Collaborative Sciences Center for ROAD SAFETY

FINAL REPORT

Developing a Taxonomy of Human Errors and Violations that Lead to Crashes

December 2, 2020

Asad Khattak, Ph.D.
Numan Ahmad
Behram Wali, Ph.D.

The University of Tennessee, Knoxville

Eric Dumbaugh, Ph.D.

Florida Atlantic University

THE UNIVERSITY OF TENNESSEE KNOXVILLE · Berkeley UNIVERSITY OF CALIFORNIA · FAU FLORIDA ATLANTIC UNIVERSITY · Duke UNIVERSITY · THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

U.S. DOT Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgement of Sponsorship

This project was supported by the Collaborative Sciences Center for Road Safety, www.roadsafety.unc.edu, a U.S. Department of Transportation National University Transportation Center promoting safety.

We gratefully acknowledge the contributions of Ms. Kinzee Clark and Mr. Zachary Jerome toward editing the report.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
CSCRS-R19		

4. Title and Subtitle: Developing a Taxonomy of Human Errors and Violations that Lead to Crashes	**5. Report Date** December 02, 2019 **6. Performing Organization Code** DUNS: 608195277
7. Author(s) Asad Khattak, PhD; Numan Ahmad, Behram Wali, PhD, Eric Dumbaugh, PhD	**8. Performing Organization Report No.**
9. Performing Organization Name and Address University of Tennessee Center for Transportation Research, Knoxville, TN Florida Atlantic University, Boca Raton, FL	**10. Work Unit No.** **11. Contract or Grant No.** Collaborative Sciences Center for Road Safety (Grant #: 69A3551747113)
12. Sponsoring Agency Name and Address University Transportation Centers Program Office of the Assistant Secretary for Research and Technology U.S. Department of Transportation 1200 New Jersey Avenue, SE, Washington, DC 20590-0001	**13. Type of Report and Period Covered** Final Report (May 2018 – Nov. 2019) **14. Sponsoring Agency Code** Office of the Assistant Secretary for Research and Technology-RDT 30

15. Supplementary Notes
Other contributors include:

16. Abstract
Driver errors and violations are highly relevant to the safe systems approach as human errors tend to be a predominant cause of crash occurrence. This study develops a deeper understanding of critical pre-crash driver errors and violations that have significant potential in reducing dangerous behaviors on roadways. A driver error and violation taxonomy (TDEV) is developed to understand the factors that contribute to crashes and it is applied using naturalistic data from detailed real-world driving and monitoring ecosystems. Specifically, data from the Naturalistic Driving study (NDS) is analyzed to understand the origin of different types of human errors, especially focusing on how they relate to roadway and built environments. Different types of human errors and violations are categorized using a perception, recognition, decision, and reaction framework and we explore how errors and violations contribute to safety-critical events in naturalistic settings. For the NDS data available to the research team, human errors and violations contributed to 93% of the observed crashes, while roadway factors contributed to 17%, vehicle factors contributed in 1%, and 4% of crashes contained unknown factors. The most common human errors were recognition and decision errors, which occurred in 39% and 34% of crashes, respectively. These two error types occurred more frequently (each contributing about 39% of crashes) when business or industrial structures were present. While the most prevalent errors in crashes and near crashes were recognition errors, performance errors such as weak judgement (8%) were strongly correlated with crash occurrence. Path analysis uncovered direct and indirect relationships between key built-environment factors, errors and violations, and crash propensity. Possibly due to their complexity for drivers, urban environments are associated with higher chances of crashes and they can induce more recognition errors, which associate with even higher chances of crashes. Finally, this report discusses implications for crash investigations.

17. Key Words	18. Distribution Statement
Human errors, violations, safety, naturalistic driving study, built environment and roadway characteristics	No restrictions. This document is available through the Collaborative Sciences Center for Road Safety (roadsafety.unc.edu), Chapel Hill, NC.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	54	

Table of Contents

Figures

Tables

Introduction

1. OVERVIEW

The social and economic costs of transportation crashes have reached a staggering $1 Trillion in the United States. A recent study showed that the critical reason (the last event in crash event chain) relates to drivers in a majority of crashes (Singh 2015). In order to reduce or eliminate traffic crashes and their social and economic consequences, the profession needs a more nuanced understanding of the reasons why crashes occur. A broad spectrum of traffic safety literature has focused on examining the key factors contributing to crash occurrence and/or injury outcomes given a crash (Frawley and Eisele 2004, Huang, Klauer, Guo et al. 2014). To better understand the contribution of these factors, studies have classified correlates of crash occurrence/injury outcomes into three major categories: driver, vehicle, and roadway-environment factors (Sabey and Staughton 1975, Treat, Tumbas et al. 1979, Singh 2015). Driving behavioral factors were found to be the most prevailing contributing factors in more than 90% of crashes (Sabey and Staughton 1975, Treat, Tumbas et al. 1979, Singh 2015). While understanding that human factors are predominant crash contributing factors, studies classified errors and developed taxonomies for driver's errors and violations focusing on different themes (Treat, Tumbas et al. 1979, Reason, Manstead et al. 1990, Petridou and Moustaki 2000, Wierwille, Hanowski et al. 2002, Stanton and Salmon 2009). Given previous taxonomies, some key driving errors include recognition errors, decision errors, and performance errors (Treat, Tumbas et al. 1979, Wierwille, Hanowski et al. 2002). Recognition and decision errors were the leading and critical errors resulting in crashes (Treat, Tumbas et al. 1979, Wierwille, Hanowski et al. 2002, Iden and Shappell 2006). Due to variations in driver characteristics and behavioral patterns, different types of roadways and environments can induce different errors and violations, potentially resulting in crashes or near-crashes (Dumbaugh and Li 2010). Behavioral factors are associated with most crashes (Frawley and Eisele 2004, Dumbaugh and Li 2010, Huang, Abdel-Aty et al. 2010, Klauer, Guo et al. 2014, Dingus, Guo et al. 2016). In particular, driver distraction was found to be one of the most critical (Klauer, Guo et al. 2014). Also, local demographics, and socio-economic conditions contribute to higher crash risks (Huang, Abdel-Aty et al. 2010). The chances of crashes are higher with higher local population density, higher traffic intensity, and urbanization (Huang, Abdel-Aty et al. 2010). Similarly, freeways were associated with lower crash risk than arterials (Huang, Abdel-Aty et al. 2010). While the literature acknowledges that different roadways and environments can have diverse impacts on crash occurrence (Dumbaugh and Li 2010), this issue is lightly addressed. Most of the aforementioned studies used traditional police crash reports, which have an element of subjectivity as the police officer typically does not observe the crash as it happens (Wali, Khattak et al. 2018). Likewise, traditional crash data do not provide objective information on pre-crash driver behavior and performance – in fact it is very difficult on part of the reporting officer to determine the exact behavior or speed profile that could have contributed to the crash. Also, other inaccuracies exist such as non-reporting of low severity (property damage) crashes Washington, Karlaftis et al. 2010). The Naturalistic Driving Study (NDS) data provide an opportunity to extract extensive real world information not only on crashes (with all severity levels) but near-crashes and baselines as well (Hankey, Perez et al. 2016). In view of the above discussion, this project aims to:

- Developing a systematic taxonomy for driving errors and violations and exploring their contribution to the occurrence of safety-critical events (i.e., crashes and near-crashes) using naturalistic driving study data.
- Exploring the pathways of errors and violations that lead to safety-critical events in diverse roadway and built environments.

To achieve the aforementioned objectives, this project first quantifies the contribution of key factors (i.e., human, vehicle, and roadway/environment) resulting in crashes. A safety matrix is developed to understand the sole as well as simultaneous contribution of human factors with vehicle and roadway/environment factors. Furthermore, the project develops a systematic taxonomy for driving errors and violations. Using extensive NDS data, and after embedding the driving errors and violations in the analysis framework, the study conducts a rigorous path analysis to explore the direct and indirect effects of different built environment factors on occurrence of safety-critical events through driving errors and violations.

RESEARCH QUESTIONS

This project addresses the following research questions:

a) What types of driving errors and violations (human factors) result in safety-critical events in naturalistic settings?
b) How do driving errors and violations vary across different roadway and built environments?
c) What are the implications of the study for how human errors and violations will impact crashes when some vehicles are human-driven while others have some level of automation?

MULTI-PRONGED APPROACH

This report describes work performed as two distinct efforts. Each effort is listed as a chapter:

Chapter 2. "A taxonomy of driving errors and violations: Evidence from the naturalistic driving study" quantifies the contribution of key factors (human, vehicle, and roadway/environment) resulting in crashes, and then develops a systematic taxonomy for driving errors and violations (TDEV) in a naturalistic environment. It classifies driver errors and violations based on their presence during the theoretically-based perception-reaction process and analyzes their contribution in safety-critical events. To empirically explore their role in diverse built environments, this study harnesses a unique and highly detailed pre-crash sensor data collected in the SHRP2 Naturalistic Driving Study (NDS).

Chapter 3. "Driver errors and violations: Pathways that lead to crashes in diverse built environments" explores pathways to uncover direct and indirect relationships between key roadway/built-environment factors, errors and violations, and crash propensity. Due to their complexity for drivers, urban environments were found to be associated with higher chances of crashes (by 7.66%), and they can induce more recognition errors, which associate with even higher chances of crashes (by 3.40% with the "total effect" amounting to 11.06%). Similar statistically significant mediating contributions of recognition errors and decision errors at school, playground, and construction zones were also observed. Other important results are discussed, along with real-world implications. The detail findings of this study are discussed in this part of the report.

RESEARCH OUTPUTS

Publications and Presentations

- Khattak, A. J., N. Ahmad, B. Wali, and E. Dumbaugh, (2020). A taxonomy of driving errors and violations: Evidence from the naturalistic driving study. Accident Analysis & Prevention, 151, 105873.
- Khattak, A. J., B. Wali and N. Ahmad (2019). A Taxonomy of Naturalistic Driving Errors and Violations and Its Variations across Different Land-Use Contexts – A Path Analysis Approach. *98th Annual Meeting of the Transportation Research Board.* National Academies, Washington D.C.
- Khattak, A. J., N. Ahmad, B. Wali. Heterogeneity in Naturalistic Driving Errors, Violations, and Crash Risk in Diverse Environmental Context. P*resented (#TRBAM-21-04104) at the 100th* Annual Meeting of Transportation Research Board at Washington, D.C (January 2021).
- Ahmad N., B. Wali, A. Khattak, and E. Dumbaugh. Driver Errors and Violations: Pathways that Lead to Crashes in Diverse Built Environments. (In-review, Accident Analysis & Prevention).

2. A Taxonomy of Driving Errors and Violations: Evidence from the Naturalistic Driving Study

Authors

Asad J. Khattak[1], Numan Ahmad[1], Behram Wali[2], Eric Dumbaugh[3]

CHAPTER SUMMARY

Driving errors and violations are identified as contributing factors in a majority of crash events. To examine the role of human factors and improve crash investigations, a systematic taxonomy of driver errors and violations (TDEV) is developed first. The TDEV classifies driver errors and violations based on their occurrence during the theoretically based perception-reaction process and analyzes their contributions in safety-critical events. To empirically explore errors and violations in diverse built environments, this study harnesses unique and highly detailed pre-crash sensor data collected in the SHRP2 Naturalistic Driving Study (NDS), containing 673 crashes, 1,331 near-crashes and 7,589 baselines (no-event). Human factors are categorized into recognition errors, decision errors, performance errors, and errors due to the drivers' physical condition or their lack of contextual experience/familiarity, and intentional violations. Built environments are classified based on roadway functional classification and land uses, e.g., residential areas, school zones, and church zones. Human errors and violations contributed to 93% of the observed crashes, while roadway factors contributed to 17%, vehicle factors contributed in 1%, and 4% of crashes contained unknown factors. The most common human errors were recognition and decision errors, which occurred in 39% and 34% of crashes, respectively. These two error types occurred more frequently (nearly 39% each) when business or industrial structures were present (but not in dense urban localities). The findings of this study reveal continued prevalence of human factors in crashes. The distribution of driving errors and violations across different roadways and environments found in this study can aid in the implementation of locality-specific countermeasures and has implications for connected and automated vehicle development, e.g., by understanding complex and unusual (fringe case) situations for safety, testing of connected and automated vehicles can be enhanced.

Author affiliations:

[1]Tickle College of Engineering, Civil & Environmental Engineering, University of Tennessee, Knoxville TN
[2]Urban Design 4 Health, 24 Jackie Circle East Rochester, NY 14612.
[3]School of Urban & Regional Planning, Florida Atlantic University, Boca Raton, FL 33431.

INTRODUCTION AND BACKGROUND

Traffic safety research classifies the pre-crash factors leading to a crash event as belonging to one or more of three factors: driver-related factors, vehicle-related factors, and environment- related factors ((Frawley and Eisele 2004, Huang, Abdel-Aty et al. 2010, Klauer, Guo et al. 2014, Ali, Ahmad et al. 2018, Ahmad, Ahmed et al. 2019; Sabey and Staughton 1975, Treat, Tumbas et al. 1979, Singh 2015). Studies using police-accident reports have found that driver-related factors are the prevailing contributing factor in crash events, associated with more than 90% of all crashes that occur (Sabey and Staughton 1975, Treat, Tumbas et al. 1979, Singh 2015). While police-reported crash data sources provide valuable information, the interpretation of the data they contain is necessarily subjective; rather than being based on direct observations of the events that immediately precede a crash, analysts are forced to infer these behaviors based on data collected after the crash event has occurred (Wali, Khattak et al. 2018, Hankey, Perez et al. 2016, Ahmad, Ahmed et al. 2019). Other inaccuracies, such as non-reporting of low severity (property damage) crashes, exist in traditional crash data (Yamamoto, Hashiji et al. 2008, Washington, Karlaftis et al. 2010, Ye and Lord 2011).

This study seeks to overcome these limitations through the use of Naturalistic Driving Survey (NDS) data. NDS data is collected from instrumented vehicles that record the behaviors of drivers, including their speed, breaking performance, and video recordings of a driver's actions, providing the ability to directly observe real-world information from crashes of all severity levels, near-crashes, and baselines (Hankey, Perez et al. 2016). In addition, the chances of losing important information that exist in traditional data protocols are minimized. The NDS data provides information on real-world driving behavior and performance, along with real-world risks and safety consequences (Dingus, Klauer et al. 2006, Carney, McGehee et al. 2015, Dingus, Guo et al. 2016, Arvin, Kamrani et al. 2019). This facilitates the examination of human factors in greater detail, e.g., errors due to driver inattention, distraction, drowsiness, and judgment-related errors, all of which contribute to crashes.

Real-world driving performance and behaviors in time-to-collision can be examined using data from numerous geo-coded (location-based) driving parameters such as speed, acceleration, time to collision, secondary tasks/durations, distractions, and eye glance behavior of the drivers being continuously captured via advanced equipment (Dingus, Klauer et al. 2006, Dingus, Guo et al. 2016). All of these advanced features make NDS databases more unique and reliable compared to conventional police crash reports. Earlier studies have classified errors and developed taxonomies for driver errors and violations through different themes (Treat, Tumbas et al. 1979, Reason, Manstead et al. 1990, Petridou and Moustaki 2000, Wierwille, Hanowski et al. 2002, Stanton and Salmon 2009). Driver errors have been classified as slips, lapses, or mistakes (Reason, Manstead et al. 1990). In another study, the classification focused on differentiating behavioral factors affecting driving capabilities from those encouraging risky driving (Petridou and Moustaki 2000). From previous taxonomies, key driving errors include recognition errors, decision errors, and performance errors (Treat, Tumbas et al. 1979, Wierwille, Hanowski et al. 2002). Recognition and decision errors were found to be the leading and critical errors resulting in crashes (Treat, Tumbas et al. 1979, Wierwille, Hanowski et al. 2002, Iden and Shappell 2006). Due to variations in driver characteristics and behavioral patterns, different roadways and environments can induce different errors and violations (Dumbaugh and Li 2010).

While previous taxonomies of driving errors and violations have led to valuable insights, none (to the best of our knowledge) have utilized real-world naturalistic driving datasets which provide detailed pre-crash driver behavior information for crashes, near-crashes, and baseline driving. Therefore, there is a need to develop a detailed and systematic taxonomy of real-world safety-critical driving errors and violations. Thus, the key objective of this study is to develop a systematic taxonomy of driving errors and violations (TDEV) which explores how different driving errors and violations vary in different roadway and environmental contexts using naturalistic driving study data. In addition, the study aims to develop a safety matrix to understand the contributions of human, vehicle, and roadway/land use environment factors based on new evidence. Importantly, this study sheds light on the risks associated with driving errors and violations made during no-event driving situations (baselines), which can potentially lead to safety-critical events (crash or near-crash) in the future. These efforts can in turn help accident investigations and provide insights into the development of automated vehicles.

METHODOLOGY

Data source

This study uses data from the Naturalistic Driving Study, which were collected as part of the 2nd Strategic Highway Research Program (SHRP 2). The NDS SHRP2 is a comprehensive data collection effort in which a variety of volunteer drivers participated. Thousands of vehicles were equipped with advanced technologies (i.e., radars, sensors, and cameras) that continuously monitored and collected data on driver performance, driver behavior, speed, acceleration, lateral and longitudinal positions, and eye glancing behavior (Dingus, Guo et al. 2016). The NDS data consists of high-frequency and high-resolution information that was collected through an onboard data acquisition system (DAS) and multiple sensors (Hankey, Perez et al. 2016). The data includes roughly 3,200 drivers from six different states including New York, Washington, Pennsylvania, North Carolina, Indiana, and Florida (Hankey, Perez et al. 2016). The data was carefully collected in the appropriate study centers located in Bloomington, Indiana; Buffalo, New York; Durham, North Carolina; Seattle, Washington; State College, Pennsylvania; and Tampa, Florida. According to the NDS reports, these study centers encompassed several counties each with more than 21,000 mi^2 contained about 7.6 million registered vehicles of all types, and had a population of approximately 6.5 million people of driving age (greater than 15 years)-see Appendix Table A1 for geographic details of counties used for NDS recruitment. Efforts were made to have a representative sample of the population living in the counties surrounding the study areas, which included both rural and urban areas. Efforts were made to obtain a representative sample of driving population by age and gender.

They seem to be fairly representative of the population from which the sample was drawn. More details are provided in relevant SHRP reports (Blatt et al., 2015). A total of 3,247 volunteer participants were recruited for the NDS (SHRP-2) showing a good distribution among the aforementioned six regions (Blatt et al., 2015). Out of these 3,247 participants, the distribution was Bloomington 7.82% (N=254), Buffalo 22.79% (N=740), Durham 16.29% (N = 529), Seattle 22.02% (N = 715), State College 8.47% (N = 275), and Tampa 22.61% (N =734) (Blatt et al., 2015)—see Table in Appendix. These participants included around 45.80% (N = 1,487) males and 50% (N = 1,624) females, which were fairly distributed among various age groups which include: minor teens (with 16-17 years of age), adult teen (18-20), young adult (21-25), adult (26-35), middle adult (36-50), mature adult (51-65), older drivers (66-75), and older drivers that are 75+. Details on age distribution are available in the reference (Blatt et al., 2015). Interestingly, there were 4.20% (N = 145) of the vehicles called AVT (Advanced vehicle Technology), equipped with advanced features such as collision avoidance radar, advanced cruise control, and electronic stability control (Blatt et al., 2015). Participants using these vehicles could be of any age or gender; for details, please refer to Blatt et al. (2015). While the NDS (SHRP2) data includes detailed information on all of the aforementioned participants, these variables were not included in the event sub-sample available for this study. Summary on these key variables for the overall NDS data can be found in the Appendix (i.e., Tables A2, A3, and A4).

This study uses a subset of the original NDS-SHRP2 data, including a total of 9,593 trips with 7,589 no-event baselines (20 to 30 seconds each), 1,331 near-crashes, and 673 crashes. The baselines (i.e., driving instances when no safety-critical event happened) in the NDS data provide an opportunity to compare them with crashes and near crashes that represent safety-critical events. Although not fully representative of exposure, the baselines can provide a sense of crash risk through comparison of pre-crash behaviors (Hankey, Perez et al. 2016). The selection of baselines is one of the key new data sources in NDS. Through baselines, researchers can get a sense of typical driving behaviors across the sample (Hankey, Perez et al. 2016). In order to select baselines, a baseline sampling method was developed by VTTI. The sampling method was presented for expert review by the Expert Technical Group (ETG) who conduct review for the NDS studies as well as from the NDS (SHRP 2) Technical Coordinating Committee (Hankey, Perez et al. 2016). In NDS, at least one baseline was selected for every driver with typical duration of 20-30 seconds. Driving speeds of 5 miles per hour were only included to mitigate the long stopping (Hankey, Perez et al. 2016). For more detail on baseline selection and specification, please refer to a study (Hankey, Perez et al. 2016).

Notably, there is sufficient variation in the baselines based on roadway and land use variables. Also, the baselines are simply a sample (20 seconds to 30 seconds) of trips where drivers did not experience any abnormal or safety-critical event. They are helpful in assessing the behavior of drivers when no safety-critical events are occurring. The variables included in the NDS data are generally classified into three classes (Hankey, Perez et al. 2016):

1. Safety-critical event variables, e.g., event nature, event severity, precipitating events, pre-incident maneuvers, and drivers' reaction.
2. Driver variables, e.g., driving behavior, type and duration of driver distraction/secondary tasks, seatbelt usage, and drivers' steering control.
3. Roadway and Land Use Variables, e.g., locality, roadway alignment, traffic flow, traffic density, and traffic control devices (Hankey, Perez et al. 2016).

It should be noted that this study uses a non-random subsample of NDS data which was accessible to us for this analysis. In the NDS, the VTTI team randomly selected baselines with the goal of selecting at least one baseline for every driver (involved in a safety critical event) with a typical duration of 20-30 seconds. Note that the number of baselines for a specific driver was proportional to his/her total driving time in the NDS experiment (Hankey, Perez et al. 2016). Driving speeds less than 5 miles per hour were excluded in order to disregard the influence of long stop times and to only consider time periods where the subject vehicle had the possibility of an at-fault crash (Hankey, Perez et al. 2016).

To understand how safety critical events (crashes and near-crashes) were identified in NDS data, we provide the definitions as per the NDS dictionary.

Crash: In NDS data, crash is considered as any contact of subject vehicle with an object (fixed or moving) at any speed which results into transfer or dissipation of kinetic energy. It also includes non-premeditated departures of the roadway in which at least one tire of the subject vehicles leaves intended travel or paved surface of roadway. Furthermore, crash is considered to have occurred if the subject vehicle strikes any other vehicle, object, roadside barrier, animal, bicyclist, or pedestrian.

Near-Crash: Near-crash is a situation where the subject vehicle, or other vehicle, cyclist, pedestrian, or animal requires a rapid evasive maneuver to escape crash. Note that rapid evasive maneuver includes accelerating, braking, steering, or combination of these three maneuvers.

Safety Matrix

A safety matrix was developed to provide insight into the prevalence of safety-critical human, vehicle, and roadway factors. The key objective for developing the safety matrix is to quantify the contributions of key factors in several combinations. Only the factors in the NDS data which data reductionists coded and clearly state that they contribute to crashes were considered. Note that the data reductionists who reviewed and analyzed the driving and video data used specific procedures. For instance, the procedures in the General Estimates System (GES) compiled by the National Highway Traffic Safety Administration were used as a reference for developing variables during data extraction from the NDS videos. This ensured that the key aspects of crashes are considered by the data reductionists and consistency in documenting variables. In this study, we use driver behavior (e.g., driving slowly and below the speed limit) and secondary tasks (given that secondary tasks contributed to crashes) to assess the contribution of human factors in crashes. Similarly, various factors in the NDS data were used to quantify the roadway and environmental factors. A detailed discussion on factors in the NDS data, which were used to derive crash contributing factors related to human, roadway/environment, and vehicle are discussed below.

Human Factors

Driver Behavior: The "Driver Behavior" factor in NDS data is defined by data reductionists as: "Driver behaviors (those that either occurred within seconds prior to the Precipitating Event or those resulting from the context of the driving environment) that include what the driver did to cause or contribute to the crash or near-crash. Behaviors may be apparent at times other than the time of the Precipitating Event, such as aggressive driving at an earlier moment which led to retaliatory behavior later." For details on driver behavior in the NDS data, please refer to Figure 2.1.

Secondary Task and Secondary Task Outcome: In some crashes no driver behavior was reported to have contributed to the crash. In such cases, we checked the "Secondary Task" and "Secondary Task Outcome" factors in the NDS data to determine whether a particular secondary task contributed to a crash. To determine whether a secondary task contributed to a crash, we evaluated the "secondary task outcome" in the NDS data which is defined by the NDS data reductionists as: "Determination of whether the Secondary Task contributed to the event sequence and severity. (Not whether the factor actually caused the event but contributed to it.)" These secondary tasks include cell phone use, talking/listening, and using hands-free communication, non-specific internal eye glance, and doing personal hygiene while driving.

Roadway and Environment Factors

The roadway and environmental factors include all of those factors which refer to roadway, weather, and visual obstructions which data reductionists state as crash contributors in NDS data. These roadway and environmental factors are separate from the "locality" factors in the NDS data which simply indicate in which roadway or land use environment the baseline or safety-critical event (near-crash or crash) was observed. The roadway and environment factors, which contributed to crashes, are derived from various factors in the NDS data including those coded as infrastructure, visual obstructions, surface conditions, and environmental conditions discussed as follows:

- **Infrastructure:** The NDS data dictionary defines these factors as possible contributing causes to the occurrence or severity of the safety-critical event (i.e., in this case a crash). Examples of such factors include roadway alignment, roadway delineation, traffic control device, and roadway sight distance.
- **Visual Obstructions:** Most of the visual obstruction factors relate to the blind spots or sight distance issues in the roadway. According to the NDS data dictionary, they had contributed to the occurrence

and severity of the safety-critical-event (i.e., in this case crash) or influenced the ability of the subject driver to effectively recognize and respond to the possible safety hazard and precipitating event. Some of the visual obstructions reported in the NDS crashes include curve or hill, inadequate roadway lighting system, parked vehicle, and other obstructions.

- **Surface Conditions:** In some crashes, surface conditions such as snowy, muddy, and oily, were reported to have affected the vehicle's coefficient of friction at the start of the precipitating event which resulted in crashes.

Vehicle Factors

This study considers the "vehicle contributing factors" available in the NDS data including faulty tires and brake system as the vehicle-related crash contributing factors. Based on NDS data reductionists, such factors might have contributed to the precipitating event or subject driver's response which resulted in crashes. Some of the factors in the "Visual Obstructions" include inadequate defrost or defog system, inadequate vehicle headlamps, and faulty headlights, which can be more appropriately categorized as vehicle factors. Hence this study classified such factors as vehicle factors, which might have contributed to the occurrence or severity of the safety-critical events (i.e., a crash).

A Systematic Taxonomy for Driving Errors and Violations

This study uses NDS data to develop an evidence-based taxonomy for driving errors and violations. These are classified using a spatio-temporal framework consisting of perception, recognition, decision, and reaction (PRDR) (Khattak et al. 2020). The PRDR process consists of four phases: perception, recognition, decision, and reaction. During the perception phase, a driver perceives the presence of a potential safety hazard in the roadway through sight. Recognition is the next phase, where a driver realizes the actual identity of the safety hazard. Next comes the decision phase in which a driver decides the course of action to overcome the potential conflict. Finally, a driver initializes the course of action (i.e., start moving foot to the brake pedal or hand to the steering). While we developed and used the TDEV for driving errors and violations in this study, it can also be used to systematically classify driving behavior in near-crashes and baselines (Figure 2.1). Brief definitions of errors and violations are provided as follows:

- Recognition errors occur when a driver fails to appropriately recognize the real situation, either due to distraction or poor judgement, requiring an evasive maneuver. This variable is coded by data reductionists (by reviewing video data) when proper attention is not maintained immediately prior to a safety-critical event. Specifically, according to the NDS data dictionary, in such cases the subject driver was reported as unable to maintain appropriate level of attention due to his/her involvement in one or more secondary tasks which occurred within seconds prior to the start of the precipitating event. The NDS data dictionary defines such situations as "the state of environment or action that began the event sequence under analysis". The recognition error variable can be indicative of the higher risk associated with distraction compared with if the subject driver was not distracted. If the driver was not distracted, then they could have effectively responded to risky situation.
- Decision errors happen when a driver is unable to decide or choose the appropriate response or action (e.g., brake or accelerate), having perceived, interpreted, and recognized the information correctly in the perception and recognition phase. For instance, some of the driver behaviors such as driving slowly below speed limit and exceeding safe speed but not the speed limit indicate inappropriate decisions of the subject drivers which the NDS data reductionists considered to have contributed to the crash occurrence or severity, which we classify as decision failures. For details on relevant driver behaviors related to decision errors, please refer to Figure 2.1.
- Driver performance errors relate to a driver's response in the reaction phase often reflecting poor lateral or longitudinal control or weak judgement of driving situations. For instance, if a driver intended to accelerate but did not accelerate enough, which resulted in a crash, then this would indicate a performance failure for the subject driver. If the driver had executed their intended course of action, i.e., accelerated enough, then they could have avoided the crash.
- Physical condition errors are motivated by a driver's physical or psychological state, e.g., drowsiness potentially leading to unsafe situations.
- Experience or exposure errors can occur when a driver is either unfamiliar with the roadway and surroundings or lacks driving experience. This variable is determined and reported by the NDS data

reductionists. Further detail about the variable can be found in the “Driver Behavior” factor provided in the NDS data dictionary.

- Violations of traffic laws can be categorized into sub-classes that include intersection-related violations (which can be intentional or unintentional), segment related violations, or simply speed related or illegal maneuvers.

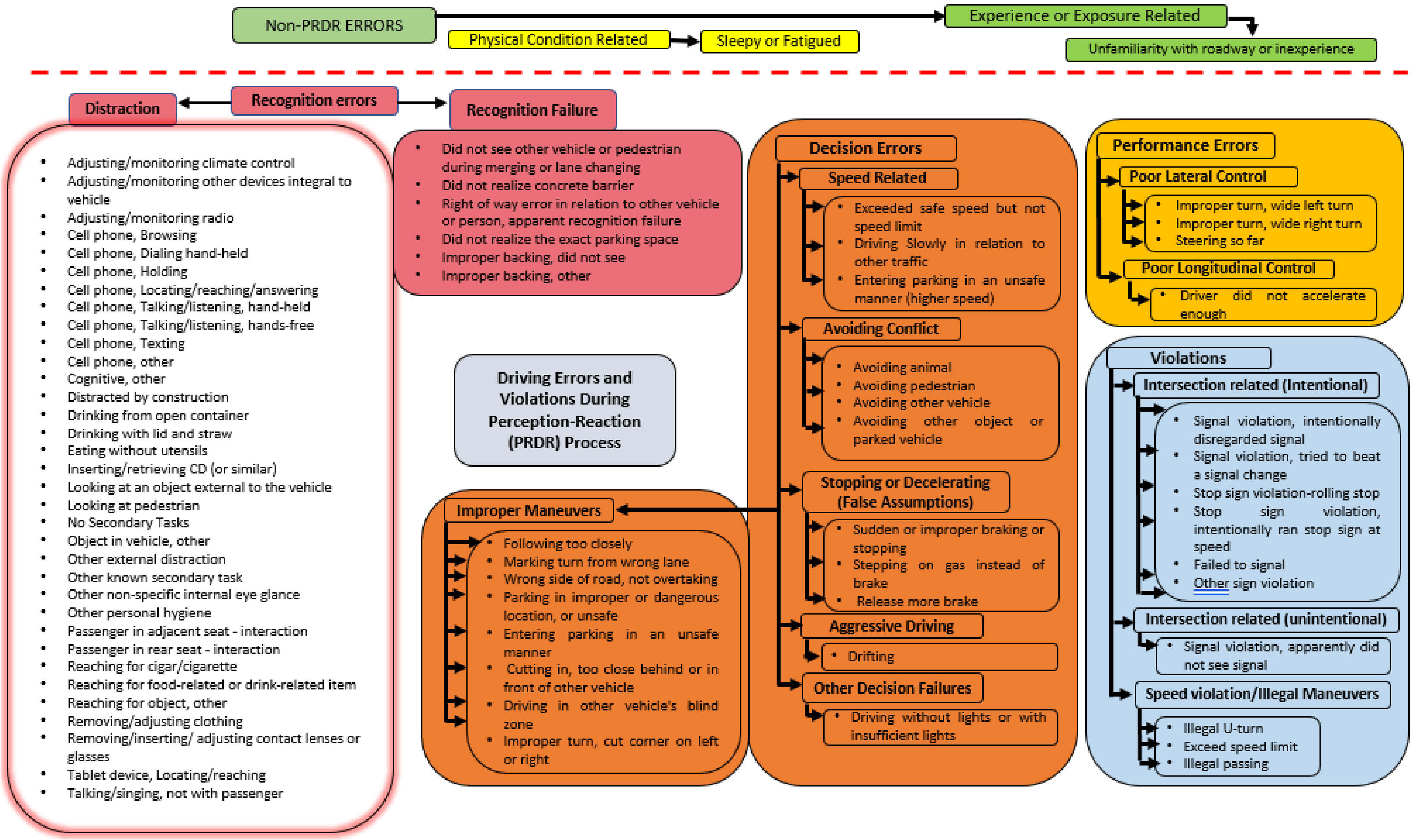

Figure 2.1. Systematic Taxonomy for Driver Errors and Violations

The classification in Figure 2.1 is based on the “Driver Behavior” factors in the NDS data which are defined by the data reductionists as: “Driver behaviors (those that either occurred within seconds prior to the Precipitating Event or those resulting from the context of the driving environment) that include what the driver did to cause or contribute to the crash or near-crash. Behaviors may be apparent at times other than the time of the Precipitating Event, such as aggressive driving at an earlier moment which led to retaliatory behavior later.” In some of the safety-critical events, “no driving behavior” was reported. In such cases, we checked for “secondary task outcome” which is defined as: “Determination of whether the Secondary Task contributed to the event sequence and severity. (Not whether the factor actually caused the event but contributed to it.)” All of the “Driver Behavior” variables in the above TDEV are reported in the NDS data. There were a few cases where no contributing factor was mentioned, i.e., none of the driver behavior, secondary task outcome, infrastructure, roadway surface condition, or vehicle obstruction were identified in the NDS data as (marked with an asterisk (*)). In such cases, the research team evaluated the narrative (detailed description) of such cases to determine if any other human, vehicle, and roadway or environment factors could be identified as contributors. After evaluating the detailed descriptions in the narratives for the few cases, the research team found driver-related contributing factors and classified them in accordance with the proposed TDEV (Figure 2.1).

Diverse Roadway Localities: NDS Data

In the NDS data available to the research team, roadway and land use environments are classified under the variable “locality” into various categories (Table 2.1). The instructions for the data reductionist from the SHRP2 researcher data dictionary for this variable is as follows: “Best description of the surroundings that influence or may influence the flow of traffic at the time of the start of the precipitating event. If there are ANY commercial buildings, indicate as business/industrial or urban area as appropriate (these categories take precedence over others except for church, school, and playground). Indicate school, church, or playground if the driver passes one of these areas (or is imminently approaching one) at the same time as the beginning of the Precipitating Event (these categories take precedence over any other categories except urban, and divided highway).” The final variable includes the following categories: interstate/bypass/divided highway with no traffic signals, bypass/divided highways with traffic signals (principal arterials), open residential (which include few houses, signifying largely undeveloped land use), moderate residential, school, church, playground, business/industrial, urban, and open country (rural) settings (Table 2.1). Importantly, the open country and open residential locations are defined to include only vegetation and one/few housing units respectively. According to the NDS dictionary, open country and open residential areas are regarded as rural and rural/semi-rural locations respectively.

Table 2.1 is presented in original format (source: NDS data dictionary) to understand the coding of various roadway localities. As defined in the table, Open Country indicates rural locations, while Open Residential (including one or few housing units) locations indicate rural or semi-rural locations are combined and considered as one category in the modeling in Chapter 3. The percentage of all events (i.e., including baselines, near-crashes, and crashes) reported on roadways within church, playground, and other locations like campground were very low; hence these three categories are collectively considered to have useful insights (also playground and campground locations somehow include similar activities which can influence driving behaviors in somehow similar way).

Table 2.1. Definition of Various Roadway Localities in the NDS Data (Source: NDS Data Dictionary)

Value	Definition	Example and Hints
Open country	Other than the roadway, there is nothing but vegetation visible during the time surrounding the Precipitating Event that is described in any of the other categories. Road is not an Interstate or a bypass/divided highway with traffic signals. (Often appears as rural roads, 2 lanes undivided.)	Includes roadways not defined as Interstate or divided highway, when no landmarks mentioned in other categories are visible.
Open Residential	Rural to semi-rural areas where there may be only one or a few houses around (i.e., farmland).	
Moderate Residential	An area where multiple houses or apartment buildings are present, but is not as dense as an Urban Locality.	e.g., residential subdivisions
Business/industrial	Any type of business or industrial structure is present, but is not as dense as an Urban Locality. (If there are also houses visible, this category takes precedence over Open residential and Moderate residential).	
Church	One or more involved vehicle passes a church building at the time of the Precipitating Event.	
Playground	One or more involved vehicle passes any type of playground or children's playing field at the time of the Precipitating Event.	If playground/field is on school grounds, code as School.
School	One or more involved vehicles passes any type of school building or is in a school zone at the time of the Precipitating Event, including adult learning institutions.	Include any training centers, universities, etc. as well as elementary and secondary schools.
Urban	Higher density area where blocks are shorter, streets are a mix of one and two way, and traffic can include buses and trams. (This category takes precedence over others when either businesses and/or residences are present.)	
Interstate/bypass/ divided highway with no traffic signals	Vehicle is travelling on an interstate, bypass, or divided highway with no traffic signals (regardless of what buildings can be seen), at the time of the Precipitating Event.	
Bypass/divided highway with traffic signals	Vehicle is travelling on a bypass or divided highway with traffic signals (no other category description is visible) at the time of the Precipitating Event. (Often appears as "Open Country", but with more lanes and/or as a divided road.)	
Other	Locality at the time of the Precipitating Event is one not described in other categories.	Ex. In campground.
Unknown	Cannot determine the Locality due to limitations in video views, lighting, visual obstructions, or limited perspective.	Ex. Part of the video is missing or there is insufficient information in the video to make a determination.

RESULTS & DISCUSSION

Most of the crashes (N = 673) in the NDS data available to the research team were not severe. The distributions and their severity based on NDS data dictionary are as follows:

- Severe Crashes, 8.92% (N=60): “Any crash that includes an airbag deployment; any injury of driver, pedal cyclist, or pedestrian; a vehicle roll over; a high Delta V; or that requires vehicle towing. Injury if present should be sufficient to require a doctor's visit, including those self-reported and those apparent from video. A high Delta V is defined as a change in speed of the subject vehicle in any direction during impact greater than 20mph (excluding curb strikes) or acceleration on any axis greater than +/-2g (excluding curb strikes).”
- Police-reportable Crashes, 13.22% (N=89): “Police-Reportable Crash. A police-reportable crash that does not meet the requirements for a Level I crash. Includes sufficient property damage that it is police reportable (minimum of $1500 worth of damage, as estimated from video). Also includes crashes that reach an acceleration on any axis greater than +/-1.3g (excluding curb strikes). If there is a police report this will be noted. Most large animal strikes and sign strikes are included here.”
- Minor Crashes, 37.59% (N=253): “Physical Contact with another Object. Most other crashes not included above are Level III crashes, defined as including physical contact with another object but with minimal damage. Includes most road departures (unless criteria for a more severe crash are met), small animal strikes, all curb and tires strikes potentially in conflict with oncoming traffic, and other curb strikes with an increased risk element (e.g., would have resulted in worse had curb not been there, usually related to some kind of driver behavior or state).”
- Low-risk Tire Strikes, 40.27% (N=271): “Tire Strike, Low Risk. Tire strike only with little/no risk element (e.g., clipping a curb during a tight turn).”

On the one hand, with a prevalence of minor and low-risk/tire strike crashes (78%), a wide range of safety-critical events are captured, most of which would not be available through police-reported crashes. On the other hand, only 22.14% of the overall crashes are police reportable, which means that the NDS data are not fully comparable with the widely available police-reported data. Simple cross-tabulations (not reported) showed that recognition errors were more frequently associated with severe crashes (51%) compared with other errors and violations.

The near crashes are defined as follows in the NDS dictionary: “Any circumstance that requires a rapid evasive maneuver by the subject vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. Near Crashes must meet the following four criteria: 1. Not a Crash. The vehicle must not make contact with any object, moving or fixed, and the maneuver must not result in a road departure. 2. Not premeditated. The maneuver performed by the subject must not be pre-meditated. This criterion does not rule out Near Crashes caused by unexpected events experienced during a pre-meditated maneuver (e.g., a premeditated aggressive lane change resulting in a conflict with an unseen vehicle in the adjacent lane that requires a rapid evasive maneuver by one of the vehicles). 3. Evasion required. An evasive maneuver to avoid a crash was required by either the subject or another vehicle, pedestrian, animal, etc. An evasive maneuver is defined as steering, braking, accelerating, or combination of control inputs that is performed to avoid a potential crash.4. Rapidity required. The required evasive maneuver must also require rapidity. Rapidity refers to the swiftness of the response required given the amount of time from the beginning of the subject’s reaction and the potential time of impact. Events classified as Near Crashes generally undergo further analysis.”

Safety Matrix

Crashes were categorized as having driver, vehicle, and roadway factors that contributed to their occurrence. Furthermore, the PDRT framework was used to characterize human factors. The roadway factors represent the data reductionist’s judgment on infrastructure-based contributing factor to the “…occurrence and severity of the event, wherein some aspect of the roadway design impacted the driver's ability to safely navigate the roadway, recognize potential safety risks, or respond effectively to the Precipitating Event.” (Virginia Tech Transportation Institute, 2019). A vehicle factor can be a “… vehicle defect or factor appeared to contribute to the occurrence of the Precipitating Event." (Virginia Tech Transportation Institute, 2019). These can include defect in tires, wheels, signals, powertrain, suspension, braking system, steering system, wipers, headlights, etc. All factors were based on the data reductionists review of the videos and other recorded information from the Data Acquisition System.

The safety matrix for the NDS crashes reveals that human factors were the sole contributing factors in 77.56% crashes (Table 2.2). The statistics indicate that human factors have a contribution in around 93% of crashes (Figure 2.2), this finding is consistent with previous studies (Sabey and Staughton 1975, Treat, Tumbas et al. 1979, Singh 2015). Roadway factors were the second most prevailing factors, contributing to around 17% of crashes (Figure 2.2). Vehicle factors were found to have the smallest impact on the number of crashes that were observed in the NDS data, with a percentage share of around 1%. This can be likely attributed to enhanced and modern vehicle technology in the NDS sample.

Table 2.2. Safety Matrix: Contribution of Human, Vehicle, and Roadway Factors

Human Factors	Vehicle Factors	Roadway Factors	Crash	
			Freq.	%
Y	Y	Y	2	0.30
Y	Y	N	5	0.74
Y	N	N	522	77.56
N	N	N	30	4.45
N	N	Y	18	2.67
N	Y	Y	0	0.00
N	Y	N	0	0.00
Y	N	Y	96	14.26
Total			673	100.0

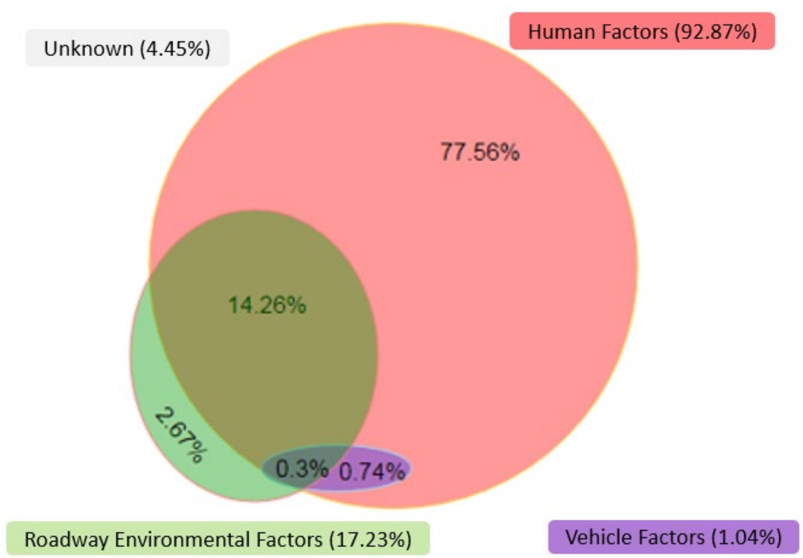

Figure 2.2. Contribution of Human, Vehicle, and Roadway Factors in Crashes

Prevalence of Driving Errors in Safety-critical Events: Evidence from Naturalistic Driving Environments

Using the NDS data, Table 2.3 provides summary statistics for driving errors and violations for safety-critical events (crashes and near-crashes) and baselines (driving when no safety-critical event occurs). The methodology section explained how the error variables were created. The NDS data statistics show that at least one driving error or violation was reported in 92.87% of crashes and 61.31% of near-crashes respectively (Table 2.3). Recognition errors were the predominant driving errors that contributed in crashes and near-crashes with a percentage share of 38.63% and 34.03%, respectively (Table 2.3). Decision errors were reported as the second leading driving error, causing 34.32% and 13.82% crashes and near-crashes respectively (Table 2.3). Statistics reveal that drivers were involved in traffic violations in 9.06% of crashes and 10.74 of near-crashes. The safety hazard associated with different types of driving errors and violations was quantified using the ratio of the percentage share for each key error type as crashes/near-crashes over percent share in baselines as presented in columns 5 and 6 (Table 2.3). The findings reveal substantial safety risks are associated with recognition errors, as the percent share of recognition errors in near-crashes and crashes were 155 and 176 times of their percent share in baselines (Table 2.3). Decision errors were reported in 2.69% (N = 204) of total baselines, and the percent share of decision errors in near-crashes and crashes were 5 and 13 times their percent share in baselines. If these decision or recognition errors and violations are constantly repeated by drivers, then they can result in a safety-critical event.

It is important to point out that the NDS data captures safety-critical events that did not necessarily result in police-reportable crashes. For example, according to the NDS data dictionary, the “impairments” factor is defined to influence the “driver behavior.” However, in only in about 1.04% of the crashes (N = 9 out of 673 crashes), the subject drivers were reported to have been impaired with alcohol or other drugs. These impairments were in turn reflected in driver error variables and classified as recognition errors (N = 4 crashes) and decision errors (N = 3 crashes).

The remarks above provide an overview of how driving errors and violations are distributed among crashes and near-crashes. Recognition errors and decision errors contribute more frequently to crashes than other types of errors. Furthermore, crashes and near-crashes can be separated into single vehicle (SV) and multi-vehicle (MV). While the contribution of a driver can be relatively straightforward in SV events, this may not be the case in MV crashes or near crashes. Specifically, in MV events, the human factors contributions may come from only the subject driver, other driver/s, both subject driver and other driver/s, or none of the drivers. Due to the nature of NDS data, this research focuses on exploring the pre-crash driving behaviors and contributions of subject drivers.

Table 2.3 shows the distribution of driving errors and violations of subject drivers across SV and MV crashes and near-crashes. Out of the 673 crashes in the NDS subsample, 512 (76.08%) and 161 (23.92%) involve single vehicles and multiple vehicles, respectively (Table 2.3). Recognition errors of the subject drivers contributed to the highest percentage (57.76%) of MV crashes compared with SV crashes (32.62%). Similarly, in near-crashes involving multiple vehicles, recognition errors (e.g., distraction) by subject drivers were the leading contributing factors resulting in 35.47% of the MV near-crashes compared with 18.97% in SV near-crashes. These findings indicate that recognition errors dominate SCEs particularly in MV crashes. Recognition failures, e.g., due to distractions of the subject drivers seems to substantially contribute to harmful events involving multiple vehicles. Decision errors were more dominant in SV crashes, resulting in 53.45% and 41.21% SV crashes and near-crashes, respectively (Table 2.3). In SV crashes and near-crashes, poor decisions are the key contributors to safety critical events. The prevalence of errors in safety critical events relative to baselines is shown in the last two columns of Table 2.3. The results indicate that recognition errors are much more likely to result in crashes, followed by performance errors. For instance, the percent contribution of recognition errors in crashes is 175.59 (= 38.63/0.22) times their percent contribution in baselines (Table 2.3). Similarly, the percent contribution of performance errors in crashes is 84.22 times their percent contribution in baselines (Table 2.3). Other errors show smaller relative contributions.

To understand the contribution of non-subject drivers in multi-vehicle crashes, the fault variable provides insights. In 45 MV crashes where driving behavior (e.g., distraction, exceed speed limit) contributed to crash occurrence, other drivers were at-fault (i.e., coded in the “Fault” variable in NDS data), indicating that 27.95% of MV crashes (N = 161) involved fault by non-subject drivers.

Table 2.3. Prevalence of Driving Errors in Safety-critical Events (SCEs)

Variable	Event Type							Prevalence of Errors in SCEs	
	Baseline (%)	Near-Crash (%)			Crash (%)			$\frac{\text{\% in Near Crash}}{\text{\% in Baseline}}$	$\frac{\text{\% in Crash}}{\text{\% in Baseline}}$
	(N =7,589)	Overall (N=1,331)	SV (N = 116)	MV (N = 1,215)	Overall (N=673)	SV (N = 512)	MV (N = 161)	Overall Near crashes (N = 1,331)	Overall Crashes (N = 673)
Driver's Errors and Violations									
No error or violation	90.12	38.69	6.90	41.73	7.13	5.08	13.66	0.43	0.08
Recognition error	0.22	34.03	18.97	35.47	38.63	32.62	57.76	154.68	175.59
Decision error	2.69	13.82	53.45	10.04	34.32	41.21	12.42	5.14	12.76
Performance error	0.09	0.68	2.59	0.49	7.58	7.42	8.07	7.56	84.22
Physical condition error	1.25	1.43	2.59	1.32	1.34	1.76	---	1.14	1.07
Experience/exposure error	0.07	0.6	1.72	0.49	1.93	2.34	0.62	8.57	27.57
Violation	5.56	10.74	13.79	10.45	9.06	9.57	7.45	1.93	1.63
Total	100%	100%	100%	100%	100%	100%	100%	---	---

Notes: SCEs indicate Safety-critical Events (near-crash, or crash). The ratio of percentage of errors in near-crashes and crashes to those in baselines are determined using values from the columns titled as baselines, overall near-crashes (N = 1,331) and overall crashes (N = 673) as presented in the two right-most columns, respectively.

Variations of Driving Errors and Violations across Localities

This section presents how driving errors and violations vary across roadway and environmental variables such as business/industrial and residential locations. Notably, local contexts may be associated with specific driving errors, and this analysis allows for the identification of hotspots of frequent errors and violations, which result in safety-critical events. The NDS subsample includes a total of 673 crashes out of which 38.63% (N = 260) and 34.32% (N = 231) crashes occurred due to recognition errors and decision errors, respectively (Table 2.3). As a next step, we were interested to see that how the two key driving errors (i.e., recognition and decision errors) are distributed across different roadway environments. We found that driving errors and violations resulting in crashes, occurred more frequently when business or industrial structures were present. Specifically, recognition and decision errors occurred more frequently with 47% (N = 121/260*100) and 52% (120/231*100) respectively, when business or industrial structures were present and thus resulted in crashes (Table 2.4). The next most frequent category was near moderate residential developments. To some extent these frequencies may reflect the exposure of the NDS sample in residential and business or industrial areas. While direct exposure of the vehicles in the sample is not available to researchers, the baselines provide a coarse surrogate for exposure. Therefore, crash percentage divided by baseline percentage is calculated to get a sense of the rate (Table 2.5). Also, business or industrial areas may be characterized by complexity of traffic, diversity of activities, and special roadway and environmental conditions that may lead to driving errors and subsequent crashes.

Table 2.4. Distribution of Driving Errors and Violations across Different Roadways and Environments in Crashes, Near-Crashes, and Baselines

Locality	Types of Driving Errors and Violations															
	No Error		Recognition		Decision		Performance		Violation		Physical Condition		Experience		Total	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
	Baselines (N = 7,589)															
Interstate	1,743	25.49	1	5.88	52	25.49	0	0.00	125	29.62	22	23.16	0	0.00	1,943	25.60
Open Country or Open Residential	654	9.56	0	0.00	25	12.25	0	0.00	31	7.35	8	8.42	0	0.00	718	9.46
Moderate Residential	1,502	21.96	6	35.29	49	24.02	3	42.86	105	24.88	25	26.32	0	0.00	1,690	22.27
School	305	4.46	0	0.00	12	5.88	1	14.29	18	4.27	5	5.26	1	20.00	342	4.51
Business/Industrial	2,214	32.37	8	47.06	51	25.00	3	42.86	119	28.20	27	28.42	3	60.00	2,425	31.95
Urban	103	1.51	0	0.00	5	2.45	0	0.00	5	1.18	1	1.05	1	20.00	115	1.52
Bypass or Divided Highways with no Traffic Signals	203	2.97	2	11.76	7	3.43	0	0.00	13	3.08	6	6.32	0	0.00	231	3.04
Others (e.g., Church, Playground and Campground)	115	1.68	0	0.00	3	1.47	0	0.00	6	1.42	1	1.05	0	0.00	125	1.65
Total	6,839	100	17	100	204	100	7	100	422	100	95	100	5	100	7,589	100
	Near-Crashes (N = 1,331)															
Interstate	104	20.19	74	16.34	36	19.57	0	0.00	26	18.18	7	36.84	3	37.50	250	18.78
Open Country or Open Residential	19	3.69	17	3.75	18	9.78	1	11.11	16	11.19	1	5.26	1	12.50	73	5.48
Moderate Residential	67	13.01	73	16.11	41	22.28	2	22.22	20	13.99	3	15.79	1	12.50	207	15.55
School	35	6.80	38	8.39	13	7.07	0	0.00	8	5.59	0	0.00	0	0.00	94	7.06
Business/Industrial	214	41.55	193	42.60	57	30.98	6	66.67	58	40.56	4	21.05	1	12.50	533	40.05
Urban	45	8.74	31	6.84	10	5.43	0	0.00	7	4.90	3	15.79	1	12.50	97	7.29
Bypass or Divided Highways with no Traffic Signals	18	3.50	10	2.21	6	3.26	0	0.00	5	3.50	1	5.26	0	0.00	40	3.01
Others (e.g., Church, Playground and Campground)	13	2.52	17	3.75	3	1.63	0	0.00	3	2.10	0	0.00	1	12.50	37	2.78

Total	515	100	453	100	184	100	9	100	143	100	19	100	8	100	1331	100
	Crashes (N = 673)															
Interstate	9	18.75	15	5.77	13	5.63	1	1.96	2	3.28	0	0.00	1	7.69	41	6.09
Open Country or Open Residential	2	4.17	11	4.23	14	6.06	2	3.92	10	16.39	1	11.11	0	0.00	40	5.94
Moderate Residential	7	14.58	54	20.77	36	15.58	17	33.33	15	24.59	4	44.44	5	38.46	138	20.51
School	2	4.17	26	10.00	17	7.36	1	1.96	5	8.20	1	11.11	1	7.69	53	7.88
Business/Industrial	24	50.00	121	46.54	120	51.95	19	37.25	19	31.15	2	22.22	6	46.15	311	46.21
Urban	2	4.17	21	8.08	18	7.79	5	9.80	6	9.84	1	11.11	0	0.00	53	7.88
Bypass or Divided Highways with no Traffic Signals	2	4.17	6	2.31	4	1.73	3	5.88	2	3.28	0	0.00	0	0.00	17	2.53
Others (e.g., Church, Playground and Campground)	0	0.00	6	2.31	9	3.90	3	5.88	2	3.28	0	0.00	0	0.00	20	2.97
Total	48	100.00	260	100.00	231	100.00	51	100.00	61	100.00	9	100.00	13	100.00	673	100.00

Table 2.5 provides the distribution of crashes, near-crashes, and baselines across different roadways and environments. The NDS data available for this study indicates that a total of 673 crashes were reported which are distributed across diverse roadways and environments. The highest percentage of crashes occurred on roadways passing through business or industrial localities and moderate residential locations making up 46.21% and 20.50% of the overall crashes. The percentage of baselines reported in these areas were also greater compared to other locations. Although the percentage of baselines reported on interstates was the second highest, the percentage of crashes on Interstates was relatively lower compared to other locations such as urban environments, school zones, and business or industrial.

Notably, the NDS data available to the team does not have information on direct exposure (e.g., vehicle miles traveled), which could help in more appropriate comparison of crashes across different roadways and environments, a coarse measure in the form of baselines is available. Therefore, the percentage of baselines in different localities is used as the best possible but admittedly coarse surrogate for exposure to compare the safety risk across a diverse set of locations (Table 2.5). Based on the percentage of crashes per percentage of baselines in a specific locality, Interstate roadways and open country/open residential areas (rural and semi-rural settings) associate with lower risks. While urban, business/industrial, and school locations seem to have higher percentage of crashes per percentage of baselines and they can be considered as high crash risk locations (Table 2.5).

TABLE 2.5. Distribution of Crashes, Near-Crashes, and Baselines across Different Roadways and Environments

VARIABLE	All Cases (N = 9,593) Percent (S.D.)	Baseline (%) (N = 7,589)	Near-Crash (%) (N = 1,331)	Crashes (%) (N = 673)	$\frac{\text{\% of Crashes}}{\text{\% of Baselines}}$
ROADWAY LOCALITY					
Interstate	23.29 (0.4226)	25.6	18.78	6.09	0.2379
Open Country or Open Residential	8.66 (0.2813)	9.46	5.48	5.94	0.6279
Moderate Residential	21.21 (0.4088)	22.26	15.55	20.50	0.9209
School	5.10 (0.2199)	4.51	7.06	7.87	1.7450
Business/Industrial	34.08 (0.4739)	31.10	40.04	46.21	1.4859
Urban	2.76 (0.1639)	1.51	7.30	7.88	5.2185
Bypass or Divided Highways with no Traffic Signals	3.00 (0.1760)	3.04	3.00	2.53	0.8322
Others (e.g., Church, Playground and Campground)	1.90 (0.1364)	1.66	2.78	2.98	1.7952
Total	100%	100%	100%	100%	

This study explores the distribution of various driving errors and violations specifically across localities, roadways and land use categories coded in the NDS data. Notably, 6.09% (N = 41) crashes occurred in localities coded as interstates (Table 2.4). The distribution of various driving errors and violations on interstates, contributing to crashes, is shown in Figure 2.3. It can be seen that recognition and decision errors contributed to 36.59% and 31.71% of the overall crashes when locality was coded as interstates. The findings are in line with a recent study which indicates that chance of making decision errors was higher on interstates (Shaon, Qin et al. 2018). Violations were reported in nearly 4.88%, while each performance and experience/exposure errors were reported to have contributed to 2.44% of the crashes (Figure 2.3). Referring to the distribution of driving errors and violations in crashes occurring on bypass/divided highways with traffic signals, 35.29% of the crashes on this roadway occurred due to recognition errors, followed by decision and performance errors which contributed to 23.53% and 17.65% respectively (Figure 2.3). Notably, performance errors were substantially higher on these bypass/divided highways with traffic signals. This finding is in agreement with recent studies (Shaon, Qin et al. 2018). This could be due to complex interaction of drivers with other roadway entities specifically at signalized intersections on principal arterials. Moreover, a good percentage (nearly 12%) of the crashes on bypass/divided highways with traffic signals occurred due to violations (Figure 2.3).

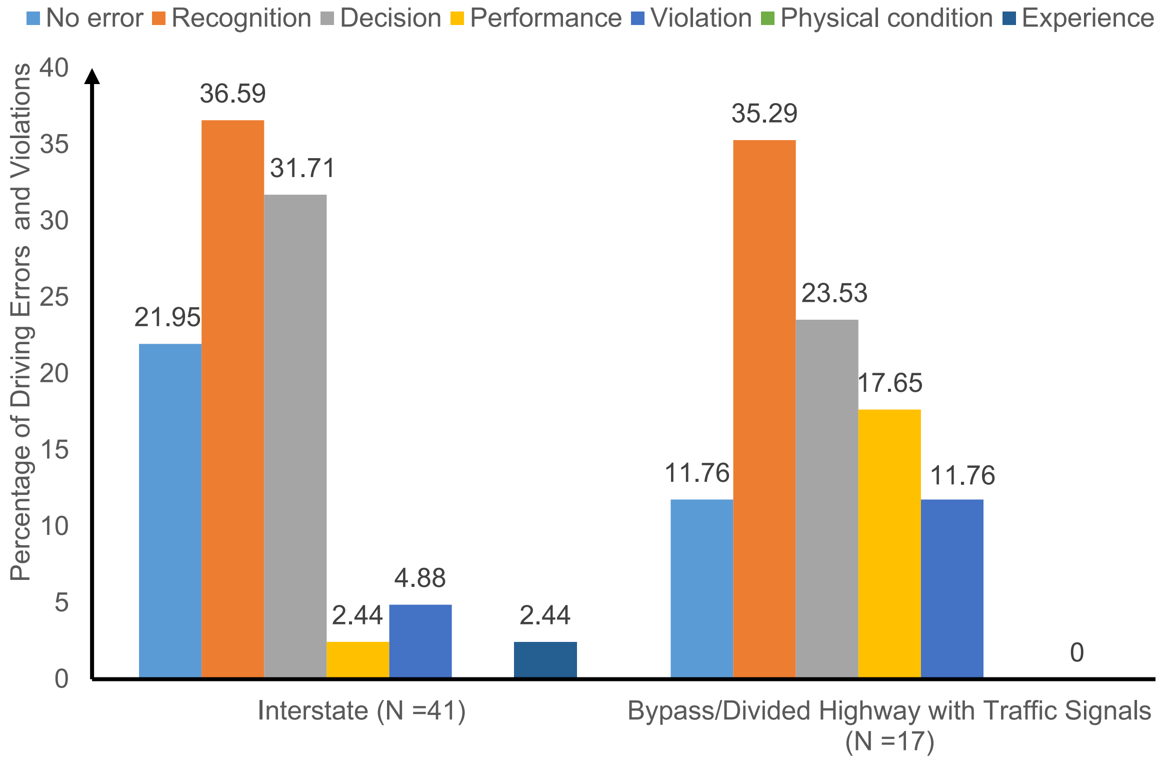

Figure 2.3. Driving Errors and Violations on Interstates and Bypass/Divided Highways with Traffic Signals (in Crashes)

A large percentage of crashes occurred in business/industrial areas (46.21%; N = 311) and moderate residential areas (20.51%; N = 138) (Table 2.4). Figure 2.4 illustrates how drivers were coded in the NDS data to make certain errors and violations in each of the moderate residential areas and business/industrial locations. In areas coded as business/industrial locations, recognition and decision errors each contributed to nearly 39% of the total crashes reported on roadways with such land uses (Figure 2.4). Violations were coded to contribute to only about 6.11% of the overall crashes occurring on roadways within business/industrial locations (Figure 2.4). Referring to moderate residential locations, the percentage of crashes occurring due to recognition errors (39.13%) was significantly higher than that of decision errors (26.09%). A significant percentage of crashes in moderate residential areas (12.32%) also occurred due to violations.

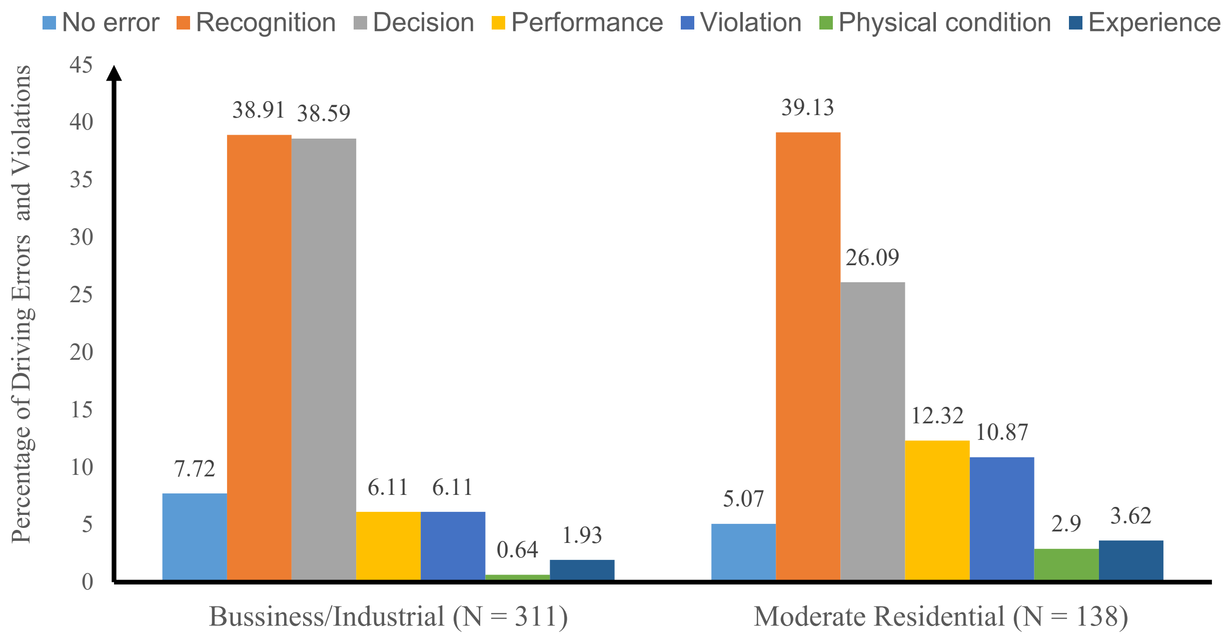

Figure 2.4. Driving Errors and Violations in Business/Industrial and Moderate Residential Areas (in Crashes)

The analysis of errors further reveals that 7.88% (N = 53) crashes were reported on roadways passing within school zones (Table 2.4). Interestingly, nearly 50% of the crashes were reported to have occurred within school zones due to recognition failures (Figure 2.5). Similarly, decision errors and violations contributed to 32.08% and 9.43% of the total crashes occurring on roadways passing through school zones, respectively (Figure 2.5). Only about 7.87% (N = 53) and 5.94% (N = 40) of the overall crashes (N=673) were coded by data reductionists to occur in urban and rural/semi-rural (open country/open residential) locations. Of the crashes that were coded to have urban areas as a key feature, 96% occurred due to driving errors. Among these, recognition and decision errors were the most prevailing errors contributing to nearly 40% and 34% of the total crashes in urban areas. In rural/semi-rural setups, the decision errors had a large contribution resulting in 35% of the crashes, followed by recognition failures which contributed to 27,5% of crashes in these areas. Importantly, violations contributed to around 25% of the crashes in rural/semi-rural locations, perhaps partly due to comparatively lower level of traffic surveillance than other locations (e.g., business/industrial and urban areas).

.

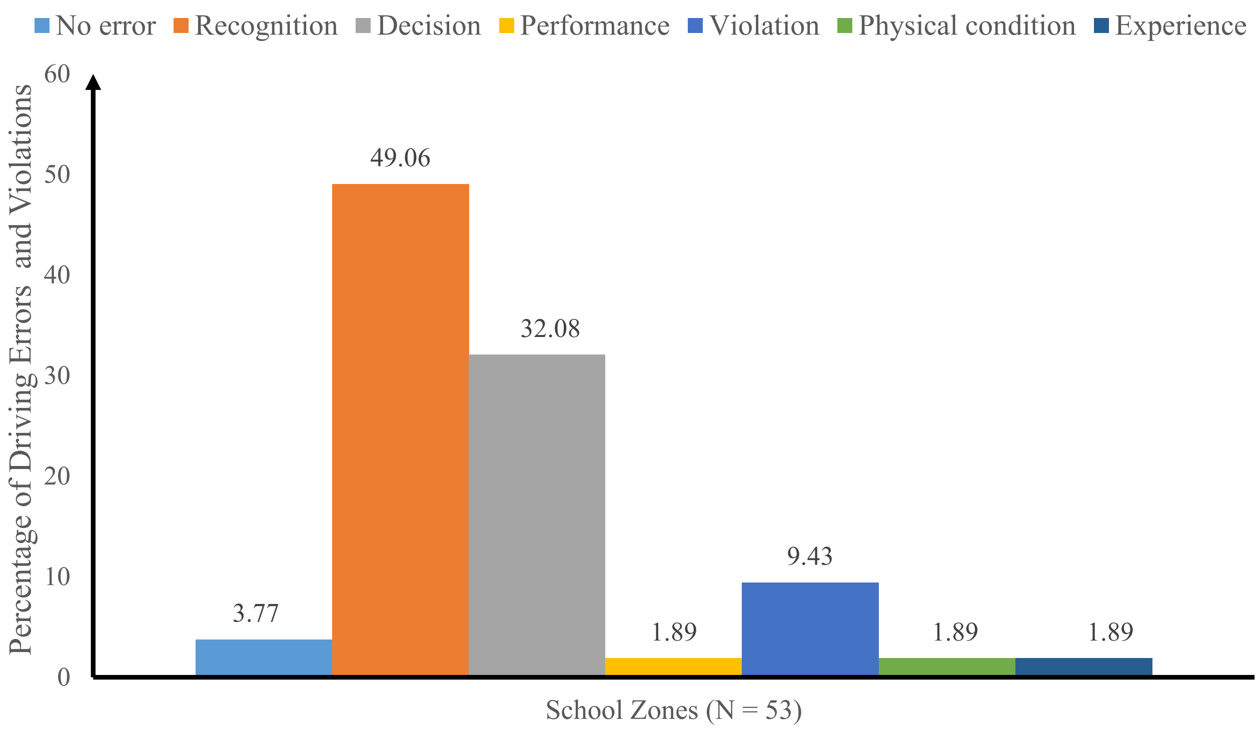

Figure 2.5. Driving Errors and Violations in School Zone Area Crashes

While this study explored the distribution of various driving errors and violations using information on various roadways and environments available in the “locality” factor in the NDS data (coded by the data reductionists, who followed procedures explained in the report). In the future, a more holistic approach can be followed where relevant variables are systematically classified to facilitate exploring the correlations of different roadways, activity centers, and land use on driving errors and violations (Figure 2.6). The roadways and environment can be divided into roadway classification, residential zones, business/commercial zones, industrial zones, and urban, suburban and rural areas. Perhaps a new activity-based land-use classification system can be used in the NDS context, along with the existing Federal Highway Administration (FHWA) roadway functional classification (FHWA 1989), and locality-based land-use classification (Sorensen 2000, Maret and Dakan 2003). To explain this further, using the locality variable, the categories are marked with (✓) or (×) indicating whether NDS data does or does not include the category (Figure 2.6). For example, a complete FHWA roadway functional classification includes rural or urban: interstates, arterials (major/minor), collectors (major or minor), and local streets (major/minor) (FHWA 1989). However, the NDS data related to locality provides information only whether a safety-critical event occurred on interstates or principal arterials. While we acknowledge that residential, industrial, and business localities can be sub-classes of built-up areas (LaGro Jr 2005), the NDS data reductionists considered the urban, residential, school, church, playground, business/industrial localities as separate locations (Figure 2.6). In short, the NDS data available to us through the locality variable does not include complete information on all categories which could completely fulfill the criteria of either the existing FHWA roadway functional classification or locality-based land-use classification or activity-based land-use classification (Figure 2.6).

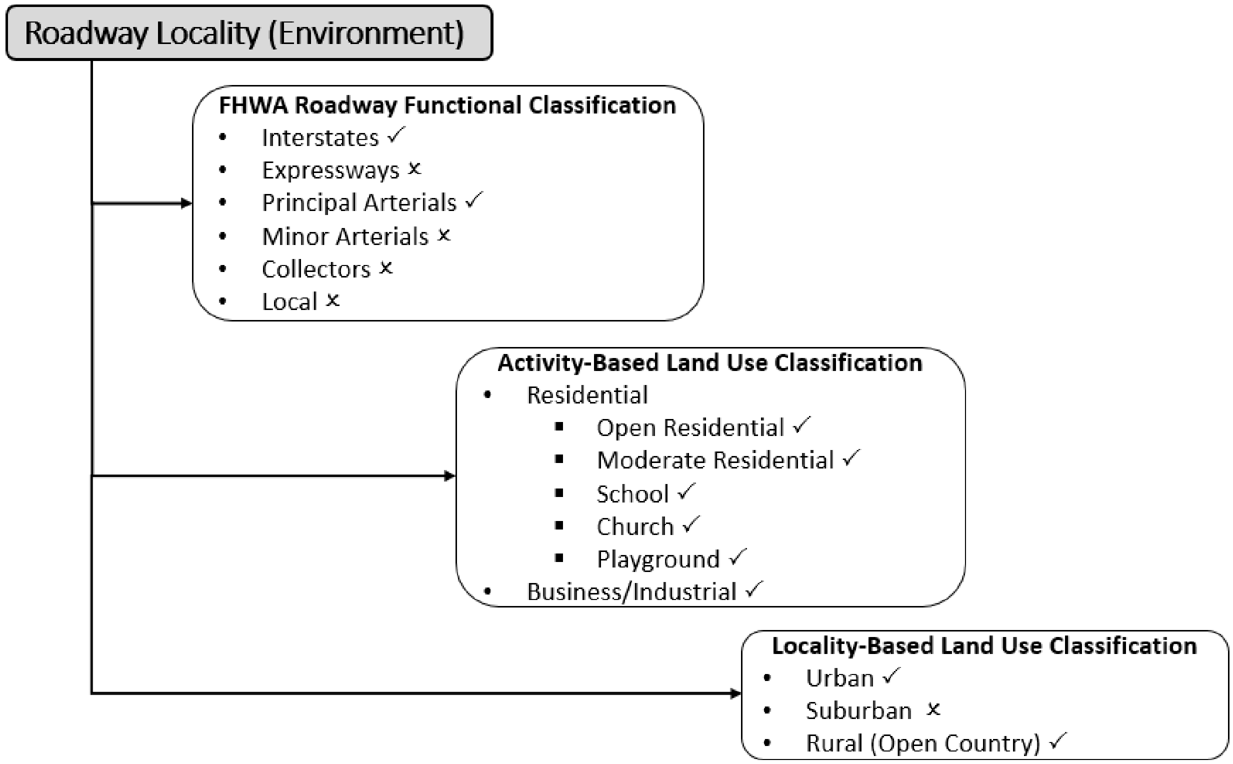

Figure 2.6. Systematic Classification for Roadway Localities

Note: All three roadway locality classifications use information from a single variable “locality” in the NDS data-hence complete information on all categories within any of the three classifications is not available. The categories marked with ✓and × indicate factors on which information was available and unavailable, respectively. All known categories within the three roadway classifications add up to 100% (N = 9,593 events).

LIMITATIONS

This study analyzed NDS data, which includes a finite number of drivers and geographical locations that do not explicitly cover all socio-geographical locations across U.S. The TDEV taxonomy proposed in this paper utilizes driving behaviors reported in the NDS data which may vary over time. In the original NDS data, the “locality” factor was classified by the data reductionists, who followed procedures explained in the report. They included information about the localities where safety-critical events occurred. These could be more systematically categorized, i.e., based on roadway as well as land use factor classifications. Unfortunately, the research team did not have access to NDS data that would provide more complete information regarding all categories of roadway and land use variables. In the future, it would be valuable if the NDS provided independent and complete data on roadway functional classification, activity-based land use, and locality-based land-use. Such data would result in a more complete picture of the distribution of driving errors across different roadways, activity centers, and land uses. Furthermore, information on driver demographics including driver age, gender, driving experience, and education are not available in NDS (SHRP2) subsample accessible to the authors. With data on driver demographics available in the future, it will be interesting to explore how they relate to driving errors, violations, and crash risk.

CONCLUSIONS

This research develops a taxonomy for human errors and violations that lead to crashes and it quantifies their contributions to crash occurrence. While it is possible to study driver behavior by examining police-reported crashes, they cannot provide the insights that naturalistic driving data provides. Using sensor-based technologies and video, the NDS data available in this study affords direct observation of pre-crash driver behavior. Using such data, this study applies a systematic taxonomy of driving errors and violations in order to explore how they vary across roadway and land-use contexts. The taxonomic framework coupled with the data analysis can be used to identify human factors that are most strongly associated with vehicle crashes and highlight the critical ones.

In the NDS data, and based on the categorization of variables as explained in the methodology section, human factors contributed to 93% of crashes; roadway factors contributed in 17% of the crashes, and in 14% of the crashes both human and roadway and environmental factors (e.g., roadway conditions, visual obstructions, and weather conditions) were present. Digging deeper into human factors, in crashes, recognition errors (39%) were most frequently reported driving errors, followed by decision errors (34%), performance errors (8%), and violations (9%). These values indicate that in crashes, recognition and decision phases are critical phases. Using information from the “locality” factor in the NDS data, driving contexts were classified by roadway functions, activity-based land use, and locality-based land-uses. While recognition errors and decision errors were the most common, they occurred more frequently (47% and 52% respectively) when business or industrial structures were present (but not in dense urban localities).

The findings reveal that recognition errors can be particularly hazardous, given their prevalence in crashes compared with their share in baselines, and they were also more frequently associated with severe crashes (51%) compared with other errors and violations. Furthermore, the percentage contribution of performance errors and violations in crashes on principal arterials and interstates are almost 7 and 2.5 times higher respectively compared with baselines. Other land uses where certain types of errors are likely to occur are identified. The findings provide a foundation upon which to build a larger transportation safety program. The analysis provides valuable insights that can be aimed at reducing transportation crashes through data-driven strategies.

The distribution of driving errors and violations across different roadways and environments found in this study can aid in the implementation of locality-specific countermeasures and has implications for connected and automated vehicle development, e.g., by understanding complex and unusual (fringe case) situations for safety, testing of connected and automated vehicles can be enhanced. In the future, connected and automated vehicles have the potential to overcome a large portion of the driving errors and violations which presently contribute to a significant percentage of crashes. More research is needed on such safety intervention programs, e.g., collision warning systems and cooperative adaptive cruise control systems to explore how they may assist during the recognition, decision, and reaction phases, and re-engage the driver in hazardous situations.

REFERENCES

Arvin, R., M. Kamrani and A. J. Khattak (2019). "The role of pre-crash driving instability in contributing to crash intensity using naturalistic driving data." Accident Analysis & Prevention **132**: 105226.

Blatt, A., J. Pierowicz, M. Flanigan, P.-S. Lin, A. Kourtellis, C. Lee, P. Jovanis, J. Jenness, M. Wilaby and J. Campbell (2015). *Naturalistic driving study: Field data collection* (No. SHRP 2 Report S2-S07-RW-1).

Carney, C., D. McGehee, K. Harland, M. Weiss and M. Raby (2015). "Using naturalistic driving data to assess the prevalence of environmental factors and driver behaviors in teen driver crashes."

Dingus, T. A., F. Guo, S. Lee, J. F. Antin, M. Perez, M. Buchanan-King and J. Hankey (2016). "Driver crash risk factors and prevalence evaluation using naturalistic driving data." In Proceedings of the National Academy of Sciences **113**(10): 2636-2641.

Dingus, T. A., S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey and S. Gupta (2006). The 100-car naturalistic driving study. Phase 2: results of the 100-car field experiment, United States. Department of Transportation. National Highway Traffic Safety Administration.

Dumbaugh, E. and W. Li (2010). "Designing for the safety of pedestrians, cyclists, and motorists in urban environments." Journal of the American Planning Association **77**(1): 69-88.

FHWA (1989). FHWA Functional Classification Guidelines. US Federal Highway Administration, Washington, DC.

Frawley, W. E. and W. L. Eisele (2004). Investigation of access point density and raised medians: Crash analysis and micro-simulation.

Hankey, J. M., M. A. Perez and J. A. McClafferty (2016). Description of the SHRP 2 naturalistic database and the crash, near-crash, and baseline data sets, Virginia Tech Transportation Institute.

Huang, H., M. A. Abdel-Aty and A. L. Darwiche (2010). "County-level crash risk analysis in Florida: Bayesian spatial modeling." **2148**(1): 27-37.

Iden, R. and S. A. Shappell (2006). A human error analysis of US fatal highway crashes 1990–2004. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Sage Publications Sage CA: Los Angeles, CA.

Khattak, A. J., Ahmad, N., Wali, B., & Dumbaugh, E. (2020). A taxonomy of driving errors and violations: Evidence from the naturalistic driving study. Accident Analysis & Prevention, 151, 105873.

Khattak, A. J., A. J. Khattak and F. M. Council (2002). "Effects of work zone presence on injury and non-injury crashes." Accident Analysis & Prevention **34**(1): 19-29.

Klauer, S. G., F. Guo, B. G. Simons-Morton, M. C. Ouimet, S. E. Lee and T. A. Dingus (2014). "Distracted driving and risk of road crashes among novice and experienced drivers." New England journal of medicine **370**(1): 54-59.

LaGro Jr, J. (2005). Land-use classification. In Encyclopedia of Soils in the Environment, Elsevier: 321-328.

Maret, I. and B. Dakan (2003). GIS and spraw management.Bulletin of the Association of French Geographer **80**(2): 220-234.

Parker, D., J. T. Reason, A. S. Manstead and S. G. Stradling (1995). "Driving errors, driving violations and accident involvement." Ergonomics **38**(5): 1036-1048.

Petridou, E. and M. Moustaki (2000). "Human factors in the causation of road traffic crashes." European Journal of Epidemiology **16**(9): 819-826.

Reason, J., A. Manstead, S. Stradling, J. Baxter and K. Campbell (1990). "Errors and violations on the roads: a real distinction?" Ergonomics **33**(10-11): 1315-1332.

Sabey, B. E. and G. Staughton (1975). "Interacting roles of road environment vehicle and road user in accidents." Ceste I Mostovi.

Shaon, M. R. R., X. Qin, Z. Chen and J. Zhang (2018). "Exploration of contributing factors related to driver errors on highway segments." Transportation Research Record **2672**(38): 22-34.

Singh, S. (2015). "Critical reasons for crashes investigated in the national motor vehicle crash causation survey." (No. DOT HS 812 115).

Sorensen, A. (2000). "Land readjustment and metropolitan growth: an examination of suburban land development and urban sprawl in the Tokyo Metropolitan Area." Progress in Planning **53**(4): 217-330.

Stanton, N. A. and P. M. Salmon (2009). "Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems." Safety Science **47**(2): 227-237.

Treat, J. R., N. Tumbas, S. McDonald, D. Shinar, R. Hume, R. Mayer, R. Stansifer and N. Castellan (1979). "Tri-level study of the causes of traffic accidents: final report. Executive summary." (No. DOTHS034353579TAC (5) Final Rpt).

Virginia Tech Transportation Institute. (2019). InSight SHRP2 NDS: Event Data Selection. Retrieved from https://insight.shrp2nds.us/data/category/events#/list

Wali, B., A. J. Khattak and J. Xu (2018). "Contributory fault and level of personal injury to drivers involved in head-on collisions: application of copula-based bivariate ordinal models." Accident Analysis & Prevention **110**: 101-114.

Washington, S. P., M. G. Karlaftis and F. Mannering (2010). Statistical and econometric methods for transportation data analysis, Chapman and Hall/CRC.

Wierwille, W. W., R. Hanowski, J. Hankey, C. Kieliszewski, S. E. Lee, A. Medina, A. Keisler and T. Dingus (2002). Identification and evaluation of driver errors: Overview and recommendations. (No. FHWA-RD-02-003,).

Yamamoto, T., J. Hashiji and V. N. Shankar (2008). "Underreporting in traffic accident data, bias in parameters and the structure of injury severity models." Accident Analysis & Prevention **40**(4): 1320-1329.

Yan, X., B. Wang, M. An and C. Zhang (2012). "Distinguishing between rural and urban road segment traffic safety based on zero-inflated negative binomial regression models." Discrete Dynamics in Nature and Society **2012**.

Ye, F. and D. Lord (2011). "Investigation of effects of underreporting crash data on three commonly used traffic crash severity models: multinomial logit, ordered probit, and mixed logit." Transportation Research Record: Journal of the Transportation Research Board (2241): 51-58.

3. Driver Errors and Violations: Pathways that Lead to Crashes in Diverse Built Environment

Authors

Numan Ahmad[1], Behram Wali[2], Asad J. Khattak[1], Eric Dumbaugh[3]

CHAPTER SUMMARY

Driving errors and violations are highly relevant to the safe systems approach as human errors tend to be a predominant cause of crash occurrence. In this study, we harness highly detailed pre-crash Naturalistic Driving Study (NDS) data 1) to understand errors and violations in crash, near-crash, and baseline (no event) driving situations, and 2) to explore pathways that lead to crashes in diverse built environments by applying rigorous modeling techniques. The "locality" factor in the NDS data provides information on various types of roadway and environmental surroundings that influence or may influence traffic flow when a precipitating event is observed. This variable was coded by data reductionists and it is used to quantify the associations of diverse environments on crash outcomes both directly and indirectly through mediating driving errors and violations. While the most prevalent errors in crashes and near-crashes were recognition errors such as failing to recognize a situation (39%) and decision errors such as not braking to avoid a hazard (34%), performance errors such as poor lateral or longitudinal control or weak judgement (8%) were the most strongly correlated with crash occurrence. Path analysis uncovered direct and indirect relationships between key built-environment factors, errors and violations, and crash propensity. Possibly due to their complexity for drivers, urban environments are associated with higher chances of crashes (by 6%), and they can induce more recognition errors, which associate with even higher chances of crashes (by 2% with the "total effect" amounting to 8%). Similar statistically significant mediating contributions of recognition errors and decision errors near school zones, business or industrial areas, and construction zones were also observed. Other important results are discussed, along with real-world implications.

Author affiliations:
[1]Tickle College of Engineering, Civil & Environmental Engineering, University of Tennessee, Knoxville TN
[2]Urban Design 4 Health, 24 Jackie Circle East Rochester, NY 14612.
[3]School of Urban & Regional Planning, Florida Atlantic University, Boca Raton, FL 33431.

INTRODUCTION AND BACKGROUND

This chapter uses the Taxonomy of Driving Errors and Violations (TDEV) detailed in Chapter 2 to apply rigorous discrete outcome-based path analysis (reflecting systems thinking) to explore how environmental conditions correlate with errors, violations and crash occurrence. A unique aspect of this study is the utilization of NDS data, which provide more objective information about pre-crash or pre-near-crash driver behaviors in diverse spatio-temporal contexts. In addition to considering crash events, "no-event" driving events are analyzed vis-à-vis near-crash (near-miss) outcomes, as such "close calls" may foreshadow future crashes. The NDS database is ideal for this analysis as it contains human errors in baseline driving as well as near-miss (near-crashes) and crash situations throughout a diverse spectrum of behavioral and roadway/environmental conditions.

Numerous studies have explored key factors contributing to crash occurrence and/or injury outcomes given a crash (Frawley and Eisele 2004, Huang, Abdel-Aty et al. 2010, Klauer, Guo et al. 2014). Several behavioral factors associate with crashes (Khattak, Khattak et al. 2002, Frawley and Eisele 2004, Dumbaugh and Li 2010, Huang, Abdel-Aty et al. 2010, Klauer, Guo et al. 2014, Dingus, Guo et al. 2016, Ali, Ahmad et al. 2018). Among them, driver distraction was found to be one of the most critical (Klauer, Guo et al. 2014). Also, local demographics and socio-economic conditions can contribute to higher crash risks (Huang, Abdel-Aty et al. 2010). The chances of crashes are higher with a higher density of local population, higher traffic intensity, and urbanization (Huang, Abdel-Aty et al. 2010). Similarly, freeways were associated with a lower crash risk than arterials (Huang, Abdel-Aty et al. 2010). From a roadway standpoint, the existence of a construction zone on a road facility increases crash risk (Khattak, Khattak et al. 2002).

While the literature acknowledges the contributions of roadway and environmental factors on crash occurrence (Dumbaugh and Li 2010), this issue is lightly addressed. Thus, there is a need to examine in greater depth how roadway and environmental (locality) factors correlate with safety-critical events. Furthermore, meager evidence exists on how roadway and environmental factors may induce certain errors and violations. Most of the literature is based on traditional police-reported crash data, which contains an element of subjectivity when it comes to pre-crash conditions (Wali, Khattak et al. 2018). Given these gaps, this study harnesses NDS data in order to examine the above issues in significant depth.

METHODOLOGY

Data Source

This study also uses the NDS data collected as a part of the 2nd Strategic Highway Research Program (SHRP2), for details please refer to Chapter 2 of this report. In order to quantify the crash risk associated with different roadway localities and driving errors and violations, this study uses a subset of original NDS-SHRP2 data including a total of 9,593 trips which include 7,589 baselines, 1,331 near-crashes, and 673 crashes. As a first step, the proposed TDEV (i.e., discussed in detail in Chapter 2) classifies driving behaviors into six key types of driving errors and violations and these six types of driving errors and violations are integrated into the original NDS data in order to achieve the study objective. We also use the three diverse classification systems to categorize the diverse set of built environments in order to have more holistic picture of their influence on crash risk through driving errors and violations.

Path Analysis Framework

Methodologically, this study hypothesizes that roadways and environments (localities) may induce certain driving errors and violations (human errors) which may result in crashes (Figure 3.1). This study classifies driving behavior using the TDEV, a driving error taxonomy (Khattak, Wali et al. 2019), and then applies a two-stage path analysis framework in order to achieve the study objectives. In the first stage, a discrete outcome model was developed for driver errors and violations with locality type, intersection influences, presence of construction zones, and secondary task durations as explanatory variables (Figure 3.1). Given the discrete nature of the errors and violations related to the response variable, a multinomial logit framework was applied. In the second stage, a discrete outcome model was developed for crash propensity with the key explanatory variables of driver errors and violations (response outcome in Stage 1), locality types (roadways and environments), intersection influences, presence of construction zones, and secondary task durations (Figure 3.1). As the crash propensity outcomes exhibit a clear ordering pattern, an ordered probit model was estimated in the second stage (Ahmad et al. 2019, Wali et al. 2020, Wali et al. 2018) (Figure 3.1). Finally, path analysis was conducted to generate combined inferences from the two models. By using the path structure shown in Figure 3.1, the systems framework allows the decomposition of complex structures embedded in the data and the estimation of relevant direct and indirect effects, as discussed earlier (Liu, Khattak et al. 2015, Zhang, Khattak et al. 2018, Kamrani, Arvin et al. 2019). In the path analysis framework, the marginal effects from both of the constituent models are used to obtain the direct, indirect, and total effects of the associated factors on the final response variable (i.e. crash outcome in this case) (Liu, Khattak et al. 2015, Zhang, Khattak et al. 2018). The error terms ($\varepsilon_{(1)}$ and $\varepsilon_{(2)}$) in the two models (i.e. as shown in Figure 3.1) are assumed to be uncorrelated in this case. (Liu, Khattak et al. 2015, Zhang, Khattak et al. 2018).

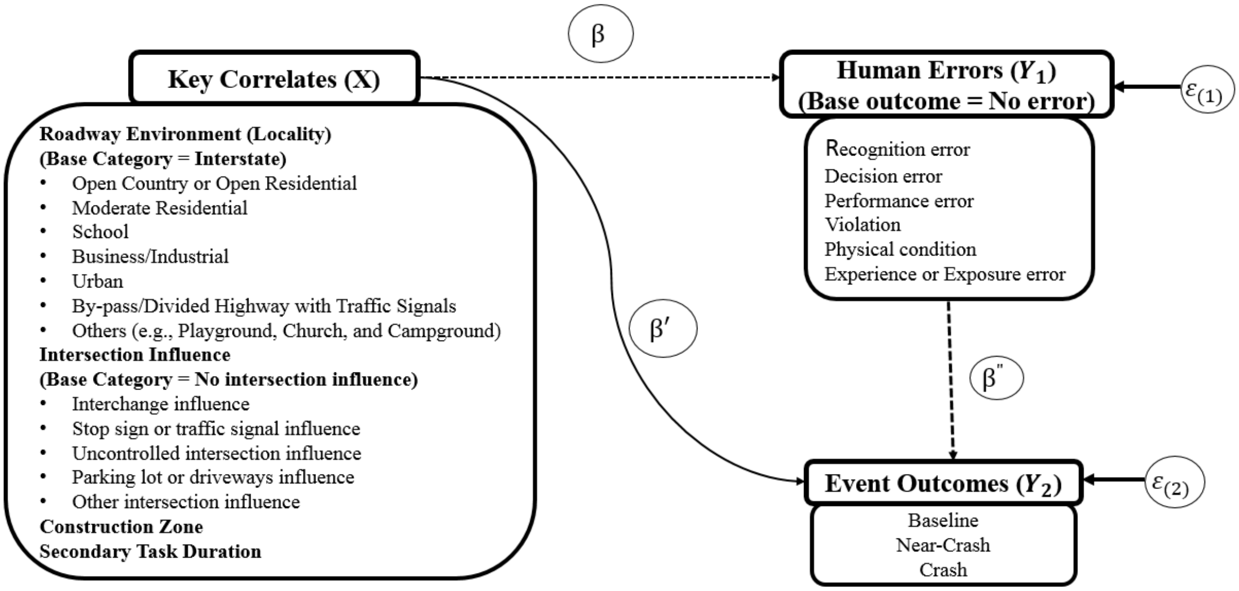

Figure 3.1. Methodological Framework

Notes: Dotted lines indicate indirect effects whereas solid lines indicate direct effects; Z_1 (Driving errors and violations) is categorical outcome which is modeled through a multinomial logit framework; Z_2 (event severity outcomes) being ordinal is modeled through an ordered probit regression.

RESULTS AND DISCUSSION

Descriptive Statistics

Table 3.1 presents summary statistics of the two response variables and other key explanatory variables. The NDS data includes 9,593 observations containing 7,589 baselines (79%), 1,331 near-crashes (14%), and 673 crashes (7%). The descriptive statistics show that in 90.12% of the baselines, no driving error or violation was observed (as expected). However, at least one driving error or violation was present in a significant portion of the near-crash and crash events, 61.31% and 92.87%, respectively (Table 3.1). In both crash and near-crash events, recognition and decision errors were the predominant error types, collectively accounting for 47.85% of the near-crashes and 72.95% of the crash events (Table 3.1). Hence, recognition errors and decision errors respectively comprise 38.63% and 34.32% of the combined crashes. At least one error or violation was present in 10% of the total baseline events (Table 3.1). This indicates that other unobserved factors, when combined with such errors, could result in safety-critical events and/or grow the negative effects of the driving errors/violations. Given that these unsafe driving errors and violations did not result in crashes/near-crashes, strategies to prevent such errors and violations need to be explored and encouraged from a behavioral standpoint.

The percentage of drivers traveling on roadways in business/industrial locations was 31.10%, 40.04% and 46.21% for baselines, near-crashes, and crashes respectively. The ratio of baseline percentage to crash percentage is 1.49, which means that such commercial areas are a clear risk factor. This study is perhaps among the first ones to identify this hazard in a substantive way (Table 3.1). Furthermore, the percentage of the NDS drivers traveling on roadways within school zones is about 5% for baselines, but 7.06% for near crashes and 7.87% for crashes, indicating a 1.74 ratio between school zone percent and baseline percent (Table 3.1). This may reflect the relatively greater potential of unsafe outcomes in school zones. Also, based on the data coded in NDS, higher risk levels are observed in urban areas. Around 25% and 20% of the subject drivers were involved in crashes due to the influence of traffic signal/stop sign and parking lot/driveways respectively (Table 3.1). The crash percentage on roadways within construction zones were 13.5% higher than the baselines observed at such locations. Of all the crashes, 4% were observed in construction zones compared with about 3% in baselines at construction zones.

Table 3.1. Descriptive Statistics of Key Variables

VARIABLE	All Cases (N = 9,593) Percent (S.D)	Baseline (%) (N = 7,589)	Near-Crash (%) (N = 1,331)	Crashes (%) (N = 673)	$\frac{\text{\% of Crashes}}{\text{\% of Baselines}}$
DRIVING ERRORS					
No Driving Errors	77.16 (0.4198)	90.12	38.69	7.13	0.0791
Recognition Errors	7.60 (0.2652)	0.22	34.03	38.63	175.5909
Decision Errors	6.45 (0.2457)	2.69	13.82	34.32	12.7584
Performance Errors	0.69 (0.0833)	0.09	0.68	7.58	84.2222
Violations	6.53 (0.2469)	5.56	10.74	9.06	1.6295
Physical Conditions	1.28 (0.1125)	1.25	1.43	1.34	1.0720
Experience or Exposure Errors	0.27 (0.0519)	0.07	0.60	1.93	27.5714
Total	100%	100%	100%	100%	
ROADWAY LOCALITY					
Interstate	23.29 (0.4226)	25.6	18.78	6.09	0.2379
Open Country or Open Residential	8.66 (0.2813)	9.46	5.48	5.94	0.6279
Moderate Residential	21.21 (0.4088)	22.26	15.55	20.50	0.9209
School	5.10 (0.2199)	4.51	7.06	7.87	1.7450
Business/Industrial	34.08 (0.4739)	31.10	40.04	46.21	1.4859
Urban	2.76 (0.1639)	1.51	7.30	7.88	5.2185
Bypass or Divided Highways with no Traffic Signals	3.00 (0.1760)	3.04	3.00	2.53	0.8322
Others (e.g., Church, Playground and Campground)	1.90 (0.1364)	1.66	2.78	2.98	1.7952
Total	100%	100%	100%	100%	---
INTERSECTION INFLUENCE					
No Intersection Influence	71.75 (0.4497)	81.42	32.83	41.01	0.5037
Interchange Influence	3.22 (0.1765)	2.33	8.41	2.97	1.2747
Stop Sign or Traffic Signal Influence	15.78 (0.3646)	12.06	32.38	24.96	2.0697
Uncontrolled Intersection Influence	3.38 (0.1807)	1.75	10.14	8.32	4.7543
Parking Lot or Driving Way Entrance/Exit Influence	4.65 (0.2106)	1.98	11.95	20.36	10.2828
Other (e.g., crosswalk, railroad crossing, roundabouts)	1.22 (0.1097)	0.46	4.28	2.38	5.1739
Total	100%	100%	100%	100%	
CONSTRUCTION ZONE INDICATOR	3.79 (0.1911)	3.40	6.01	3.86	1.1353
SECONDARY TASK DURATION* (Min = 0; Max = 24.1)	2.0918 (2.719)	1.75 (2.16)	3.28 (3.83)	3.58 (4.19)	---

Modelling Results and Discussion

Both models are systematically derived to include the most important variables (locality type, intersection influence, secondary task duration, presence of construction zone, and errors/violations) on the basis of statistical significance, specification parsimony, and theoretical justification. A 95% confidence criterion was used for variables in either model, except for locality variables which were left in the model as per 90% confidence criteria given their conceptual importance and for the sake of completeness. The estimation results of multinomial logit and ordered probit models along with marginal effects are presented in Table 3.2 and 3.3 respectively. The marginal effects can be interpreted as an increase or a decrease in the probability of observing a specific outcome in the case an indicator variable switches from 0 to 1 or with a unit increase in the case of continuous variable (keeping all variables at their mean values). When multiplied by 100, the marginal effects can be interpreted as a percent change in the chance of observing a specific outcome (Ahmad et al. 2019, Wali et al. 2020). The results of path analysis, quantifying the direct effects of explanatory factors (such as locality type) each on crash propensity and human errors, as well as indirect effects of key locality-related factors on crash propensity through its mediation paths over human errors are computed (detail results available from authors) and simply illustrated in Figure 3.2. The key findings from the individual models and path analysis are briefly discussed below.

Driver Error Model

The multinomial logit model of driving errors and violations considered no error or violation as a base outcome. For the sake of completeness, all the explanatory variables were kept in the functions of all driving errors. However, for the sake of simplicity and understanding, the marginal effects are provided for only those variables which were found to have significant association with driving errors and violations (Table 3.2). The variables, significantly correlated with the various types of driving errors and violations, belong to roadways and environments, especially the influence of intersections (using no intersection influence as the base category), construction zones, and secondary task durations (Table 3.2).

Interstate/bypass/divided highway with no traffic signals was used as the base category to explore the correlations of all locality-related factors with driving errors and violations. According to the driver error model, secondary task duration was positively correlated with recognition, decision, performance, physical conditions, and experience errors. The marginal effects for all significant variables are computed which measures the change in the probability of the response variables (the six error types) relative to the base outcome (no driving errors/violation) with a unit increase in a specific explanatory variable (Table 3.2). The goodness-of-fit of the driver error model, as estimated by the McFadden Pseudo R2, was found to be 0.1121, indicating a modest model fit with the data. The driver error model quantifies the changes in the probability of recognition errors and decision errors within the localities characterized by moderate residential, school zones, urban areas, business/industrial areas, and other localities (that include church, playground, and campground). The probability of recognition and decision errors increases by 5% and 3% in school zones compared to safety-critical events on interstates, respectively. This finding is reasonable in the sense that school zones typically include complex movements and vulnerable road users. As a result, the chances of a driver failing to recognize and decide correctly in hazardous situations increases (Gregory, Irwin et al. 2014). Furthermore, the modeling results indicate that, compared to a safety-critical event happening on interstate highways, the chance of recognition and decision errors on roadways in urban areas (presumably on local roads) is higher by 8%, 7% respectively. This is reasonable due to complex and highly densified traffic conditions and consistent with other studies (Huang, Abdel-Aty et al. 2010). Referring to business/industrial areas, modeling results reveal that the chance of making recognition and decision errors are higher by 3% and 2% on roadways passing through business or industrial areas compared to when interstate was coded as the locality factor, respectively. The higher traffic volume, complex activities, and increased presence of commercial drivers can negatively affect drivers' recognition and decision processing capabilities in such areas.

Table 3.2. Driver Error Model Results (Multinomial Logit Model): Marginal Effects for Only Significant Variables

Independent Variables	Type of Driving Error and Violations											
	Recognition		Decision		Performance		Violation		Physical Condition		Experience/ Exposure	
	Coeff.	ME	Coeff.	ME	Coeff.	ME	Coeff.	ME	Coeff.	ME	Coeff.	ME
Roadway Locality (Base outcome = Interstate)												
Open Country/Open Residential (Rural/Semi-Rural)	---	---	0.5503[a]	0.0328	---	---	---	---	---	---	---	---
Moderate Residential	0.3115[b]	0.0130	0.3627[a]	0.0168	2.5362[a]	0.0094	---	---	---	---	---	---
School	0.8722[a]	0.0478	0.6745[a]	0.0345	---	---	-0.3544[b]	-0.0324	---	---	---	---
Business/Industrial	0.4884[a]	0.0251	0.3626[a]	0.0177	2.0974[a]	0.0058	-0.5115[a]	-0.0369	---	---	---	---
Urban	1.3698[a]	0.0800	1.1656[a]	0.0665	3.1258[a]	0.0133	---	---	0.9352[b]	0.0127	2.2329[a]	0.0065
Bypass or Divided Highway with traffic signals	---	---	---	---	2.8566[a]	0.0132	---	---	0.8554[a]	0.0138	---	---
Others (e.g., church, playground, & Campground)	0.8005[a]	0.0417	0.6259[a]	0.0309	2.8218[a]	0.0117	---	---	---	---	---	---
Intersection Influence (Base outcome = No Intersection Influence)												
Interchange influence	1.3505[a]	0.0725	1.2575[a]	0.0893	---	---	---	---	---	---	2.4834[a]	0.0152
Stop sign or traffic signal influence	1.3089[a]	0.0722	0.3759[a]	0.0086	0.7947[a]	0.0030	1.0765[a]	0.0730	---	---	---	---
Uncontrolled intersection influence	1.6459[a]	0.0917	1.2141[a]	0.0686	2.0881[a]	0.0176	1.0002[a]	0.0504	---	---	2.2721[a]	0.0103
Parking lot or driveways influence	2.0741[a]	0.1138	1.9886[a]	0.1523	2.8664[a]	0.0324	1.3485[a]	0.0633	---	---	---	---
Others intersection influence	2.3620[a]	0.1716	1.4822[a]	0.0805	---	---	1.3335[a]	0.0697	---	---	---	---
Construction zone indicator	0.8073[a]	0.0387	---	---	---	---	0.7905[a]	0.0419	---	---	1.7906[a]	0.0043
Secondary Task Duration	0.3407[a]	0.0184	0.0678[a]	0.0019	0.0853[b]	0.0002	---	---	-0.2763[a]	-0.0037	0.1448[a]	0.0003
Summary Statistics												
Number of observations	9593											
Pseudo-R^2	0.1121											
Log Likelihood at 0	-8226.7559											
Log Likelihood at β	-7304.5218											
AIC	14789.0400											
BIC	15434.2300											

Notes: ME = marginal effects, which predict the change in the probability of observing a response outcome with a unit change in continuous explanatory variable (or a switch from 0 to 1 for indicator variable); Base outcome in the multinomial logit model is "No error"; and β'_1, β'_2, β'_3, β'_4, β'_5, and β'_6 indicates the association of the explanatory variables with recognition, decision, performance, violation, physical condition related, and experience/exposure errors. The superscripts i.e. "a" and "b" indicate the significance levels of explanatory variables in the model ("a" and "b" shows that 95% and 90% confidence criteria were met respectively. (-) indicates not applicable as all these variables were tried in the model but did not meet the significance criteria (hence, for simplicity kept out of model).

Crash Outcome Model

The estimation results along with the marginal effects for each explanatory variable obtained from crash outcome model are presented in Table 3.3. The estimation results shed light on the associations of human error with crash outcomes, i.e., crash propensity. One key advantage of the SHRP2 NDS data is that the effect of any specific contributing factor on crash propensity can be interpreted relative to the baseline (no-event driving). The crash outcome model includes a total of fourteen conceptually relevant factors having significant correlations with crash propensity (Table 3.3). The key finding is that all the error types (recognition, decision, and performance errors, physical condition, and experience errors) and violations, compared to no driving errors (base outcome), are significantly correlated with crash propensity. Other significant variables in the crash outcome model relate to roadways and environments (locality), intersection influence, construction zone presence, and secondary task duration. The overall goodness-of-fit of the crash outcome model, as evaluated via McFadden Pseudo-R^2, had a value of 0.384, which indicates a relatively good fit (Table 3.3). The performance errors have a strong correlation with crash propensity, associated with higher chance of crashes by 55%. Similarly, the chance of crash outcome is 5% higher if a driver violates traffic laws. In contrast, driver violations were found to associate with higher chances of a near-crash by a greater magnitude than for a crash outcome (see marginal effects in Table 3.3). This may be because drivers are aware that they are engaging in traffic violations and adjusting their operating behavior in response. Also, the chance of both crash (6%) and near-crash (8%) is higher in an urban environment compared to interstate/bypass/divided highway with no traffic signals. This can be due to highly complex and densified traffic conditions. This finding is consistent with previous studies based on police-reported crashes (Huang, Abdel-Aty et al. 2010). Furthermore, compared to the base, the chance of a crash is higher by 2% and 1%, on roadways passing close to school zones and business/industrial areas, respectively. These findings are reasonable because of the presence of vulnerable road users and potential distraction in such areas compared to driving on interstates. Similar interpretation applies to other important correlates such as intersection influence, secondary task duration, and existence of construction zone.

Path Analysis

The correlations of various roadway and environmental factors (localities) and driving errors/violations on crash outcomes are illustrated in Figure 3.2. The indirect association of various roadways and environments with crash occurrence through mediating driving errors and violations are also illustrated in Figure 3.3. In short, the variables that possess statistically significant parameter estimates in the driver error model (Table 3.2) are indirectly associated with crash outcome, a finding that cannot be extracted using the traditional modeling approach. Moreover, it was found that all variables which showed significant direct correlation with crash outcome were also found to have significant indirect correlation with crash outcome through one or more of the six mediating driving errors and violations. For instance, compared to interstate/bypass/divided highways with no traffic signals, the urban locality shows that drivers traveling in an urban locality have 8% higher chance of making recognition errors (compared to no driving error or violation), and as a direct effect, compared to no driving error or violation, drivers involved in recognition errors are associated with a 29% higher chance of getting into a crash (Figure 3.2). As such, the indirect effect of urban locality on crash outcome is 8% x 29% = 2%. Therefore, the total effect of urban locality on crash outcome becomes 8% (6%+2%), which is greater than the direct effect of 6% (Table C1 in Appendix). This implies that while urban locality, compared to interstate roadways, may correlate with increases in the chance of observing a crash outcome, the actual increase is even more when we simultaneously consider the effect of urban locality on chance of recognition errors and the effect of the latter on chance of crash outcome. The path analysis results indicate that all key correlates were both directly and indirectly associated with crash propensity through one or more mediating driving errors and violations. For instance, the path analysis results indicate that the presence of a school zone increases the chance of recognition and decision errors (compared to no driving error or violation) by 5% and 3%, respectively (Table C1 in Appendix), while recognition and decision errors increase the chance of a crash by 29% and 24%, respectively (Figure 3.2). Compared to travel on interstate roadways, the presence of a school zone directly associates with increases the chance of a crash by 2% (Figure 3.2). However, path analysis indicates that the overall effect of school zones on crash propensity is larger; the increased chance of crash occurrence through recognition and decision errors is 3% and 3% respectively (detailed results of path analyses are not shown here and can be provided by authors on request) as can be seen that school zones is significantly correlated with the two driving errors which are significantly correlation with crash outcome (Figure 3.2). Detailed results of the path analysis are provided in Table C1 in the Appendix of this report. Overall, rigorous path analysis framework uncovers important information about driving errors that moves us closer to more targeted interventions in complex environments (Liu et al. 2015, Zhang et al. 2018).

Table 3.3. Crash Outcome Model Results (Ordered Probit Model)

Independent variables	Coeff.	t-stat	p-value	Marginal Effects		
				Baseline	Near-Crash	Crash
Drivers Errors (Base Outcome = No Driving Errors)						
Recognition Errors	2.1886	39.63	<0.001	-0.6482	0.3554	0.2928
Decision Errors	1.9964	36.97	<0.001	-0.5866	0.3510	0.2356
Performance Errors	2.9181	17.83	<0.001	-0.8185	0.2697	0.5488
Violations	0.9973	17.70	<0.001	-0.2337	0.1844	0.0493
Physical Conditions	0.9153	7.30	<0.001	-0.2076	0.1660	0.0416
Experience or Exposure	2.2860	10.10	<0.001	-0.6771	0.3529	0.3241
Roadway Locality (Base Outcome = Interstate)						
Open Country or Open Residential (Rural/Semi-Rural)	0.0505*	0.64	0.524	---	---	---
Moderate Residential	-0.0045*	-0.07	0.942	---	---	---
School	0.2534	3.06	0.002	-0.0414	0.0240	0.0174
Business/Industrial	0.1914	3.39	0.001	-0.0305	0.0177	0.0129
Urban	0.6913	7.48	<0.001	-0.1323	0.0769	0.0553
Bypass or Divided Highway with traffic signals	0.1832	1.70	0.089	-0.0291	0.0169	0.0123
Others (e.g., church, playground, and Campground)	0.2357	1.92	0.055	-0.0382	0.0221	0.0161
Intersection Influence (Base Outcome = No intersection influence)						
Interchange influence	0.8957	10.71	<0.001	-0.1826	0.1123	0.0703
Stop sign or traffic signal influence	0.6471	14.40	<0.001	-0.1203	0.0747	0.0456
Uncontrolled intersection influence	1.0042	13.72	<0.001	-0.2126	0.1299	0.0827
Parking lot or driveways influence	1.0828	16.54	<0.001	-0.2353	0.1430	0.0923
Others intersection influence	0.8342	7.01	<0.001	-0.1663	0.1026	0.0638
Construction zone indicator	0.1409	1.74	0.082	-0.0225	0.0128	0.0096
Secondary task duration	0.0269	4.66	<0.001	-0.0043	0.0024	0.0018
Thresholds						
cut1	1.8904	38.83	<0.001			
cut2	3.1489	53.39	<0.001			
Summary Statistics						
Number of observations	9593					
Pseudo-R^2	0.3840					
Log Likelihood at 0	-6195.4121					
Log Likelihood at β	-3816.5870					
AIC	7677.1740					
BIC	7834.8870					

Note: ME indicates the marginal effects which predict the probability of observing a response outcome with unit change in continuous explanatory variable (or a switch from 0 to 1 for indicator variable). The lowest level for event severity is "Baseline"; and Base level for drivers' errors is "No error". The superscript (*) in Table 3.3 indicates these variables were not found to be significant as per 90% or 95% confidence criteria. In final path analysis, only significant variables were considered (other than *).

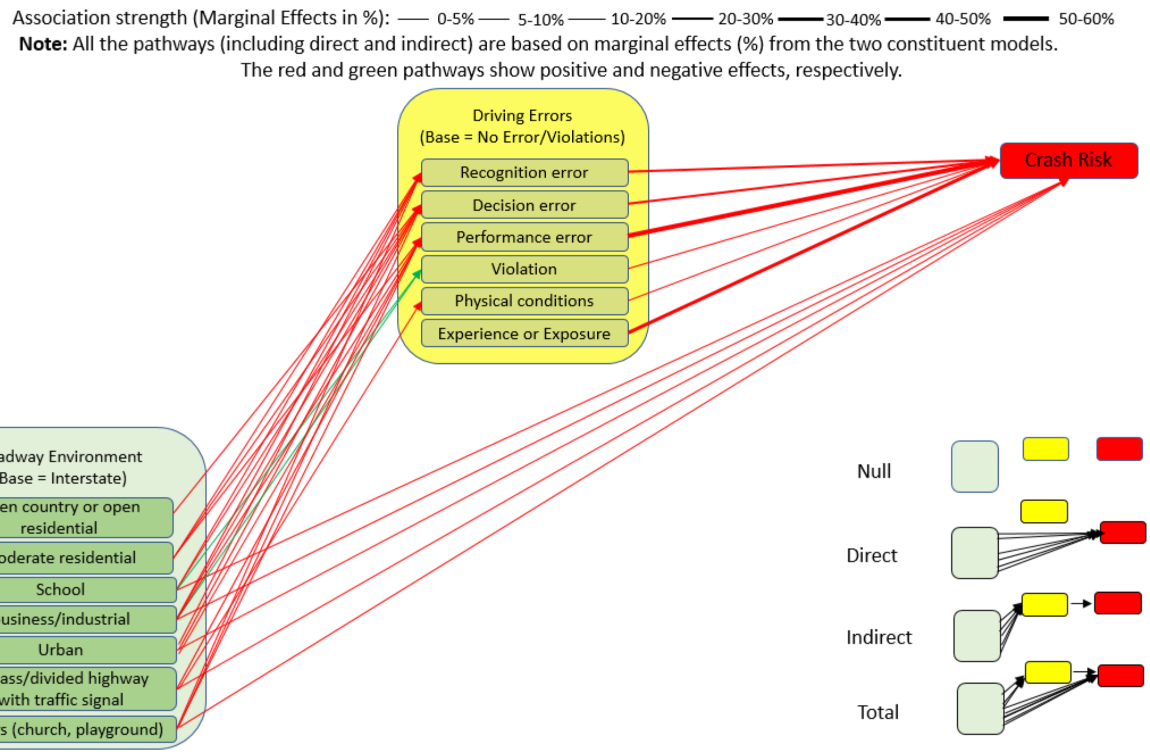

Figure 3.2. Direct and Indirect Effects of Roadway and Built Environment on Crash Occurrence Mediated by Driver Errors

Note: Detailed results about direct and indirect effects of various roadway environments on crash outcomes can be found in Table C1 in the Appendix.

LIMITATIONS

This study analyzed the NDS data containing extensive information on driver performance and behavior, roadway and environment factors, and pre-crash maneuvers, still there could be other factors (e.g., drivers' characteristics, and local demographics) which might be significantly correlated with crash propensity. Hence, it would be interesting to account for all such unobserved factors in the modeling framework, which could have significant influence on crash propensity. This study is based on two independently estimated models (equations), which represent the standard path analytic framework (as shown in Figure 3.1). This method is appropriate because driving errors mediate the relationship between variables of interest such as land use and safety critical outcomes. Similar models can be estimated where the error terms (residuals) correlate with each other, e.g., seemingly unrelated regressions. The analysis reported does not perform regressions where the error terms in the two models are tested for correlations. In the future, a full-information joint approach can be used to account for potential correlation between the error (residual) terms. This can be done by using the conditional (recursive) mixed process procedure in STATA (Roodman, 2011).

This analysis utilized SHRP2 NDS data which included specific drivers and geographical regions that do not explicitly cover all of the socio-geographical regions across the U.S. Furthermore, TDEV—the human error taxonomy uses driver behaviors while classifying driving errors and violations which may vary with time. Also, as mentioned before, information on important variables such as driver demographics was not available to the authors, which is a topic for future research.

We have explored the contribution of speed and speed variations to crash outcomes in a related study (Ahmad et al. 2021). However, this study has not explored the role of speed or speed variations in the analysis presented.

To categorize roadway and land use factors, a more holistic approach would be to apply appropriate roadway-locality related classifications, e.g., roadway functional classes (interstate, expressways, arterials (major/minor), collectors (major/minor), and local roads), activity centers (specifically separately categorizing business from industrial), and locality (urban, suburban, and rural). The information available from a single variable called “locality” in the NDS data only partially fulfills the research needs. In the future, it will be appropriate to either collect separate and complete information on roadway functional classes (interstate, expressways, arterials (major/minor), collectors (major/minor), and local roads), activity centers (specifically separately categorizing business from industrial), and locality (urban, suburban, and rural), or provide geo-codes so that such information can be obtained. Such efforts can help with even deeper understanding of the variations in driving errors and violations and crash risks across contexts.

CONCLUSIONS

This research contributes to the literature by conceptualizing the direct relationships between key roadways and environmental (roadway locality) factors and crash propensity, as well as the indirect relationships between roadways and environmental factors and crash propensity through mediating errors and violations. In order to achieve these objectives, the study capitalizes on emerging sensor-based technologies, telematics, video and radar surveillance to obtain an objective and nuanced understanding of key crash contributing factors. Along these lines, the study harnesses data from the Naturalistic Driving Study where driving behaviors in thousands of baseline, crash, and near-crash events in diverse spatio-temporal contexts are analyzed and compared. In addition to considering crash events, no-event driving is analyzed vis-à-vis near-crash outcomes, as such “close calls” may foretell actual future crashes.

A taxonomy based on the perception-reaction to hazardous situations classified driving errors and violations and analyzed them. Next, this study uses information from the “locality” factor in the NDS data to quantify the direct as well as indirect effects of different roadway and environmental variables on crash outcome through different types of driving errors and violations.
The results of the driver error model indicate that several roadway and environmental factors are statistically significantly associated with driver errors. The most prevalent types of errors are recognition errors and decision errors, which are correlated with moderate residential, school zones, urban areas, business or industrial areas, at interchanges stop signs and signalized intersections. Modeling results indicate that the duration of secondary tasks is correlated with recognition errors, decision errors, performance errors, and violations.

To model roadway safety outcomes, a discrete outcome model was estimated. The model examines crash propensity as a function of mediating error types/violations and roadway-environmental contexts. Importantly, while urban locality associate with higher chances of observing a crash by 6%, path analysis showed that they can induce recognition errors and the effect of the latter on the chance of crash outcome, i.e., the “total effect” of urban locality amounts to 8% on the chances of having a crash. Additionally, school zones and business (commercial) or industrial land uses were found to have a substantial correlation with crash propensity both directly and indirectly through mediating recognition errors and decision errors. By examining detailed behavior of drivers, this study is perhaps among the first ones to uncover business/commercial areas as hazardous and risky in a substantive way.

Indeed, the analysis indicates that complex path structures should be explored in line with the systems approach. This avoids the possibility of concluding that key factors have only direct association with crash outcomes, while in reality almost all factors (e.g., roadways and environments, intersection influence, presence of construction zone, and secondary task duration) can and were found to have significant indirect correlations with crash outcome through driving errors and violations.

The study does not advocate for specific government policies but provides information that can be used for future formulation of safety policies and research. The data and methods discussed in the paper should allow for replication of the study. As a part of future work, it will be interesting to investigate how driver errors and violations may change with some control of the driving task being given to connected and automated vehicles. Similarly, it would be interesting to explore the effects of the roadway-environment on injury outcomes (given a crash). To this end, a practice and research-relevant theme could be to develop a knowledge base of emerging intelligent transportation system technologies that could help prevent the prevalent safety-critical errors and violations observed in naturalistic driving setups.

REFERENCES

Ahmad, N., Arvin, R., and Khattak, A. J. (2021). Exploring Pathways from Driving Errors and Violations to Crashes: The Role of Speed Volatility. To be presented (#TRBAM-21-02508) at the Transportation Research Board 100th Annual Meeting Transportation Research Board. Washington D.C., National Academy of Sciences.

Ahmad, N., Ahmed, A., Wali, B., & Saeed, T. U. (2019, May). Exploring factors associated with crash severity on motorways in Pakistan. In Proceedings of the Institution of Civil Engineers-Transport (pp. 1-10). Thomas Telford Ltd.

Ahmad, N., Ahmed, A., & Shah, A. A. (2019). Is Effective Enforcement of Seatbelt Law Helpful in Reduction of Road Crash Fatalities: An Exploratory Empirical Analysis using Aggregate Data. In Proceedings of the Pakistan Academy of Sciences: A. Physical and Computational Sciences, 56(4), 31-41.

Arvin, R., M. Kamrani and A. J. Khattak (2019). The role of pre-crash driving instability in contributing to crash intensity using naturalistic driving data. Accident Analysis & Prevention **132**: 105226.

Carney, C., D. McGehee, K. Harland, M. Weiss and M. Raby (2015). Using naturalistic driving data to assess the prevalence of environmental factors and driver behaviors in teen driver crashes.

Dingus, T. A., F. Guo, S. Lee, J. F. Antin, M. Perez, M. Buchanan-King and J. Hankey (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. In Proceedings of the National Academy of Sciences **113**(10): 2636-2641.

Dingus, T. A., S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey and S. Gupta (2006). The 100-car naturalistic driving study. Phase 2: results of the 100-car field experiment, United States. Department of Transportation. National Highway Traffic Safety.

Dumbaugh, E. and W. Li (2010). Designing for the safety of pedestrians, cyclists, and motorists in urban environments. Journal of the American Planning Association **77**(1): 69-88.

Frawley, W. E. and W. L. Eisele (2004). Investigation of access point density and raised medians: Crash analysis and micro-simulation.

Gregory, B., J. D. Irwin, I. J. Faulks and E. Chekaluk (2014). Speeding in school zones: violation or lapse in prospective memory?. Journal of Experimental Psychology: Applied **20**(3): 191.

Hankey, J. M., M. A. Perez and J. A. McClafferty (2016). Description of the SHRP 2 naturalistic database and the crash, near-crash, and baseline data sets, Virginia Tech Transportation Institute.

Huang, H., M. A. Abdel-Aty and A. L. Darwiche (2010). County-level crash risk analysis in Florida: Bayesian spatial modeling. **2148**(1): 27-37.

Kamrani, M., R. Arvin and A. J. Khattak (2019). The Role of Aggressive Driving and Speeding in Road Safety: Insights from SHRP2 Naturalistic Driving Study Data.

Khattak, A. J., A. J. Khattak and F. M. Council (2002). Effects of work zone presence on injury and non-injury crashes. Accident Analysis & Prevention **34**(1): 19-29.

Khattak, A. J., B. Wali and N. Ahmad (2019). A Taxonomy of Naturalistic Driving Errors and Violations and Its Variations Across Different Land-Use Contexts – A Path Analysis Approach. Transportation Research Board 98th Annual Meeting Transportation Research Board. Washington D.C., National Academy of Sciences.

Klauer, S. G., F. Guo, B. G. Simons-Morton, M. C. Ouimet, S. E. Lee and T. A. Dingus (2014). Distracted driving and risk of road crashes among novice and experienced drivers. New England Journal of Medicine **370**(1): 54-59.

Liu, J., A. J. Khattak, S. H. Richards and S. Nambisan (2015). What are the differences in driver injury outcomes at highway-rail grade crossings? Untangling the role of pre-crash behaviors. Accident Analysis & Prevention **85**: 157-169.

Roodman, D. (2011). Fitting fully observed recursive mixed-process models with CMP. The STATA Journal, 11(2), 159-206.

Singh, S. (2015). "Critical reasons for crashes investigated in the national motor vehicle crash causation survey." (No. DOT HS 812 115).

Wali, B., Khattak, A. J., & Ahmad, N. (2020). Injury severity analysis of pedestrian and bicyclist trespassing crashes at non-crossings: A hybrid predictive text analytics and heterogeneity-based statistical modeling approach. Accident Analysis & Prevention, 150, 105835.

Wali, B., A. J. Khattak and N. Ahmad (2019). "Examining correlations between motorcyclist's conspicuity, apparel related factors and injury severity score: Evidence from new motorcycle crash causation study." Accident Analysis & Prevention **131**: 45-62.

Wali, B., A. J. Khattak and J. Xu (2018). Contributory fault and level of personal injury to drivers involved in head-on collisions: application of copula-based bivariate ordinal models. Accident Analysis & Prevention **110**: 101-114.

Wali, B., Ahmed, A., & Ahmad, N. (2018, August). An ordered-probit analysis of enforcement of road speed limits. In Proceedings of the Institution of Civil Engineers-Transport (Vol. 171, No. 4, pp. 225-234). Thomas Telford Ltd.

Zhang, M., A. J. Khattak, J. Liu and D. Clarke (2018). A comparative study of rail-pedestrian trespassing crash injury severity between highway-rail grade crossings and non-crossings. Accident Analysis & Prevention **117**: 427-438.

APPENDIX

Table A1. Geographic size of study center areas (Source: Blatt et al. 2015)

Study Center	County	Percentage of County in Study	Water Area (sq. mi.)	Land Area (sq. mi.)	Total Area (sq. mi.)	Notes
Bloomington	Brown	81.3%	3.8434	253.688	257.531	Values provided are for the 39 zip codes for the primary recruiting area. Secondary recruiting area included additional parts of these counties and significant areas of Marion and Bartholomew Counties.[a]
	Dubois	2.6%	0.0454	11.466	11.512	
	Greene	55.1%	1.5957	299.085	300.681	
	Johnson	19.0%	0.5141	60.684	61.198	
	Lawrence	94.8%	2.7608	425.851	428.611	
	Martin	52.6%	2.5698	176.483	179.052	
	Monroe	97.8%	16.8120	385.588	402.399	
	Morgan	67.5%	4.2562	272.241	276.497	
	Orange	85.6%	6.9348	342.327	349.262	
	Owen	64.6%	0.9396	249.718	250.658	
	Putnam	0.7%	0.0000	3.350	3.350	
	Primary total (sq. mi.)		40.2718	2,480.481	2,520.751	
	Total of all primary and secondary (towns, cities, unincorporated areas)				3,800.000	
Buffalo	Erie	100%	183	1,043	1,227	
	Total (sq. mi.)		183	1,043	1,227	
Durham	Chatham	49.5%	23.36	330.81	354.1	Values provided are for targeted zip codes in each county.
	Wake	51.3%	12.21	427.41	439.62	
	Orange	14.0%	0.20	55.90	56.10	
	Durham	89.1%	7.92	257.73	265.65	
	Granville	3.3%	0.58	17.01	17.59	
	Johnston	0.3%	0.06	2.10	2.16	
	Hartnett	1.6%	0.02	9.72	9.74	
	Total (sq. mi.)		44	1,101.00	1,145.00	
Seattle[b]	King	100%	191.3	2,115.57	2,306.87	
	Pierce	100%	136.93	1,669.51	1,806.44	
	Snohomish	100%	109.03	2,087.27	2,196.30	
	Total (sq. mi.)		437.00	5,872.00	6,310.00	
State College	Blair	100%	1	526	527	
	Cambria	100%	5	688	693	
	Centre	100%	4	1,108	1,112	
	Clearfield	100%	7	1,147	1,154	
	Clinton	100%	7	891	898	
	Huntingdon	100%	15	874	889	
	Juniata	100%	2	392	394	
	Mifflin	100%	3	412	415	
	Snyder	100%	1	332	333	
	Union	100%	0	317	317	
	Total (sq. mi.)		45	6,687	6,732	
Tampa	Hillsborough	100%	215	1,051	1,266	
	Pasco	100%	123	745	868	
	Total (sq. mi.)		338	1,796	2,134	

[a] Bloomington detailed data for 39 primary zip codes plus total area for all primary and secondary regions. The Census Bureau has developed approaches for giving zip codes approximate areas, which allows a size estimate to be provided based on zip code tabulation areas (ZCTA); see http://www.census.gov/geo/reference/zctas.html. Accessed April 22, 2014.

[b] State of Washington Office of Financial Management. 2011. *Census 2010 Redistricting Data [P.L. 94-171] for Washington, County Summary, Table 1: Population and Housing.* http://www.ofm.wa.gov/pop/census2010/data.asp. Accessed Nov. 7, 2013.

Table A2. Distribution of Recruits Contacted and Participants based on Geographical Locations (Source: Blatt et al. 2015)

Study Center	Total Recruits in MCS[a]	Total Recruits Contacted	Total Participants[b]	Percentage Contacted Who Became Participants
Bloomington	967	480	254	52.9%
Buffalo	3,444	2,211	740	33.5%
Durham	2,885	2,885	529	18.3%
Seattle	3,629	2,451	715	29.2%
State College	1,166	717	275	38.4%
Tampa	4,267	2,948	734	24.9%
Total	16,358	11,692	3,247	27.8%

[a] Not counting duplicate entries for drivers entered more than once in MCS because they switched vehicles.
[b] Totals include primary, additional primary, and AVT participants who were in the study at least 1 day (VTTI 2014).

Table A3. Design of Sample with Target & Actual Cell Values: Driver Age (Source: Blatt et al. 2015)

Gender	Age Range (years)	Age-Range Description	Planned Primary Participants[a]	Actual Participants[b]	Delta (Actual – Planned)
M	16–17	Minor teen	172	119	–53
M	18–20	Adult teen	172	237	65
M	21–25	Young adult	172	245	73
M	26–35	Adult	172	158	–14
M	36–50	Middle adult	172	156	–16
M	51–65	Mature adult	172	157	–15
M	66–75	Younger older driver	172	166	–6
M	76+	Older older driver	172	249	77
F	16–17	Minor teen	172	143	–29
F	18–20	Adult teen	172	289	117
F	21–25	Young adult	172	348	176
F	26–35	Adult	172	150	–22
F	36–50	Middle adult	172	165	–7
F	51–65	Mature adult	172	182	10
F	66–75	Younger older driver	172	148	–24
F	76+	Older older driver	172	199	27
Not specified			0	1	1
Any Advanced Vehicle Technology (AVT)			350	135	–215
Total			3,102	3,247	145

[a] Campbell (2010).
[b] Includes 3,200 primary and 47 additional primary drivers in study at least 1 day. Data provided by VTTI (2014).

Table A4. Participants by Age Group and Gender for All Test Sites (Source: Blatt et al. 2015)

Age Group (years)	Gender	Total Participants for All Study Centers[a] Primary Driver	Additional Primary Driver	Total by Gender	Total by Age Group	Secondary Driver[b]
16–17	Male	109	10	119	262	0
	Female	140	3	143		0
18–20	Male	233	4	237	526	4
	Female	284	5	289		4
21–25	Male	241	4	245	593	8
	Female	345	3	348		8
26–35	Male	156	2	158	308	12
	Female	148	2	150		17
36–50	Male	153	3	156	321	16
	Female	161	4	165		15
51–65	Male	154	3	157	339	19
	Female	181	1	182		23
66–75	Male	166	0	166	314	19
	Female	148	0	148		10
76+	Male	248	1	249	448	4
	Female	197	2	199		6
AVT	Both	135	0	135	135	0
Not specified		1	0	1	1	44
Total		3,200	47	3,247	3,247	209

[a] Primary participants and secondary drivers with at least 1 day in the study are included.
[b] Secondary drivers are not counted toward "participant" age group total. Only secondary drivers with consent date and reference image are included in secondary driver totals. Note that age and gender are available for 79% of secondary drivers (if designations are unavailable, drivers are included in "not specified").

TABLE B1. Deriving Key Driving Errors and Violations from Driver Behavior Variable in NDS (SHRP2) Data

Driver Behavior (in NDS Data)	Definition (NDS Data Dictionary)	Taxonomy for Driving Errors and Violations					
		Recognition Error	Decision Error	Performance Error	Violations	Physical Condition	Experience or Exposure Error
None (or No Additional Driver Behaviors)	Subject vehicle driver engages in no apparent behavior(s) related to causing or contributing to the crash or near-crash.	No	No	No	No	No	No
Distracted	Subject vehicle driver is not maintaining acceptable attention to the driving task due to engagement in one or more secondary tasks. Includes times when any Secondary Task has an Outcome that is not No. This is a subjective judgment call by the video analyst indicating whether any secondary tasks the driver might be involved in (Variables 32, 36, 40) contributed to the crash or near crash (Variables 35, 39, 43). NOTE: This category is excluded from Baseline analysis.	Yes	No	No	No	No	No
Drowsy, sleepy, asleep, fatigued	Subject vehicle driver exhibits obvious signs of being asleep or tired, or is actually asleep while driving, degrading performance of the driving task. This should also be coded as Drowsy, sleepy, asleep, fatigued under Driver Impairment.	No	No	No	No	Yes	No
Exceeded speed limit	Subject vehicle traveling at a speed greater than the posted speed limit (not in a work zone). In Variable Speed Zones, this is relative to the speed limit in effect at the time of the event.	No	No	No	Yes	No	No
Exceeded safe speed but not speed limit	Subject vehicle traveling at a speed close to or under the posted speed limit, but still too fast to maintain a safe driving environment given current environmental conditions (e.g., weather, traffic, lighting). (Not in a work zone.)	No	Yes	No	No	No	No
Driving slowly: below speed limit	Subject vehicle traveling at a speed much lower than the posted speed limit when higher speeds are appropriate.	No	Yes	No	No	No	No
Driving slowly in relation to other traffic: not below speed limit	Subject vehicle traveling much slower than other vehicles in traffic stream (but not substantially below the posted speed limit).	No	Yes	No	No	No	No

TABLE B1. Deriving Key Driving Errors and Violations from Driver Behavior Variable in NDS (SHRP2) Data (Continued)

Passing on right	Subject vehicle deliberately passes another vehicle in the lane immediately to the right of the other vehicle.	No	No	No	Yes	No	No
Illegal passing	Subject vehicle passes another vehicle in an unsafe or illegal manner (other than on the right).	No	No	No	Yes	No	No
Other improper or unsafe passing	Subject vehicle passes another vehicle in an improper manner not included in previous categories.	No	Yes	No	No	No	No
Cutting in, too close in front of other vehicle	Subject vehicle enters lane of another vehicle too closely to the front of that vehicle.	No	Yes	No	No	No	No
Cutting in, too close behind other vehicle	Subject vehicle enters lane of another vehicle too closely to the back of that vehicle.	No	Yes	No	No	No	No
Making turn from wrong lane	Subject vehicle turns left or right from a lane not intended for making that turn.	No	Yes	No	No	No	No
Did not see other vehicle during lane change or merge	Subject vehicle enters a lane or merges into a lane without being aware of another vehicle close by that is already traveling in that lane.	Yes	No	No	No	No	No
Driving in other vehicle's blind zone	Subject vehicle is traveling close to another vehicle in such a way that the driver of the other vehicle is not expected to be able to see it. Subject vehicle must maintain this relative position for at least 5 seconds.	No	Yes	No	No	No	No
Aggressive driving, specific, directed menacing actions	Subject vehicle driver is driving in a purposefully aggressive manner, with actions intended for a specific recipient.	No	Yes	No	No	No	No
Aggressive driving, other	Driver is driving in an aggressive manner not described in previous categories. Includes reckless and "sporty" driving.	No	Yes	No	No	No	No
Wrong side of road, not overtaking	Subject vehicle is traveling on the wrong side of the road with no intent of passing or overtaking another vehicle.	No	Yes	No	No	No	No
Following too closely	Subject vehicle is traveling at an unsafe distance (too close) behind the lead vehicle.	No	Yes	No	No	No	No
Failed to signal	Subject vehicle failed to properly signal its intent by not signaling at all. Applies to planned maneuvers, not sudden evasive maneuvers.	No	No	No	Yes	No	No

TABLE B1. Deriving Key Driving Errors and Violations from Driver Behavior Variable in NDS (SHRP2) Data (Continued)

Improper signal	Subject vehicle failed to properly signal its intent by signaling incorrectly. Use with planned maneuvers, not sudden evasive maneuvers.	Yes	No	No	No	No	No
Improper turn, wide right turn	Subject vehicle turned right from the initial travel path, unnecessarily encroaching into the left adjacent lane or median.	No	No	Yes	No	No	No
Improper turn, cut corner on right turn	Subject vehicle turned right from the initial travel path, unnecessarily encroaching into the right adjacent lane or shoulder/curb.	No	Yes	No	No	No	No
Improper turn, wide left turn	Subject vehicle turned left from the initial travel path, unnecessarily encroaching into the right adjacent lane or shoulder/curb.	No	No	Yes	No	No	No
Improper turn, cut corner on left	Subject vehicle turned left from the initial travel path, unnecessarily encroaching into the left adjacent lane or median.	No	Yes	No	No	No	No
Improper turn, other	Subject vehicle turned left or right from the initial travel path in an unsafe manner not described in previous categories.	No	Yes	No	No	No	No
Improper backing, did not see	Subject vehicle traveled in reverse without obtaining a proper view of the surroundings behind the vehicle.	Yes	No	No	No	No	No
Improper backing, other	Subject vehicle traveled in reverse in an unsafe manner not described in previous categories.	Yes					
Improper start from parked position	Subject vehicle moved from a parked position in an unsafe manner.	No	Yes	No	No	No	No
Disregarded officer or watchman	Subject vehicle driver did not notice or obey an officer of the law or traffic guard serving to provide guidance in traffic flow and the driving task.	No	No	No	Yes	No	No
Signal violation, apparently did not see signal	Subject vehicle driver did not notice and thus disobeyed (or nearly disobeyed) a traffic signal.	No	No	No	Yes	No	No
Signal violation, intentionally disregarded signal	Subject vehicle driver saw a traffic signal but purposefully disregarded its instruction. (If driver was trying to beat a yellow light before it phased into red, code "Signal violation, tried to beat signal change".)	No	No	No	Yes	No	No
Signal violation, tried to beat signal change	Subject vehicle driver accelerated or continued at a speed intended to pass through an intersection before the traffic signal turned red.	No	No	No	Yes	No	No

TABLE B1. Deriving Key Driving Errors and Violations from Driver Behavior Variable in NDS (SHRP2) Data (Continued)

Stop sign violation, apparently did not see stop sign	Subject vehicle driver did not notice and thus disobeyed or nearly disobeyed a stop sign.	No	No	No	Yes	No	No
Stop sign violation, intentionally ran stop sign at speed	Subject vehicle driver saw a stop sign but purposefully drove through the intersection at a speed greater than 15 mph.	No	No	No	Yes	No	No
Stop sign violation, "rolling stop"	Subject vehicle driver did not come to a complete stop at a stop sign (minimum speed was below 15 mph, but above 0 mph).	No	No	No	Yes	No	No
Other sign (e.g., Yield) violation, apparently did not see sign	Subject vehicle driver did not notice and thus disobeyed a traffic sign (other than a stop sign).	No	No	No	Yes	No	No
Other sign (e.g., Yield) violation, intentionally disregarded	Subject vehicle Driver saw a traffic sign (other than a stop sign) but purposefully disobeyed that sign.	No	No	No	Yes	No	No
Other sign violation	Subject vehicle driver disobeyed a traffic sign in a manner not described in previous categories.	No	No	No	Yes	No	No
Non-signed crossing violation	Subject vehicle driver proceeded through a non-signed intersection in an unsafe manner.	No	No	No	Yes	No	No
Right-of-way error in relation to other vehicle or person, apparent recognition failure	Subject vehicle driver made the incorrect decision regarding who had the right-of-way (his/her own vehicle or another vehicle or pedestrian) due to a misunderstanding of the situation.	Yes	No	No	No	No	No
Right-of-way error in relation to other vehicle or person, apparent decision failure	Driver made the incorrect decision regarding who had the right-of-way (his/her own vehicle or another vehicle or pedestrian) due to improper analysis of the situation.	No	Yes	No	No	No	No
Right-of-way error in relation to other vehicle or person, other or unknown cause	Subject vehicle driver made incorrect decision regarding who had the right-of-way (his/her own vehicle or another vehicle or pedestrian) for an unknown reason or for reasons not described in previous categories.	No	Yes	No	No	No	No
Sudden or improper braking	Subject vehicle braked suddenly or in an unsafe manner in the roadway, but did not come to a complete stop (i.e., speed indicator did not drop to zero).	No	Yes	No	No	No	No
Sudden or improper stopping on roadway	Subject vehicle stopped (speed indicator dropped to zero) without ample warning or in an unsafe manner in the roadway.	No	Yes	No	No	No	No

TABLE B1. Deriving Key Driving Errors and Violations from Driver Behavior Variable in NDS (SHRP2) Data (Continued)

Parking in improper or dangerous location	Subject vehicle parked (stopped with the intent of remaining stopped) in a location not intended for parking.	No	Yes	No	No	No	No
Speeding or other unsafe actions in work zone	Subject vehicle traveling at a speed greater than the posted speed limit, specifically while driving in a work zone.	No	No	No	Yes	No	No
Failure to dim headlights	Subject vehicle traveling with high beams activated on headlights, without dimming the lights when appropriate.	No	Yes	No	No	No	No
Driving without lights or with insufficient lights	Subject vehicle traveling with no headlights on (or insufficient headlights) when the situation requires such lighting for safety.	No	Yes	No	No	No	No
Avoiding pedestrian	Subject vehicle driver behaved in a manner intended to avoid conflict with a pedestrian.	No	Yes	No	No	No	No
Avoiding other vehicle	Subject vehicle driver behaved in a manner intended to avoid conflict with another vehicle.	No	Yes	No	No	No	No
Avoiding animal	Subject vehicle driver behaved in a manner intended to avoid conflict with an animal.	No	Yes	No	No	No	No
Apparent unfamiliarity with roadway	Subject vehicle driver behaved in an unsafe manner, apparently due to an unfamiliarity with the surrounding traffic situation or locality.	No	No	No	No	No	Yes
Apparent unfamiliarity with vehicle	Subject vehicle driver behaved in an unsafe manner, apparently due to an unfamiliarity with the vehicle.	No	No	No	No	No	Yes
Apparent general inexperience driving	Subject vehicle driver behaved in an unsafe manner, apparently due to lack of experience with the driving task.	No	No	No	No	No	Yes
Use of cruise control contributed to late braking	Subject vehicle driver delayed applying brake pedal because the cruise control was activated, resulting in an unsafe situation.	No	Yes	No	No	No	No
Unknown	Cannot determine the behavior(s) engaged in by the subject vehicle driver due to limitations in video views, lighting, visual obstructions, or limited perspective.	---	---	---	---	---	---

Note: The first two columns include driver behaviors and their definitions as per the NDS (SHRP2) dictionary and are copied from NDS (SHRP2) data dictionary.

TABLE B2: Deriving Key Driving Errors and Violations from Secondary Task and Secondary Task Outcome in NDS (SHRP2) Data

Secondary Task	Secondary Task Outcome*	Taxonomy for Driving Errors and Violations					
		Recognition Error	Decision Error	Performance Error	Violations	Physical Condition	Experience or Exposure Error
Cell phone, Talking/listening, hands-free	Yes	Yes	No	No	No	No	No
Cognitive, other	Yes	Yes	No	No	No	No	No
Other non-specific internal eye glance	Yes	Yes	No	No	No	No	No
Other personal hygiene	Yes	Yes	No	No	No	No	No

Note: The Secondary Task Outcome (*) is a factor/variable in the NDS data which shows whether a particular secondary task contributed to the occurrence or severity of safety critical events (i.e., crash or near-crash). After classifying the driver behaviors (i.e., presented in Table A1) into key six driving errors and violations, we checked the secondary task and secondary task outcome to see if there could be some safety critical events (for which no driver behavior was reported in the NDS data) but could have occurred due to involvement of subject drivers in secondary tasks (given that it contributed to safety critical events: determined from secondary task outcome).The secondary tasks reported in A2 do not represent all of the secondary tasks (i.e., there are many more in the NDS data), but only those which were reported in safety critical events where no driver behavior was reported and these particular secondary tasks contributed to safety critical events.

TABLE B3: Deriving Key Driving Errors and Violations from Other Driver Behaviors Extracted from Narratives in NDS (SHRP2) Data

Definition	Taxonomy for Driving Errors and Violations					
	Recognition Error	Decision Error	Performance Error	Violations	Physical Condition	Experience or Exposure Error
Does not realize concrete barrier	Yes	No	No	No	No	No
Misjudgment of space or situations	Yes	No	No	No	No	No
Drifting	No	Yes	No	No	No	No
Steered too far	No	No	Yes	No	No	No
Inappropriate acceleration/deceleration	No	No	Yes	No	No	No
Illegal U-turn	No	No	No	Yes	No	No

Note: There were some safety critical events for which there was no Driver Behavior, Secondary Task (that has contributed to crash or near-crash), roadway environment factors (e.g., surface conditions, visual obstructions, infrastructure factors), and vehicle factors were reported. Hence, we evaluated the narrative (detailed description of the situation) in the NDS (SHRP2) data and found other Driver Behaviors factors (as presented in Table A3).

Table C1. Direct, Indirect, and Total Effects of Explanatory Variables (Path Analysis)

Independent Variables	Direct Effect on Crash (β') (%)	Effects on Driving Errors (%)						Indirect Effects on Crash (%)						Total Effect through Each Error Type on Crash (%)					
		β_1	β_2	β_3	β_4	β_5	β_6	β_1* β''_1	β_2* β''_2	β_3* β''_3	β_4* β''_4	β_5* β''_5	β_6* β''_6	A	B	C	D	E	F
Driver Errors (Base outcome = No driving errors)																			
Recognition Errors (β''_1)	29.28	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Decision Errors (β''_2)	23.56	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Performance Errors (β''_3)	54.88	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Violations (β''_4)	4.93	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Physical Condition Errors (β''_5)	4.16	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Experience/Exposure Errors (β''_6)	32.41	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Roadway Locality and Environment (Base outcome = Interstate)																			
Open Country or Open Residential	---	---	3.28	---	---	---	---	---	0.77	---	---	---	---	---	0.77	---	---	---	---
Moderate Residential	---	1.30	1.68	0.94	---	---	---	0.38	0.40	0.51	---	---	---	0.38	0.40	0.51	---	---	---
School	1.74	4.78	3.45		-3.24	---	---	1.40	0.81		-0.16	---	---	3.14	2.55	1.74	1.58	1.74	1.74
Business/Industrial	1.29	2.51	1.77	0.58	-3.69	---	---	0.74	0.42	0.32	-0.18	---	---	2.02	1.70	1.61	1.10	1.29	1.29
Urban	5.53	8.00	6.65	1.33	---	1.27	0.65	2.34	1.57	0.73	---	0.05	0.21	7.87	7.10	6.26	5.53	5.58	5.74
Bypass/Divided Highway with traffic signals	1.23	---	---	1.32	---	1.38	---	---	---	0.72	---	0.06	---	1.23	1.23	1.95	1.23	1.29	1.23
Others (church, playground, & Campground)	1.61	4.17	3.09	1.17	---	---	---	1.22	0.73	0.64	---	---	---	2.83	2.34	2.25	1.61	1.61	1.61
Intersection Influence (Base outcome = No Intersection Influence)																			
Interchange influence	7.03	7.25	8.93	---	---	---	1.52	2.12	2.10	---	---	---	0.49	9.15	9.14	7.03	7.03	7.03	7.52
Stop sign or traffic signal influence	4.56	7.22	0.86	0.30	7.30	---	---	2.11	0.20	0.16	0.36	---	---	6.67	4.76	4.72	4.92	4.56	4.56
Uncontrolled intersection influence	8.27	9.17	6.86	1.76	5.04	---	1.03	2.68	1.62	0.97	0.25	---	0.33	10.95	9.88	9.23	8.52	8.27	8.60
Parking lot or driveways influence	9.23	11.38	15.23	3.24	6.33	---	---	3.33	3.59	1.78	0.31	---	---	12.56	12.82	11.01	9.54	9.23	9.23
Others intersection influence	6.38	17.16	8.05	---	6.97	---	---	5.03	1.90		0.34	---	---	11.40	8.27	6.38	6.72	6.38	6.38
Construction zone indicator	0.96	3.87	---	---	4.19		0.43	1.13	---	---	0.21	---	0.14	2.09	0.96	0.96	1.17	0.96	1.10
Secondary Task Duration	0.18	1.84	0.19	0.02	---	-0.37	0.03	0.54	0.05	0.01	---	-0.02	0.01	0.72	0.23	0.20	0.18	0.17	0.19

Notes: **β'** is the direct effect of explanatory variables on crash outcome (obtained from Ordered probit model); $\boldsymbol{\beta}_1$ through $\boldsymbol{\beta}_6$ indicates the effects (obtained from Multinomial logit model) of explanatory factors on recognition, decision, performance, violations, physical condition related, and experience/exposure errors respectively; whereas A through F indicates the total effects corresponding to each of the explanatory factors, i.e., $A = \boldsymbol{\beta}' + \boldsymbol{\beta 1} * \boldsymbol{\beta}''\mathbf{1}$; $B = \boldsymbol{\beta}' + \boldsymbol{\beta 2} * \boldsymbol{\beta}''\mathbf{2}$; $C = \boldsymbol{\beta}' + \boldsymbol{\beta 3} * \boldsymbol{\beta}''\mathbf{3}$; $D = \boldsymbol{\beta}' + \boldsymbol{\beta 4} * \boldsymbol{\beta}''\mathbf{4}$; $E = \boldsymbol{\beta}' + \boldsymbol{\beta 5} * \boldsymbol{\beta}''\mathbf{5}$; $F = \boldsymbol{\beta}' + \boldsymbol{\beta 6} * \boldsymbol{\beta}''\mathbf{6}$; (---) indicates not applicable.

Collaborative Sciences Center for
ROAD SAFETY

730 Martin Luther King Jr. Blvd.
Suite 300
Chapel Hill, NC 27599-3430
info@roadsafety.unc.edu

www.roadsafety.unc.edu