70-89	0.0044	-	<0.0001
90 & older	-	-	-

Table 11 displays the odds ratios for each of the performed MLR tests partitioned by driver age. In regards to GPS speed, throttle position, and yaw rate variables, the odds ratios for all ages across all three tests were similar and close to 1.0. This means that the odds ratio did not predict any specific driver behavior and was not useful. On the other hand, the odds ratios for both lateral and longitudinal acceleration provided different results. The younger drivers (ages 16-29) had very high odds ratios when examining Group 1 and Group 2 at 48.123 and 120.35 respectively. These odds ratios were significant, if a driver between the ages 16-29 increased his/her lateral acceleration by just 1-unit, that driver was 48 times more likely to be talking/listening on the phone and 120 times more likely to be texting. According to this result, lateral acceleration was a valid surrogate measure for which to identify whether younger drivers were engaged in either of these tasks. Longitudinal acceleration also seemed to be a good predictor for two specific secondary tasks, this time for drivers ages 30-49 and 70-89. The odds ratio found in the Control vs. Group 2 test for ages 30-49 was 171.78. For ages 70-89, the odds ratio was >999 when examining the Control vs. Group 1 test. Therefore, the driver was almost 1000 times more likely to be classified as talking/listening on the phone for drivers 70-89 and 171 times more likely to be classified as texting/dialing with a 1-unit increase in his/her longitudinal acceleration. Older driver's longitudinal acceleration seemed to be a better indicator of certain types of driver distraction, while lateral acceleration was a better indicator for younger drivers.

Table 11
Odds Ratio results of MLR tests partitioned by age

Variable	Driver Age	Control vs. Group 1 Test	Control vs. Group 2 Test	Control vs. Group 3 Test
GPS Speed	16-29	1.007	0.998	1
	30-49	1.005	1.008	1.018
	50-69	1.039	-	1.009
	70-89	0.958	-	0.996
	90 & older	-	-	-
Lateral Acceleration	16-29	48.123	120.35	0.004
	30-49	0.038	<0.001	<0.001
	50-69	0.578	-	<0.001
	70-89	<0.001	-	0.17
	90 & older	-	-	-
Longitudinal Acceleration	16-29	0.282	0.007	0.239
	30-49	0.231	171.78	<0.001
	50-69	0.003	-	1.269
	70-89	>999.999	-	0.428
	90 & older	-	-	-
Throttle Position	16-29	0.998	1.03	1.002
	30-49	1.013	0.998	1.09
	50-69	1.007	-	1.007
	70-89	0.945	-	1.004
	90 & older	-	-	-
Yaw Rate	16-29	0.962	0.944	1.059
	30-49	1.069	1.162	1.213
	50-69	1.08	-	1.116
	70-89	1.076	-	1.031
	90 & older	-	-	-

Results of MLR based on Driver Gender

The MLR test was performed for the Control group vs. Group 1, Control group vs. Group 2 and Control group vs. Group 3 based on the participant's gender. All R-square values were less than 0.05 which means very little of the variance could be accounted for by the dependent variables. The p-values for the Hosmer and Lemeshow test were significant for all but one of the gender partitioned MLR tests: most of the predicted estimates did not statistically match those observed. However, for the female drivers within the Control vs. Group 1 test the Hosmer and Lemeshow p-value was not significant at 0.1116 for alpha equal to 0.05. Due to this result, the odds ratio for the women included in this test can be deemed more credible. The Hosmer and Lemeshow p-values for each of the tests are in Table 12.

Table 12
Hosmer and Lemeshow p-values for MLR tests partitioned by gender

Driver Gender	Control vs. Group 1	Control vs. Group 2	Control vs. Group 3
Male	<0.0001	<0.0001	<0.0001
Female	0.1116	<0.0001	<0.0001

The odds ratios for each of the MLR tests partitioned by gender are in Table 13. A similar trend was seen when examining the odds ratios based on gender as to that obtained for the age. The GPS speed, throttle position, and yaw rate odds ratios were all close to 1.0 regardless of the gender or the secondary task that was examined. The odds ratio for lateral acceleration, on the other hand, was the highest for both Group 1 and Group 2 tests. For the Control vs. Group 1 test, regardless of the gender, the odds ratio was near 14, making both a male and female driver about 14 times more likely to be classified as talking/listening with an increase in his/her lateral acceleration. However, when the texting/dialing task (Group 2) was tested, males and females were around 179 and 4 times, respectively, more likely to be classified as such when increasing their lateral acceleration. It appears the lateral acceleration performance variable was useful for predicting texting/dialing as well as talking/listening for all drivers, but especially good at predicting such male drivers.

Finally, when testing whether the driver was talking/listening on the phone, the odds ratio for longitudinal acceleration for females was around 5, while this same value for males was almost 0. This means an increase in longitudinal acceleration in female drivers made them more likely to be talking/listening, however the same conclusion could not be drawn for male drivers in the same scenario.

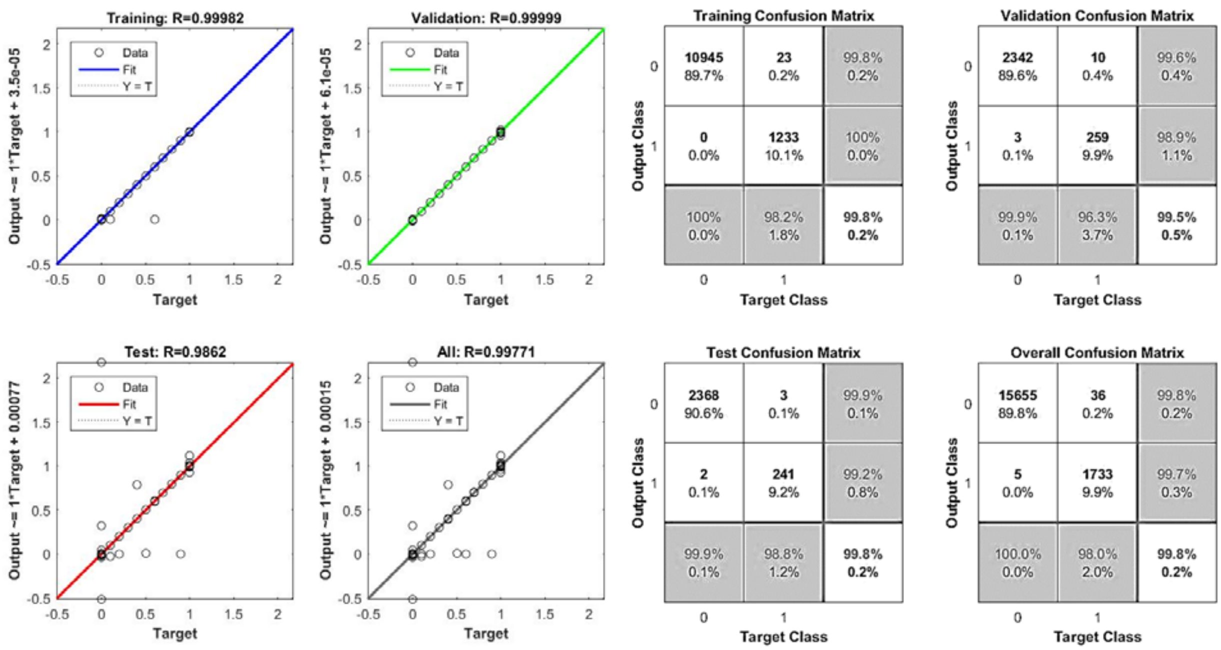
Table 13
Odds ratio results of MLR tests partitioned by gender

Variable	Driver Gender	Control vs. Group 1	Control vs. Group 2	Control vs. Group 3
GPS Speed	Male	1.012	0.995	1.006
	Female	1.006	1.003	0.999
Lateral Acceleration	Male	14.998	178.668	0.002
	Female	13.446	4.197	0.011
Longitudinal Acceleration	Male	0.011	0.095	0.318
	Female	5.427	0.039	0.144
Throttle Position	Male	1.035	1.021	1.012
	Female	0.981	1.019	1.013
Yaw Rate	Male	0.991	0.916	1.054
	Female	0.978	1.028	1.075

Neural Network Modeling Results

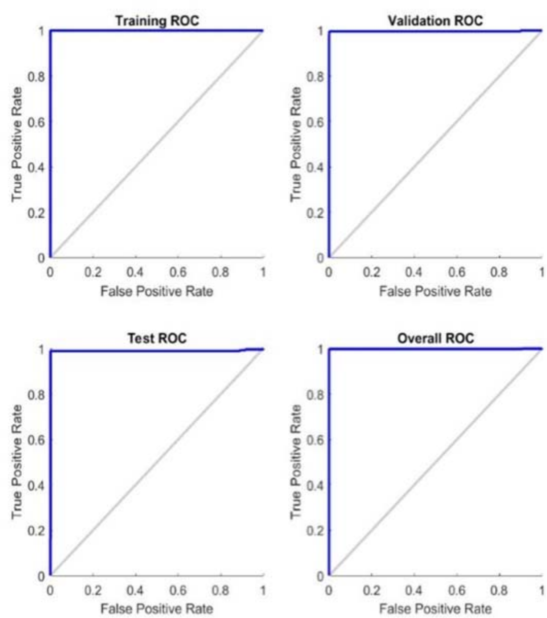
The regression curves, confusion matrix, and Receiver Operating Characteristic (ROC) curves were used as performance attributes of the three ANN models. The curves were obtained for the training, validation, and testing results. The following sections present a discussion of the results obtained for each of the three models (Calling, Texting, and Passenger Interaction).

Cell Phone Calling. Figure 25(a) shows the regression results for the training, validation, and testing datasets of the cellphone calling secondary task. Ideally, each point should fall on the 45-degree, “Y=T” line, which represents outputs = targets. The closer the output points to the 45-degree line, the more accurate the detections are. The colored lines represent the best linear fitting for the model. The best case scenario for these lines matches the “Y=T” line. As shown in Figure 25(a) the two lines are very close and almost indistinguishable. This implies a good fit of the developed model. This is reflected by the correlation coefficient (R) between the model output and the target results. R value is an indication of the relationship between the outputs and targets. If $R = 1$, this indicates that there is an exact linear relationship between outputs and targets. If R is close to zero, then there is no linear relationship between outputs and targets. The correlation coefficient values for the training, validation, and testing datasets are close to 0.99, which indicates a high correlation between the predicted and observed values.



(a) Regression Results

(b) Confusion Matrix



(c) Receiver Operating Characteristic (ROC) curves

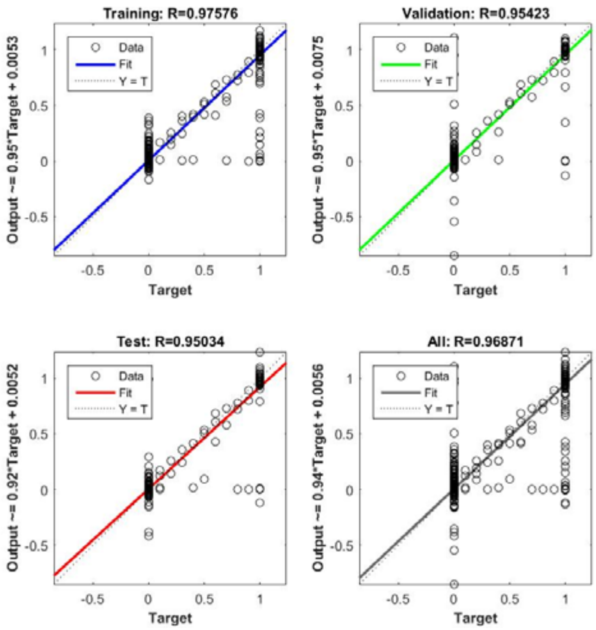
Figure 25
Detection results for the cellphone calling model

The results are confirmed by the confusion matrices shown in Figure 25(b). In each matrix, the "target class" at the bottom represents the correct class that each observation should be placed in. The "output class" on the vertical side shows the ANN-model classification results. If the output class is identical to the target class, this observation is determined as accurate/correct detection, and will be located accordingly. For each column, the number in each block shows the number of observations that ANN accurately or wrongly classified. The bottom row and the right-most column, shown in gray, provide the accurate rate (top value in percentage) and error rate (bottom value in percentage) for each class. As shown in the matrix, calling model has an outstanding overall classification sensitivity of 98% with a false negative rate of 1.8%, 3.7% and 1.2% for the training, validation, and testing datasets, respectively.

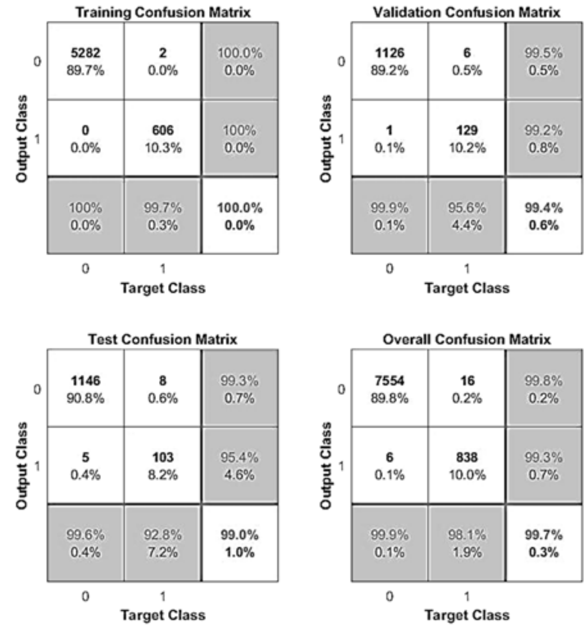
The confusion matrix can also be used to extract two important attributes: sensitivity and specificity. Sensitivity (or true positive rate) is defined as an observation classified by the model as "1" when the target is "1." Specificity (true negative rate) is defined as an observation classified by the model as "0" when the target value is "0". As shown in the matrix, the model has high sensitivity and specificity values of 98% and 100%, respectively. To further evaluate the performance of the ANN model, the ROC curves are used. The closer the curve is to the 45-degree diagonal of the ROC space and the less the area between the curves and the 45-degree line is, the less accurate the model is. According to the ROC plots in Figure 25c, the model detection performance is outstanding since the covered area under the blue lines is almost 100% of the total area above the 45-degree line.

Cell Phone Texting. The results depicted in Figure 26 show that the Texting detection model does not perform as well as the Calling model; yet, it is still a promising performance. As shown in Figure 26(a), the correlation coefficient values for the training, validation, and testing datasets are 0.98, 0.95, and 0.95, respectively. The regression curves clearly show that the fitting and target lines are discernible which indicates lower performance compared to the Calling detection model.

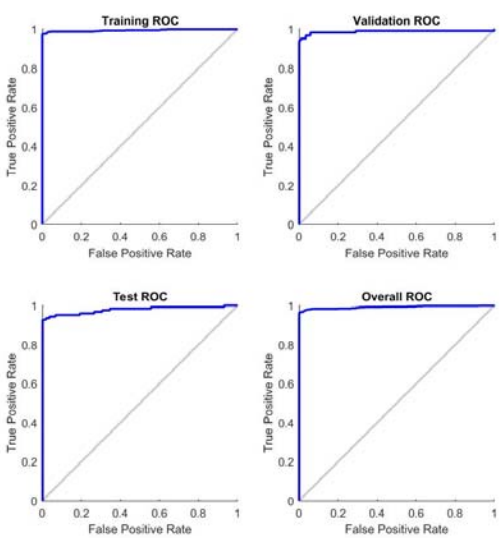
The confusion matrix in Figure 26(b) also shows promising results for the Texting detection model. The overall model sensitivity was 98.1% with individual sensitivity rates of 99.7%, 95.6%, and 92.8% for the training, validation, and testing datasets, respectively. By looking at the validation results, Texting model has very slightly lower accuracy than Calling; yet, Texting model also has a very good detection performance. These results are supported by the covered area in the ROC curves shown in Figure 26(c). The figures show a very slightly less covered area for texting compared to calling.



(a) Regression Results



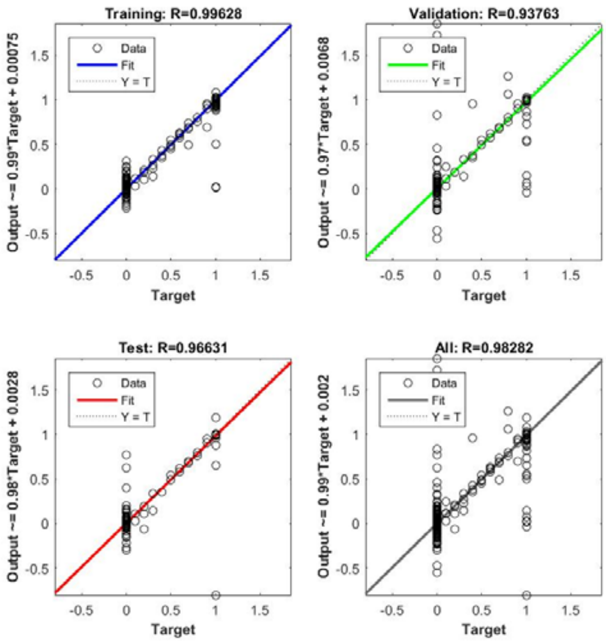
(b) Confusion Matrix



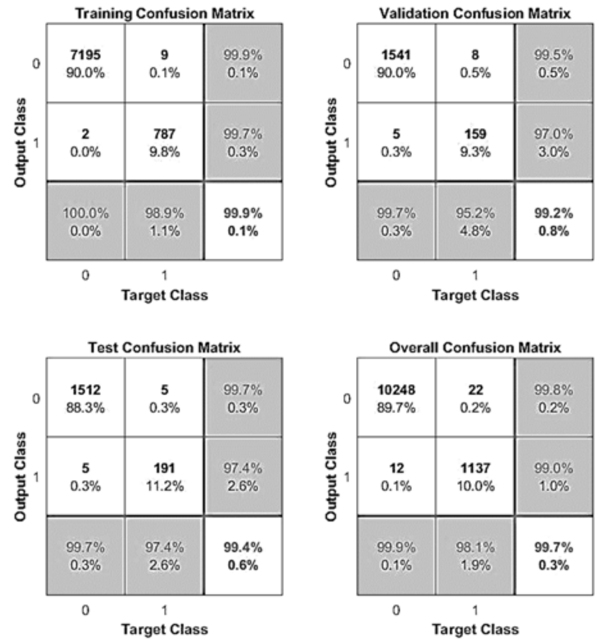
(c) Receiver Operating Characteristic

Figure 26
Detection results for the texting model

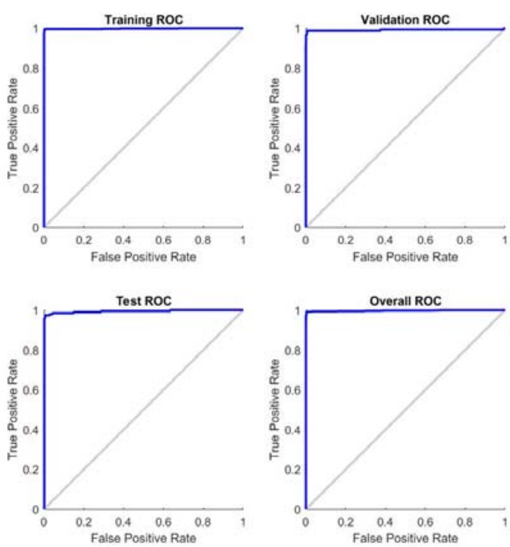
Passenger Interaction. For the Passenger Interaction detection model, the results show that this model is the least accurate in its predictions compared to texting and calling. Figure 27(a) shows a correlation coefficient for the Passenger Interaction model of around 0.9828 and a detection accuracy rate of 99.7% as shown in Figure 27(b). These values are marginally lower than those for Texting and Calling models. According to the literature, the amount of cognitive distraction associated with passenger interactions is the least compared to texting and calling *[29]*. This means that the changes in the associated driver behavior are not as significant as those associated with either texting or calling. Yet, the detection performance of the model is still considered outstanding with a sensitivity value of 98.1%. By looking at the ROC curves in Figure 27(c), it is clear that the covered area under the blue lines is almost 100%.



(a) Regression Results



(b) Confusion Matrix



(c) Receiver Operating Characteristic

Figure 27
Detection results for the Passenger Interaction model

CONCLUSIONS

The objectives of this research were to conduct a thorough exploration of the SHRP 2 NDS data to identify appropriate performance measures that can be used as surrogate measures of distraction, and to outline a methodology for developing a crash index. The time series and events NDS data were used to develop the models. GPS speed, lateral and longitudinal acceleration, throttle position, and yaw rate were the driving performance measures of drivers' engagement in one of three defined groups of secondary tasks: talking/listening on hand-held phone, texting/dialing on hand-held phone, and interacting with the adjacent passenger. The time series nature of the input used provided more robust data than data typically used in distracted driving studies, as they can show minor changes in the driving pattern associated with the different secondary tasks. To identify the appropriate distracted driving surrogate measures, statistical secondary task detection models were developed. The input data used in the detection models included the time series data of the aforementioned five performance variables.

Multiple logistic regression was used to determine the odds of a driver being engaged in one of the secondary tasks, given their corresponding driving performance data. The results indicated that none of the models computed had strong R-square values. However, according to the outcomes observed when analyzing tasks in Group 1 (talking/listening on the phone) and Group 2 (texting/dialing on the phone), a change in lateral acceleration seemed to be an indicator the driver was in one of those two groups. In the test analyzing adjacent passenger interaction, or Group 3, little was revealed in the results, an indication that the five performance measures did not significantly increase or decrease when that interaction took place. These results indicated that only the lateral acceleration could be used a surrogate measure to detect distracted driving resulting from secondary tasks in Group 1 and Group 2.

When considering the driver's age, a similar trend was observed. Driver's acceleration was a good indicator of the participant either talking/listening or texting/dialing. Older driver's longitudinal acceleration seemed to be a better indicator of distracted behavior resulting from these secondary tasks, while lateral acceleration was a better indicator for younger drivers. When the data was analyzed based on driver gender, the lateral acceleration performance variable proved to be useful for predicting texting/dialing as well as talking/listening for all drivers. More so, longitudinal acceleration proved to be useful for predicting distracted driving behavior resulting from the same secondary tasks for female drivers.

When analyzing the data using neural networks, researchers obtained different and more credible results as compared to the logistic regression analysis. The developed neural network models proved the five performance measures (speed, lateral acceleration, longitudinal acceleration, yaw rate, and throttle position) to be important as surrogate measures for distracted driving behavior. Using the five measures, the results showed that all three neural network models have accurate detection capabilities. This was obvious in terms of accuracy rates for the three models which exceeded 95%. Expressing the selected five driving performance attributes in terms of their standard deviation also proved to be an effective approach for capturing the driving pattern associated with each secondary task. These results show that the selected driving performance attributes were effective in detecting the associated secondary tasks with driving behavior. In other words, if detailed information of speed, longitudinal acceleration, lateral acceleration, throttle position, and yaw rate could be gathered, ANN model could detect the likelihood of driver's engagement in calling, texting, or interacting with a passenger with a considerably high accuracy.

In summary, this research showed how useful the SHRP2 NDS data could be for distracted driving studies. Although the statistical analysis results of this research cannot be taken as credible in most cases, they showed that the high-quality and high-resolution data available on the SHRP 2 NDS database can provide useful insight on detecting distracted driving. Identifying the right surrogate measures and the use of a more suitable analysis tool that can recognize nonlinear patterns in driving behavior can help in detection of distracted driving behavior. Therefore, the neural network modeling was deployed to analyze the five performance measures. Unlike the multiple logistic regression, the neural network analysis identified the five-time series measures as important surrogate measures of distracted driving behavior. The developed neural network models also showed high accuracy in detecting drivers' engagement in secondary tasks. The proposed crash index outline can also provide an insight on quantifying the crash risk associated with distracted driving behavior.

RECOMMENDATIONS

Based on the findings of this study, several recommendations are made as follows:

- The proposed crash index in this study is only an outline, yet it shows good potential for quantification of crash risk associated with distracted driving behavior. Thus, statistical analysis can be performed on the SHRP 2 NDS events and socioeconomic characteristics data to provide a clear insight on how crash index can be quantified.

- The provided summary on state regulations of cellphone use in this study can be further investigated using SHRP2 NDS data. The available data in the NDS database can be used to identify whether further recommendations can be made about the available state regulations.

- The available data can also be used in conjunction with the Roadway Information Database to evaluate the crash risk associated with distracted driving behavior at different roadway facility types.

- Artificial Intelligence (AI) proved itself as a promising tool to analyze the nonlinear pattern in drivers' behavior and detect drivers' engagement in secondary tasks. A possible extension for this would be investigating whether AI can be used to detect the type of secondary task drivers are engaged in. The accurate secondary task detection achieved by the neural networks, triggered extending the research in this area to identification of secondary task type using artificial intelligence and machine learning tools.

- The available data can also be used to investigate changes in driving pattern before crashes and near-crashes take place when specific secondary tasks are performed.

ACRONYMS, ABBREVIATIONS, AND SYMBOLS

CTRE	Center for Transportation Research and Education
CTS	Center for Transportation Studies
DAS	Data Acquisition System
DOTD	Louisiana Department of Transportation and Development
FCW	Forward Collision Warnings
FHWA	Federal Highway Administration
g/s	Gravitational Force per Second
GEV	Generalized Extreme Function
GPS	Geographic Positioning System
IRB	Institutional Review Board
km/h	Kilometer per Hour
LTRC	Louisiana Transportation Research Center
MLE	Maximum Likelihood Estimators
MLR	Multiple Linear Regression
NDS	Naturalistic Driving Study
NHTSA	National Highway Traffic Safety Administration
RID	Roadway Information Database
SHRP 2	Second Strategic Highway Research Program
TTC	Time to Collision
UTC	University Transportation Centers
VTTI	Virginia Tech Transportation Institute

REFERENCES

1. National Highway Traffic Safety Administration. Traffic Safety Facts. Research Note DOT HS 812012, U.S. Department of Transportation, 2014.

2. Transportation Research Board of the National Academy of Sciences. The 2nd Strategic Highway Research Program Naturalistic Driving Study Dataset. Available from the SHRP 2 NDS InSight Data Dissemination. 2013. https://insight.shrp2nds.us. Accessed Dec 2016.

3. Harbluk, J. L., Noy, Y. I., Trbovich, P. L., and Eizenman, M. "An On-Road Assessment of Cognitive Distraction: Impacts on Drivers." Visual Behavior and Braking Performance." Accident Analysis and Prevention, Vol. 39, No.2, 2007, pp. 372-379.

4. Narad, M., Garner, A. A., Brassell, A. A., Saxby, D., Antonini, T. N., O'Brien, K. M., Epstein, J. N. "Impact of Distraction on the Driving Performance of Adolescents with and without Attention-Deficit/Hyperactivity Disorder." JAMA pediatrics, Vol.167, No.10, 2013, pp. 933-938.

5. Hosking, S. G., Young, K., and Regan, M. A. "The Effects of Text Messaging on Young Novice Driver Performance." Monash University, Australia, 2005.

6. Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., and Dingus, T. A. "Distracted Driving and Risk of Road Crashes Among Novice and Experienced Drivers." New England journal of medicine, Vol. 370, No. 1, 2014, pp. 54-59.

7. Elander, J., West, R., and French, D. "Behavioral Correlates of Individual Differences in Road-Traffic Crash Risk: An Examination of Methods and Findings." Psychological bulletin, Vol. 113, No. 2, 1993, pp. 279.

8. Pöysti, L., Rajalin, S., and Summala, H. "Factors Influencing the Use of Cellular (mobile) Phone During Driving and Hazards While Using It." *Accident Analysis and Prevention*, Vol. 37, No. 1, 2005, pp. 47-51.

9. Strayer, D. L., Drews, F. A., and Crouch, D. J. "A Comparison of the Cell Phone Driver and the Drunk Driver." Human factors: *The Journal of the Human Factors and Ergonomics Society*, Vol. 48, No. 2, 2006, pp. 381-391.

10. Kaptein, N.A., J. Theeuwes and R. Van der Horst. "Driving Simulator Validity: Some Considerations." In Transportation Research Record: Journal of the Transportation Research

Board, No. 1550, Transportation Research Board of the National Academies, Washington, D.C., 1996. pp. 30-36.

11. Virginia Tech Transportation Institute. “100-Car Naturalistic Driving Study.” Vulnerable Road User Safety. 2013. https://www.vtti.vt.edu/research/vrus/projects/100car/100car.html. Accessed Apr 2014.

12. Jonasson, J., and Rootzen, H. “Internal Validation of Near-Crashes in Naturalistic Driving Studies: a Continuous and Multivariate Approach.” *Accident Analysis and Prevention*, Vol. 62, 2014, pp. 102-109.

13. Bagdadi, O. Assessing Safety Critical Braking Events in Naturalistic Driving Studies. *Transportation Research Part F: Traffic Psychology and Behavior*, Vol. 16, 2013, pp. 117-126.

14. Guo, F., and Fang, Y. “Individual Driver Risk Assessment using Naturalistic Driving Data.” *Accident Analysis & Prevention*, Vol. 61, 2013, pp. 3-9.

15. Sangster, J., Rakha, H. and Du, J., “Application of Naturalistic Driving Data to Modeling of Driver Car-Following Behavior.” In *Transportation Research Record* 2390, National Research Council, Washington, DC, 2013, pp.20-33.

16. Wong, J. T., & Huang, S. H. “Attention Allocation Patterns in Naturalistic Driving.” *Accident Analysis & Prevention*, Vol. 62, 2013, pp. 102-9.

17. Montgomery, J., Kusano, K. D., and Gabler, H C. “Age and Gender Differences in Time to Collision at Braking from the 100-car Naturalistic Driving Study.” *Traffic injury prevention* Vol. 15, 2014.

18. Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D. and Ramsey, D.J., The Impact of Driver Inattention on Near-Crash/Crash Risk: An analysis using the 100-car naturalistic driving study data. DOT HS 810 594, NHTSA NVS-331, 2006.

19. Hallmark, S.L., Oneyear, N., Tyner, S., Wang, B., Carney, C. and McGehee, D., “Analysis of Naturalistic Driving Study Data: Roadway Departures on Rural Two-Lane Curves.” SHRP 2 Report S2-S08D-RW-1, 2015.

20. Hutton, J.M., Bauer, K.M., Fees, C.A. and Smiley, A., “Analysis of Naturalistic Driving Study Data: Offset Left-Turn Lanes.” No. SHRP 2 Report S2-S08B-RW-1, 2015.

21. Hallmark, S., McGehee, D., Bauer, K.M., Hutton, J.M., Davis, G.A., Hourdos, J., Chatterjee, I., Victor, T., Bargman, J., Dozza, M. and Rootzen, H., “Initial Analyses from the SHRP 2 Naturalistic Driving Study: Addressing Driver Performance and Behavior in Traffic Safety.” No. SHRP 2 Safety Project S08, 2013.

22. Victor, T., Dozza, M., Bärgman, J., Boda, C.N., Engström, J., Flannagan, C., Lee, J.D. and Markkula, G., “Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk.” No. SHRP 2 Report S2-S08A-RW-1, 2015.

23. Dingus, T. A., F. Guo, S. Lee, J. F. Antin, M. Perez, M. Buchanan-King, and J. Hankey. “Driver Crash Risk Factors and Prevalence Evaluation Using Naturalistic Driving Data.” Proceedings of the National Academy of Sciences, 2016.

24. Codjoe, J., McCarter, K., and Ishak, S. “Multivariate Analysis of Distracted Driving.” In Transportation Research Board. 94th Annual Meeting of the Transportation Research Board. Washington, D.C., No. 15-2246. 2015.

25. Meyers, Lawrence S., Glenn Gamst, and A. J. Guarino. Data Analysis using SAS Enterprise Guide. Cambridge University Press, 2009.

26. Wilson, S. “Choosing Between Logistic Regression and Discriminant Analysis.” *Journal of the American Statistical Association*, Vol. 73, No. 364, 1978, pp. 699-705.

27. Karp, Andrew H. Getting started with PROC LOGISTIC. Proceedings of the 25th SAS Users Group International Conference, Cary, NC. 2000.

28. Newsom. “Overview: Logistic and OLS Regression Compared.” 2013.

29. Strayer, D.L., J.M. Cooper, J. Turrill, J. Coleman, N. Medeiros-Ward, and F. Biondi. “Measuring Cognitive Distraction in the Automobile.” AAA Foundation for Traffic Safety, 2013.

APPENDIX A

All Data Available for Each Category Within the NDS

Data Category	Topic	Subtopic
Drivers	Summary Statistics on Drivers	by Age Group
		by Age Group and Gender
		Data collected by Driver Age Group
		Data collected by Driver Age Gender
	Driver Demographic Questionnaire	
	Driving History Survey	
	Driving Knowledge Survey	
	Visual/Cognitive Tests	
	Conner's Continuous Performance Test	
	Clock Drawing Assessment	
	Physical Strength Tests	
	Barkley's ADHD Screening Test	
	Risk Perception Questionnaire	
	Risk Taking Questionnaire	
	Sensation Seeking Scale Survey	
	Driver Behavior Questionnaire	
	Medical Conditions & Medications	
	Sleep Habits Questionnaire	
	Medical Conditions and Medications - Exit	
	Driver Exit Interview	
Vehicles	Vehicles by Vehicle Classification	
	Vehicles by Model Year	
	Vehicles by Beginning Mileage	
	Vehicles Active by Calendar Month	
	Data Collected by Vehicle	
	Data Collected by Vehicle Classification	
	Vehicle Detail Table	
Trips	Trip Summary Table	
	Time Series	
	Data Collected by Trip Start Hour of Day	
	Data Collected by Day of Week	
	Maximum Deceleration	
	Maximum Speed	
	Maximum Deceleration by Vehicle Classification	

Data Category	Topic	Subtopic
	Maximum Speed by Vehicle Classification	
	Maximum Deceleration by Age Group	
	Maximum Speed by Gender	
	Maximum Deceleration by Data Collection Site	
	Maximum Speed by Data Collection Site	
	Travel Density Map for Florida	
	Travel Density Map for Indiana	
	Travel Density Map for New York	
	Travel Density Map for North Carolina	
	Travel Density Map for Pennsylvania	
	Travel Density Map for Washington	
Events	Post-Crash Interview	
	Event Detail Table	
Query Builder	User can select variables and conditions to submit to query	

APPENDIX B

List of All Secondary Task Options Available within the NDS Dataset

1	No Secondary Task
2	Talking/Singing audience unknown
3	Dancing
4	Reading
5	Writing
6	Passenger in adjacent seat - interaction
7	Passenger in rear seat - interaction
8	Child in adjacent seat - interaction
9	Child in rear seat - interaction
10	Moving object in vehicle
11	Insect in vehicle
12	Pet in vehicle
13	Object dropped by driver
14	Reaching for object, other
15	Object in vehicle, other
16	Cell phone, holding
17	Cell phone, Talking/listening hand-held
18	Cell phone, Talking/listening, hands-free
19	Cell phone, Texting
20	Cell phone, Browsing
21	Cell phone, Dialing hand-held
22	Cell phone, Dialing hand-held using quick keys
23	Cell phone, Dialing hands-free using voice-activated software
24	Cell phone, Locating/reaching/answering
25	Cell phone, other
26	Tablet device, Locating/reaching
27	Tablet device, Operating
28	Tablet device, Viewing
29	Tablet device, Other
30	Adjusting/monitoring climate control
31	Adjusting/monitoring radio
32	Inserting/retrieving CD (or similar)
33	Adjusting/monitoring other devices integral to vehicle
34	Looking at previous crash or incident
35	Looking at pedestrian
36	Looking at animal
37	Looking at an object external to the vehicle

38	Distracted by construction
39	Other external distraction
40	Reaching for food-related or drink-related item
41	Eating with utensils
42	Eating without utensils
43	Drinking with lid and straw
44	Drinking with lid, no straw
45	Drinking with straw, no lid
46	Drinking from open container
47	Reaching for cigar/cigarette
48	Lighting cigar/cigarette
49	Smoking cigar/cigarette
50	Extinguishing cigar/cigarette
51	Reaching for personal body-related item
52	Combing/brushing/fixing hair
53	Applying make-up
54	Shaving
55	Brushing/flossing teeth
56	Biting nails/cuticles
57	Removing/adjusting clothing
58	Removing/adjusting jewelry
59	Removing/inserting/adjusting contact lenses or glasses
60	Other personal hygiene
61	Other non-specific internal eye glance
62	Other known secondary tasks
63	Unknown type (secondary task present)
64	Unknown

APPENDIX C

Summary of Data Removed During Editing Phase of Analysis

Variable	Group	Original Total	Outside Acceptable Limits		Missing Data	Outliers		New Total
			Criteria	# Removed	# Removed	Criteria	# Removed	
GPS Speed	0	32262	< 0	700	63	> 150	8	31491
	1	3483	< 0	146	0	n/a	0	3337
	2	1678	< 0	66	0	n/a	0	1612
	3	9518	< 0	140	0	n/a	0	9378
Lateral Acceleration	0	32262	-999	90	63	<-0.5 or >0.5	6	32103
	1	3483	-999	0	0	≤ -0.31 or >0.4	8	3475
	2	1678	-999	0	0	<-0.18 or >0.30	6	1672
	3	9518	-999	21	0	≤-0.5 or >0.4	4	9493
Longitudinal Acceleration	0	32262	-999	46	63	<-0.4 or >0.3	8	32145
	1	3483	-999	0	0	<-0.3 or ≥0.25	9	3474
	2	1678	-999	0	0	<-0.25 or ≥0.25	14	1664
	3	9518	-999	21	0	≤-0.5 or >0.4	4	9493
Throttle Position	0	32262	< 0	7512	63	>70	33	24654
	1	3483	< 0	1082	0	≥45	13	2388
	2	1678	< 0	563	0	>50	15	1100
	3	9518	< 0	2477	0	≥85	9	7032
Yaw Rate	0	32262	< -100	43	63	≤-35 or >30	16	32140
	1	3483	< -100	21	0	≤-30 or >25	3	3459
	2	1678	< -100	0	0	≤-20 or >20	10	1668
	3	9518	< -100	21	0	≤-30 or ≥30	6	9491

This public document is published at a total cost of $250
42 copies of this public document were published in this first printing at a cost of $250. The total cost of all printings of this document including reprints is $250. This document was published by Louisiana Transportation Research Center to report and publish research findings as required in R.S. 48:105. This material was duplicated in accordance with standards for printing by state agencies established pursuant to R.S. 43:31. Printing of this material was purchased in accordance with the provisions of Title 43 of the Louisiana Revised Statutes.