Report S2-L10-RR-1

Feasibility of Using In-Vehicle Video Data to Explore How to Modify Driver Behavior That Causes Nonrecurring Congestion

SHRP2 RELIABILITY RESEARCH

SHRP 2
STRATEGIC HIGHWAY RESEARCH PROGRAM
Accelerating solutions for highway safety, renewal, reliability, and capacity

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

The Second

STRATEGIC HIGHWAY RESEARCH PROGRAM

SHRP 2 Report S2-L10-RR-01

Feasibility of Using In-Vehicle Video Data to Explore How to Modify Driver Behavior That Causes Nonrecurring Congestion

H. Rakha, J. Du, S. Park, F. Guo, Z. Doerzaph, and D. Viita
Virginia Tech Transportation Institute
Blacksburg, Virginia

G. Golembiewski, B. Katz, N. Kehoe, and H. Rigdon
Science Applications International Corporation
McLean, Virginia

TRANSPORTATION RESEARCH BOARD

WASHINGTON, D.C.
2011
www.TRB.org

Subscriber Categories

Data and Information Technology
Highways
Operations and Traffic Management
Safety and Human Factors

The Second Strategic Highway Research Program

America's highway system is critical to meeting the mobility and economic needs of local communities, regions, and the nation. Developments in research and technology—such as advanced materials, communications technology, new data collection technologies, and human factors science—offer a new opportunity to improve the safety and reliability of this important national resource. Breakthrough resolution of significant transportation problems, however, requires concentrated resources over a short time frame. Reflecting this need, the second Strategic Highway Research Program (SHRP 2) has an intense, large-scale focus, integrates multiple fields of research and technology, and is fundamentally different from the broad, mission-oriented, discipline-based research programs that have been the mainstay of the highway research industry for half a century.

The need for SHRP 2 was identified in *TRB Special Report 260: Strategic Highway Research: Saving Lives, Reducing Congestion, Improving Quality of Life,* published in 2001 and based on a study sponsored by Congress through the Transportation Equity Act for the 21st Century (TEA-21). SHRP 2, modeled after the first Strategic Highway Research Program, is a focused, time-constrained, management-driven program designed to complement existing highway research programs. SHRP 2 focuses on applied research in four areas: Safety, to prevent or reduce the severity of highway crashes by understanding driver behavior; Renewal, to address the aging infrastructure through rapid design and construction methods that cause minimal disruptions and produce lasting facilities; Reliability, to reduce congestion through incident reduction, management, response, and mitigation; and Capacity, to integrate mobility, economic, environmental, and community needs in the planning and designing of new transportation capacity.

SHRP 2 was authorized in August 2005 as part of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU). The program is managed by the Transportation Research Board (TRB) on behalf of the National Research Council (NRC). SHRP 2 is conducted under a memorandum of understanding among the American Association of State Highway and Transportation Officials (AASHTO), the Federal Highway Administration (FHWA), and the National Academy of Sciences, parent organization of TRB and NRC. The program provides for competitive, merit-based selection of research contractors; independent research project oversight; and dissemination of research results.

SHRP 2 Report S2-L10-RR-1

ISBN: 978-0-309-12898-8
LOCCN: 2011932569

Copyright Information

Note: SHRP 2 report numbers convey the program, focus area, project number, and publication format. Report numbers ending in "w" are published as web documents only.

Notice

The project that is the subject of this report was a part of the second Strategic Highway Research Program, conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council.

The members of the technical committee selected to monitor this project and to review this report were chosen for their special competencies and with regard for appropriate balance. The report was reviewed by the technical committee and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the Governing Board of the National Research Council.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board, the National Research Council, or the program sponsors.

The Transportation Research Board of the National Academies, the National Research Council, and the sponsors of the second Strategic Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

SHRP 2 Reports

Available by subscription and through the TRB online bookstore:

www.TRB.org/bookstore

Contact the TRB Business Office:
202-334-3213

More information about SHRP 2:

www.TRB.org/SHRP2

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

SHRP 2 STAFF

ACKNOWLEDGMENTS

This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials. It was conducted in the second Strategic Highway Research Program (SHRP 2), which is administered by the Transportation Research Board of the National Academies. The project was managed by Ralph Hessian, Visiting Professional for SHRP 2 Reliability.

Virginia Tech Transportation Institute (VTTI) is the primary contractor for this study and is supported by subcontracts through Science Applications International Corporation (SAIC). Dr. Hesham A. Rakha, Director of the Center for Sustainable Mobility (CSM) at VTTI, is the Principal Investigator for this study. The other authors of this report are Dr. Jianhe Du, Senior Research Associate, CSM; Dr. Sangjun Park, Research Associate, CSM; Dr. Feng Guo, Assistant Professor, Statistics Department, Virginia Tech; Dr. Zachary Doerzaph, Senior Research Associate, Center for Vehicle-Infrastructure Safety (CVIS) at VTTI; Derek Viita, Research Associate, CVIS; and Ahmed Amer, Hao Chen, and Ismail Zohdy, graduate student assistants at VTTI.

Authors from SAIC are Gary Golembiewski, Research Psychologist; Dr. Bryan Katz, Senior Transportation Engineer; Nicholas Kehoe, Research Engineer; and Heather Rigdon, Research Psychologist.

FOREWORD

William Hyman, *SHRP 2 Senior Program Officer, Reliability*

This research report—a product of the Reliability focus area of the second Strategic Highway Research Program (SHRP 2)—presents findings on the feasibility of using existing in-vehicle data sets, collected in naturalistic driving settings, to make inferences about the relationship between observed driver behavior and nonrecurring congestion. General guidance is provided on the protocols and procedures for conducting video data reduction analysis. In addition, the report includes technical guidance on the features, technologies, and complementary data sets that researchers should consider when designing future instrumented in-vehicle data collection studies. Finally, a new modeling approach is advanced for travel time reliability performance measurement across a variety of traffic congestion conditions.

Traffic congestion continues to grow on the nation's highways, increasing the concerns of transportation agencies, the business community, and the general public. Congestion includes recurring and nonrecurring components. Recurring congestion reflects routine day-to-day delays during specific time periods where traffic demand exceeds available roadway capacity. Road users come to expect these daily traffic patterns, and they adjust their travel plans accordingly to achieve timely arrivals. Nonrecurring congestion results from random incidents, such as crashes, weather, and work zones, that cause unexpected extra delays. Road users are frustrated by these unexpected delays, which can make for unreliable arrival times at their destinations. The SHRP 2 Reliability research objective focuses on reducing nonrecurring congestion through incident reduction, management, response, and mitigation. Achieving this objective will improve travel time reliability for both people and freight.

Human factors contribute to traffic operating conditions and safety performance on public roads. This research seeks to better understand how driver behavior influences the primary causes of nonrecurring congestion and to identify countermeasures to modify these behaviors. The research team identified domestic and international candidate studies on driver behavior conducted in recent years that captured video driver behavior data sets and other supplementary data. Evaluation criteria were established to determine the key dimensions of feasibility for selecting the best candidate studies to investigate in detail.

The research results provide the foundation for recommendations on the feasibility of using existing data sets for this research purpose; general guidance on the proper protocols, procedures, and facilities to conduct video data reduction; and technical guidance on features, technologies, and complementary data—all of which should be considered in designing future in-vehicle video data collection studies to explicitly examine driver behavior and the impacts on nonrecurring congestion.

CONTENTS

Color versions of the figures and tables in this report are available online:
www.trb.org/Main/Blurbs/165281.aspx.

Executive Summary

Nonrecurring congestion is traffic congestion due to nonrecurring causes, such as crashes, disabled vehicles, work zones, adverse weather events, and planned special events. According to data from the Federal Highway Administration (FHWA), approximately half of all congestion is caused by temporary disruptions that remove part of the roadway from use, or "nonrecurring" congestion. These nonrecurring events dramatically reduce the available capacity and reliability of the entire transportation system. The objective of this project is to determine the feasibility of using in-vehicle video data to make inferences about driver behavior that would allow investigation of the relationship between observable driver behavior and nonrecurring congestion to improve travel time reliability. The data processing flow proposed in this report can be summarized as (1) collect data, (2) identify driver behavior, (3) identify correctable driver behavior, and (4) model travel time reliability, as shown in Figure ES.1.

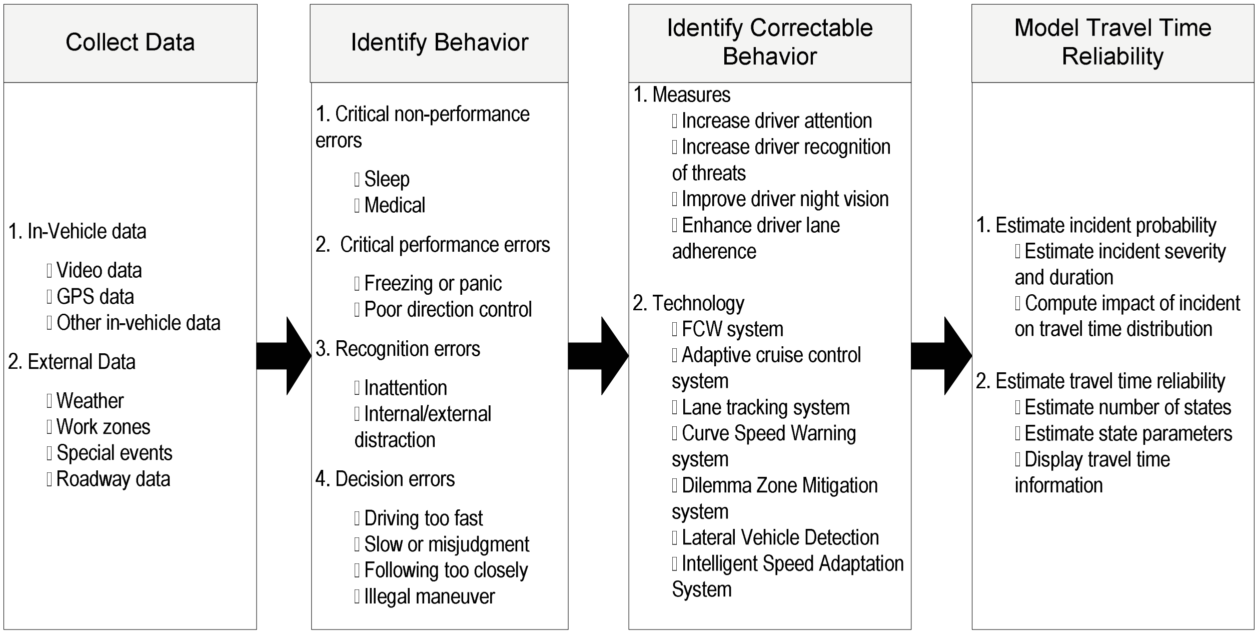
Introduction

Key domestic and international studies in which in-vehicle video cameras were used to collect data were investigated in this study. The research team reviewed video, kinematic, and external data collected in each candidate data set. To quantitatively assess the qualification of candidate data sets, dimensions of feasibility were defined to evaluate the data sets with respect to legal restrictions, comprehensiveness, video data quality, in-vehicle data quality, linkages to external data, and format and structure. A list of qualified data sets was identified for further data reduction and analysis.

The original research goals, data reduction process, and data formats of these studies were examined by the research team. The video data were manually reviewed, and additional data reduction was conducted to identify contributing factors to crashes and near crashes using video data and supplementary data. If the events were caused by inappropriate driver behavior or driver inattention, then countermeasures were suggested to prevent these safety-related events.

In modeling travel time reliability, the team reviewed models used by other researchers. A multimode distribution of travel time was then proposed to model travel time variations. A statistical method to model travel time reliability was developed that considers two factors: the probability of encountering congestion and the probability that certain estimated travel times will be experienced when congestion occurs.

Potential problems and risks associated with existing data, including kinematic data, video data, reduced data, and external data, were identified during the data reduction and analysis phase. To facilitate future data collection efforts, including those related to in-vehicle data, other external data were proposed to improve the efficiency of data collection, reduction, and analysis.



FCW = forward crash warning.

Figure ES.1. Data processing flow.

Findings

According to data reduction results, most crashes or near crashes are caused by driver inattention and errors. These events might be prevented if appropriate instrumentation were installed to warn drivers in a timely manner. The following factors imply that the proposed systems can resolve driver inattention and errors and thus prevent a collision:

1. In the Road Departure Crash Warning System (RDCWS) Field Operational Test (FOT), the largest contributing factor to freeway crashes and near crashes was decision errors, including driving too fast or too slowly, following too closely, and misjudging a gap; more than 85% of the events were caused by this driver-related factor. For events that occurred on arterials, the same pattern followed. The next largest category for both road types was recognition errors, including inattention, inadequate surveillance, and other types of distraction; more than 5% of events were ascribed to this category.
2. In the 100-Car Study, the largest contributing factor category for crashes was driver recognition errors, totaling 32% of the events. The second largest category was decision errors, counting 28% of the total. The largest and second largest contributing factor categories for near crashes were decision errors and recognition errors at 29% and 26%, respectively.
3. In the Drowsy Driver Warning System (DDWS) FOT, a different pattern was exhibited. The most frequent critical reason for crashes was an object in the roadway, which constituted 57% of the events. The next largest groups were the driver-related factors of recognition errors and performance errors; each had more than 14% of the cases related to driver factors. In tire strike cases, the majority were attributed to environment-related factors; more than 64% of the events were ascribed to this category. For near crashes, recognition errors and decision errors constituted 31% and 18%, respectively.
4. In the Naturalistic Truck Driving Study (NTDS), the most frequent critical factor for crashes was an object in the roadway, followed by the driver-related factors of recognition errors, decision errors, and performance errors; each constituted 20% of the total cases. Not surprisingly, almost all (75%) the tire strikes involved some type of improper turn. The second and third largest categories of contributing factor for crashes were driver performance errors and decision errors, respectively. For near crashes, the most frequent factor was driver-related recog-

nition errors; more than 40% of near crashes resulted from inattention or distraction. Among these near crashes, almost one-quarter involved the subject driver's not seeing the other vehicle during a lane change or merge.

Countermeasures were proposed to correct the driver behaviors that caused these events. Almost all the crashes in the RDCWS FOT study have the potential to be prevented if one or multiple countermeasures are applied; 91% of the near crashes in that study are correctable. In the 100-Car Study, almost 40% of the crashes can or are likely to be prevented, and more than 80% of the near crashes can or are likely to be prevented given reasonable countermeasures. In the two truck studies, all the crashes, tire strikes, and near crashes are preventable using appropriate countermeasures.

To model nonrecurring congestion related to modifying driver behavior, it is ideal to find a substantial number of crashes that result in changes in traffic conditions. The congestion, therefore, can be monitored and modeled. According to the data reduction results, not enough realizations of such crashes occurred. Other supplemental data sets were used to construct the statistical model. The travel time data in the 100-Car Study were used only to validate the multimode distribution of travel time that is being proposed; the model is valid regardless of the source of the travel time data. The travel time reliability model results provide a better fit to field data compared with traditional unimodal travel time model results. The reliability model is also more flexible and provides superior fitting to travel time data compared with single-mode models. It provides a direct connection between the model parameters and the underlying traffic conditions and can be directly linked to the probability of incidents. Thus, it can capture the impact of nonrecurring congestion on travel time reliability.

Conclusions

The team explored the identified data sets to discuss the various issues associated with video and other supplementary data collection and data reduction, to propose models for travel time reliability, and to identify potential problems in data. From the analysis of the naturalistic data sources, this study demonstrates the following:

1. It is feasible to identify driver behavior before near crashes and crashes from video data collected in a naturalistic driving study and to thus infer the causes of those events.
2. Recommendations can be made to change driver behavior and, therefore, prevent crashes and near crashes or reduce the frequency of such events.
3. Naturalistic data are useful to identify impacts of crashes on traffic conditions. Given the small sample of crashes and the fact that the data acquisition system (DAS) does not gather data when the engine is off, it is not possible to study the impact of incidents on travel time reliability. When effectively integrated with external data sources, which is extremely feasible given an accurate time and location stamp in the data set, naturalistic data can be highly efficient in recognizing the relationship between modifying driver behavior and nonrecurring congestion.
4. Increased coordination with weather and traffic volume data is required to determine when nonrecurring congestion exists, as well as to determine what driver actions are a result of these nonrecurring events.
5. It is possible to analyze naturalistic driving data to characterize typical levels of variability in travel times and to develop measures for quantifying travel time reliability.

Limitations of Existing Data Sets

The team reviewed multiple naturalistic driving data sets involving video data that are currently available. In analyzing driver behavior, as is the case with this research effort, high-quality video data is required. In general, all existing data sets are satisfactory in terms of video quality because driver behavior can be clearly viewed regarding decision errors, performance errors, inattention, and recognition errors. The following limitations still exist:

1. Some studies had fewer video cameras installed compared with other studies. For example, the Automotive Collision Avoidance System (ACAS) FOT and the RDCWS FOT conducted by the University of Michigan Transportation Research Institute (UMTRI) had only two video cameras: one facing the driver and the other facing the front view. In these cases, the data sets are limited because traffic conditions beside and behind the subject vehicles are not available. The video frequencies of these UMTRI studies were set relatively low because the original research purposes were not driver-behavior oriented. Consequently, the causal factors of safety-related events are not viewable.
2. Image glare was a typical problem with video data. Some data sets have issues with glare that sometimes make it difficult to make judgments regarding driver behavior.
3. Accidental cable unplugging or malfunction caused incompleteness or errors in data. Although linear interpolation can solve some of the missing data problems, in many cases such problems were not easily detected or corrected.
4. Driver identification is an issue worthy of attention. In a naturalistic driving study, it is not uncommon for the equipped car to be driven by drivers other than the appointed participant. Although the video data can be manually viewed afterward to differentiate drivers in data reduction, it is more efficient if an automatic identification process can be used to tag each trip recorded with driver information so that the data analysis can avoid unnecessary biases.
5. Existing data sources lack a sufficient sample size of crash events to identify changes in driver behavior and the impact of these changes on nonrecurring congestion. The data collection effort in SHRP 2 Safety Project S07, In-Vehicle Driving Behavior Field Study, will offer a unique data set that can be used for further analysis.

Recommendations

To improve the quality of video data in future data collection efforts of this kind (i.e., designed to investigate the reduction of nonrecurring congestion through modifying driver behavior), there are several recommendations.

First, the procedure to recruit participants needs to be carefully designed. It is ideal to include a comprehensive population of drivers ranging evenly across every age, income, and occupation category. When recruiting participants, it is crucial to make it clear that driver information is vital for the research. To better identify drivers, two methods can be used:

1. A formal statement needs to be included in the contract to make the signer the exclusive driver of the vehicle.
2. A touch-screen device can be installed onboard to collect information before and after each trip. The touch-screen equipment can be designed so that a customized interface will be displayed to the driver to input trip-related information by selecting certain check boxes. The before trip information-collecting interface may consist of a list of the first names of household members for the driver to select from as passengers, a list of trip purposes, weather conditions when the trip started, and any information about why the driver selected the time of departure. The after trip information-collecting interface may include an "original trip purpose changed" option, a "route choice changed" option, and a "crashes happened en route" option. Necessary hardware can be designed to connect the input touch screen with the engine so that the driver can start the engine only after the information is input. To ensure safety while driving, the device should be disabled while the vehicle is in motion to prevent driver distraction. One concern with this approach is that it reminds drivers that they are being monitored and thus may deem the study nonnaturalistic.

Second, to serve the research purpose, certain data are more important than others. The following four categories are imperative:

1. Basic onboard equipment should include devices that collect the following data: video; vehicle network information (speed, brake pedal, throttle, turn signal); global positioning system (GPS) data (latitude, longitude, heading); *X, Y,* and *Z* acceleration; distances between the subject and surrounding objects; lane location information (*X, Y, Z*); driver behavior (seat belt usage, lights on or off); and yaw rate.
2. The video cameras should shoot at least five views: front, back, right, left, and the driver. The resolution should be high enough to identify ongoing traffic conditions, weather conditions, and the driver's hand movements and facial expressions. Correction of sun glare to improve video quality is available when needed.
3. The frequency setting should be high enough that the video is continuous, the acceleration and deceleration of the vehicles clearly recorded, and the reaction times recorded and measured. The recommended minimum frequency for GPS devices is 1 Hz and for all other equipment, 10 Hz.
4. To improve the versatility of the data so that the data can be used in other related research, vehicle performance parameters such as engine speed, throttle position, and torque should be recorded.

Third, the data collection system needs to run for an additional 10 minutes after the engine is turned off in case an accident occurs. During the data reduction, data collection usually halted as soon as the driver stopped the vehicle. Because it is important to observe the traffic conditions being affected by a safety-related event, additional data are required after a driver turns off the engine. One concern is that if some malfunction to the subject vehicle occurs (in case of an accident), gathering data may cause a safety hazard. This issue needs further investigation.

Fourth, to improve linking vehicle data with external data, it is ideal to standardize the format for time and location information. For vehicle data, the synchronized GPS clock should be used rather than local computer time for better connection of the data with external traffic, crash, work zone, and weather data. For external data, some states have their database built on the milepost system. The conversion of mileage post locations to a standard latitude and longitude should be conducted ahead of time.

Fifth, because a limited number of crashes—especially severe accidents that affected traffic conditions—occurred in all the candidate data sets, certain adjustments are needed to create a statistically significant database. A longer data collection effort or more drivers involved in the study would be ideal. For example, SHRP 2 Safety Project S07, In-Vehicle Driving Behavior Field Study (a 2,500-Car Study), which will soon be conducted, is a quality candidate. Another solution is simulation, which can be used to compensate for data shortage.

Sixth, additional analysis of existing data is required to study typical levels of variability in driver departure times and in trip travel times and the level of variability in driver route choices. A characterization of this behavior is critical in attempting to quantify and develop travel time reliability measures and to understand the causes of observed travel time reliability. The data may be augmented with tests on a driving simulator to study the impact of travel time reliability on driver route-choice behavior.

Finally, although numerous studies have used video cameras to gather data, an ideal starting point is a compiled data source list that summarizes existing video-involved studies with specifications of data collected, limitations of data usage, and access issues. Such a list will help prevent redundancy in future investigation efforts.

CHAPTER 1

Introduction

SHRP 2 was established in 2006 to investigate the underlying causes of highway crashes and congestion in a short-term program of four interrelated focus areas: Safety (significantly improve highway safety by understanding driving behavior in a study of unprecedented scale); Renewal (develop design and construction methods that cause minimal disruption and produce long-lived facilities to renew the aging highway infrastructure); Reliability (reduce congestion and improve travel time reliability through incident management, response, and mitigation); and Capacity (integrate mobility, economic, environmental, and community needs into the planning and design of new transportation capacity). This report results from Project L10, part of the Reliability research of SHRP 2.

Nonrecurring congestion is traffic congestion that results from nonrecurring causes, such as crashes, disabled vehicles, work zones, adverse weather events, and planned special events. According to data from the Federal Highway Administration (FHWA), approximately half of all congestion is caused by temporary disruptions that remove part of the roadway from use, or nonrecurring congestion. These nonrecurring events dramatically reduce the available capacity and reliability of the entire transportation system. Three main causes of nonrecurring congestion are incidents ranging from a flat tire to an overturned hazardous material truck (25% of congestion), work zones (10% of congestion), and weather (15% of congestion). Accidents resulting in fatalities and injuries may occur if a driver behaves in an inappropriate or less-than-optimal manner in response to such factors as incident scenes, work zones, inclement weather, roadside distractions, and queues of vehicles.

In-vehicle video, along with other data, can potentially provide insight regarding how to modify driver behavior that contributes to nonrecurring congestion. The objective of this project is to determine the feasibility of using in-vehicle video data to make inferences about driver behavior that would allow investigation of the relationship between observable driver behavior and nonrecurring congestion to improve travel time reliability. The successful execution of this research effort requires an in-depth understanding of existing research using video camera data and research about the modeling of travel time reliability.

The team investigated key domestic and international studies that used in-vehicle video cameras to collect data. After an initial screening of candidate data sets, the feasibility of using existing data was assessed by assigning scores to five general areas: access to the data; comprehensiveness of the data; types of vehicle performance data collected; ability to link to operational, traffic control, work zone, and environmental data; and data structure and format. The evaluation results generated a list of potential data sets. The data sets were examined in further detail, determining criteria to identify crashes and near crashes, contributing factors to these safety-related events, and countermeasures to prevent these events. The team then reviewed literature in both the traffic engineering and human factors arena on safety impacts of driver behavior and travel time reliability. A new statistical method was also developed to model travel time reliability.

To provide constructive suggestions for the next stage of research, potential problems and risks were identified and reviewed to determine strategies to address these shortcomings. Issues worthy of notice for future data collection efforts and the general guidelines for video data reduction and analysis are also discussed.

The current report conveys the results of this effort, beginning with an introduction of the existing video-based driver performance databases (Tasks 1 and 2, as described in Chapter 2). Chapter 3 details each data set and identifies qualified data sources by assigning scores to the data set, including comprehensiveness, types of data, and data structure (Tasks 2 and 3). Chapter 4 details issues about data storage and the data reduction conducted by the original research team (Task 4). Chapter 5 summarizes step-by-step guidelines to analyze video data for studying driver behavior in relation to nonrecurring congestion (Tasks 1 through 4). Chapter 6 introduces the statistical models that were developed to assess travel time reliability (Task 5). Chapter 7 discusses potential problems, risks, and limitations in the data sets. Finally, Chapter 8 summarizes the data reduction results for each candidate data set and proposes recommendations for future research efforts (Tasks 6 and 7).

Reference

1. Federal Highway Administration, Office of Operations. Reducing Non-Recurring Congestion. http://ops.fhwa.dot.gov/program_areas/reduce-non-cong.htm. Accessed May 12, 2011.

CHAPTER 2

Existing Studies Using In-Vehicle Video Cameras

Digital video cameras have rapidly evolved since the 1990s. Tapeless video cameras make it possible to use relatively small hardware equipment to record and save large-sized data. The quality of images has greatly improved and the editing process is simplified, allowing nonlinear editing systems to be widely deployed on desktop computers. Even though image sizes are small, it is easy to recognize the movements of targets and the ongoing background environment and acquire the information needed.

Because of their advanced capability, digital cameras have recently been used in multiple transportation safety research projects to capture drivers' behaviors in a naturalistic driving environment. Table 2.1 lists the studies discussed in this report. Because of difficulties associated with older data collection methods (e.g., videocassette tapes), emphasis was placed on research conducted during recent years. A short description of each project follows Table 2.1.

Project 1: Sleeper Berth

Conducted by the Virginia Tech Transportation Institute (VTTI), this study examined drivers' sleeping habits and drowsiness with respect to crash risk. Two tractor trailers were instrumented and loaned to trucking companies for their extended runs. Data from 41 drivers were used. In total, 47 males and 9 females were involved in the study. The average age was 43, with ages ranging from 28 to 63. On average, the drivers had 13 years of driving experience. Data collection runs lasted up to 240 h (6 to 10 days).

Continuous video data were recorded on four channels: driver's face, forward roadway, left rear, and right rear. Data were saved only for predefined critical incidents. If a specified kinematic trigger was activated, the data acquisition system (DAS) saved video and parametric data for 90 s before and 30 s after the event. Predefined events included:

- Steering wheel moved faster than 3.64 rad/s;
- Lateral acceleration was greater than 0.3 *g*;
- Longitudinal acceleration was greater than 0.25 *g*;
- Critical incident button was pressed;
- Vehicle crossed solid lane border;
- Time to collision (TTC) of 4 s or less;
- PERCLOS (percent eyelid closure) of 8% for 1 min;
- Driver subjectively assessed drowsiness as "extremely fatigued or difficult to stay awake" or did not respond;
- Lane departure followed by a steering event (disabled if turn signal was on); and
- A baseline data file, collected every 45 to 75 min.

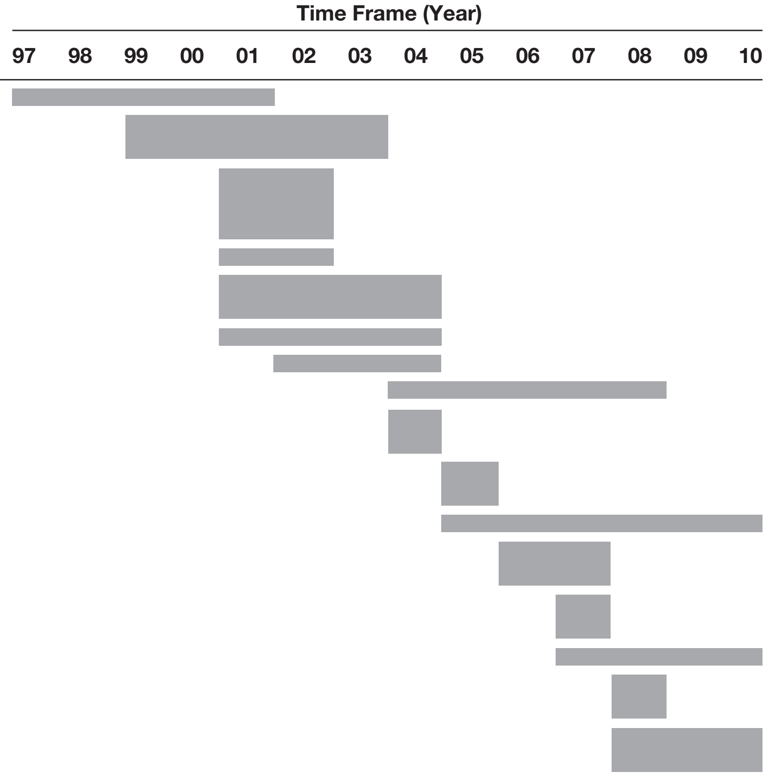
In addition to the video-recorded data, the data collected included vehicle network information (speed, accelerator, brake pedal, and steering wheel); environmental monitoring (temperature, illumination, vibration, and noise in decibels); *X*, *Y*, and *Z* acceleration; and lane orientation using a SafeTRAC lane tracking system, as well as some data generated after the data reduction, such as eyeglance behavior and road type and geometry (*1–3*).

Project 2: Automotive Collision Avoidance System Field Operational Test

The Automotive Collision Avoidance System (ACAS) FOT was led by General Motors (GM) under a cooperative agreement with the U.S. Department of Transportation (DOT). The FOT involved exposing a fleet of 11 ACAS-equipped Buick LeSabre passenger cars to 12 months of naturalistic driving by drivers from southeastern Michigan. The ACAS included a forward crash warning (FCW) and an adaptive cruise control (ACC) system. The FOT's goal was to determine the feasibility of the ACAS for widespread deployment from the perspective of driving safety and driver acceptance. Ninety-six drivers participated, resulting in 137,000 mi of driving. Results indicated that the ACC was widely accepted by drivers, but the acceptance of the FCW was mixed (due to false alarms) and not found to be significantly related to the FCW alert rate.

Table 2.1. List of Studies Using In-Vehicle Cameras

Study (Institute That Conducted the Research)
1. Sleeper Berth (VTTI[a])
2. Automotive Collision Avoidance System Field Operational Test (UMTRI[b])
3. Quality of Behavioral and Environmental Indicators Used to Infer the Intention to Change Lanes (Chemnitz University of Technology and INRETS)[c]
4. Lane Change Field Operational Test (VTTI)
5. Road Departure Crash Warning System Field Operational Test (UMTRI)
6. The 100-Car Study (VTTI)
7. Drowsy Driver Warning System Field Operational Test (VTTI)
8. Naturalistic Truck Driving Study (VTTI)
9. Naturalistic Driving Performance During Secondary Tasks (UMTRI)
10. Effect of In-Vehicle Video and Performance Feedback on Teen Driving Behavior (Iowa)
11. Naturalistic Teen Driving Study (VTTI)
12. Cooperative Intersection Collision Avoidance System for Violations (CICAS-V)—Infrastructure (VTTI)
13. Pilot Study to Test Multiple Medication Usage and Driving Functioning (NHTSA[d])
14. Older Driver Field Operational Test (ongoing study) (VTTI)
15. Cooperative Intersection Collision Avoidance System for Violations (CICAS-V)—Pilot Field Operational Test (VTTI)
16. Volvo Driving Behavior Field Operational Test (ongoing study) (Volvo and SAFER[e])



[a] Virginia Tech Transportation Institute
[b] University of Michigan Transportation Research Institute
[c] French National Institute for Transport and Safety Research (Institut National de Recherche sur les Transports et leur Sécurité)
[d] National Highway Traffic Safety Administration
[e] SAFER Vehicle and Traffic Safety Centre, Chalmers University of Technology, Sweden

Driver image data were collected for 8 s (4 s before and 4 s after the event) when the ACAS was activated. Alerts consisted of ACC engagements, collision warnings, hard braking, and hard steering. In addition to the video data, 250 separate data signals were collected, including data for trip time history, trip transition, trip summary (duration and counts), trigger summary, buffered Controller Area Network (CAN) packets, video time and image, and audio data.

The data set also has a valid time stamp and global positioning system (GPS) data that can be used to link to weather and other environmental data, such as traffic count and work zones (*4*).

Project 3: Quality of Behavioral and Environmental Indicators Used to Infer the Intention to Change Lanes

The study was conducted by the Chemnitz University of Technology (Chemnitz, Germany) and the Institut National de Recherche sur les Transports et leur Sécurité (INRETS) in Bron, France. It focused on the examination and comparison of selected behavioral and environmental indicators that predict the intention to change lanes. Data were collected on a multilane motorway between Bron and the Lyon International Airport (in both directions), with a length of 31 mi in an area of central France around the city of Lyon. The study included 22 participants aged 24 to 58; among them, 10 were female and 12 were male. Their driving experience ranged from 2 to 39 years and annual driving distance ranged between 1,243 and 31,075 mi. Participants were required to drive a Renault Scenic equipped to record and synchronize sensor data and videos on this route.

Video signals were recorded from five sources: (1) stereovision camera with radar for distance estimation to obstacles; (2) front view; (3) rear view; (4) view from the left outside mirror down to the road surface; and (5) view of the participant's head with indications of the eye tracker.

Data other than video data included speed, acceleration, deceleration, yaw rate, driver's eye movement, steering wheel position, pedal use, turn signal, and inclination. Environmental data that included the distance to the car ahead and GPS data were also gathered (*5*).

Project 4: Lane Change Field Operational Study

The main purpose of this VTTI study was to examine lane change behavior. The study monitored the commutes of 16 participants for approximately 1 month. Drivers commuted on Route 460 through the New River Valley or on Interstates 81 and 581 in the Roanoke Valley. Commutes were 25 or more miles in each direction to and from work.

Data would begin recording when a vehicle reached 35 mph and stopped recording when the vehicle slowed to 25 mph. In all, 24,000 vehicle miles of travel (VMT) data were collected. More than 8,000 lane changes were identified in the full data set and then classified by urgency and severity. Approximately 500 of the more urgent lane change events were analyzed in depth.

Video data were recorded on five channels: driver's face, forward roadway, rear view, and two side views. Data were saved using 8-mm videotapes. Besides the video data, the vehicle network information collected speed, accelerator, brake pedal, steering wheel, and turn signal data, as well as lateral acceleration, radar-collected information (one front and two rear sensors), and reduced data, such as eyeglance behavior and road type and geometry (*6*).

Project 5: Road Departure Crash Warning System Field Operational Test

The project was conducted under a cooperative agreement between U.S. DOT and the University of Michigan Transportation Research Institute (UMTRI) and its partners: Visteon Corporation and AssistWare Technologies. This FOT was designed to assess a Road Departure Crash Warning System (RDCWS). Two areas were assessed: safety-related changes in driver performance that may have been attributed to the system and levels of driver acceptance in key areas. Testing involved 11 passenger sedans equipped with the RDCWS and a DAS that collected performance, video, and audio data. Seventy-eight drivers participated for 4 weeks each, and the resulting data set captured 83,000 mi of driving. Analysis showed that drivers improved lane-keeping and reduced lane excursions while the RDCWS was active. Driver acceptance of the system was relatively positive, especially for the lateral drift component of the system.

Two video cameras were mounted on the vehicle's A-pillar: one forward-looking and one aimed at the driver. The inside camera also had a set of infrared light-emitting diodes (LEDs) that provided nighttime illumination of the driver's face. The images of the driver's face camera were captured in three modes: at 0.5 Hz when the data system was on, during an RDCWS alert captured for 8 s (4 s before and 4 s after the event), and a batch of 50 driver images spaced every 0.2 s for 5 min at the beginning of the trip and every 5 min thereafter.

In addition to the video data, roughly 400 separate data signals were collected, including data for vehicle and driver information, vehicle position, heading and motion, driver control inputs, RDCWS driver displays, RDCWS intermediate values, roadway environment, RDCWS and subsystem diagnostic information, RDCWS radar data, and audio data.

The data set has a valid time stamp that can be used to link to weather data and to link valid GPS data to other environmental data, such as traffic count and work zones (*7*).

Project 6: The 100-Car Study

The 100-Car Study was undertaken by VTTI, which collected large-scale naturalistic driving data from 100 vehicles in northern Virginia for approximately 18 months (12 to 13 months per vehicle). Drivers were given no special instructions, and the majority (78 out of 100) drove their own vehicles. The resulting data set has 6.4 terabytes (TB) of approximately 2 million VMT from 241 primary and secondary driver participants with a 12- to 13-month data collection period for each vehicle. The 8,295 incidents recorded included 15 police-reported crashes, 67 other crashes, and 761 near crashes. A variety of crash risk factors were analyzed.

Continuous video was collected on four channels at 30 Hz: driver's face, instrument panel (over driver's right shoulder), forward roadway, and rear roadway. The final data set contains 43,000 h of video. Vehicle network information (speed, brake pedal, throttle, and turn signal); GPS (latitude, longitude, and heading); and *X*, *Y*, and *Z* acceleration were also collected. Forward radar and rear radar were used to collect surrounding information. Data reduction generated other data, such as driver status, traffic flow, vehicle status, seat belt usage, and road type and geometry.

Because of a malfunction in the GPS subsystem, the time data are unreliable. Consequently, it is not possible to link some environmental data from external databases, such as work zone data and traffic condition data. The weather variable that has been coded in the reduced data set is available (*8–9*).

Project 7: Drowsy Driver Warning System Field Operational Test

The Drowsy Driver Warning System (DDWS) study was conducted by VTTI. The main purpose was to examine the effectiveness of a mechanism that alerted drivers that they were about to fall asleep (monitored using a PERCLOS meter). VTTI instrumented 34 trucks with an experimental warning system, video cameras, and a DAS. The final data set included 2.3 million VMT, 12 TB of data, and 46,000 h of driving.

Continuous video collected data on four channels at 30 Hz: driver's face, forward roadway, left rear, and right rear. Additional data included vehicle network information (speed, brake pedal, throttle, and turn signal), GPS (latitude, longitude, and heading), lateral and longitudinal acceleration, forward radar-collected data, and sleep quantity (measured by an activity wristwatch).

After the data reduction, 16 crashes and 136 near crashes were identified. Data were reduced to identify events based on such information as:

- Lateral acceleration;
- Longitudinal acceleration;
- Lane deviation;
- Normalized lane position; and
- Forward TTC.

The following events were identified: baseline driving epoch, crash, crash: tire strike (defined as any physical contact of tires with other objects on the road), near crash (evasive maneuver), crash-relevant conflict (evasive maneuver), crash-relevant conflict (no evasive maneuver), or nonconflict. Other variables, such as seat belt usage, date, time, light, weather, work zone, roadway condition, and traffic density, were also coded. The status of the vehicle and driver before events was coded as well (*10*).

Project 8: Naturalistic Truck Driving Study

Conducted by VTTI, the Naturalistic Truck Driving Study (NTDS) attempted to examine the crash risk factors of commercial truck drivers. VTTI instrumented eight tractor trailers to monitor truck-driving behavior. The data set includes 735,000 VMT data, which amounts to 6.2 TB and 14,500 h of driving. Almost 3,000 critical events, such as crashes, illegal maneuvers, and unintentional lane deviations, were analyzed.

Continuous video data were collected on four channels: driver's face, forward roadway, left rear, and right rear. Additionally, the final data set has vehicle network information (speed, brake pedal, throttle, and turn signal), GPS (latitude, longitude, and heading), lateral and longitudinal acceleration, forward and rear radar-collected data, and sleep quantity (measured by an activity wristwatch). Data were reduced based on the following triggers:

- Longitudinal acceleration (LA);
- Swerve;
- TTC;
- Lane deviation;
- Critical incident button; and
- Analyst identified.

Events identified from the data reduction included crash, crash: tire strike, near crash, crash-relevant conflict, unintentional lane deviation, and illegal maneuver. After the data reduction, five crashes, 61 near crashes, 1,586 crash-relevant conflicts, 1,215 unintentional lane deviations, and 5,069 baselines were identified (*11*).

Project 9: Naturalistic Driving Performance During Secondary Tasks

The purpose of the study, which was conducted by UMTRI, was to determine the frequency and conditions under which drivers engage in secondary behaviors and to explore the relationship these behaviors might have to driving performance. Data from 36 drivers involved in a naturalistic driving study were divided into three age-groups and analyzed to determine the frequency and conditions under which drivers engage in secondary behaviors, such as eating, drinking, and using a cellular phone. Mean ages for drivers in this study were 25.1, 45.6, and 64.2 for the younger, middle, and older age-groups, respectively. The data collected were also analyzed to explore the relationship these behaviors might have to driving performance.

A video camera was mounted to the inside of the vehicle's A-pillar and captured 5-s images of the driver's face at 10 frames/s at 5-min intervals. Researchers examined a representative sample of 18,281 video clips from the FOT. The sample was not associated with any RDCWS alerts, represented driving at least 25 mph, and included drivers with at least 10 qualifying video clips. Researchers coded 1,440 5-s video clips of the drivers' faces for the occurrence of specific secondary behaviors and the duration of glances away from the forward scene.

Other performance data from instrumented vehicles were used to calculate the variability of the steering angle, the mean and the variability of lane position, the mean and the variability of throttle position, and the variability of speed (*12*).

Project 10: Effect of In-Vehicle Video and Performance Feedback on Teen Driving Behavior

The study was conducted with 26 participants from a high school in rural Iowa. Study periods consisted of a 9-week baseline period followed by 40 weeks of video collection and feedback and 9 weeks of video collection without immediate feedback. The study found that teen drivers showed a statistically significant decrease in triggering behaviors between the feedback and nonfeedback conditions, possibly indicating that drivers became aware of their unsafe driving behaviors and learned to improve their driving.

The study used a DriveCam camera mounted to the windshield underneath the rearview mirror. The DriveCam is a palm-sized device that integrates two cameras (in-cab and forward view) and a wireless transmitter. Video data are continuously buffered 24 h per day but only write to memory when a threshold in latitudinal or longitudinal force is exceeded. Twenty seconds of data (10 before and 10 after each "event trigger") were recorded. Event triggers included any event that exceeded *g*-forces of .055 for lateral movement or 0.50 for longitudinal movement. If an event occurred, the drivers were given immediate feedback.

In this data set, weather was coded as clear or cloudy; fog; rain; mist; snow, sleet, or hail; or smoke or dust. Because no GPS data were collected, location-related environmental data cannot be linked to the data. However, the data were reduced such that extensive environmental data (e.g., traffic condition, work zones, and driver behavior data) are coded in the reduced database by reductionists (*13–14*).

Project 11: Naturalistic Teen Driving Study

The Naturalistic Teen Driving Study (NTNDS) was conducted by VTTI. The primary purpose of the study was to evaluate and quantify crash risk among teen drivers. VTTI installed DASs in 42 cars primarily driven by newly licensed teenage drivers in the New River Valley area of Virginia. Naturalistic driving data of the teens and a parent of each teen were collected during the course of 18 months. The resulting data set has 500,000 VMT, amounting to 5.1 TB of data.

Continuous video was collected on four channels: driver's face, instrument panel (over driver's right shoulder), forward roadway, and rear roadway. Two additional cameras would periodically activate for a few seconds at a time. These cameras provided views of the vehicle's entire cabin (blurred to protect passenger identities) and the lap area of the back seat.

Other data, such as GPS (latitude, longitude, and heading); *X, Y,* and *Z* acceleration; forward radar-collected data; and video-based lane tracking data, as well as reduced data (e.g., eyeglance behavior, time-of-day and ambient lighting, road type and geometry, and traffic density), were available in the resulting database (*15*).

Project 12: Cooperative Intersection Collision Avoidance System for Violations Infrastructure

During the Cooperative Intersection Collision Avoidance System for Violations (CICAS-V), a VTTI and Collision Avoidance Metrics Partnership (CAMP) collaborative project, the first study was an infrastructure-based effort monitoring signalized and stop-controlled intersections. The study was undertaken to model stopping behavior and the kinematic factors that could lead to intersection violations.

Continuous video was collected. Stop-controlled intersections generally had one camera focused on one particular approach. Signalized intersections had four channels of video, one for each approach. In total, 1.5 TB of video and radar data were collected. Other data collected included lateral speed, lateral and longitudinal acceleration, lateral and longitudinal range, and lane tracking for approaching objects (*16–17*).

Project 13: Pilot Study to Test Multiple Medication Usage and Driving Functioning

The study was performed by TransAnalytics for the National Highway Traffic Safety Administration (NHTSA). Its purpose was to explore the relationship between polypharmacy and driving functioning through separate but related research activities. Driver performance evaluations, brake response time, and functioning screening measures were conducted for the study sample; drug profiles were documented through a "brown bag" review by a licensed pharmacist. Field evaluation occurred on routes in residential communities in the Hockessin, Delaware, and Parkville, Maryland, vicinities.

Two miniature video cameras were used: one for the driver's face view and one for the forward road view. Cameras were used in the field study of driver performance of 44 older adults. Additionally, cameras were used in private cars of a subsample of five individuals. The video data included the Advanced System Format (ASF) with 704 × 496 resolutions and a 12-Hz frame rate. Each trip was recorded in 10- to 100-s snippets (depending on the amount of motion in the video), which were later combined and rendered in postprocessing to produce single clips for subsequent video coding analysis. Recorders were set to start recording automatically when powered on and to stop recording when no motion was detected in the driver face view camera for at least 30 s.

Other data, such as driving speed, brake response time, GPS, onboard diagnostics (including vehicle speed, throttle position, and engine speed), and date and time, were also recorded. The lane-changing behavior of the drivers was manually recorded by researchers in this study (*18*).

Project 14: Older Driver Field Operational Test

The purpose of the FOT, which is being conducted by VTTI, is to study older drivers' driving behavior. The data collection process is still ongoing. The estimated resulting data set should have 4,867 h of video data and 2.5 TB of video and binary data collected from 131,400 vehicle miles of travel.

Four cameras are used to collect driving data: forward, rear, driver's face, and instrument panel over the driver's shoulder. Other data collected include latitude and longitude acceleration, forward radar-collected data, lanetracker data that tracks lane changing, GPS location, and acceleration.

Project 15: Cooperative Intersection Collision Avoidance System for Violations Pilot Field Operational Test

The Cooperative Intersection Collision Avoidance System for Violations (CICAS-V) study was conducted by VTTI. This was the second study performed during the CICAS-V project. It was a pseudonaturalistic field test of a collision warning system for both effectiveness and driver acceptance.

Continuous video was collected on four channels: driver's face, instrument panel (over the driver's right shoulder), forward roadway, and rear roadway. The study collected 214 gigabytes (GB) of data, which amounted to 194 h of data. Other information, including vehicle network information (speed, brake pedal, throttle, and turn signal); GPS (latitude, longitude, and heading); *X, Y,* and *Z* acceleration; forward and rear radar-collected data; and reduced data, such as eyeglance behavior and map-matching variables, were also available. For applicable intersections, only the distance to stop bar, time to intersection crossing, lane number, and signal phase were also gathered (*19*).

Project 16: Volvo Driving Behavior Field Operational Test

The Swedish manufacturer Volvo is conducting an ongoing study to compile a variety of data about driving behavior. The research project is part of the European Union (EU) project called EuroFOT, in which Volvo Cars and the SAFER Vehicle and Traffic Safety Centre at Chalmers University of Technology are engaged. The overall goal is to develop a safer, cleaner, and more efficient road transportation system in Europe. The study started in May 2008 and is expected to last 3 years. Approximately 100 Volvo V70 and XC70 cars will be involved in the data collection.

Cameras to record the driver's head and eye movements, as well as the view of the road and behind the car are installed in the car to collect video data. A data logger will also be used to record the information from the safety features in the car. Systems to be tested include Collision Warning with Auto Brake (CWAB), ACC, Lane Departure Warning System (LDWS), Driver Alert Control (DAC), and Blind Spot Information System (BLIS).

Concluding Remarks

As demonstrated in this chapter, there have been significant efforts to gather naturalistic driver behavior using video and other sensor systems. These data sources will be analyzed in more detail in Chapter 3.

References

1. Dingus, T., V. Neale, S. Garness, R. Hanowski, A. Keisler, S. Lee, M. Perez, G. Robinson, S. Belz, J. Casali, E. Pace Schott, R. Stickgold, and J. A. Hobson. *Impact of Sleeper Berth Usage on Driver Fatigue.* Report RT-02-050. FMCSA, 2001.
2. *Impact of Sleeper Berth Usage on Driver Fatigue: Final Report.* Report FMCSA-MCRT-02-070. FMCSA, 2002.
3. Neale, V. L., G. S. Robinson, S. M. Belz, E. V. Christian, J. G. Casali, and T. A. Dingus. *Impact of Sleeper Berth Usage on Driver Fatigue, Task 1: Analysis of Trucker Sleep Quality.* Report DOT MC 00 204. Office of Motor Carrier Safety, FHWA, 1998.
4. University of Michigan Transportation Research Institute. *Automotive Collision Avoidance System Field Operational Test Report: Methodology and Results.* Report DOT HS 809 900. NHTSA, 2005.
5. Henning, M., O. Georgeon, and J. Krems. The Quality of Behavioral and Environmental Indicators Used to Infer the Intention to Change Lanes. *Proc., 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design,* Stevenson, Wash., 2007, pp. 231–237.
6. Lee, S. E., E. C. B. Olsen, and W. W. Wierwille. *A Comprehensive Examination of Naturalistic Lane Changes.* Report DOT HS 809 702. NHTSA, 2004.
7. University of Michigan Transportation Research Institute. *Road Departure Crash Warning System Field Operational Test: Methodology and Results.* NHTSA, 2006.
8. Dingus, T. A., S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, and R. R. Knipling. *The 100-Car Naturalistic Driving Study, Phase II: Results of the 100-Car Field Experiment.* Report DOT HS 810 593. NHTSA, 2006.
9. Neale, V. L., S. G. Klauer, R. R. Knipling, T. A. Dingus, G. T. Holbrook, and A. Petersen. *The 100-Car Naturalistic Driving Study, Phase 1: Experimental Design.* Report DOT HS 809 536. NHTSA, 2002.
10. Hanowski, R. J., M. Blanco, A. Nakata, J. S. Hickman, W. A. Schaudt, M. C. Fumero, R. L. Olson, J. Jermeland, M. Greening, G. T. Holbrook, R. R. Knipling, and P. Madison. *The Drowsy Driver Warning System Field Operational Test, Data Collection: Final Report.* Report DOT HS 811 035. NHTSA and Virginia Tech Transportation Institute, Blacksburg, Va., 2008.
11. Blanco, M., J. S. Hickman, R. L. Olson, J. L. Bocanegra, R. J. Hanowski, A. Nakata, M. Greening, P. Madison, G. T. Holbrook, and D. Bowman. *Investigating Critical Incidents, Driver Restart Period, Sleep Quantity, and Crash Countermeasures in Commercial Vehicle Operations Using Naturalistic Data Collection.* FMCSA, 2008.
12. Sayer, J., J. Devonshire, and C. Flanagan. Naturalistic Driving Performance During Secondary Tasks. Presented at 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Stevenson, Wash., 2007.
13. McGehee, D., M. Raby, C. Carney, J. D. Lee, and M. L. Reyes. Extending Parental Mentoring Using an Event-Triggered Video Intervention in Rural Teen Drivers. *Journal of Safety Research,* Vol. 38, 2007, pp. 215–222.

14. McGehee, D. V., C. Carney, M. Raby, J. D. Lee, and M. L. Reyes. The Impact of an Event-Triggered Video Intervention on Rural Teenage Driving. Presented at 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Stevenson, Wash., 2007.
15. Lerner, N., J. Jenness, J. Singer, S. G. Klauer, S. Lee, M. Donath, M. Manser, and M. Ward. *An Exploration of Vehicle-Based Monitoring of Novice Teen Drivers: Draft Report.* Virginia Tech Transportation Institute, Blacksburg, Va., 2008.
16. Doerzaph, Z. R., V. L. Neale, J. R. Bowman, and K. I. Wiegand. *Live Stop-Controlled Intersection Data Collection.* Virginia Transportation Research Council, Charlottesville, Va., 2007.
17. Doerzaph, Z. R., V. L. Neale, J. R. Bowman, D. C. Viita, and M. A. Maile. *Cooperative Intersection Collision Avoidance System Limited to Stop Sign and Traffic Signal Violations: Subtask 3.2 Interim Report.* NHTSA, 2008.
18. Staplin, L., K. H. Lococo, K. W. Gish, and C. Martell. *A Pilot Study to Test Multiple Medication Usage and Driving Functioning.* Report DOT HS 810 980, NHTSA, 2008.
19. Neale, V. L., Z. R. Doerzaph, D. C. Viita, J. R. Bowman, T. Terry, R. Bhagavathula, and M. A. Maile. *Cooperative Intersection Collision Avoidance System Limited to Stop Sign and Traffic Signal Violations: Subtask 3.4 Interim Report.* NHTSA, 2008.

CHAPTER 3

Dimensions of Data Feasibility

To determine the feasibility of using data from projects based on in-vehicle video data, the accuracy of two components of the data is critical. The components are (1) the video data and (2) parametric data, such as GPS data, radar-detected data, and complementary data. It is preferable to have continuously recorded and high-frequency data that are capable of showing the details of drivers' maneuvers, environment, and vehicle status. The existence of complementary data and the availability of, for example, flow data, crash data, and weather data need to be investigated. An accurate time and location stamp that enables a proper link between vehicle data and complementary data is necessary to the vehicle and complementary data sources. Figure 3.1 shows the conceptual relationship between the data sets to be analyzed and the potential travel time reliability improvement measures.

Some of the studies listed in Table 2.1 are readily suitable to serve the research purpose of this study. The quality of vehicle data and complementary data of each data set are discussed individually in this section.

Quality of Vehicle Data

The following data sets were eliminated from the candidate list after examining the data size, quality, and availability.

Project 1: Sleeper Berth

The Sleeper Berth project instrumented two tractor trailers with cameras and DASs. Data from only 41 drivers were used in this study. Because of data acquisition limitations at the time the data were gathered, this data set is event-triggered. Consequently, it is not considered in the current study.

Project 3: Quality of Behavioral and Environmental Indicators Used to Infer the Intention to Change Lanes

This data set contains only 22 participants, and the data were collected in France. Because of the small sample size, potential differences in driver behavior relative to that in North America, and the challenge of using data from another country, this data set is considered unsuitable.

Project 4: Lane Change FOT

Because this is an older data set, only the most urgent lane change events have been converted to digital video. Significant effort and time would be required to digitize the data. Consequently, this data set is deemed unsuitable for the current study.

Project 9: Naturalistic Driving Performance During Secondary Tasks

This study used a small subset of data collected during the RDCWS FOT conducted by UMTRI. Because the RDCWS FOT is included in this research effort and is discussed in detail later, this subset of the data set is not considered further.

Project 10: Effect of In-Vehicle Video and Performance Feedback on Teen Driving Behavior

Instead of continuously recording data, the video cameras in this study were trigger-activated. Therefore, this data set is eliminated from further consideration.

Project 12: CICAS-V Infrastructure

The CICAS-V infrastructure study collected almost 1.5 TB of video and radar data for approaching vehicles at several instrumented signalized intersections. Driver image data were not collected in this study. The videos were installed at the intersections to capture vehicle movements and thus are not in-vehicle video data. Consequently, this data set is not considered further.

Project 13: Pilot Study to Test Multiple Medication Usage and Driving Functioning

According to the consent forms for the study, NHTSA is not allowed to share the video data with other parties.

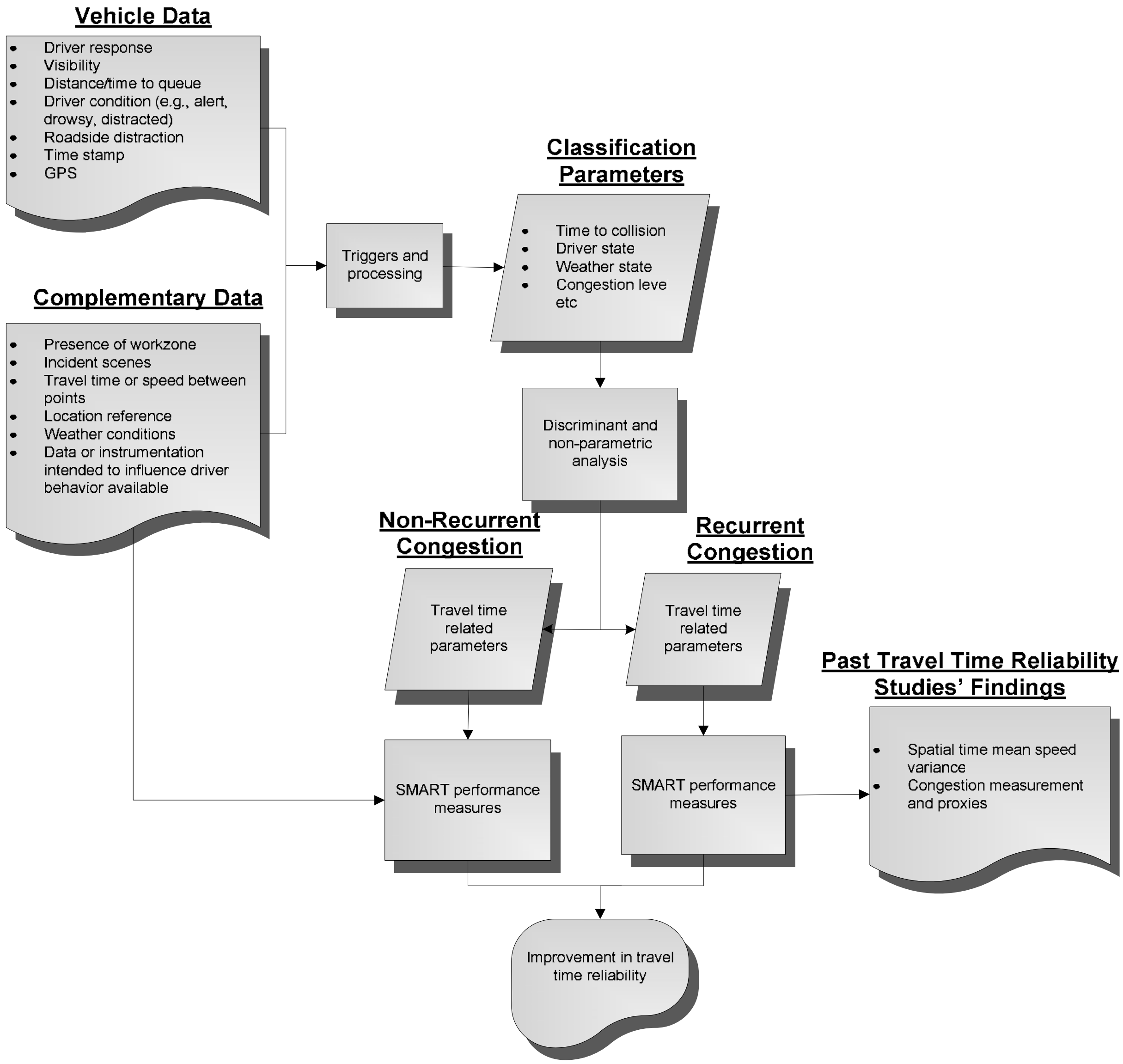


Figure 3.1. Relationship between project data sets and travel time reliability improvement.

Consequently, the team would not have access to the data, and thus it is not considered further.

Project 14: Older Driver FOT

Data collection for this study is under way. It is anticipated that the data should be available in early 2010. Although the resulting data set has great potential to be used for future studies, it cannot be used in this study, given that the data are still being gathered.

Project 15: CICAS-V Pilot FOT

The CICAS-V pilot study has a relatively small data set. Only 87 drivers were recruited for this study, which involved driving an experimental vehicle on a predetermined short route (approximately 40 mi). The driving data were not collected using a completely naturalistic method, and the length and roadways of the route are limited; therefore, this data set is excluded from the analysis.

Project 16: Volvo Driving Behavior FOT

This study commenced in May 2008 and is expected to last for 3 years. It will still be ongoing by the time this research effort ends. Therefore, this study is excluded.

Data Sets After Initial Screening

After the initial filtering, the resulting valid data sets are listed in Table 3.1. As can be seen from the table, data for Project 2 were collected in southeastern Michigan. For Project 5, data

Table 3.1. Locations of Candidate Studies

Candidate Data Set Dimensions	MI	DC	DE	MD	NJ	NY	PA	VA	WV
Project 2. ACAS FOT	✔								
Project 5. RDCWS FOT	✔								
Project 6. 100-Car Study		✔		✔				✔	
Project 7. DDWS FOT		✔	✔	✔	✔	✔	✔	✔	✔
Project 8. NTDS		✔	✔	✔	✔	✔	✔	✔	✔
Project 11. NTNDS								✔	

were collected in southeastern Michigan, including Detroit and surrounding suburbs and rural areas, resulting in 2,500 h of video data. For Project 6, data were collected in the Washington, D.C., northern Virginia, and Maryland areas, resulting in 43,000 h of video data. For Project 7, drivers were recruited in Roanoke, South Boston, Stuarts Draft, and Cloverdale, Virginia, as well as in Charlotte, North Carolina. Data were collected for long-haul (cross-country) trips, as well as for overnight express (out-and-back) operations throughout the Mid-Atlantic area, resulting in 46,000 h of video data. For Project 8, drivers were recruited at trucking centers in Charlotte, Kernersville, and Henderson, North Carolina, and in Roanoke, Gordonsville, and Mount Crawford, Virginia. Data were collected for trucking runs throughout the Mid-Atlantic region, resulting in 14,500 h of video data. For Project 11, data were collected in the New River and Roanoke Valleys in Virginia, resulting in 16,644 h of driving data, including 10,754 h of teen driving data.

Quality of External Data

Using in-vehicle video data to help assess the role of driver behavior in nonrecurring congestion requires analyzing not only the vehicle and driver data but also the complementary data. For instance, it has been documented that weather affects driving behavior and performance and leads to longer following distances, thereby decreasing throughput at intersections and resulting in longer travel time. Another factor that has been shown to affect crash risk is the occurrence of a prior incident. Driver behavior in the vicinity of traffic control devices also contributes to nonrecurring congestion. Finally, previous studies in the Los Angeles conurbation have shown that more vehicle-hours of delay result from extraordinary and accidentally occurring traffic disturbances (nonrecurring) than from regularly occurring network overloading during typical daily peak hours (recurring).

Although some of the complementary data can be obtained from data reduction (e.g., the traffic condition is a variable recorded by data reductionists while they were viewing video data to describe the surrounding traffic), the availability of weather, traffic condition, crash, and work zone data in related states is investigated to provide a consistent data set across the studies and to avoid potential subjective bias brought by data reductions.

Weather data can be reliably obtained by acquiring data from a nearby weather station. Figure 3.2 shows the locations of weather stations in the related states. As can be seen from the map, weather stations are densely located and thus there is a high possibility of linking a vehicle location to a nearby weather station through GPS data. Out of 592 weather stations, 253 are either Automated Surface Observing System (ASOS) stations or Automated Weather Observing System (AWOS) stations.

Weather stations using ASOS are located at airports. ASOS is supported by the Federal Aviation Administration (FAA), the National Weather Service (NWS), and the Department of Defense (DOD). The system provides weather observations that include temperature, dew point, wind, altimeter setting, visibility, sky condition, and precipitation. Five hundred sixty-nine FAA-sponsored and 313 NWS-sponsored ASOS stations are installed at airports throughout the United States. The weather reports by ASOS that could be used in this study are of METAR type (Aviation Routine Weather Reports) and contain precipitation type, precipitation intensities (in./h), and visibility readings (m). ASOS visibility measurements are performed at 30-s increments using a forward scatter sensor to compute 1-min average extinction coefficients (sum of the absorption and scattering coefficients). For this purpose, a photocell that identifies the time of day (day or night) is used to select the appropriate equation for use in the procedure. ASOS computes a 1-min average visibility level that is used to compute a 10-min moving average (MA) visibility level. This value is then rounded down to the nearest reportable visibility level. The system uses precipitation identification sensors to determine the type of precipitation (rain, snow, or freezing rain). Precipitation intensity is recorded as liquid-equivalent precipitation accumulation measurements using a Heated Tipping Bucket (HTB) gauge. The HTB has a resolution of 0.01 in. and an accuracy of ±0.02 in., or 4% of the hourly total, whichever is greater.

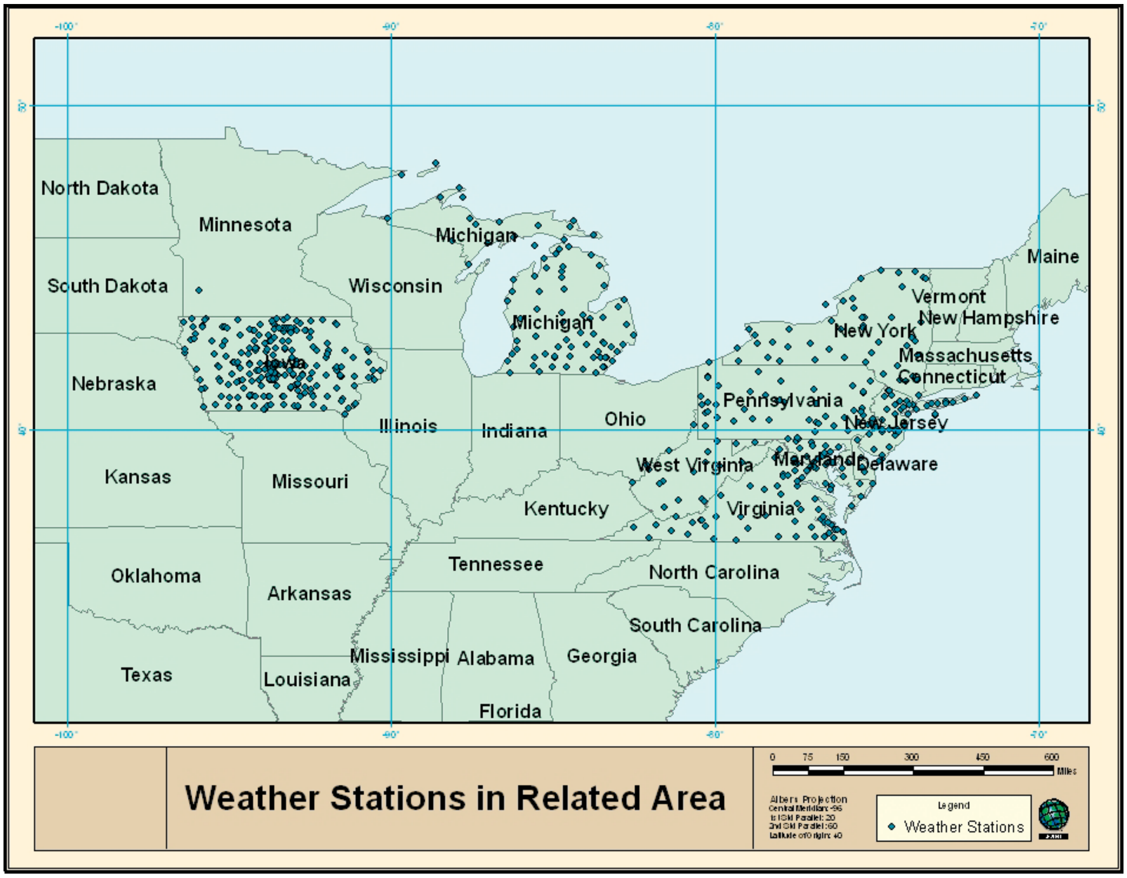


Figure 3.2. Weather station locations.

AWOS is one of the oldest automated weather stations and predates ASOS. It is a modular system utilizing a central processor that receives input from multiple sensors. Operated and controlled by the FAA, state and local governments, and some private agencies, the AWOS reports weather information at 20-min intervals but does not report special observations for rapidly changing weather conditions. Depending on the different varieties, AWOS observes different indices. The most common type, AWOS-III, observes temperature and the dew point in degrees Celsius, wind speed and direction in knots, visibility, cloud coverage and ceiling up to 12,000 ft, and altimeter setting. Additional sensors, such as for freezing rain and thunderstorms, have recently become available.

Traffic count data are available in all the states that were studied. The state of Virginia has extensive locations of traffic count stations. There are more than 72,000 stations, 470 of which collect continuous data. A subset from the Virginia Count Station list—the traffic count locations in the city of Richmond—is shown in Figure 3.3. The traffic count stations in West Virginia are shown in Figure 3.4. There are approximately 60 permanent count stations in West Virginia, and one-third of the state is counted each year in a Coverage Count Program. In Pennsylvania, there are 116 continuous stations out of 30,661 count stations. Figure 3.5 shows a subset of the traffic count stations in Pittsburgh. Figure 3.6 shows traffic count stations in Delaware.

A sample of the longitudinal and latitudinal information of one traffic count station on link ID 507002 is shown in Table 3.2. Table 3.3 demonstrates a sample of the raw traffic counts collected by that station at 15-min intervals by vehicle class in the state of Virginia. Traffic conditions (e.g., traffic density and level of service) can be inferred from the counts. The longitudinal and latitudinal fields can then be used to link in-vehicle GPS data and traffic count data. Some states have location information for the stations listed as mileposts and street names that can be digitized when necessary.

Crash data are readily available for every state, although some have stringent data privilege requirements. Some states, such as New Jersey and Michigan, have online crash databases from which the information can be downloaded. The District of Columbia DOT coded their crashes with work zone information if there was a work zone in the surrounding area when the crash happened. Table 3.4 provides a crash sample from Washington, D.C. Because of space limitations, only

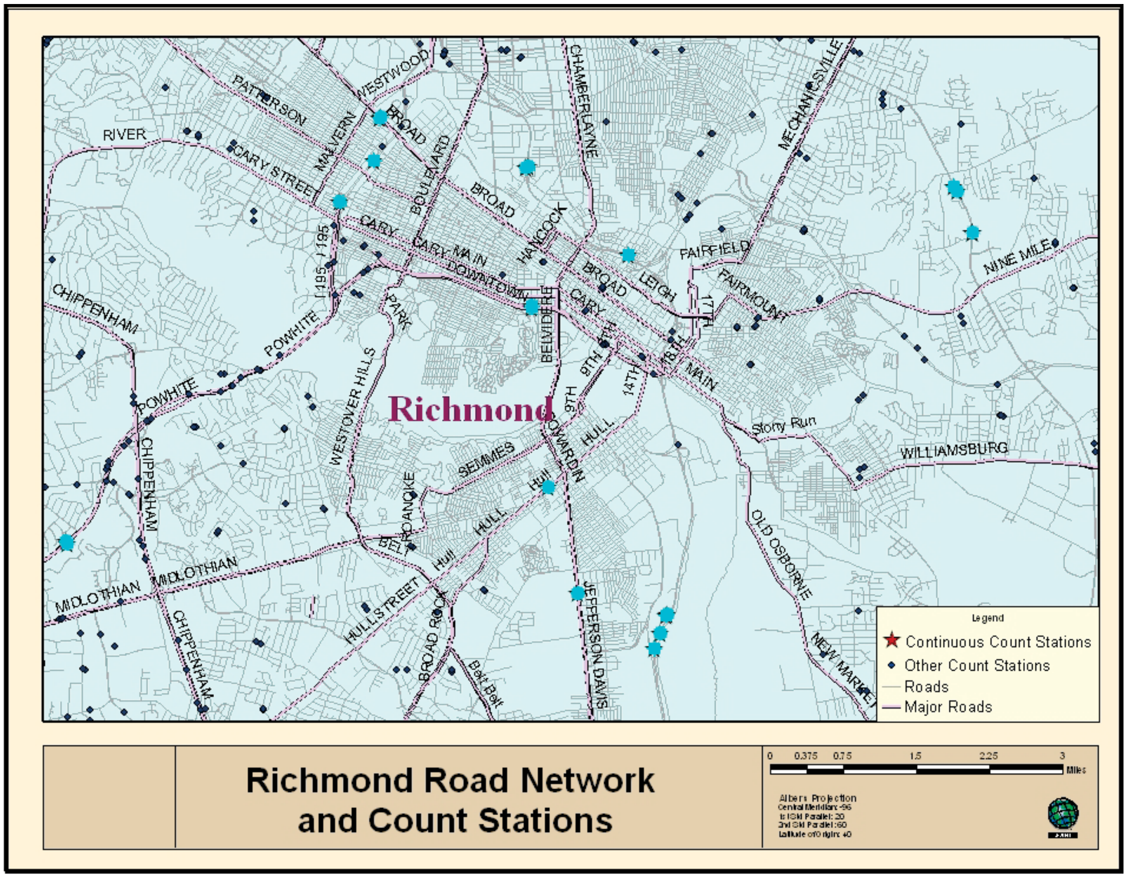


Figure 3.3. Traffic count stations in Richmond, Va.

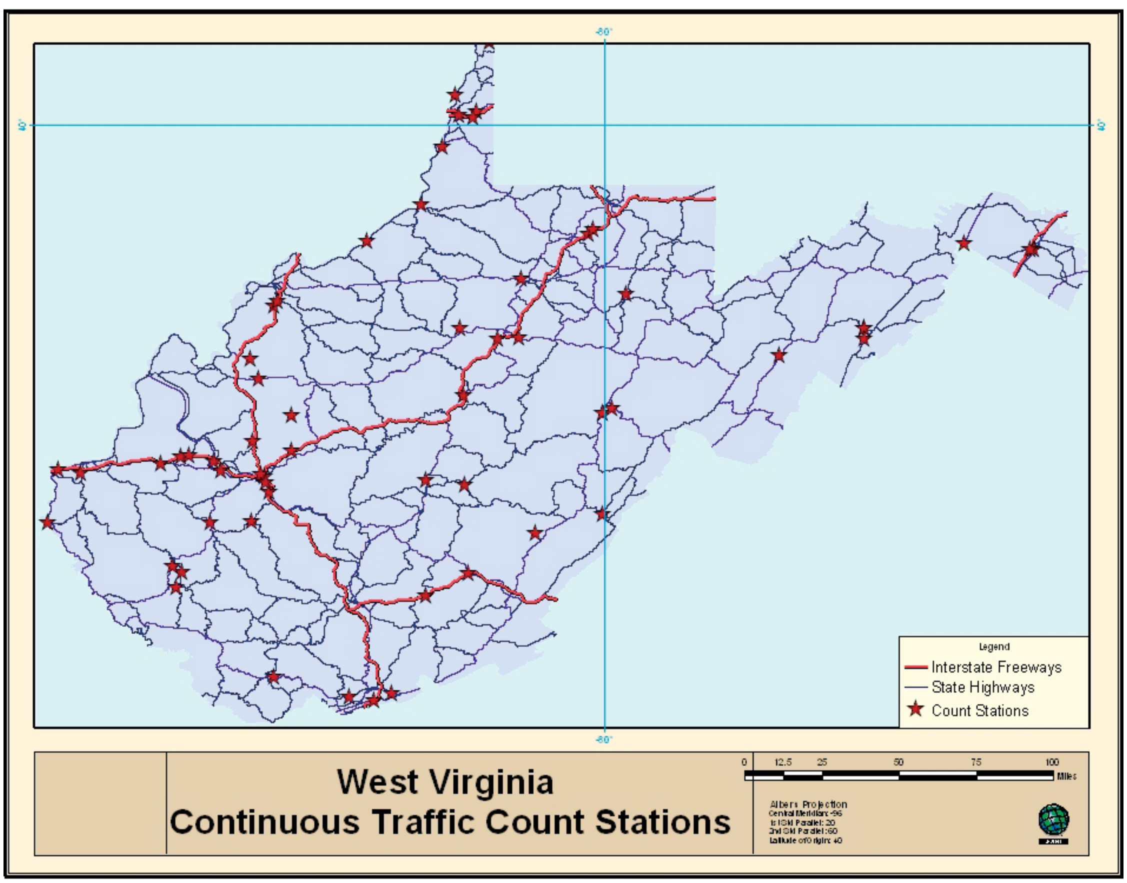


Figure 3.4. Traffic count stations in West Virginia.

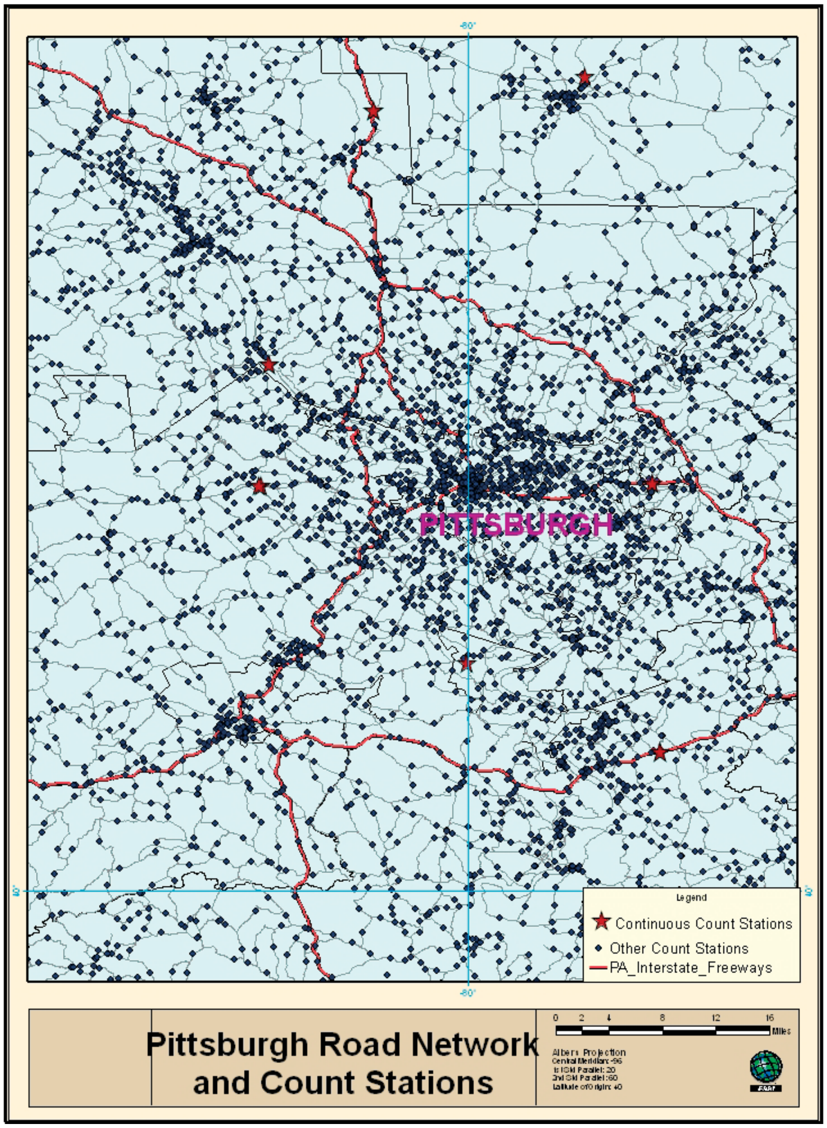


Figure 3.5. Traffic count stations in Pittsburgh, Pa.

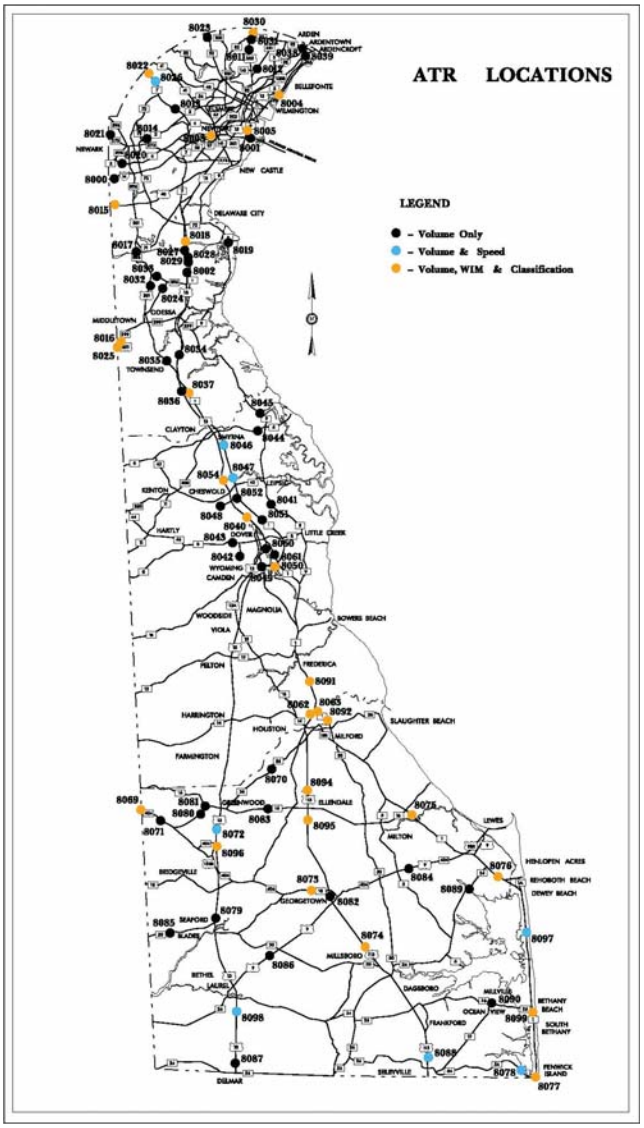
some of the variables in the original database are listed here. Other information, such as road condition, light condition, weather, and the sobriety of involved drivers, is listed in the original database.

Crash and traffic volume data are typically saved in a database, but work zone data usually are not. This is especially true for completed road work. Most states have ongoing projects recorded in the database, but completed projects are not stored. A few states have incomplete data or data that are not in sufficient condition for use. For example, as of August 2008 in the state of Virginia, the 511 system has used a program called VA Traffic to log information. Before August 2008, there was limited free-text-based information recording the scheduled start and end of road work. For the state of Pennsylvania, the work zone data are only available for projects that occur on state highways. Table 3.5 provides a sample of the work zone data from West Virginia.

Table 3.6 summarizes the availability of complementary data in related states. Online sources from which data can be downloaded are listed in the footnotes.

Evaluation of Candidate Data Sets

To help determine the feasibility of candidate databases, a multidimensional criterion is established for the data sources. These dimensions include comprehensiveness, video data quality, vehicle data, linkages, and data format and structure. Table 3.7 provides a detailed explanation and definition for each feasibility dimension.



Color version of this figure: www.trb.org/Main/Blurbs/165281.aspx.

Figure 3.6. Traffic count stations in Delaware.

Each candidate database is evaluated on each dimension to demonstrate the suitability for further analysis. At the same time, the legal restriction for each data set is examined. Certain data sets have IRB restrictions, meaning that the data collected in that study are restricted for external usage and need to be eliminated from the candidate pool. Some studies collected video data at a lower frequency and, therefore, are not suitable for this study.

To quantitatively evaluate each qualified candidate database, a composite feasibility score is computed to reflect the database's strengths and weaknesses, as displayed in Table 3.8. Each dimension receives a seven-point scale score, ranging from 1, representing a low score (e.g., small sample and data available only in raw form), to 7, representing a high score (e.g., large, representative sample and data in reduced form). In computing the score, each feasibility category is assigned a weight so that the sum of weights across all categories totals 100 (e.g., the weight for the feasibility category comprehensiveness is 15). Within each feasibility category the various measures that constitute a category are assigned a score so that the sum of scores within a category is 10. For example, the comprehensiveness category includes four measures: (1) driver population, (2) types of roadways, (3) types of trips, and (4) types of vehicles. These measures are assigned a weight of 4, 4, 1, and 1, respectively. The weights are used to compute a weighted average score for each feasibility category. The feasibility category scores are then used to compute a weighted average overall score between 1 and 7. The quality of video data is vital to this project. Some dimensions, such as whether the driver's face and hand movements can be clearly seen or whether the surrounding vehicles can be accurately located by the radar sensors, receive greater weights to emphasize the importance of those data to this study.

To further illustrate the scoring methodology, the procedure is demonstrated using the 100-Car data set. The score for the comprehensiveness category is computed as the weighted average of the four measures that constitute this category as

$$\frac{6\times4+7\times4+6\times1+7\times1}{10}=6.50 \tag{1}$$

This computation is repeated for each of the remaining feasibility categories (video data quality, vehicle data, linkages, and data format and structure). The overall score is then computed as the weighted average of the various category scores as

$$\frac{6.50\times15+6.20\times40+6.11\times20+2.67\times20+7.00\times5}{100}=5.6 \tag{2}$$

Table 3.2. Location of Traffic Station

Link ID	Counter	Sensor	Physical Location	Latitude	Longitude
507002	1	1	0.205 mi from Country Club	37.21934	–80.4056

Table 3.3. Sample of Traffic Count Data in Virginia

Link ID	Direction	Lane	Start Date and Time	Interval	Class Quality	Class 15
507002	1	1	3/15/2007 7:00	15	1	4
507002	1	1	3/15/2007 7:15	15	1	6
507002	1	1	3/15/2007 7:30	15	1	20
507002	1	1	3/15/2007 7:45	15	1	26
507002	1	1	3/15/2007 8:00	15	1	20
507002	1	1	3/15/2007 8:15	15	1	26
507002	1	1	3/15/2007 8:30	15	1	32
507002	1	1	3/15/2007 8:45	15	1	20
507002	1	1	3/15/2007 9:00	15	1	22
507002	1	1	3/15/2007 9:15	15	1	8
507002	1	1	3/15/2007 9:30	15	1	10
507002	1	1	3/15/2007 9:45	15	1	10
507002	1	1	3/15/2007 10:00	15	1	6
507002	1	1	3/15/2007 10:15	15	1	4
507002	1	1	3/15/2007 10:30	15	1	7
507002	1	1	3/15/2007 10:45	15	1	11
507002	1	1	3/15/2007 11:00	15	1	8
507002	1	1	3/15/2007 11:15	15	1	6
507002	1	1	3/15/2007 11:30	15	1	10
507002	1	1	3/15/2007 11:45	15	1	18
507002	1	1	3/15/2007 12:00	15	1	4
507002	1	1	3/15/2007 12:15	15	1	16
507002	1	1	3/15/2007 12:30	15	1	18
507002	1	1	3/15/2007 12:45	15	1	12
507002	1	1	3/15/2007 13:00	15	1	18
507002	1	1	3/15/2007 13:15	15	1	9

Table 3.4. Crash Sample Data from Washington, D.C.

Date	Time	Report	Type	Street	Block	No. of Vehicles	No. of Injuries	No. of Passengers in Car 1	No. of Passengers in Car 2
01/22/07	19:20	1/22/07	Injury	Good Hope Rd.	2300	2	1	1	2
01/22/07	19:20	1/22/07	Prop. Damage	Benning Rd.	3330	1	0	1	0
01/23/07	12:20	1/23/07	DC Property	Benning Rd.	4500	2	0	1	1
08/12/07	11:25	8/12/07	Injury	Southern Ave.	4400	2	1	1	0
08/12/07	14:00	8/12/07	Hit and Run	57th St.	100	2	0	1	0

Table 3.5. Sample Work Zone Data from West Virginia

County	Route	Project Description	Miles	Start Date	Completion Date
Braxton	WV 4	Replace Drainage, Gassaway–Sutton Road	0.68	9/25/2008	9/28/2008
Braxton	I-79	Reset Bearings, Sutton/Gassaway Bridge	0.22	5/28/2008	7/28/2008
Braxton	I-79	Resurfacing, County 19/26–Flatwoods	2.76	6/28/2008	10/31/2008
Braxton	I-79	Resurfacing, Flatwoods–Burnsville Road	3.74	5/28/2008	10/31/2008
Braxton	I-79	Resurfacing, WV 5–Burnsville	3.95	6/28/2008	10/31/2008
Brooke	WV 2	Beech Bottom–Wellsburg Road	4.32	7/28/2008	10/28/2008
Brooke	WV 2	Follansbee–Coketown Road	2.09	7/1/2008	10/28/2008

Table 3.6. Environmental Data of Candidate Studies

State	Traffic Count	Work Zone Log	Crash Log
Virginia	Yes	Yes[a]	Yes
Delaware	Yes	No	Yes
Maryland	Yes	No	Yes
Washington, D.C.	Yes[b]	Yes[c]	Yes
New York[d]	Yes[e]	Yes	Yes
New Jersey	Yes[f]	No	Yes[g]
Pennsylvania[h]	Yes	Yes	Yes
West Virginia	Yes[i]	Yes[j]	Yes[k]
Michigan	Yes	No	Yes

[a]Data before August 2008 incomplete.
[b]Data available at http://ddot.dc.gov/DC/DDOT/About+DDOT/Maps/Traffic+Volume+Maps. According to DDOT, the data were collected using portable counters every 3 years and are converted to Annual Average Daily Traffic (AADT) as shown on the map.
[c]Data available online at http://app.ddot.dc.gov/information_dsf/construction/index.asp?pro=COM&wardno=1.
[d]Online real-time data available at www.511ny.org/traffic.aspx.
[e]Online data available at http://gis.nysdot.gov/tdv/.
[f]Online data available at www.state.nj.us/transportation/refdata/roadway/traffic_counts/.
[g]Online data available at www.state.nj.us/transportation/refdata/accident/rawdata01-03.shtm.
[h]Online real-time data available at www.dot7.state.pa.us/TravelerInformation/.
[i]AADT data online at www.wvdot.com/3_roadways/rp/TA%20Traffic%20files/SecBcounts.htm.
[j]Real-time work zone information online www.transportation.wv.gov/highways/traffic/Pages/roadconditions.aspx. Historical work zone data available for the past 2–3 years.
[k]www.transportation.wv.gov/highways/traffic/Pages/default.aspx.

The results show that the NTDS and the NTNDS score the highest, with high scores for video, vehicle, and format structure measures. Although these data sets are limited in terms of population or vehicle coverage, the weights assigned to these categories were lower; thus the final score is still considered high relative to the other studies. The next two studies are the 100-Car Study and the DDWS FOT. Noteworthy is the fact that the truck and teen driver studies scored low in the comprehensiveness category because they are restricted either to specific vehicle types or specific driver populations. Given that the focus of this project is to investigate the feasibility of using video data to characterize driver behavior before nonrecurring congestion, this feasibility category is not assigned a high weight.

All the data sources are accessible to the research team, although some limitations may apply. For example, the teen study conducted by VTTI, which studied minors, would require special procedures before any data mining could be conducted. Specifically, data for this study are strictly limited to VTTI researchers, and data reduction can be conducted only in a separate laboratory in which data reductionists cannot be seen by other personnel. Special instructions should be given to reductionists regarding what to do if they find sensitive data or if they meet participants socially and other such conduct instructions. The resulting qualified data sets are listed in Table 3.8. As can be seen from the table, Projects 6, 7, 8, and 11 score relatively higher than the other projects overall. These projects collected data with fewer flaws in video data. Postprocessing of that data will require fewer monetary and human resources.

Table 3.7. Definitions of Dimensions

Feasibility Dimensions	Definitions
Institutional Review Board (IRB)	Do the consent forms used in the study allow for the data to be released to third parties?
Comprehensiveness	
Driver population	Does the data set contain a sample of heterogeneous drivers (e.g., teens, older adults, novices)?
Types of roadways	Are multiple driving locations (freeways, urban settings, rural roads) represented in the data set?
Types of trips	Does the data set contain different trips that occurred at different times during the day (e.g., peak and nonpeak)?
Types of vehicles	Does the data set contain multiple vehicle types?
Video Data Quality	
Driver's hand and foot movements captured	Can the driver's hands and feet be clearly seen in the video?
Driver's face captured	Does the video resolution allow for eyeglance reduction? Is it possible to see driver–passenger interactions and other sources of distraction?
Front view	Do the camera views allow verification of interaction between the vehicle and other vehicles in sight?
Side view	Do camera views outside the vehicle allow the researcher to see what the driver is responding to by the side of the vehicle?
Rear view	Will the following vehicle be seen in the video?
Vehicle Data	
Lane location for each target	Are radar data available for calculating lane locations for other targets?
Projected collision time	Is TTC available in the data set or can it be calculated by data reductionists?
Speed	Is vehicle speed available?
Headway	Is headway available (either distance or time headway)?
Accelerometer	Is acceleration measured?
Braking	Is braking behavior recorded in the data?
GPS	Are GPS data available to identify the vehicle's location?
Lane-changing behavior	Is lane-changing behavior coded in the data or in the reduced data?
Lateral placement	Is the lateral placement of the vehicle measured?
Linkages	
Ability to link to environmental data	Is it possible to link the data to environmental data (such as weather) using the time and location information?
Ability to link to operational data	Will it be possible to link the vehicle data with the surrounding operational data, such as traffic volume or congestion?
Ability to link to special-event data	Is the time stamp valid to link the data to a surrounding special event?
Ability to link to incident data	Are the crash data available to be linked to the data sets?
Ability to link to traffic control devices	Can any traffic control devices (e.g., traffic light, stop sign, yield sign) be linked to the data set?
Ability to link to work zone data	Can work zone data be linked to the data set?
Data Format and Structure	
Sampling rate suitability	Is the sampling rate of the data collection sufficient to understand driver behavior and traffic conditions outside the vehicle?
Event or continuous	Does the data set contain continuous driving behavior, or just segments that are event-triggered?
Reduced or raw	Is the data set already in a format that would allow for efficient analysis (reduced), or are the data only available in a raw format?

Table 3.8. Scale Scores of Candidate Studies

Feasibility Dimensions (Score)	Weight	Project 2: ACAS Field Operational Test (FOT)	Project 5: Road Departure Crash Warning System FOT	Project 6: 100-Car Study	Project 7: Drowsy Driver Warning System FOT	Project 8: Naturalistic Truck Driving Study	Project 11: Naturalistic Teen Driving Study
Legal Restrictions (0/1)		**1**	**1**	**1**	**1**	**1**	**1**
Comprehensiveness	**15**	**6.40**	**6.40**	**6.50**	**3.40**	**3.80**	**5.30**
Driver population	4	7	7	6	4	5	3
Types of roadways	4	7	7	7	4	4	7
Types of trips	1	7	7	6	1	1	6
Types of vehicles	1	1	1	7	1	1	7
Video Data Quality	**40**	**3.40**	**3.40**	**6.20**	**7.00**	**7.00**	**6.20**
Driver's hand and foot movements captured	2	1	1	7	7	7	7
Driver's face captured	2	7	7	7	7	7	7
Front view	2	7	7	7	7	7	7
Side view	2	1	1	3	7	7	3
Rear view	2	1	1	7	7	7	7
Vehicle Data	**20**	**7.00**	**5.33**	**6.11**	**6.00**	**7.00**	**6.33**
Lane location for each target	1	7	4	5	7	7	5
Projected collision time	1	7	1	7	7	7	7
Speed	1	7	7	7	7	7	7
Headway	1	7	1	7	7	7	7
Accelerometer	1	7	7	7	7	7	7
Braking	1	7	7	7	7	7	7
GPS	1	7	7	5	4	7	7
Lane-changing behavior	1	7	7	5	7	7	5
Lateral placement	1	7	7	5	1	7	5
Linkages	**20**	**5.5**	**5.5**	**2.67**	**3.17**	**5.50**	**5.83**
Ability to link to environmental data	2	7	7	2	3	7	7
Ability to link to operational data	2	7	7	2	3	7	7
Ability to link to special-event data	2	4	4	2	4	4	4
Ability to link to incident data	2	7	7	2	3	7	7
Ability to link to traffic control devices	2	6	6	6	3	6	6
Ability to link to work zone data	2	2	2	2	3	2	4
Data Format and Structure	**5**	**5.20**	**5.20**	**7.00**	**7.00**	**7.00**	**7.00**
Sampling rate suitability	1	7	7	7	7	7	7
Event or continuous	6	7	7	7	7	7	7
Raw or reduced	3	1	1	7	7	7	7
Overall Score (1–7)		**5.1**	**4.7**	**5.6**	**5.5**	**6.2**	**6.1**

CHAPTER 4

Develop a Methodology for Analyzing Data

A list of data elements has been identified from the qualified data sets in Chapter 3. From the various data sets, data elements that need to be considered could include video and vehicle kinematic data, in-vehicle and surrounding environmental data, vehicle- and infrastructure-based data, and raw and reduced data. Processing and analyzing the data requires addressing data storage issues, data storage configuration, reduction, and the computing necessary to analyze each data set. Because the candidate studies identified earlier are conducted by VTTI or UMTRI, the data computation capability of these two institutes is discussed briefly.

Data Storage and Computation Requirements

Naturalistic data collection studies that include video data usually generate large files that require professional management. Consequently, research data at VTTI are stored on the Virtual Library System (VLS) Storage Area Network (SAN) that operates within a dedicated private network. These data are isolated from VTTI's operational network and all other networks, including the web, by high-end firewall hardware managed by VTTI's Information Technology Group (ITG). All network connections within VTTI are high-speed Gigabit Ethernet. The data sets can be accessed using a dedicated, high-speed structured query language (SQL) server and by means of special application servers using Microsoft Sequel Server, MatLab, or SAS. Figure 4.1 shows the data center at VTTI.

The data center has the following features:

- Emergency power provided by an on-site diesel generator that supplies backup power for the data center, emergency lighting, and a telecommunications closet;
- External wide area network (WAN) speeds equal to an OC-3 (45 Mbps, approximately 30 times that of a T-1 connection);
- A dedicated climate control system with a backup contingency system;
- Remote monitoring alarm system for indication of fire, smoke, intrusion, power outage, climate control failure, hardware failure, water presence, and high temperature;
- Elevated flooring system;
- High physical security with steel-reinforced structural walls;
- Limited personnel accessibility; and
- Two 600-ft^2 secure data reduction laboratories that house high-end Dell workstations.

VTTI's Data Services Center server room houses the following equipment:

- 250+ Dell-branded business-class desktops, laptops, and laboratory workstations;
- 15+ high-availability, high-performance Dell PowerEdge servers;
- More than 60 TB of redundant high-speed storage with short-term expandability exceeding 100 TB;
- Redundant optical SAN switch/routers;
- A large-capacity backup system that includes a tape library capable of handling 12 GB of data per min; and
- Network connections that are all high-speed Gigabit Ethernet (VTTI was a pioneer in implementing this technology for every network portal).

The computing requirements necessary for data manipulation and analysis are a result of the data size, data storage configuration, reduction needs, and analysis requirements. The VTTI team created codes that extract subsets of data for analysis purposes in multiple software environments, including MatLab, SAS, and SQL. An example of a flowchart for a MatLab function to extract a subset of data on a defined highway section is provided in Figure 4.2.

Besides the existing commercial software, VTTI developed proprietary data-viewing software, the Data Analysis and Reduction Tool (DART), to allow synchronized viewing of driver performance (parametric) data and video and audio

Figure 4.1. Environmentally controlled and secured VTTI data center.

streams. This system allows researchers and data reductionists to work directly with large databases while providing ease of use and improved accuracy. As shown in Figure 4.3, reductionists can select specific variables and customize the interface of the illustration. While the video is playing, the software will draft charts for those variables along the time axis synchronized with the video. When multiple targets are sensed by radar units, the charts are color coded for better viewing.

Similar to the data storage and calculation capability of VTTI, UMTRI has developed its data collection, storage, and computation capability over the years. UMTRI has developed large, driver-vehicle databases since the mid-1990s. By the end of 2006, approximately 1 million vehicle miles of data had been collected. The data archive at UMTRI is maintained on an Internet-accessible system for on-demand access by

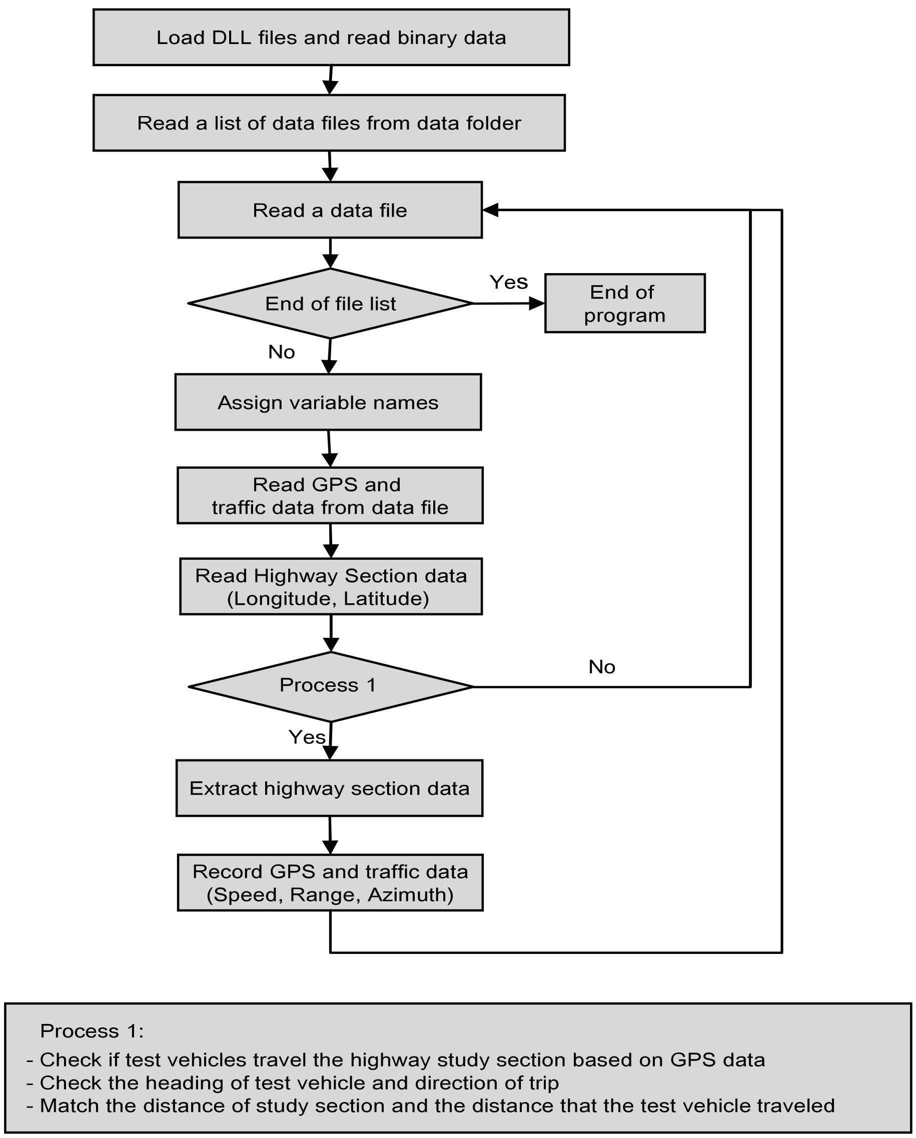


Figure 4.2. Sample flowchart of a MatLab function.

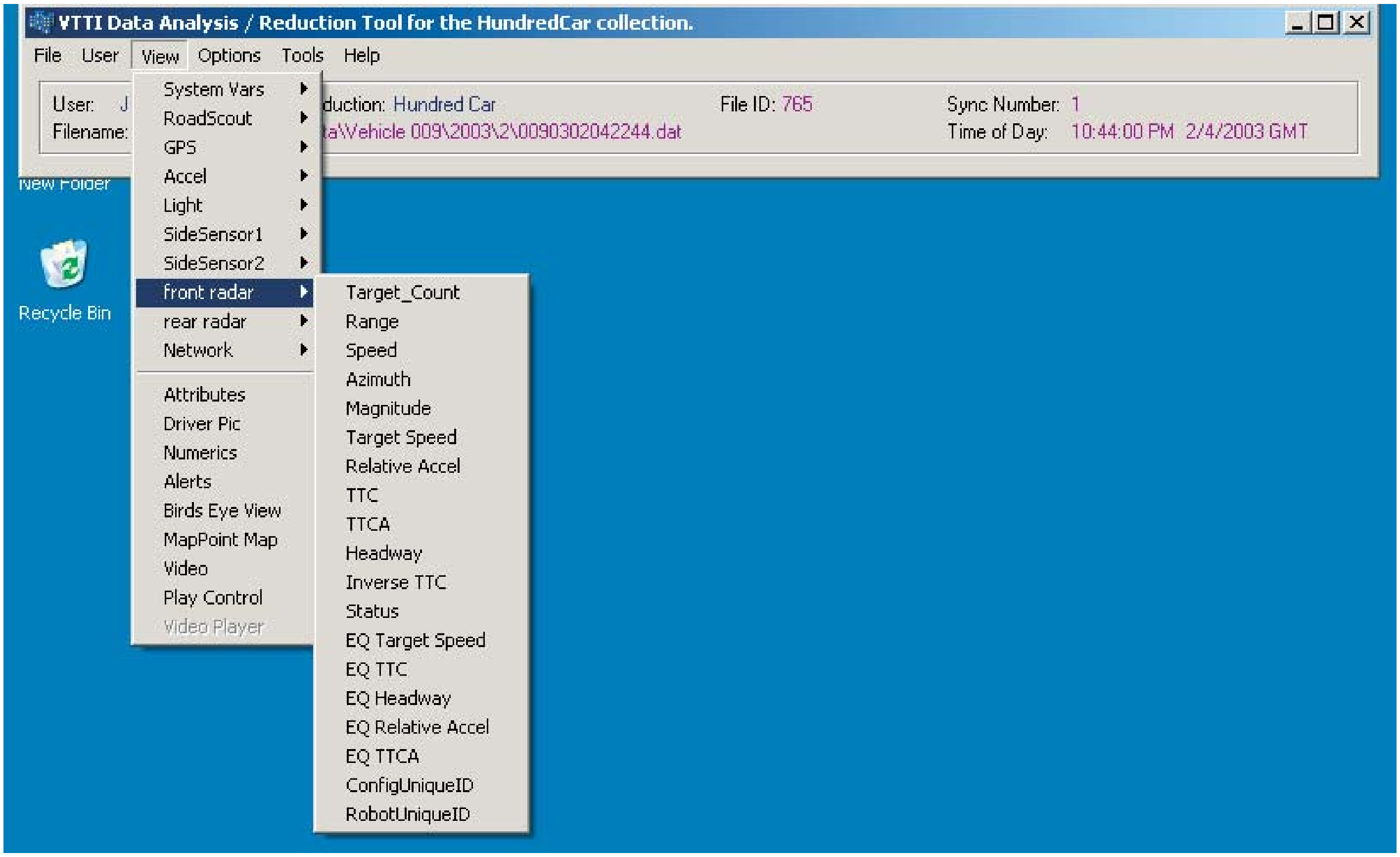


Figure 4.3. DART developed by VTTI.

registered users. Acquired data sets go through rigorous data quality assurance and validation procedures and can be reformatted for use in standard analytical systems (e.g., ADAAS, SAS, and SPSS) and on multiple software platforms. Two servers hosted UMTRI data when it was first collected in 2004 through 2005. The first server is a Dell PowerEdge 4600 with a 12-GB system drive, a 546-GB RAID 10 drive containing database data files, and a 124-GB RAID 1 drive containing database log files. The server runs Windows 2003 Server SP2 and Microsoft SQL Server 2005. The video files are stored on a separate Dell PowerEdge 4600 with a 1.2-TB logical drive. The servers are among several servers housing data from six FOTs and dozens of other projects that are located in a small, secure server room. UMTRI uses data structures that are intended to provide fast access so that custom tools allow users to access video in any project within 5 s. Direct access to the data and video is available for authorized users located within UMTRI, throughout the university, and within selected research partners, including access to all tools from remote locations.

The descriptions provided summarize the data computation and storage abilities of VTTI and UMTRI. Data-processing capability does not have to achieve such levels to conduct video data–related research. Data storage and reduction will be satisfactory to make a judgment about nonrecurring congestion caused by driver behavior as long as the equipment can (1) secure the data so that only authorized personnel can access it; (2) be large enough to accommodate the video data, radar data, and associated vehicle and external data; (3) organize a relational database so that the data can be queried and retrieved efficiently; and (4) allow synchronized playback of the video data together with other vehicle and external data.

Data Reduction and Crash and Near-Crash Detection

As stated, data collection studies that include video data usually generate large files. The 100-Car Study raw data set, including both video and vehicle kinematic data, requires 7 TB of data storage. Reduced data sets require much less storage capacity. The quality of reduced data sets should be sufficient for the purposes of this study.

With the exception of the NTNDS, which is still involved in an ongoing data reduction effort (see Table 3.7), the other three VTTI data sets and the two UMTRI FOT studies have been reduced and are discussed in detail in this chapter. The data reduction methodology, criteria to identify crashes and near crashes, and the associated weather and traffic conditions at the point that crashes and near crashes occurred are enumerated.

Project 2: Automotive Collision Avoidance System Field Operational Test

The original purpose of the study was to test the effectiveness of the FCW and ACC systems developed and integrated by GM and Delphi. The FCW delivers a warning when there is a forward crash possibility, and the ACC system maintains a safe headway. In addition to these systems, a DAS developed by UMTRI was installed on the test vehicles. The DAS deployed in this study organized data by trips. The system started to collect data each time the vehicle ignition was turned on and stopped when the engine was stopped. After each trip, files comprising the time history, transition, trip summary, triggered summary, raw CAN, video (each record contains a time and the bitmap image), and audio were recorded in the DAS.

The video system had two cameras: a forward camera and a face camera. The forward scene was recorded continuously at 1 Hz, and the driver's face was sampled at 5 Hz for 4 s every 5 min. Exposure video was saved simultaneously. One of the following triggers enabled the video system to capture a retrospective clip, 8 s of video data (5 s before the event and 3 s after for the forward camera; 4 s before and 4 s after for the face camera), and saved the data to disk. These were activated by (1) a driver comment button; (2) Conventional Cruise Control (CCC) and ACC engagement events; (3) FCW events; and (4) hard-braking events.

The files recorded by the triggers were sent to UMTRI by way of cell modem for the purpose of preparing and facilitating the participant debriefing. When the car was returned to UMTRI, software was used to process and view data, as shown in Figure 4.4 (*1*). Correspondingly, there were two phases of data reduction. The first phase was completed by the DAS as the vehicles were being driven by the participants. The onboard data-processing system generated CAN message lists and histograms for multiple variables, as well as an inventory file with all triggers. The second phase involved the processing of data by analysts after the vehicle was returned to UMTRI. The resulting data set of the ACAS FOT contained approximately 164 GB of data. Invalid trips were filtered out by discarding trips with any of the following attributes: zero distance; maximum speed less than 20 mph; system health histogram not okay; critical okay less than 90%; frozen threat index; radar malfunction; or non-FOT subject driving.

While the video data were examined, environmental factors (e.g., weather and traffic density), target characteristics, driver behavior, alert classification, and the driving scenario were identified and classified. A total of 634 alerts were manually examined wherever the target speed was estimated to exceed 2 m/s. Specifically, the weather was coded as 0 for "dry," 1 for "wet," and 2 for "snow covered." As defined in the data reduction dictionary, the road was classified as wet if it was wet from snow but not snow covered. Also, any moisture on the road was classified as wet; there did not need to be standing water. A snow-covered classification would have included ice-covered if it was observed. If any portion of the road, including turn lanes, were covered in snow, the classification was snow covered. The traffic density was coded using the number of visible traffic counts in the front camera, as shown in Figure 4.5 (*1*), in which "sparse" describes the top figure for which smoothed traffic count is less than 1.5 targets, "moderate" describes the middle figure for which smoothed traffic count is between 1.5 and 4.0 targets, and "dense" describes the bottom figure for which smoothed traffic count is greater than 4.0 targets. The reduced data dictionary appears in Appendix A (*1*).

Although the reduced data set included a description of secondary tasks (e.g., cell phone use, eating, drinking, smoking, and grooming), hand location, and eye location, it did not specify if the behavior was a contributing factor to the event. The team explored the possibility of conducting additional data reduction to identify contributing factors but realized that the RDCWS FOT conducted by UMTRI (Project 5) has a similar but more sophisticated data structure compared with that for this project; that data set fits better with the current research goal. Also, the data transferring of this project has issues because of a waiver signed by the participants to allow for secondary usage of the data. To complete the additional data reduction for the UMTRI studies in a timely manner, it was decided that the team would focus on the data collected from Project 5, and the results of data reduction performed by the team are discussed in Chapter 8.

Project 5: Road Departure Crash Warning System Field Operational Test

The research goal was to study the effectiveness and acceptance of the RDCWS, in which a Lane Departure Warning System (LDWS), a Curve Speed Warning System (CSWS), and a DAS were equipped onboard to provide warnings and collect data. The LDWS used a camera to observe visual features that delineate lane and road edges and a set of onboard radar units to modulate the warnings when potentially dangerous objects were sensed. The CSWS used GPS technology and a digital map to anticipate curve location and radius. The CSWS issued an alert when the speed was detected to be unsafe for the curve. The DAS, developed and managed by UMTRI, collected data from the RDCWS CAN bus, two radar buses, two video streams, an audio stream, and several other instruments. Data variables collected in this study are listed in Table 4.1 (*2*).

The two video cameras—one capturing the driver's face and the other recording the forward view—recorded video

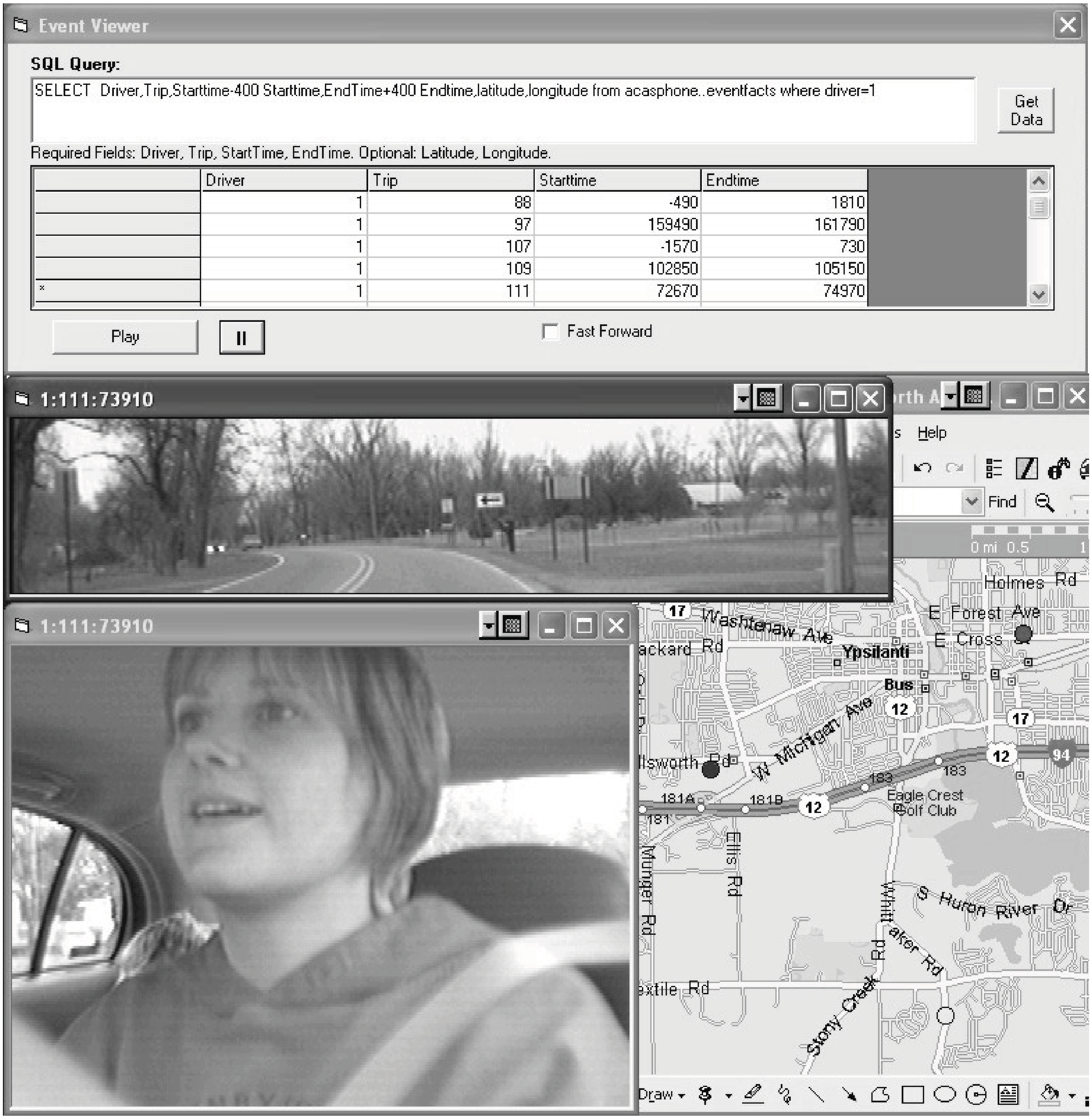


Figure 4.4. Software interface used in Project 2.

data at the frequencies shown in Table 4.2. Specifically, the DAS captured a 5-s video clip every 5 min from the face camera, regardless of the driving situations or driver activity, creating a random sample of driver activity. The triggers (RDCWS alerts generated by the LDWS, CSWS, or a driver comment button pressed by the driver) sent a signal to the video system, and the cameras then recorded data at a higher frequency (as shown in Table 4.2) for 4 s before the event and 4 s after the alert.

To reduce the data, analysts assigned values to 24 variables based on a review of the video and numerical data from the alerts. Twelve of the variables were associated with the circumstances (e.g., road types and curves), and 12 variables addressing driver activities before, during, and after the alert (e.g., driver's distractions, including glance direction) were coded for the events. Appendix B contains a summary of the data dictionary (*2*), and Figure 4.6 shows a screenshot of the video data collected in this project. Although the frequency

Figure 4.5. Traffic density definitions in data reduction in Project 2.

Table 4.1. Data Collected in Project 5

Data	Sources
Vehicle and driver identifications	Preset values
Vehicle position, heading, and motion: speed, yaw rate, accelerations, pitch and roll angle and rates, GPS (differential and nondifferential)	CAN bus and FOT sensors
Driver control inputs: steering wheel angle, throttle input, brake switch, turn signal, headlamp state, cruise control state and set speeds, LDWS and CSWS sensitivity settings	CAN bus
RDCWS driver displays: LDWS and CSWS alerts and levels, availability icons	CAN bus
RDCWS intermediate values: lane position, warning thresholds, road geometry estimates, threat locations, vehicle-centered object map	CAN bus
Roadway environment: road type and attributes, urban and rural settings	Onboard digital map via CAN bus, plus postprocessing, Highway Performance Monitoring System (HPMS) database
RDCWS and subsystem health and diagnostics information, as well as subsystem version numbers	CAN bus
RDCWS radar data: forward radar data, side radar data	CAN bus
Video: forward driving scene and driver face views	LDWS camera, FOT sensors
Audio from driver comment button: dictated messages from driver	FOT sensors

Table 4.2. Video Camera Configurations in Project 5

Video	Nominal Rate (Hz)	Triggered Rate (Hz)	Pretrigger Window(s)	Posttrigger Window(s)
Forward video	2	10	4	4
Face video	0.5	10	4	4
Face video—exposure	5	n/a	0	5

set for the video cameras was relatively low, the data collected in this study are accurate, comprehensive, and efficiently reduced.

Project 6: 100-Car Study

The data reduction performed in this project can be summarized in two steps. First, events were identified using predefined trigger criteria values that resulted in a low miss rate and a high false alarm rate to avoid missing valid events. These criteria were derived after a sensitivity analysis. The specifics of the trigger values are described in Table 4.3.

Second, the video data for all the events identified were reviewed by data reductionists. The reductionists focused on a 90-s epoch for each event (60 s before and 30 s after the event) to validate the event, determine severity, and code the event for a data reduction dictionary.

The severities of the valid events were determined based on various criteria, and all required variables were recorded and edited in MySQL databases. According to the severities, the valid events were classified into five categories: nonconflict, proximity events, crash-relevant, near crash, and crash. Nonconflicts were "any event that increases the level of risk associated with driving but does not result in a crash, near-crash, or incident. . . . Examples include: driver control error without proximal hazards being present; driver judgment error, such as unsafe tailgating or excessive speed; or cases in which drivers are visually distracted to an unsafe level." Proximity events were "any circumstance resulting in extraordinarily close proximity of the subject vehicle to any other vehicle, pedestrian, cyclist, animal, or fixed object where, due to apparent unawareness on the part of the driver(s), pedestrians, cyclists or animals, there is no avoidance maneuver or response. Extraordinarily close proximity is defined as a clear case where the absence of an avoidance maneuver or response is inappropriate for the driving circumstances (including speed, sight distance, etc.)." Crash-relevant events were "any circumstance that requires a crash avoidance response on the part of the subject vehicle or any other vehicle, pedestrian, cyclist, or animal that is less severe than a rapid evasive maneuver, but greater in severity than a 'normal maneuver' to avoid a crash. A crash avoidance response can include braking, steering, accelerating, or any combination of control inputs.

Figure 4.6. Video data collected in Project 5.

Table 4.3. Event Triggers in Project 6

Trigger Type	Description
Lateral acceleration	Lateral motion equal to or greater than 0.7 *g.*
Longitudinal acceleration	Acceleration or deceleration equal to or greater than 0.6 *g.*
	Acceleration or deceleration equal to or greater than 0.5 *g* coupled with a forward TTC of 4 s or less.
	All longitudinal decelerations between 0.4 and 0.5 *g* coupled with a forward TTC value less than or equal to 4 s; corresponding forward range value at the minimum TTC is not greater than 100 ft.
Event button	Activated by the driver pressing a button on the dashboard when an event occurred that he or she deemed critical.
Forward TTC	Acceleration or deceleration equal to or greater than 0.5 *g* coupled with a forward TTC of 4 s or less.
	All longitudinal decelerations between 0.4 and 0.5 *g* coupled with a forward TTC value less than or equal to 4 s; corresponding forward range value at the minimum TTC is not greater than 100 ft.
Rear TTC	Any rear TTC trigger value of 2 s or less that also has a corresponding rear range distance less than or equal to 50 ft and any rear TTC trigger value in which the absolute acceleration of the following vehicle is greater than 0.3 *g.*
Yaw rate	Any value greater than or equal to a plus and minus 4-degree change in heading (i.e., vehicle must return to the same general direction of travel) within a 3-s window.

A ‘normal maneuver’ for the subject vehicle is defined as a control input that falls inside of the 99 percent confidence limit for control input as measured for the same subject.” Near crashes were defined as “any circumstance that requires a rapid, evasive maneuver by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. A rapid, evasive maneuver is defined as a steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle capabilities. As a guide: subject vehicle braking greater than 0.5 *g*, or steering input that results in a lateral acceleration greater than 0.4 *g* to avoid a crash, constitutes a rapid maneuver.” Crashes were defined as events with “any contact with an object, either moving or fixed, at any speed, in which kinetic energy is measurably transferred or dissipated. Includes other vehicles, roadside barriers, objects on or off the roadway, pedestrians, cyclists or animals” (*3–4*).

The variables defined in the reduced data can be categorized into one of the following homogeneous groups: general information, event variables, contributing factors, surrounding factors, driver state variables, and driver information of the second vehicle involved. The variables are described in Table 4.4.

Table 4.4. List of Variables in Reduced Data in Project 6

Classification	List of Variables
General information	Vehicle number, epoch number, event severity, trigger type, driver subject number, onset of precipitating factor, and resolution of the event
Event variables	Event nature, incident type, pre-event maneuver, judgment of Vehicle 1 maneuver before event, precipitating factor, evasive maneuver, and vehicle control after corrective action
Contributing factors	Driver behavior: Driver 1 actions and factors relating to the event, Driver 1 physical or mental impairment, Driver 1 distracted, willful behavior, driver proficiency, Driver 1 drowsiness rating, Driver 1 vision obscured by, and vehicle contributing factors
Environmental factors: driving environment	Weather, light, windshield wiper activation, surface condition, and traffic density (level of service)
Driving environment: infrastructure	Kind of locality, relation to junction, traffic flow, number of travel lanes, traffic control, and alignment
Driver state variable	Driver 1 hands on wheel, occupant safety belt usage, Driver 1 alcohol use, fault assignment, average PERCLOS, and Driver 1 eyeglance reconstruction
Driver/Vehicle 2	Number of other vehicles/person(s), location of other vehicle(s)/persons, Vehicle/Person 2 type, Vehicle 2 maneuver, Driver/Vehicle 2 corrective action attempted, Driver/Vehicle 2 physical or mental impairment, and Driver 2 actions and factors relating to crash or incident

As listed in Table 4.4, event nature was one of the variables coded in the reduced data set. Event nature usually was decided by reductionists based on predefined criteria together with subjective judgments. According to the definition, the valid events were classified into five categories according to their severity: nonconflict, proximity event, crash-relevant, near crash, and crash. These categories are described previously in this report.

Among the safety-related events, crashes are relatively easy to identify. Near crashes and crash-relevant events are not as straightforward. Many safety applications would greatly benefit from a reliable, purely quantitative definition of a near crash and crash-relevant event. Because a near crash is more severe than a crash-relevant event and could have developed into a real crash, it is important to differentiate a near crash from a crash-relevant event. During data reduction, it was found that strictly applying a fixed number as a threshold value in acceleration or speed to identify near crashes inherently generates some "noise." For example, a TTC of 0.1 s occurs regularly on interstates during congestion and is normal, whereas the same TTC happening on a rural road at night is significantly different. Also, because of the varied driving habits of different drivers, it is hard to apply one uniform number. Some drivers apply more frequent aggressive brakes simply because they are comfortable with rapid decelerations. Therefore, both qualitative and quantitative criteria should be incorporated. As a guideline in this study, a subject vehicle braking greater than 0.5 *g* or a steering input that results in a lateral acceleration greater than 0.4 *g* to avoid a crash constitutes a candidate rapid maneuver, as defined in near crashes. Combined with the subjective judgment of experienced reductionists, this maneuver is decided by reductionists as a near crash or crash-relevant event. A similar policy for identifying near crashes was applied to most VTTI studies with appropriate adjustments made for threshold values. Altogether, 69 crashes and 761 near crashes were identified in the 100-Car data set. The crashes and near crashes were parsed into the following 18 conflict categories:

- Conflict with a lead vehicle;
- Conflict with a following vehicle;
- Conflict with oncoming traffic;
- Conflict with a vehicle in an adjacent lane;
- Conflict with a merging vehicle;
- Conflict with a vehicle turning across the subject vehicle path in the same direction;
- Conflict with a vehicle turning across the subject vehicle path in the opposite direction;
- Conflict with a vehicle turning into the subject vehicle path in the same direction;
- Conflict with a vehicle turning into the subject vehicle path in the opposite direction;
- Conflict with a vehicle moving across the subject vehicle path through the intersection;
- Conflict with a parked vehicle;
- Conflict with a pedestrian;
- Conflict with a pedal cyclist;
- Conflict with an animal;
- Conflict with an obstacle or object in the roadway;
- Single-vehicle conflict;
- Other (specify); and
- Unknown conflict.

Table 4.5 shows the level of service (LOS) at the time the crash or near crash happened. Preferably, the LOS both before and after the event can be identified so that the travel time reliability caused by incidents can be analyzed and modeled. (Details of modeling travel time reliability are discussed in Chapter 6.) Once a car was involved in an accident and the engine was turned off, the onboard data collection system would be turned off and would cease collecting data. Therefore, the LOS after events usually was not available. Table 4.5 lists only the LOS before the events occurred.

The weather conditions associated with safety-related events varied. Clear weather was associated with the same percentage of 100-Car Study crashes (78%) and near crashes (78%). The second most associated weather factor was rain. Drivers had a slightly higher percentage of crashes (12%) than near crashes (8%) that occurred during rainy days. Cloudy weather was associated with more near crashes (13%) than crashes (9%). Only one 100-Car Study crash occurred with snow as an associated factor.

Project 7: Drowsy Driver Warning System Field Operational Test

Data reduction started with the identification of potential events using the DART. A 90-s epoch was created for each event, which included 1 min before the trigger and 30 s after. The automatic scanning resulted in an event database. The triggers and values used in identifying critical events are listed in Table 4.6.

Data reductionists reviewed the video data for the identified events to validate them. Invalid events for which sensor readings were spurious because of a transient spike or some other false-positive classification were filtered out. Valid events were classified as conflicts or nonconflicts. Conflicts were further classified into four categories based on the severity: crash, crash: tire strike, near crash, and crash-relevant. Nonconflicts were events with valid threshold values but did not create safety-significant traffic events. Verified valid conflicts were categorized based on the following descriptions. Crashes are classified as any contact with an object, either moving or fixed, at any speed in which kinetic energy is measurably

Table 4.5. LOS for Crashes and Near Crashes by Conflict Type

	Crashes				Near Crashes			
Conflict type		LOS A	LOS B	LOS C+		LOS A	LOS B	LOS C+
Single	24	20	3	1	48	42	3	3
Lead	15	5	5	5	380	78	241	61
Following	12	4	7	1	70	16	39	15
Obstacle	9	6	3	0	6	1	5	0
Animal	2	2	0	0	10	10	0	0
Turning across opposite direction	2	1	1	0	27	12	15	0
Adjacent vehicle	1	0	1	0	115	32	69	14
Parking	4	3	1	0	5	3	2	0
Across path through intersection	**NA**				27	11	15	1
Oncoming					27	14	12	1
Other					2	1	1	0
Pedestrian					6	4	1	1
Turning across in same direction					3	2	1	0
Turning in same direction					28	16	12	0
Merging					6	2	4	0
Unknown					1	1	0	0

transferred or dissipated. Near crashes (evasive maneuver) are classified as any circumstance that requires a rapid, evasive maneuver, including steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle capabilities, by the subject vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. Near crashes (no evasive maneuver) are classified as any circumstance that results in extraordinary proximity of the subject vehicle to any other object. Extraordinary proximity is defined as a clear case in which the absence of an avoidance maneuver or response is inappropriate for the driving circumstances (e.g., speed and sight distance). Crash-relevant conflicts (evasive maneuvers) are assessed similar to near crashes (evasive maneuvers). Longitudinal decelerations of $-0.35\ g$ or greater are reviewed to assess whether they qualify as crash-relevant conflicts (or near crashes); those with decelerations of $-0.50\ g$ or

Table 4.6. Trigger Values to Identify Critical Incidents in Project 7

Trigger Type	Description
Longitudinal acceleration (LA)	(1) Acceleration or deceleration greater than or equal to $\lvert 0.35\ g \rvert$. Speed greater than or equal to 15 mph. (2) Acceleration or deceleration greater than or equal to $\lvert 0.5\ g \rvert$. Speed less than or equal to 15 mph.
Time to collision (TTC)	(3) A forward TTC value of less than or equal to 1.8 s, coupled with a range of less than or equal to 150 ft, a target speed of greater than or equal to 5 mph, a yaw rate of less than or equal to $\lvert 4°/s \rvert$, and an azimuth of less than or equal to $\lvert 0.8° \rvert$. (4) A forward TTC value of less than or equal to 1.8 s, coupled with an acceleration or deceleration greater than or equal to $\lvert 0.35\ g \rvert$, a forward range of less than or equal to 150 ft, a yaw rate of less than or equal to $\lvert 4°/s \rvert$, and an azimuth of less than or equal to $\lvert 0.8° \rvert$.
Swerve (S)	(5) Swerve value of greater than or equal to 3. Speed greater than or equal to 15 mph.
Critical incident (CI) button	(6) Activated by the driver pressing a button, located by the driver's visor, when an incident occurred that he or she deemed critical.
Analyst identified (AI)	(7) Event identified by a data reductionist viewing video footage; no other trigger listed above identified the event (e.g., LA and TTC).

greater are always coded as crash-relevant conflicts or near crashes. Crash-relevant conflicts (no evasive maneuver) are classified similar to near crashes (no evasive maneuver). Here a TTC of 1 s is used to identify near crashes. Cases with a TTC between 1 and 2 s relied on subjective judgment (*5*). Relevant information was coded in a data dictionary that includes more than 50 variables (Appendix C). Figure 4.7 provides the flowchart showing the process of data reduction.

After data reduction, 915 safety-relevant events were identified; of these, there were 14 crashes, 14 tire strikes, 98 near crashes, and 789 crash-relevant conflicts. A random sample of 1,072 baseline epochs was also selected to represent normal driving. Each baseline epoch is 60 s long. The data dictionary was also applied to those baseline epochs. The purpose of selecting baseline events was to provide a comparison with the events in that the baseline events represent normal driving. Figure 4.8 shows a screenshot of the video data in this study (*5*).

In characterizing traffic density at the time the event happened, a detailed description was provided to the data reductionists to assist in assigning the LOS. According to the report from Project 7 (*5*), six levels of traffic density are defined plus a status of unknown or unable to determine:

1. LOS A/Free flow: Individual users are virtually unaffected by the presence of others in the traffic. Freedom to select desired speeds and to maneuver within the traffic stream is extremely high. The general level of comfort and convenience provided to the motorist, passenger, or pedestrian is excellent;
2. LOS B/Flow with some restrictions: In the range of stable traffic flow, but the presence of other users in the traffic stream begins to be noticeable. Freedom to select desired speeds is relatively unaffected, but there is a slight decline in the freedom to maneuver within the traffic stream from LOS A because the presence of others in the traffic stream begins to affect individual behavior;
3. LOS C/Stable flow: Maneuverability and speed are more restricted. In the range of stable traffic flow but marks the beginning of the range of flow in which the operation of individual users becomes significantly affected by the interactions with others in the traffic stream. The selection of speed is now affected by the presence of others, and maneuvering within the traffic stream requires substantial vigilance on the part of the user. The general level of comfort and convenience declines noticeably at this level;
4. LOS D/Unstable flow: Temporary restrictions substantially slow driver. Represents high-density and unstable traffic flow. Speed and freedom to maneuver are severely restricted, and the driver or pedestrian experiences a generally poor level of comfort and convenience. Small increases in traffic flow will generally cause operational problems at this level;
5. LOS E: Vehicles are unable to pass and there are temporary stoppages. Represents operating conditions at or near the capacity level. All speeds are reduced to a low but relatively uniform value. Freedom to maneuver within the traffic stream is extremely difficult, and it is generally accomplished by forcing a vehicle or pedestrian to yield to accommodate such maneuvers. Comfort and convenience levels are extremely poor, and driver or pedestrian frustration is generally high. Operations at this level are usually unstable because small increases in flow or minor perturbations within the traffic stream will cause breakdowns;

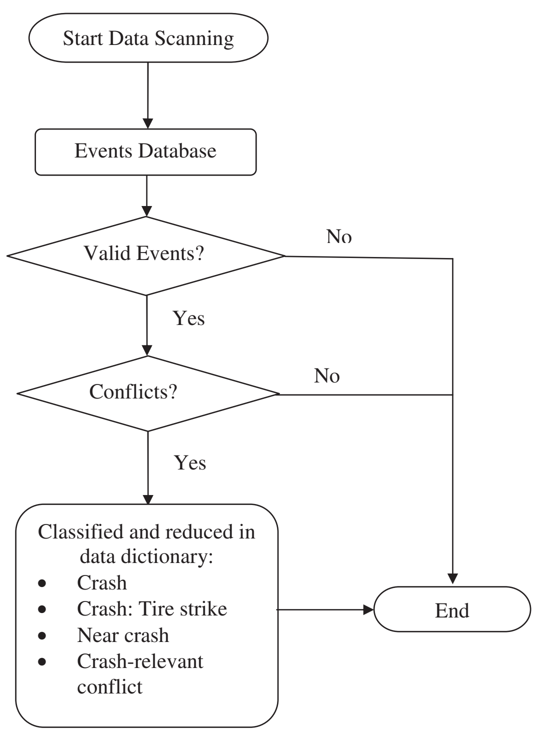


Figure 4.7. Flowchart for data reduction in Project 7.

Figure 4.8. Screenshot from Project 7.

Table 4.7. Level of Service for Crashes and Near Crashes

Traffic Density	No. of Crashes	% of Crashes	No. of Crashes: Tire Strikes	% of Crashes: Tire Strikes	No. of Near Crashes	% of Near Crashes
LOS A	13	92.9%	9	64.3%	61	62.2%
LOS B	1	7.1%	1	7.1%	21	21.4%
LOS C	0	0.0%	4	28.6%	11	11.2%
LOS D	0	0.0%	0	0.0%	1	1.0%
LOS E	0	0.0%	0	0.0%	2	2.0%
LOS F	0	0.0%	0	0.0%	2	2.0%
Unknown	0	0.0%	0	0.0%	0	0.0%
Total	14	100.0%	14	100.0%	98	100.0%

6. LOS F: Forced traffic flow condition with low speeds and traffic volumes that are below capacity; queues form in particular locations. This condition exists whenever the amount of traffic approaching a point exceeds the amount that can traverse the point. Queues form behind such locations. Operations within the queue are characterized by stop-and-go waves, and they are extremely unstable. Vehicles may progress at reasonable speeds for several hundred feet or more and then be required to stop in a cyclic manner. LOS F is used to describe the operating conditions within the queue, as well as the point of the breakdown. In many cases, operating conditions of vehicles or pedestrians discharged from the queue may be quite good. It is the point at which arrival flow exceeds discharge flow that causes the queue to form, and LOS F is an appropriate designation for such points.

Table 4.7 shows the numbers and percentages of crashes, crashes: tire strikes, and near crashes associated with each LOS.

Weather conditions were coded as "No adverse conditions," "Rain," "Sleet," "Snow," "Fog," "Rain and fog," "Sleet and fog," "Other," and "Unknown." Table 4.8 shows the numbers and percentages of crashes, crashes: tire strikes, and near crashes associated with each weather condition.

Project 8: Naturalistic Truck Driving Study

Data reduction for the NTDS involved two main steps. Step 1 was to identify events of interest. The DART was used to find events of interest by scanning the data set for notable actions, including hard braking, quick steering maneuvers, short TTCs, and lane deviations. Table 4.9 displays the various trigger threshold values (*6*). VTTI researchers developed the values based on data reduction experience obtained from the 100-Car Study. A 75-s epoch was created for each trigger comprising 1 min before the trigger and 15 s after the trigger. The result of the automatic scan was an event data set that

Table 4.8. Weather Condition When Events Happened

Weather	No. of Crashes	% of Crashes	No. of Crashes: Tire Strikes	% of Crashes: Tire Strikes	No. of Near Crashes	% of Near Crashes
No adverse conditions	11	78.6%	14	100.0%	91	92.9%
Rain	2	14.3%	0	0.0%	7	7.1%
Sleet	0	0.0%	0	0.0%	0	0.0%
Snow	0	0.0%	0	0.0%	0	0.0%
Fog	1	7.1%	0	0.0%	0	0.0%
Rain and fog	0	0.0%	0	0.0%	0	0.0%
Sleet and fog	0	0.0%	0	0.0%	0	0.0%
Other	0	0.0%	0	0.0%	0	0.0%

Table 4.9. Trigger Values Used to Identify Critical Incidents in Project 8

Trigger Type	Definition	Description
Longitudinal acceleration (LA)	Hard braking or sudden acceleration	Acceleration or deceleration greater than or equal to $\lvert 0.20\ g \rvert$. Speed greater than or equal to 1 mph (1.6 km/h).
Time to collision (TTC)	Amount of time (in seconds) it would take for two vehicles to collide if one vehicle did not perform an evasive maneuver.	A forward TTC value of less than or equal to 2 s, coupled with a range of less than or equal to 250 ft, a target speed of greater than or equal to 5 mph (8 km/h), a yaw rate of less than or equal to $\lvert 6°/s \rvert$, and an azimuth of less than or equal to $\lvert 0.12° \rvert$.
Swerve (S)	Sudden jerk of the steering wheel to return the truck to its original position in the lane.	Swerve value of greater than or equal to 2 rad/s2. Speed greater than or equal to 5 mph (8.05 km/h).
Lane deviations (LD)	Any time the truck aborts the lane line and returns to the same lane without making a lane change.	Lane tracker status equals abort. Distance from center of lane to outside of lane line less than 44 in.
Critical incident (CI) button	Self-report of an incident by the driver.	Activated by the driver pressing a button by the driver's visor when an incident occurred that he or she deemed critical.
Analyst identified (AI)	Event identified by the analyst but not by a trigger.	Event that was identified by a data analyst viewing video footage; no other trigger listed above identified the event (e.g., LA and TTC).

included both valid and invalid events waiting to be further identified in step 2.

Step 2 involved a manual inspection of these potential events of interest by data reductionists to filter out invalid events. Figure 4.9 shows a screenshot of the video data. Valid events were further classified into one of six safety-critical events: crash, crash: tire strike, near crash, crash-relevant conflict, unintentional lane deviation, and illegal maneuver. Table 4.10 summarizes the definitions of these event types (*6*).

The details of each valid event were coded by reductionists using an established coding directory by watching the associated video data and answering questions in a pull-down menu in the DART. Invalid events were eliminated when sensor readings were spurious because of a transient spike or some other anomaly (i.e., false positive). Appendices C and D provide the data dictionary for event and environmental criteria, respectively.

Figure 4.9. Screenshot of video data in Project 8.

Most of the events happened during smooth traffic conditions and nonadverse weather conditions. Tables 4.11 and 4.12 show the details of the LOS and weather categories, respectively. The traffic and weather conditions at the instant the crashes and near crashes occurred should not be used alone to decide if these factors have an impact on the likelihood of the safety-related events. The percentage of heavy traffic and adverse weather conditions occurring in baseline epochs should be compared with the event epochs. For example, heavy traffic occurs in 10% of the baseline epochs but is present in 20% of the event epochs. Although 20% is not a high percentage, it is a factor worthy to note.

Project 11: Naturalistic Teen Driving Study

Data reduction procedures in this study included three tasks: initial data reduction, straight road segment data reduction, and event data reduction. The initial data reduction involved recording general information about the driver and passengers of each trip. A trip was defined as a continuous data collection from the start of the engine of a participant's vehicle to its turn-off. The recorded variables included participant ID, number of passengers, age of passengers,

Table 4.10. Event Types in Project 8

Event Type	Description
Crash	Any contact with an object, either moving or fixed, at any speed.
Crash: tire strike	Any contact with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated when the contact occurs only on the truck's tire. No damage occurs during these events (e.g., a truck is making a right turn at an intersection and runs over the sidewalk or curb with a tire).
Near crash	Any circumstance that requires a rapid, evasive maneuver (e.g., hard braking, steering) by the subject vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash.
Crash-relevant conflict	Any circumstance that requires a crash-avoidance response on the part of the subject vehicle, any other vehicle, pedestrian, cyclist, or animal that was less severe than a rapid evasive maneuver (as defined above) but more severe than a normal maneuver. A crash-avoidance response can include braking, steering, accelerating, or any combination of control inputs.
Unintentional lane deviation	Any circumstance in which the subject vehicle crosses over a solid lane line (e.g., onto the shoulder) where there is not a hazard (guardrail, ditch, or vehicle) present.
Illegal maneuver	Any circumstance in which either the subject vehicle or the other vehicle performs an illegal maneuver, such as passing another vehicle across the double yellow line or on a shoulder. In many cases, neither driver performs an evasive action.

Table 4.11. LOS for Crashes and Near Crashes in Project 8

Traffic LOS	No. of Crashes	% of Crashes	No. of Crashes: Tire Strikes	% of Crashes: Tire Strikes	No. of Near Crashes	% of Near Crashes
LOS A	3	60.0%	4	50%	23	37.7%
LOS B	2	40.0%	3	37.5%	22	36.1%
LOS C	0	0.0%	0	0.0%	13	21.3%
LOS D	0	0.0%	0	0.0%	1	1.6%
LOS E	0	0.0%	1	12.5%	2	3.3%
LOS F	0	0.0%	0	0.0%	0	0.0%
Unknown	0	0.0%	0	0.0%	0	0.0%
Total	5	100.0%	8	100.0%	61	100.0%

Table 4.12. Weather Condition for Crashes and Near Crashes in Project 8

Weather	No. of Crashes	% of Crashes	No. of Crashes: Tire Strikes	% of Crashes: Tire Strikes	No. of Near Crashes	% of Near Crashes
No adverse conditions	5	100.0%	8	100.0%	56	91.8%
Rain	0	0.0%	0	0.0%	5	8.2%
Sleet	0	0.0%	0	0.0%	0	0.0%
Snow	0	0.0%	0	0.0%	0	0.0%
Fog	0	0.0%	0	0.0%	0	0.0%
Rain and fog	0	0.0%	0	0.0%	0	0.0%
Sleet and fog	0	0.0%	0	0.0%	0	0.0%
Other	0	0.0%	0	0.0%	0	0.0%
Total	5	100.0%	8	100.0%	61	100.0%

Table 4.13. Trigger Types in Project 11

Trigger Type	Description
Longitudinal deceleration (LD)	Less than or equal to –0.65 *g*
Lateral acceleration	Greater than or equal to 0.75 *g* or less than or equal to –0.75 *g*
Forward time to collision (FTTC)	• FTTC less than or equal to 4 s with a deceleration less than or equal to –0.6 *g*. • FTTC less than or equal to 4 s paired with a deceleration less than or equal to –0.5 *g* and forward range of less than or equal to 100 ft.
Yaw rate	Vehicle swerves ±4 degrees/s to ±4 degrees/s within a window of 3.0 s
Longitudinal acceleration (LA)	Greater than or equal to 0.5 *g,* returning to 0.1 *g* within 0.2 s
Critical incident (CI) button	Boolean response
Speeding trigger	In excess of 70 mph but not traveling on an interstate

and time of day. For the straight road segment reduction, data reductionists recorded a set of variables for 22 previously selected straight road segments that were chosen based on the frequency of travel. The recorded variables included driver behavior and performance in an attempt to analyze driver engagement in secondary tasks and the driver's eye-scanning patterns. For the event reduction, potential events were initially identified by the kinematic trigger values, and trained reductionists reviewed the video data to confirm their validity. The details of the trigger values are described in Table 4.13.

Valid events were further classified by data reductionists into three categories—crash, near crash, and judgment error—based on their severity. A crash is defined as "any contact with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated. Includes other vehicles, roadside barriers, objects on or off the roadway, pedestrians, cyclists, or animals." A near crash is defined as "any circumstance that requires a rapid evasive maneuver by the subject vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. A rapid evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle capabilities."

Judgment error is defined as "any circumstance where the teen driver purposefully or inadvertently creates a safety-relevant situation due to either general inexperience or performance error. Examples include those events where the teen drivers engage in 'horseplay' or overreact to surrounding traffic" (*7*).

For crash and near crash events, a series of variables were coded during the data reduction process. The variables are broadly classified into several homogeneous groups as vehicle, event, environmental, driver's state, and second vehicle variables, as summarized in Table 4.14.

Figure 4.10 illustrates some sample screenshots for the NTNDS. The left screenshot is a snapshot from the continuous video data captured by four continuous camera views monitoring the driver's face and the driver side of the vehicle, the forward view, the rear view, and an over-the-shoulder view for the driver's hands and surrounding areas. In the right screenshot, the bottom photos show sample periodic still shots by two other cameras in the interior vehicle cabin, as well as the lap area of the rear passenger seat.

Data reduction for this study is ongoing, and not much can be presented in this report. Because this study is the most recent naturalistic study accomplished by VTTI, all the equipment involved is the most updated and highly accurate. The data reduction protocol is refined so that the threshold values selected are more reasonable, and data reductionists have acquired more experience from previous studies. The resulting data set will be beneficial in the next stage of research for this study.

Table 4.14. List of Variables in Reduced Data in Project 11

Classification	Examples
Vehicle variables	Vehicle ID, vehicle type, owned or shared, and VMT
Event variables	Nature of event or crash type, pre-event maneuver, precipitating factors, corrective action or evasive maneuver, contributing factors, types of inattention, driver impairment
Environmental variables	Weather, ambient lighting, road type, traffic density, relation to junction, surface condition, traffic flow
Driver's state	Hands on wheel, seat belt usage, fault assignment, eyeglance
Driver of vehicle 2	Vehicle 2 body style, maneuver, corrective action attempted

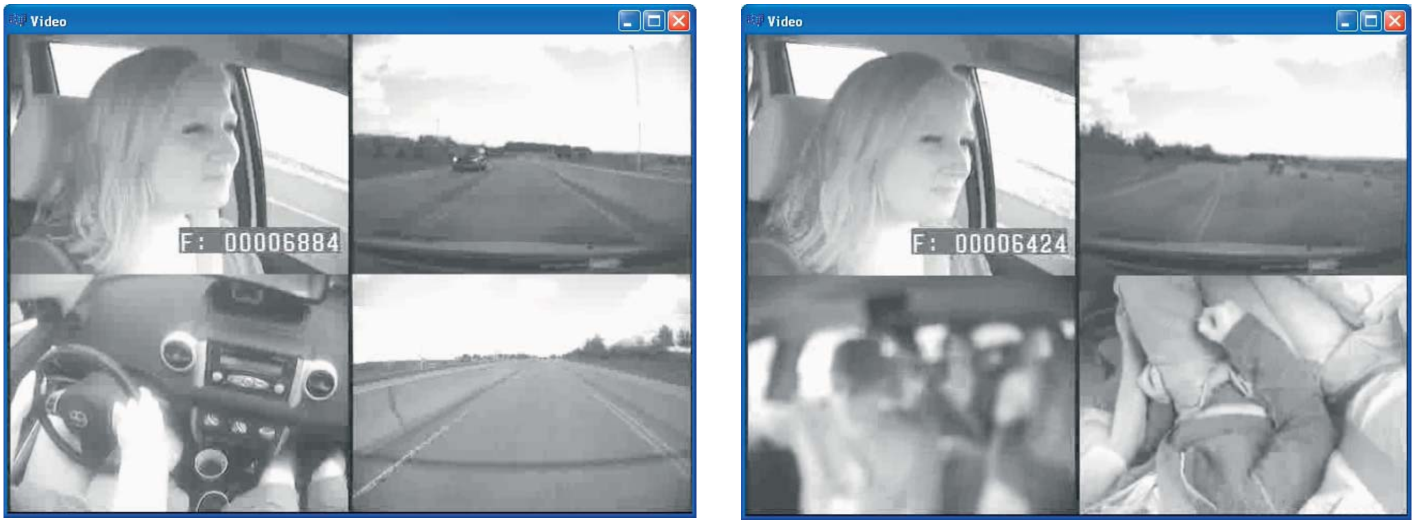


Figure 4.10. Screenshots of video data in Project 11.

References

1. University of Michigan Transportation Research Institute. *Automotive Collision Avoidance System Field Operational Test Report: Methodology and Results.* Report DOT HS 809 900. NHTSA, 2005.
2. University of Michigan Transportation Research Institute. *Road Departure Crash Warning System Field Operational Test: Methodology and Results.* NHTSA, 2006.
3. Dingus, T. A., S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, and R. R. Knipling. *The 100-Car Naturalistic Driving Study, Phase II: Results of the 100-Car Field Experiment.* Report DOT HS 810 593. NHTSA, 2006.
4. Neale, V. L., S. G. Klauer, R. R. Knipling, T. A. Dingus, G. T. Holbrook, and A. Petersen. *The 100-Car Naturalistic Driving Study, Phase 1: Experimental Design.* Report DOT HS 808 536. NHTSA, 2002.
5. Hanowski, R. J., M. Blanco, A. Nakata, J. S. Hickman, W. A. Schaudt, M. C. Fumero, R. L. Olson, J. Jermeland, M. Greening, G. T. Holbrook, R. R. Knipling, and P. Madison. *The Drowsy Driver Warning System Field Operational Test, Data Collection: Final Report.* Report DOT HS 811 035. NHTSA and Virginia Tech Transportation Institute, Blacksburg, Va., 2005.
6. Blanco, M., J. S. Hickman, R. L. Olson, J. L. Bocanegra, R. J. Hanowski, A. Nakata, M. Greening, P. Madison, G. T. Holbrook, and D. Bowman. *Investigating Critical Incidents, Driver Restart Period, Sleep Quantity, and Crash Countermeasures in Commercial Vehicle Operations Using Naturalistic Data Collection.* FMCSA, 2008.
7. Lerner, N., J. Jenness, J. Singer, S. G. Klauer, S. Lee, M. Donath, M. Manser, and M. Ward. *An Exploration of Vehicle-Based Monitoring of Novice Teen Drivers: Draft Report.* Virginia Tech Transportation Institute, Blacksburg, Va., 2008.

CHAPTER 5

General Guidelines for Video Data Analysis

Because all the data sets discussed in this report used video cameras to record driver behavior, certain rules need to be followed. All the naturalistic data sets identified in this project are protected by the VT IRB, a committee that approves, monitors, and reviews research involving humans with the aim of protecting the rights and welfare of the study participants. These approvals have been granted to the team members. In video-based data sets, IRB protocols are restrictive regarding who may access the data because of the inherent personal information (e.g., faces of drivers or location of residences). Access is often limited to key personnel associated with the original data collection project, making it difficult for an outside entity to gain access to an institution's data.

It is straightforward for the VTTI team to access and use in-house data, as well as outside data that have IRB approval, but special issues need to be addressed when conducting research using video data because human beings are involved. In case any institute performs a similar task using raw video data and other raw kinematic vehicle data, the following procedures are necessary.

Assuming the IRB protocol permits access to a contractor, the contractor must be capable of complying with IRB requirements. At a minimum, IRB approval from the entity maintaining the data set is required; approval from the contractor's IRB may also be necessary. In some instances, the original IRB agreement will not permit sharing of any personally identifiable data. In these cases, the contractor must perform the analyses without the personally identifiable data (e.g., without video). Once IRB approval is granted, the institution performing analysis on video-based data sets must possess the capability to work efficiently with data sets from outside sources. The contracting institution must have a secure computing facility dedicated to the storage, retrieval, processing, and analysis of data. Permissions must be established so that the videos and any associated data files for the project may be accessible only to researchers and other personnel involved in the project.

Each data set to be analyzed for this project typically has two data types associated with it. When applicable, video files will exist in a certain format, such as MPEG2, Audio Video Interleaved (AVI), or MP4. Additionally, each video file will be associated with a parametric data file, such as a comma-separated values (CSV) file, SQL database, or binary data file. This parametric data file contains information collected from the vehicle sensors, such as velocity, acceleration, time of day, and radar data. The contracting institution must have the capability to access and manipulate the various data formats provided.

For researchers to determine congestion factors, several tools are required to view and examine a video-based data set. The institution performing the analysis must have access to and expertise with the tools necessary to perform manipulation of parametric data. Commercially available tools, such as MatLab and SAS, are commonly used to transform data, identify events of interest, and perform statistical analyses of large data sets. Once events of interest are identified (e.g., crashes and near crashes), the data reduction process begins.

Data reduction is the process of manually viewing video and parametric data to validate events and derive additional measures. This process requires an institution that has facilities and tools for viewing synchronized video and parametric data, as well as a tool that links parametric data to video data and provides an interface for entering additional data into the data set. Because no tool designed specifically for this purpose is commercially available, the contracting institution will need to have the capability to develop a custom solution. Some research groups (e.g., VTTI) have developed software packages for data reduction that could easily be modified to meet the needs of this project.

Finally, a team of data reductionists will need to be trained on proper treatment of video data involving human subjects and the software used to reduce and analyze data. Reductionists should be educated on the goals of the project, the nature of the events, and the protocol for coding variables in a robust manner. They work independently to interpret events and record additional variables into the database. Examples of variables obtained through data reduction include weather, traffic conditions, and roadway type. Input from reductionists should be monitored and supervised by a reduction manager, and measures of inter-rater and intra-rater reliability should be used to ensure that quality control is maintained. The contracting institution should have experience with data reduction, as well as facilities within which data reduction can be performed.

The reduced data provide a rich source of information to relate driver behavior to surrounding roadway, weather, and traffic conditions. Driver status information (e.g., secondary tasks, driver distraction, driver apparent impairment, and driving habits) is available from the video data. Vehicle status information (e.g., vehicle speed, acceleration and deceleration, and lane orientation) is available as part of the in-vehicle kinematic data. A proper linkage of driver and vehicle variables to external variables (traffic, weather, work zone, and accident data) through location and time information will enable an in-depth study of the relationship between driver behavior and travel time reliability.

General Guidelines for Video Data Reduction

According to the general guidelines discussed earlier, the most challenging effort related to data reduction involves video data. Having conducted multiple naturalistic driving studies involving video data, VTTI has developed a mature procedure to organize, reduce, and analyze large-scale naturalistic driving data. According to different research objectives, the threshold values selected may vary from one data reduction to another, but the basic principle is consistent and the procedure is uniform. The step-by-step description of video data reduction in each project has been detailed individually in Chapter 4. Following is a summary of the general guidelines for video data reduction.

There are three primary steps in video data reduction. The first is to scan the raw data to identify potential safety events. VTTI developed its own software (DART) that reads and calculates numerical data to detect if events of interest have occurred. Therefore, the first step in data reduction is to run the event trigger program. The variables examined include velocity, acceleration (longitudinal or lateral), swerve, TTC (calculated from range and range rate between the subject vehicle and other vehicles detected by radar), and yaw rate. The software reads through the raw numeric data and extracts an epoch of data when it detects a parameter exceeding the trigger threshold set by researchers in advance. Different lengths of the epoch and threshold values can be adopted according to different attributes of the subject vehicle (e.g., heavy commercial trucks or light cars) and study goals. Details of the trigger threshold values for VTTI studies are listed in Tables 4.3, 4.6, and 4.9 in Chapter 4.

The second step is to check the validity of the potential event epochs generated in step 1 through visual inspection. As described in Figure 4.7 (*1*), invalid events and valid events (including conflicts and nonconflicts) are differentiated. Invalid events are false alarms in which sensor readings are spurious because of a transient spike or other anomaly. Valid events are events in which recorded dynamic-motion values occur and are verifiable in the video and other sensor data from the event. Valid events are further parsed as conflicts and nonconflicts. Conflicts are occasions in which a crash, near crash, incident, or conflict happens. Nonconflicts are occasions in which the trigger values satisfy the thresholds but the driver makes a maneuver and is in complete control of the vehicle. Examples can be a hard brake without any obvious reasons or a high swerve value from a lane change. In most cases, nonconflicts reflect aggressive driving habits or styles. In step 2, valid events are labeled for the next step.

The final step is to apply the data dictionary to the validated events. In this step, a data dictionary created by the researcher in advance is used to extract useful information from the raw data. Data reductionists code results while they watch the video by choosing correct answers from pull-down menus for each research question in the data dictionary. Variables included in the data dictionary should reflect the interests of investigators and the research objectives. This is the step in which the identification of driver behavior that affects congestion is extracted. Reductionists can plot real-time charts of any selected variable in the DART—such as range, velocity, and acceleration—simultaneously with the play of the associated video clips. When studying the relationship between driver behavior and nonrecurring congestion, the LOSs and traffic density before and after the event, as well as the contributing factors, are important. The LOS or traffic density before and after can be used in a simulation environment or in statistical modeling (details of travel time modeling are included in the previously submitted Task 5 Report) so that the events that generate congestion can be identified. Contributing factors, especially those that are driver- and vehicle-related, are judged by reductionists as correctable and avoidable or otherwise. For example, data reductionists decide if a crash or near crash was caused by one of the following reasons: (1) the driver was conducting a secondary task, such as talking on a cell phone, dining, adjusting the radio, or talking to passengers; (2) the driver was inattentive to the road, either distracted by something inside or outside the vehicle or simply daydreaming; (3) the driver was driving under the influence (such as drugs, alcohol, or medicine); (4) the driver made a wrong decision; or (5) the driver has questionable driving habits or driving skills. Simultaneously, other possible

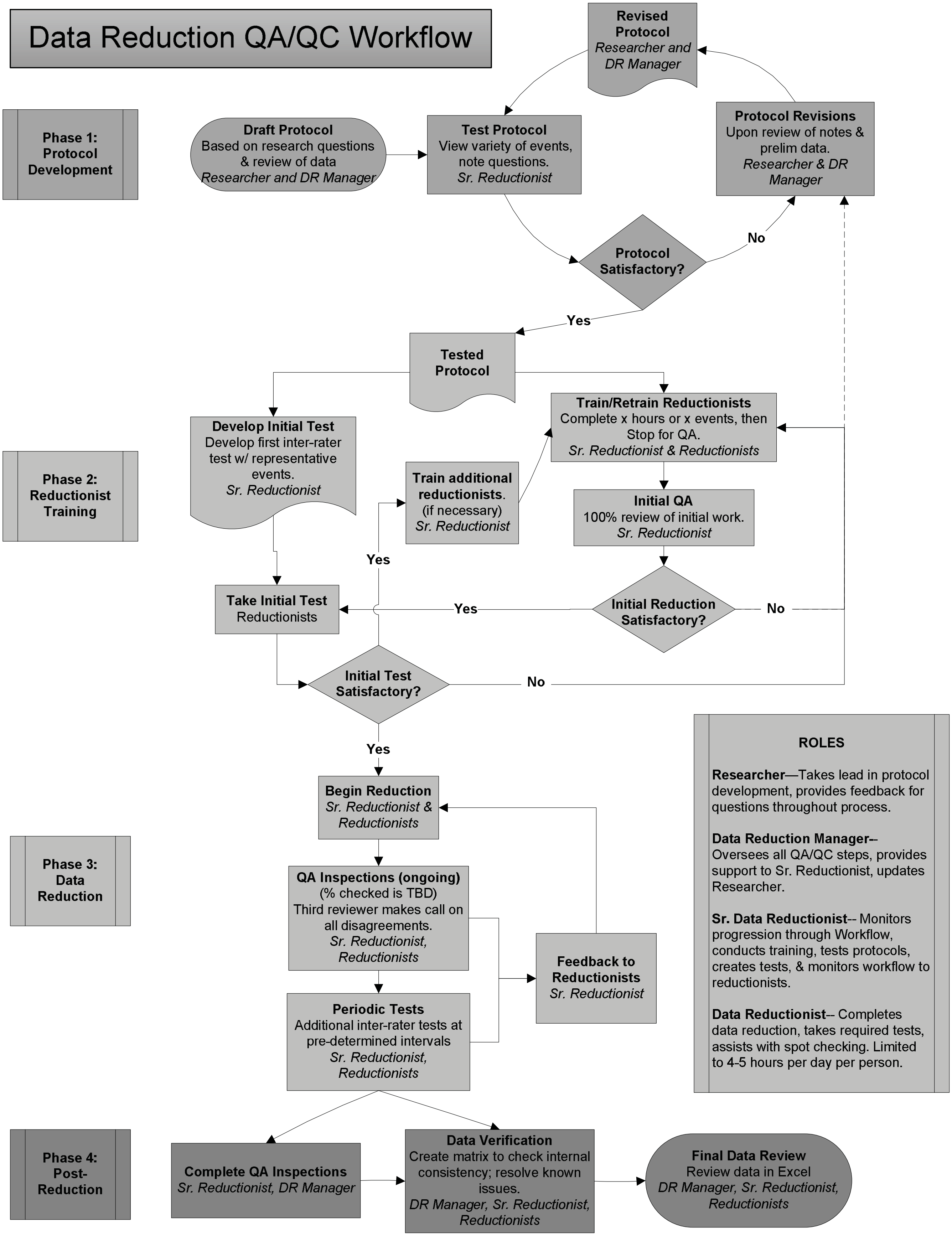


Figure 5.1. Quality assurance and quality control flowchart at VTTI.

contributing factors, such as weather, traffic flow, traffic density, traffic control, road profile, surface condition, and lighting condition, are identified. If available from the video clip, the status of other involved vehicles can also be examined and coded. A similar procedure is applied to reduce baseline epochs. The baseline epochs were used as a benchmark in which driver behavior can be compared with that in a safety-related event. According to the design of the researchers, a certain number of baseline epochs of selected length are decided. For example, the baseline epochs in Project 7 were 60 s long with a total of 1,072 epochs, whereas in Project 8 the baseline epoch was 30 s long and 456 were selected.

To ensure high-quality data reduction, a quality control procedure needs to be established. On the one hand, differences between data reductionists should be minimized. On the other hand, data reductionists should ensure that their judgments are consistent over time. Accordingly, inter-rater and intra-rater reliability are tested regularly. Before a data reduction effort starts, several test events are selected and coded by expert analysts (e.g., the principal investigator [PI] of the project or a data reduction director who has extensive experience with data reduction). Next, each reductionist is asked to code the same events. Their coding results are compared with those of the experts, and discrepancies are noted and discussed. The validation results help to determine (1) if reductionists are correctly coding the events; (2) if certain reductionists have systematic bias on certain questions; and (3) if the data dictionary is rationally composed and easily understood. Associated training, supervision, and modification of the data dictionary are applied if needed. For the intra-rater reliability test, one example event is selected for each category representing crash, near crash, proximity, and crash-relevant events. Data reductionists are required to code these events to the dictionary at the beginning of the process. They code the same event again after a period of time (varying from a week, a month, or a year, depending on the full length of the data reduction effort). The results are compared with the original reduction results. If there are differences, reductionists are notified to review their work and make necessary adjustments.

In summary, three-step data reduction is effective in processing large-scale video data and other numeric data. The selection of threshold values is always a compromise between exhaustiveness and false events. Lower trigger values will capture the maximum number of potential events, but the trade-off is a higher chance of false-positive events, nonconflict events, and less severe conflicts. Similarly, a higher trigger value will result in a higher percentage of valid events but will generate some omissions. A careful examination of threshold values by experts is highly recommended before data reduction starts.

For studies such as Project 2 and Project 5, which were aimed at testing onboard alarm systems, data reduction is relatively more straightforward. When an alert was triggered or the driver comment button was pushed, the main system logged the triggered summary file, and the video system was notified and captured a retrospective clip of video data with transition counts, histograms, errors, and other trip summary information recorded to a trip-summary log. Because data were already organized by trips and alarms during collection, the scanning step to identify potential events from numerical data is unnecessary. Data reductionists can start data reduction from viewing epochs of video data of events and then coding variables to the data dictionary. Data associated with the reduced data include demographic information (driver's age-group), lighting condition, road type, and traffic density. Additionally, numerous secondary behaviors (behaviors besides driving) are coded and include such actions as conversing, adjusting the radio, speaking on a cell phone, and brushing one's hair. The same modeling method for estimating travel time reliability and judging if driver behavior is correctable can be applied. Figure 5.1 summarizes a typical data reduction process (*2*).

References

1. Hanowski, R. J., M. Blanco, A. Nakata, J. S. Hickman, W. A. Schaudt, M. C. Fumero, R. L. Olson, J. Jermeland, M. Greening, G. T. Holbrook, R. R. Knipling, and P. Madison. *The Drowsy Driver Warning System Field Operational Test, Data Collection: Final Report.* Report DOT HS 811 035. NHTSA and Virginia Tech Transportation Institute, Blacksburg, Va., 2005.
2. Lerner, N., J. Jenness, J. Singer, S. G. Klauer, S. Lee, M. Donath, M. Manser, and N. Ward. *An Exploration of Vehicle-Based Monitoring of Novice Teen Drivers: Draft Report.* Virginia Tech Transportation Institute, Blacksburg, Va., 2008.

CHAPTER 6

Measuring Travel Time Reliability

Previous chapters discussed data sources; the next step is to characterize the impact of incidents on travel time reliability. This chapter introduces the development of travel time reliability performance measures. A set of specific, measurable, achievable, results-oriented, and timely (SMART) performance measures needs to be developed. The measures should provide valuable information for a traveler's decision making and facilitate the management of transportation systems.

Travel time reliability is a measure of the stability of travel time and is subject to fluctuations in flow and capacity. Typical quantifiable measures of travel time reliability that are widely used include the 90th or 95th percentile travel times (the amount of travel time required on the heaviest travel days), the buffer index (extra time required by motorists to reach their destination on time), the planning time index (total travel time that is planned, including the buffer time), and the frequency with which congestion exceeds some expected threshold. These measures are typically applied in cases of recurring congestion. The same measures can be applied in cases of nonrecurring congestion. For example, on an urban freeway, traffic backups may not be strictly caused by a recurring bottleneck (e.g., lane drop) but in many cases may be caused by nonrecurring bottlenecks (e.g., incidents).

Literature Review

Literature on modeling the relationship between crash reduction and travel time reliability improvement is scant. However, there are significant publications that deal with how to measure travel time reliability, as well as how to estimate travel time delays caused by incidents. This section provides an overview of these studies.

Research on Modeling Travel Time Reliability

Existing travel time reliability measures have been created by different agencies. Based on the ways the measurements were calculated, travel time reliability can be classified into empirical and practical-based or mathematical and theoretical-based (*1*).

Empirical and Practical Measures

FHWA recommended four travel time reliability measures: 90th or 95th percentile travel time, a buffer index (the buffer time that most travelers add to their average travel time as a percentage, calculated as the difference between the 95th percentile and average travel time divided by average travel time), the planning time index (calculated as the 95th percentile divided by the free-flow travel time), and the frequency that congestion exceeds some expected threshold (*2*).

In the FHWA report *Monitoring Urban Freeways in 2003: Current Conditions and Trends from Archived Operations Data* (*3*), several reliability measures were discussed (e.g., statistical range measures, buffer measures, and tardy-trip indicator measures). The recommended measures in this report include percentage of variation (the amount of variation is expressed in relation to the average travel time in a percentage measure), misery index (length of delay of only the worst trips), and buffer time index (the amount of extra time needed to be on time for 95% of the trips).

Florida's Mobility Performance Measures Program proposed using the Florida Reliability Method. This method was derived from Florida DOT's definition of reliability of a highway system as the percentage of travel on a corridor that takes no longer than the expected travel time (median travel time) plus a certain acceptable additional time (a percentage of the expected travel time) (*4*).

Monitoring and Predicting Freeway Travel Time Reliability Using Width and Skew of Day-to-Day Travel Time Distribution (*5*) examined travel time data from a 6.5-km eastbound carriageway of the A20 freeway in the Netherlands between 6 a.m. and 8 p.m. for the entire year of 2002. Data were binned into 15-min intervals. Two reliability metrics (skew and width of the travel time distribution) were created as reliability measures: $\lambda^{skew} = (T90–T50)/(T50–T10)$ and $\lambda^{var} = (T90–T10)/T50$, where *TXX* denotes the *XX* percentile values. Plots of these two measures showed an analogous behavior to the relationship between traffic stream flow and density, although the authors argued that the general pattern may change if the binning size is set larger (greater than 45 min). The preliminary conclusion is that for $\lambda^{skew} \approx 1$ and $\lambda^{var} \leq 0.1$ (free flow is expected most of the time), travel time is reliable; for $\lambda^{skew} << 1$ and $\lambda^{var} >> 0.1$ (congested), longer travel times can be expected in most cases; the larger the λ^{var}, the more unreliable travel times may be classified. For $\lambda^{skew} >> 1$ and $\lambda^{var} \geq 0.1$, congestion may set in or dissipate, meaning that both free-flow and high travel times can be expected. The larger the λ^{skew}, the more unreliable travel times may be classified. An indicator UI_r for unreliability, standardized by the length of a roadway, was proposed in this research using λ^{skew} and λ^{var}. A reliability map was created for each index: λ^{skew}, λ^{var}, UI_r, as well as a commonly used index $UI_r^{alternative}$, calculated as the standard deviation divided by mean. It was found that the UI_r takes both congestion and transient periods as periods with unreliable travel times, whereas λ^{skew} and λ^{var} each cover one situation and $UI_r^{alternative}$ shows much less detail. The author incorporated the indices to predict long-term travel times using a Bayesian regularization algorithm. The results showed a comparable travel time prediction to the observed travel times.

Mathematical and Theoretical Measures

The report *Using Real-Life Dual-Loop Detector Data to Develop New Methodology for Estimating Freeway Travel Time Reliability* (*6*) used real-life dual-loop detector data collected on I-4 in Orlando, Florida, on weekdays in October 2003 to fit four different distributions—lognormal, Weibull, normal, and exponential—for travel times for seven segments. Anderson-Darling goodness-of-fit statistic and error percentages were used to evaluate the distributions, and lognormal produced the best fit to the data. The developed lognormal distribution was used to estimate segment and corridor travel time reliabilities. The reliability in this paper is defined as follows: A roadway segment is considered 100% reliable if its travel time is less than or equal to the travel time at the posted speed limit. This definition is different from many existing travel time reliability definitions in that it puts more emphasis on the user's perspective. The results showed that the travel time reliability for segments was sensitive to the geographic locations where the congested segments have a higher variation in reliability.

New Methodology for Estimating Reliability in Transportation Networks (*7*) defined link travel time reliability as the probability that the expected travel time at degraded capacity is less than the free-flow link travel time plus an acceptable tolerance. The authors suggest that the reliability for a network is equal to the product of the reliabilities of its links. As a result, the reliability of a series system is always less than the reliability of the least reliable link. Therefore, the multipath network system should have a reliability calculated as

$$R_s = 1e\prod_{J=1}^{W}\left(1 = R_{path\,J}\right)$$

where *J* is the path label and *W* is the number of possible paths in the transportation network. The authors independently tested different degraded link capacities for a hypothetical network (with four nodes and five links) with capacity reduction at 0%, 10%, 20%, 30%, 40%, and 50%. It was found that a small reduction in link capacity would cause only a small or no variation in travel time reliability.

In *A Game Theory Approach to Measuring the Performance Reliability of Transport Networks* (*8*), the network reliability in two dimensions is defined as: connectivity and performance reliability. According to the author, the measurement of network reliability is a complex issue because it involves the infrastructure, as well as the behavioral response of users. He defines the reliability of a network as acceptable expected trip costs even for users who are extremely pessimistic. A game theoretical approach is described in this paper to assess network reliability. A sample network composed of nine nodes and 12 links with six possible paths from one origin to one destination was used. A nonparametric and noncooperative model was developed. The model assumes that network users do not know with certainty which of a set of possible link costs they will encounter, and the evil entity imposing link costs on the network users does not know which route the users would choose. The problem was formulated as a linear program with path choice probabilities as the primal variable and link-based scenario probabilities as the dual variables. Because this requires path enumeration, it is not feasible for large networks. Alternatively, a simple iterative scheme based on the method of successive averages (MSA) was proposed. The author believes that the expected trip cost for pessimistic trip makers offers a quality measure for network reliability.

Trip Travel-Time Reliability: Issues and Proposed Solutions (*9*) proposed five methods to estimate path travel time variances from its component segment travel time variances to estimate travel time reliability measures. To test these five methods and the assumption of travel time normality, field data collected on a section of I-35 running through San Antonio, Texas, were used. The field data included 4 months of automatic vehicle identification (AVI) data obtained from the TransGuide System at all 54 AVI stations from June 11, 1998, to December 6, 1998. Travel times for detected vehicles were estimated and invalid data were filtered out. Besides the

San Antonio network, a fictitious freeway network was created and modeled using INTEGRATION simulation software. The analysis found that, under steady-state conditions, a lognormal distribution describes the travel times better than other distributions. To evaluate the performance of the five proposed methods, AVI data over a 20-day period were analyzed in addition to the simulated data. The results were consistent for real-world and simulated data. The method that computes the trip coefficient of variation (CV) as the mean CV over all segment realizations (Method 3) outperformed other methods using the field data, whereas the method that estimates the median trip CV over all realizations *j* (Method 4) and the method that estimates the path CV as the midpoint between the maximum and minimum CV over all realizations *j* (Method 5) produced the best results. With the simulated data, Method 4 performed the best. When some localized bottlenecks were introduced to the system, Method 1 performed well and Methods 3 and 4 generated reasonable results.

Research on Estimating Delays from Incidents

Travel-Time Reliability as a Measure of Service (*10*) used travel time data and incident records collected from a 20-mi corridor along I-5 in Los Angeles, California, to describe the travel time variation caused by incidents. It was found that both the standard deviation and the median travel time were larger when there were incidents. The research suggested that the 90th percentile travel time is a meaningful way to combine the average travel time and its variability.

The I-880 Field Experiment Analysis of Incident Data (*11*) conducted the I-880 study to evaluate the effectiveness of the Freeway Service Patrol (FSP) implemented at a particular freeway section. The data sources were the field observations made by probe-vehicle drivers traveling the freeway section with an average headway of 7 min, as well as the incident data collected by the California Highway Patrol computer-aided dispatch system, officers' records, tow-truck company logs, and FSP records. In this field study, incident response, clearance times, and durations depend on the incident type and severity and the availability of incident management measures. Average response time decreased after the implementation of FSPs.

New Model for Predicting Freeway Incidents and Incident Delays (*12*) constructed a new model called IMPACT using incident data from six metropolitan areas to predict incidents and delays based on aggregate freeway segment characteristics and traffic volumes. There are four submodels in IMPACT: an incident rate submodel to estimate the annual number of peak and off-peak incidents by type; an incident severity submodel to estimate the likelihood that incidents block one or more lanes; an incident duration submodel to estimate how long it takes to clear the incident; and a delay submodel to estimate the delays caused by the incidents. Seven standard incident types were adopted and studied. The peak and off-peak incident rates for average annual daily traffic over capacity (AADT/C) ≤ 7 were similar across all incident types. Magnitudes of the peak period rates are sensitive to the degree of congestion. Some rates decline with increasing AADT/C, whereas others have a U-shaped relationship. Based on the findings of previous studies, IMPACT estimates the capacity lost because of incidents.

Estimating Magnitude and Duration of Incident Delays (*13*) developed two regression models for estimating freeway incident congestion and a third model for predicting incident duration. According to the authors, factors that affect the impact of nonrecurring congestion on freeway operation include incident duration, reduction in capacity, and demand rate. Two sets of data were collected in this study for 1 month before (from February 16, 1993, to March 19, 1993) and 1 month after (from September 27, 1993, to October 29, 1993) implementing FSP. The first set covered such incident characteristics as type, severity, vehicle involved, and location. The second set was related to such traffic characteristics as 30-s speed, flow, and occupancy at freeway mainline stations and at on-and-off ramp stations upstream and downstream of the incident location. Two models were developed to estimate incident delay. The first depicts incident delay as a function of incident duration, traffic demand, capacity reduction, and number of vehicles involved. The second predicts the cumulative incident delay as a function of incident duration, number of lanes affected, and number of vehicles involved. Model 1 outperformed model 2. The incident duration prediction model uses log transformation of duration. This model can predict 81% of incident durations in a natural log format as a function of six variables: number of lanes affected, number of vehicles involved, dummy variable for truck, dummy variable for time of day, log of police response time, and dummy variable for weather.

Quantifying Incident-Induced Travel Delays on Freeways Using Traffic Sensor Data (*14*) applied a modified deterministic queuing theory to estimate incident-induced delays using 1-min aggregated loop detector data. The delay was computed using a dynamic traffic volume–based background profile, which is considered a more accurate representation of prevailing traffic conditions. Using traffic counts collected by loop detectors upstream and downstream of the accident location, the research team developed a curve for arrival and departure rates for a specific location. The area between the two curves was used to compute the total system delay. To validate the algorithm, VISSIM software was used to construct some incident scenarios. Before conducting the simulation analysis, the model parameters were calibrated by matching in-field loop detector and simulated traffic counts. Data collected on SR-520 at the Evergreen Point Bridge were fed to

VISSIM software to simulate the incident-induced delay (IID). Even though most of the IIDs estimated by the algorithm were smaller than the IID obtained from the simulation models, they were reasonably close with an average difference of 15.3%. The proposed algorithm was applied to two sample corridors: the eastbound section of SR-520 (Evergreen Point Bridge) and the I-405 northbound section between mileposts 1.68 and 15.75. The results validated the algorithm, and the estimated delay was comparable to field data.

Proposed Modeling Methodology

Although the introduced existing measures attempt to quantify travel time reliability, they fail to distinguish between congested and noncongested conditions. Consequently, a more sophisticated model is needed to quantitatively measure travel time reliability and, at the same time, reflect the underlying traffic conditions that affect travel time reliability. In achieving these objectives, the team proposes the use of a novel multistate travel time reliability modeling framework to model travel times under complex traffic conditions (*15*). This chapter provides an overview of the proposed approach.

According to the model, traffic could be operating in either a congested state (caused by recurrent and nonrecurrent events) or an uncongested state. Travel time variability in a noncongested state is primarily determined by individual driver preferences and the speed limit of the roadway segment. Alternatively, travel time for the congested state (recurring or nonrecurring) is expected to be longer with larger variability compared with free-flow and uncongested states. The multistate model is used in this research to quantitatively assess the probability of traffic state and the corresponding travel time distribution characteristics within each state. A finite multistate model with K component distributions has the density function shown in Equation 3,

$$f(T|\lambda,\theta)=\sum_{k=1}^{K}\lambda_k f_k(T|\theta_k) \tag{3}$$

where T is the travel time; $f(T|\lambda,\theta)$ is the density function of the distribution for T, representing the distribution of travel time in the corresponding state; $\lambda = (\lambda_1, \lambda_2, \ldots \lambda_K)$ is a vector of mixture coefficients and $\sum_{k=1}^{K} \lambda_k = 1$; $\theta = (\theta_1, \ldots, \theta_K)$ is a matrix of model parameters for each component distribution; $\theta_k = (\theta_{k1} \ldots, \theta_{kI})$ is a vector of model parameters for the *kth* component distribution that determines the characteristics of the *kth* component distribution; and $f_k(T|\theta_k)$ is the density function for the *kth* component distribution corresponding to a specific traffic condition. Depending on the nature of the data, the component distributions $f_k(.)$ can be modeled using normal, lognormal, or Weibull distributions.

The distribution of travel time under a certain traffic condition corresponds to a specific component distribution in Equation 3. For instance, travel time in free-flow conditions can be reasonably assumed to be generated from a single-mode distribution. For a given period, multiple traffic conditions might exist, and the overall distribution of travel time for this period will follow a mixture distribution. The multistates could be a result of differing traffic conditions spatially (local bottlenecks at different sections), temporally (during the peak buildup or decay), or both. The multistate model has the advantage of better model-fitting in these multiple states and provides a novel approach for interpreting the results.

The *kth* component $f_k(T|\theta_k)$ in Equation 3 represents the distribution of travel time corresponding to a specific traffic condition. The parameter vector θ_k determines the characteristics of the *kth* component distribution. The parameter λ_k represents the probability of each state and has a significant implication in travel time reliability reporting, which is discussed later.

A specific example is a two-component normal distribution, as shown in Equation 4. With different combinations of mean and variance, as can be seen in Figure 6.1, the model can theoretically generate any form of distribution that fits any specific traffic conditions and travel time distributions.

$$\begin{aligned} f(T|\lambda,\mu_1,\mu_2,\sigma_1,\sigma_2) &= \lambda\frac{1}{\sqrt{2\pi}\sigma_1}e^{\frac{(T-\mu_1)^2}{2\sigma_1^2}} \\ &\quad + (1-\lambda)\frac{1}{\sqrt{2\pi}\sigma_2}e^{\frac{(T-\mu_2)^2}{2\sigma_2{}^2}} \end{aligned} \tag{4}$$

where λ is the mixture coefficient for the first component distribution, which is a normal distribution with mean μ_1 and standard deviation σ_1; the probability for the second component distribution is $1 - \lambda$, and the parameters for the second normal distribution are μ_2 and σ_2. Figure 6.1 (*15*) shows the density curves of a two-component normal mixture distribution. The parameters for the two-component distribution are $\mu_1 = 10$, $\sigma_1 = 5$, and $\mu_2 = 35$, $\sigma_2 = 10$, respectively. The plot shows the variation in the mixture distribution as a function of variations in λ. The model can accommodate multiple modes as commonly observed in travel time data. It is flexible enough to capture a wide range of patterns. In theory, the mixture distribution can approximate any density function. The mixture model is calibrated using the expectation and maximization (EM) method instead of maximum likelihood methods because the data have multiple modes.

To verify the multistate distribution of travel time proposed above, the team randomly examined data from 10 drivers in the 100-Car Study data set. The home and work addresses provided by drivers were geocoded to a geographic information system (GIS) road network database; all the trips made by that

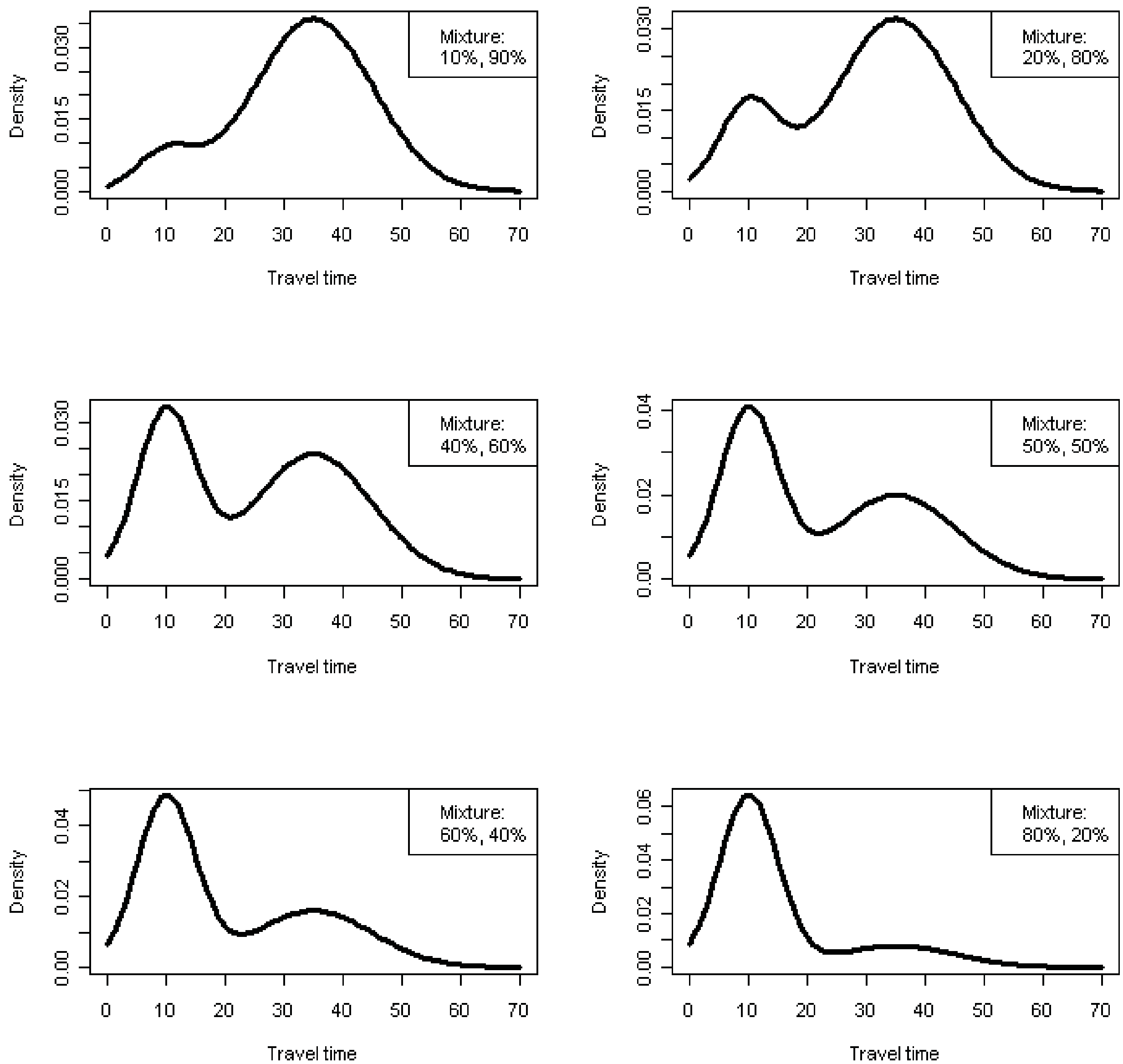


Figure 6.1. Mixture distribution and multimode travel time distribution.

driver were mapped to this database to visualize the trips. Figure 6.2 shows the home-to-work trips made by one participant. Travel times were then extracted from the relational database and plotted in histograms. As shown in Figure 6.3, the distribution of travel time for work-to-home trips by that participant is double-mode, which is in accordance with the assumption of the travel time model proposed by the team in the previous section. The start and end points of the trips have been eliminated from the picture to follow IRB rules.

Model Interpretation and Travel Time Reliability Reporting

The multistate model provides a platform to relate the parameters of the mixture distribution with the underlying traffic conditions. In particular, the mixture parameter λ_k in Equation 3 represents the probability that a particular travel time follows the *kth* component distribution, which corresponds to a particular traffic condition, as discussed earlier. This provides a mechanism for travel time reliability reporting. A novel two-step travel time reliability reporting method is thus proposed. The first step is to report the probability of each state as indicated by the mixture parameter λ_k. From a statistical standpoint, λ_k represents the mixture probability of each component distribution. The interpretation of this probability from the transportation perspective depends on the sampling mechanism. The sampling mechanism refers to how trips were selected for analysis. Two types of sampling schemes—proportional sampling and fixed-size sampling—could be used, as discussed in this section.

Figure 6.2. Home-to-work trip visualization.

The number of travel time observations for a given period depends on traffic conditions. Typically, the number of trips per unit time is larger for congested periods when compared with trips in a free-flow state. In a proportional sampling scheme, the number of trips is proportional to the number of trips for any given interval. For example, in a 10%, proportional sampling approach, 10 trips are selected from every 100 trips. For proportional sampling, the probability λ_k can be interpreted from both macro- and micro-level perspectives. From the macro-level perspective, this corresponds to the percentage of vehicles in traffic state k; for example, the percentage of drivers that experience congested traffic conditions. This interpretation can be used to quantitatively assess system performance from a traffic management perspective. The λ_k can also be interpreted from a micro-level perspective. Because the relative frequency (percentage) can also be interpreted as the probability for individuals, the probability λ_k also represents the probability that a particular traveler will travel in state k in a given period. This is most useful for personal trip prediction.

In a fixed-size sampling scheme, a fixed number of trips are sampled for a given period regardless of the total number of trips during this period. For example, 30 trips will be sampled every 10 min. The λ_k under a fixed sample scheme represents the proportion of the total duration where traffic is in the kth condition. For example, a λ value of 80% for the congested component implies that, out of 60 min, the traffic is in

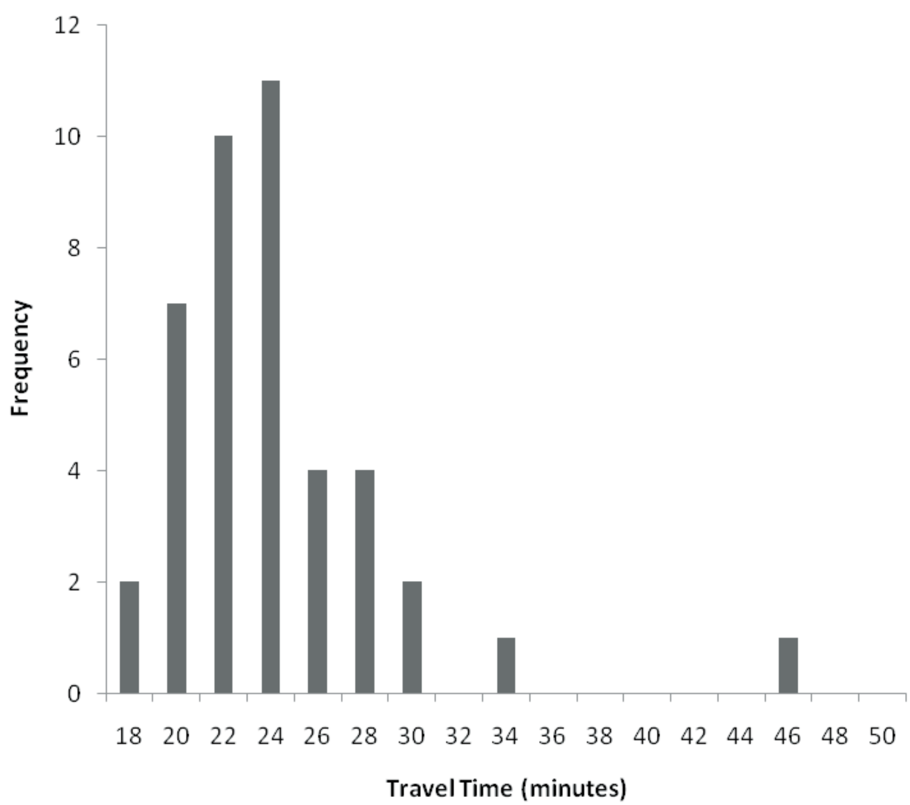


Figure 6.3. Home-to-work travel time histogram.

a congested state for a total of 0.8 × 60 min. = 48 min. The fixed-size sampling scheme also provides useful information for individual travelers, such as the proportion of time traffic will be in a congested condition.

The multistate model provides a convenient travel time reliability analog to the well-accepted weather forecasting example. The general population is familiar with the two-step weather forecasting approach (e.g., "the probability of rain tomorrow is 80%, with an expected precipitation of 2 in. per hour"). The same method can be used in travel reliability forecasting (e.g., "the probability of encountering congestion in the morning peak along a roadway segment is 67%, with an expected travel time of 30 min"). Travel time under each state can be reported by using well-accepted indices such as the percentile and the misery index, which can be readily calculated from each component distribution.

This two-step reporting scheme provides rich information for both travelers and traffic management agencies. By knowing the probability of a congested or incident state and the expected travel time in each state, an individual traveler can make better travel decisions. For instance, in the case of an important trip in which the traveler must arrive at his or her destination at a specific time, the traveler can make a decision based on the worst-case scenario and estimate the required starting time from that scenario. For a flexible trip, the traveler can pick a starting time with a lower probability of encountering a congested state. For traffic management agencies, the proportion of trips in a congested state and the travel time difference between the congested state and the free-flow state can provide critical information on the efficiency of the overall transportation system. This can also provide an opportunity to quantitatively evaluate the effects of congestion alleviation methods.

Model Testing

To demonstrate and validate the interpretation of the mixture multistate model, simulation was conducted using INTEGRATION software along a 16-mi expressway corridor (I-66) in northern Virginia. The validation was not conducted using the in-vehicle data for several reasons. First, the time stamps in the 100-Car data set were not GPS times and had errors up to several hours. Second, the in-vehicle data do not provide the ground truth conditions, given that the data are only available for the subject vehicle and not for all vehicles within the system. However, the simulation environment provides similar probe data with additional information on the performance of the entire system and detailed information on any incidents that are introduced.

An origin-destination (O-D) matrix was developed representing the number of trips traveled between each O-D pair using field loop detector traffic volume measurements. The goal of the simulation effort was twofold. The first objective was to demonstrate that the model parameters estimated are comparable to the characteristics of each traffic state. The second objective was to demonstrate model interpretation under two alternative sampling schemes: proportional sampling and fixed-size sampling.

Two O-D demands, a congested state and an uncongested state (scaling down from the congested matrix), were used to simulate temporal variations. Specifically, a database of 1,000 simulation runs, 500 of them with high demand and 500 with low demand, was constructed for various travel demand levels. The mixture scenarios were generated by sampling from the 1,000 time units of simulation output. The simulated travel times were mixed at fixed mixture levels of 10%, 25%, 50%, and 75% of time units in a congested stage. The mixed travel times were fitted to the two-state model, a mixture of two normal distributions. The fitting results demonstrated that the two-state model provides a better match to the simulated travel times when compared with a unimodal model, as shown in Figure 6.4 (*16*).

The results also showed that for proportional sampling in which the number of trips sampled in a given period is proportional to the total number of trips in that period, under high-congestion scenarios (75% of time units in congested state), the model underestimates the true proportion and overestimates the variance of the travel time in the free-flow state. The reason for the bias is that a single normal distribution cannot sufficiently model the travel time in the congested state when the percentage is high. This problem can be resolved by introducing a third component or by using alternative component distributions (e.g., lognormal or gamma). For fixed-size sampling in which a fixed number of trips is sampled for any given period, the model does reflect the characteristics of travel time under different traffic conditions. The parameters of the component distribution can be estimated satisfactorily, and the interpretation of the mixture parameters depends on the sampling scheme.

The multistate model was then applied to a data set collected from I-35 near San Antonio, Texas. The traffic volume near downtown San Antonio varied between 89,000 and 157,000 vehicles per day. The most heavily traveled sections were near the interchange with I-37, with average daily traffic counts between 141,000 and 169,000 vehicles, and between the southern and northern junctions with the Loop 410 freeway, with average daily traffic counts between 144,000 and 157,000. Vehicles were tagged with radio frequency (RF) sensors, and the travel time for each tagged vehicle was recorded whenever vehicles passed any pair of AVI stations. A two-component model and a three-component model were fitted to the morning peak travel time. The Akaike information criterion (AIC) was used to compare these models. The smaller the AIC value, the better fitted the model is. The results, shown in Table 6.1,

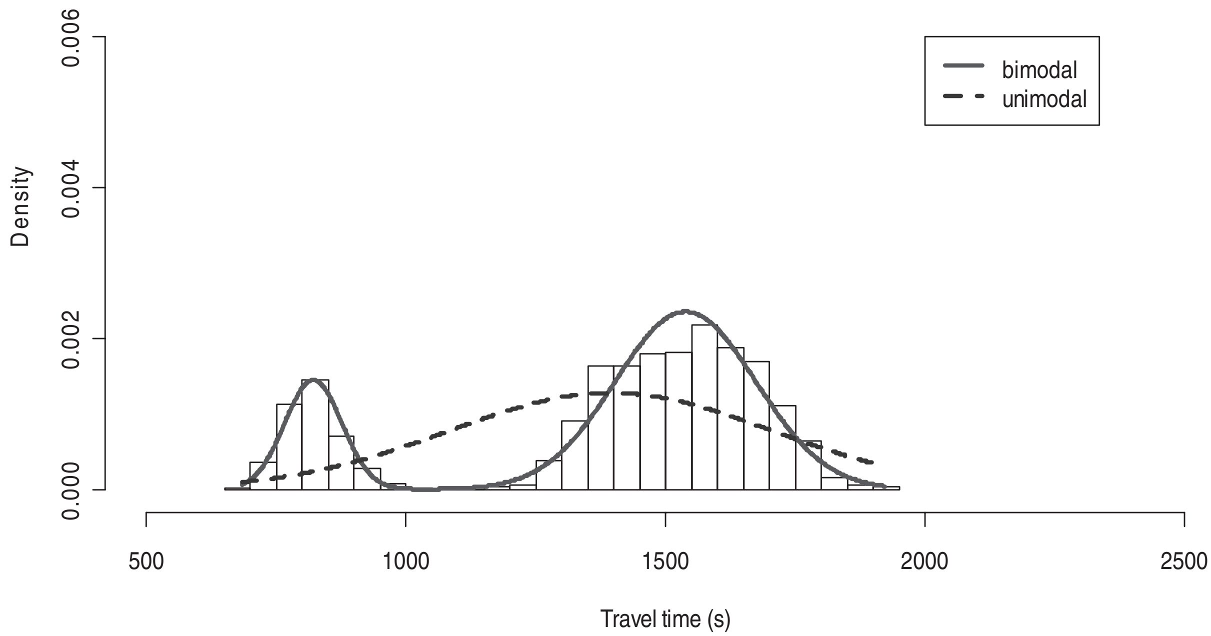


Figure 6.4. Comparison between unimodal and bimodal models.

demonstrate that the three-component model provides substantially better model-fitting than does the two-component model. The travel time reliability reporting listed in Table 6.1 clearly expresses the travel time information needed by travelers and decision makers. It not only reports the probability of encountering a congested state but also reports the expected travel time under that state.

The multistate travel time reliability model is more flexible and provides superior fitting to travel time data compared with traditional single-mode models. The model provides a direct connection between the model parameters and the underlying traffic conditions. It can also be directly linked to the probability of incidents and thus can capture the impact of nonrecurring congestion on travel time reliability.

Table 6.1. Mixture Normal Model-Fitting for Morning Peak Travel Time

	Two-Component Model			Three-Component Model		
	Mixture Proportion λ	Mean μ	Standard Deviation σ	Mixture Proportion λ	Mean μ	Standard Deviation σ
Comp. 1	33%	588	38	0.33	588	38
Comp. 2	67%	1089	393	0.59	981	230
Comp. 3	NA	NA	NA	0.08	1958	223
Log likelihood	–3567			–3503		
AIC	7144			7020		
Travel time reliability reporting	1. The probability of encountering congestion is 67%. If congestion is encountered, there is a 90% probability that travel time is less than 1,592 s. 2. The probability of encountering a free-flow state is 33%. If congestion is encountered, there is a 90% probability that travel time is less than 637 s.			1. There is a 59% probability of encountering congestion. If congestion is encountered, there is a 90% probability that travel time is less than 1,276 s. 2. There is a 33% probability of encountering free-flow conditions. In this case, there is a 90% probability that travel time is less than 637 s. 3. The probability of encountering an incident is 8%. In this case, there is a 90% probability that travel time is less than 2,244 s.		

Conclusions and Discussion

The proposed multistate model provides a better fit to field data compared with traditional unimodal travel time models. The distribution of field travel time is tested to be multimode. As demonstrated in the last row of Table 6.1, the reliability measures generated from the proposed model are specific and measurable. Specifically, the two model parameters—the probability of encountering congestion and the probability of an expected travel time—are both specific and measurable. The proposed travel time reliability reporting is achievable because the model can be developed using in-vehicle, loop detector, video, or other surveillance technologies. Running this model is not time-consuming, so it can provide timely information. Consequently, the proposed model provides valuable information to assist travelers in their decision-making process and facilitates the management of transportation systems.

Travel time reliability can be enhanced by modifying driver behavior to reduce incidents. The proposed model is designed to model travel time reliability and congestion before and after incident-induced congestion. The events have been viewed by data reductionists and designated as correctable or preventable by modifying driver behavior. Ideally, the data will also incorporate a data set with sufficient peak hour travel time data with and without the influence of safety-related events. It is relatively easy to capture correctable driver behavior with the aid of the data reduction tool developed by VTTI. The challenge is to collect travel time data before and after these events. The original plan was to use the in-vehicle naturalistic data collected in the candidate studies by VTTI and related external data using in-vehicle time and location information, but using such data is much more complicated and infeasible in most cases.

To develop travel time reliability models, a large data set was required, preferably with numerous trips sharing common origin and destination. The team planned to extract work-to-home trips from the 100-Car data set but realized that this plan had to be abandoned. The 100-Car Study, like most naturalistic driving studies, provided incentives to participants. As stated in the 100-Car Study report,

> One hundred drivers who commuted into or out of the Northern Virginia/Washington, DC metropolitan area were initially recruited as primary drivers to have their vehicles instrumented or receive a leased vehicle for this study. Drivers were recruited by placing flyers on vehicles as well as by placing newspaper announcements in the classified section. Drivers who had their private vehicles instrumented received $125 per month and a bonus at the end of the study for completing necessary paperwork. Drivers who received a leased vehicle received free use of the vehicle, including standard maintenance, and the same bonus at the end of the study for completing necessary paperwork. Drivers of leased vehicles were insured under the Commonwealth of Virginia policy. (*17*)

Because participants could receive monetary compensation or free use of leased vehicles, they were, to some extent, self-selected. Data reduction revealed that a relatively large portion of the subjects were students, hourly workers, or from other low-income populations. Consequently, relatively fewer home-to-work or work-to-home trips were collected, resulting in a limited size of trips at regular peak hours collected.

Additionally, there are other limitations. For example, the instrumented cars were sometimes driven by other drivers instead of the participant who signed up for the data collection. Consequently, the trips collected by GPS reflect multiple drivers' travel patterns. Instead of making regular trips sharing starting and ending points, some data sets illustrated a rather complicated travel pattern in which the trips recorded are relatively scattered, covering an expanded road network.

Another limitation is that, in the 100-Car Study, the computer time was used instead of a synchronized time, which resulted in some errors in time stamp. Consequently, even though the team does have a high-quality travel time database collected and maintained by the state of Virginia, it is hard to link the in-vehicle data with such external travel time data.

The statistical model proposed in this chapter used other travel time data rather than candidate data sources because of some limitations of those data. If future data collection is carefully designed with the recommendations the team proposes (as discussed in Chapter 8 of this report), the data will be significantly improved to serve this research goal.

References

1. Chalumuri, R. S., T. Kitazawa, J. Tanabe, Y. Suga, and Y. Asakura. Examining Travel Time Reliability on Han-Shin Expressway Network. *Eastern Asia Society for Transportation Studies,* Vol. 7, 2007, pp. 2274–2288.
2. Office of Operations, Federal Highway Administration. Travel Time Reliability: Making It There On Time, All The Time. http://ops.fhwa.dot.gov/publications/tt_reliability/brochure/index.htm. Accessed May 17, 2011.
3. Lomax, T., D. Schrank, S. Turner, and R. Margiotta. *Selecting Travel Time Reliability Features.* Texas Transportation Institute and Cambridge Systematics, Inc., May 2003. http://tti.tamu.edu/documents/474360-1.pdf. Accessed May 17, 2011.
4. Florida Department of Transportation. Florida Mobility Measures. www.dot.state.fl.us/planning/statistics/mobilitymeasures. Accessed May 17, 2011.
5. Van Lint, J. W. C., and H. J. Van Zuylen. Monitoring and Predicting Freeway Travel Time Reliability: Using Width and Skew of Day-to-Day Travel Time Distribution. *Transportation Research Record: Journal of the Transportation Research Board, No. 1917,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 54–62. http://trb.metapress.com/content/n76607qk003v26l1/fulltext.pdf. Accessed May 17, 2011.
6. Emam, E. B., and H. Al Deek. Using Real-Life Dual-Loop Detector Data to Develop New Methodology for Estimating Freeway Travel Time Reliability. *Transportation Research Record: Journal of the Trans-*

portation Research Board, No. 1959, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 140–150. http://trb.metapress.com/content/m0lv677211m12710/fulltext.pdf. Accessed May 17, 2011.

7. Al-Deek, H., and E. B. Emam. New Methodology for Estimating Reliability in Transportation Networks with Degraded Link Capacities. *Journal of Intelligent Transportation Systems,* Vol. 10, No. 3, 2006, pp. 117–129.
8. Bell, M. G. H. A Game Theory Approach to Measuring the Performance Reliability of Transport Networks. *Transportation Research Part B,* Vol. 34, 2000, pp. 533–545.
9. Rakha, H., I. El Shawarby, and M. Arafeh. Trip Travel-Time Reliability: Issues and Proposed Solutions. *Journal of Intelligent Transportation Systems,* Vol. 14, No. 4, 2010, pp. 232–250.
10. Chen, C., A. Skabardonis, and P. Varaiya. Travel-Time Reliability as a Measure of Service. *Transportation Research Record: Journal of the Transportation Research Board, No. 1855,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 74–79. http://trb.metapress.com/content/74t6691220058954/fulltext.pdf. Accessed May 17, 2011.
11. Skabardonis, A., K. Petty, R. Bertini, P. Varaiya, H. Noeimi, and D. Rydzewski. I-880 Field Experiment: Analysis of Incident Data. *Transportation Research Record 1603,* TRB, National Research Council, Washington, D.C., 1997, pp. 72–79. http://trb.metapress.com/content/0356m4h4v853681q/fulltext.pdf. Accessed May 17, 2011.
12. Sullivan, E. New Model for Predicting Freeway Incidents and Incident Delays. *Journal of Transportation Engineering,* Vol. 123, No. 4, 1997, pp. 267–275.
13. Garib, A., A. E. Radwan, and H. Al Deek. Estimating Magnitude and Duration of Incident Delays. *Journal of Transportation Engineering,* Vol. 123, No. 6, 1997, pp. 459–466.
14. Wang, Y., M. Hallenbeck, and P. Cheevarunothai. *Quantifying Incident-Induced Travel Delays on Freeways Using Traffic Sensor Data.* Washington Department of Transportation, 2008.
15. Guo, F., H. Rakha, and S. Park. Multistate Model for Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board, No. 2188*, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 46–54. http://trb.metapress.com/content/87m631594x815745/fulltext.pdf. Accessed May 17, 2011.
16. Park, S., H. Rakha, and F. Guo. Calibration Issues for Multistate Model of Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board, No. 2188*, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 74–84. http://trb.metapress.com/content/a6447w6uv7684011/fulltext.pdf. Accessed May 17, 2011.
17. Dingus, T. A., S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, and R. R. Knipling. *The 100-Car Naturalistic Driving Study, Phase II: Results of the 100-Car Field Experiment.* Report DOT HS 810 593. NHTSA, 2006.

CHAPTER 7

Potential Problems and Issues in Data Reduction

Because of the limitation of hardware or software design, data elements in candidate data sets may or may not be sufficiently accurate and verifiable for the analysis of driver behavior associated with crashes and near crashes and travel time reliability. The derivation of data elements from raw data sets may not meet the needs of the study purpose and may require modifications. In some cases, the data element may not be accurate but can be easily transformed. An example of such a case is when an accelerometer box is installed backward in a vehicle. In this case, the data can be quickly converted through a simple mathematical procedure without compromising data integrity. In other cases, the method to collect the data may be inaccurate for a particular purpose (e.g., GPS-calculated vehicle speed compared with speed measured directly from the CAN bus). Similarly, it may be that some portions of the data are useful but others are not. Data need to be reviewed to determine the suitability for nonrecurrent congestion research purposes. Potential problems, risks, and limitations for data collection and reduction are discussed in this section.

Overall Data Collection

A common problem associated with naturalistic studies is the proper identification of drivers. Normally, the data collection process tries to ensure the exclusive use of an equipped vehicle by the assigned driver whose demographic information has been recorded. It is not unusual that the equipped vehicle would be used by family members or friends. The different driving habits and behaviors of a second driver and the mismatched demographic information can bias the data. An elaborate scheme to correctly match driver information with the actual driver can improve the situation.

Multiple pieces of equipment are on board the test vehicles. Different devices can be chosen for different research purposes. The DAS adopted in Project 2 and Project 5 is shown in Figure 7.1 (*1*). The basic arrangement of the DAS in Project 7 by VTTI is illustrated in Figure 7.2 (*2*). A similar setup was used in the other three VTTI projects. These differences in equipment result in variations in types of data collected and the associated data storage and computation requirements. Customized data computation software and hardware are needed for individual studies. UMTRI developed a system called Debrief View, as shown in Figure 4.4, to conduct data reduction. VTTI developed DART software, which visualizes and reduces data, as shown in Figure 4.3.

Even with similar types of equipment, different settings might apply according to the varied research purposes. For example, although all the candidate studies recorded video data at a predefined stable frequency (i.e., continuous), the frequency set by each study was not the same. Project 2 and Project 5 saved data at a relatively low frequency (as shown in Table 4.2) unless an event trigger was sent to the DAS to initialize a continuous video recording for 8 s. The fact that the purpose of these studies was to test the effectiveness of an ACA-RDCWS warrants this lower frequency. The disadvantage, as stated in the UMTRI reports, was that several alerts of very short duration that occurred below the 10-Hz recording rate may be omitted (*1*). On the contrary, most VTTI studies collected continuous video data at approximately 30 Hz. Details of drivers' behavior are essential to the studies conducted by VTTI. A relatively higher video recording frequency will provide researchers with a better opportunity to closely observe drivers' driving habits and distractions. A higher frequency generates data sets that are larger in size and brings more challenge to data reduction. When postprocessing and analyzing the same type of data from different sources, as was done for this study, special attention should be paid to the differences in data collection rates. Conversion or inference of data may be necessary.

Another common problem is data dropping out; for example, GPS outages because of urban canyons. As shown in Figure 7.3, when high buildings in downtown areas block out satellite signals, the resulting GPS data collected has gaps between the points. Postprocessing such GPS data to trace the real route traveled by the vehicle usually leads to an error, as

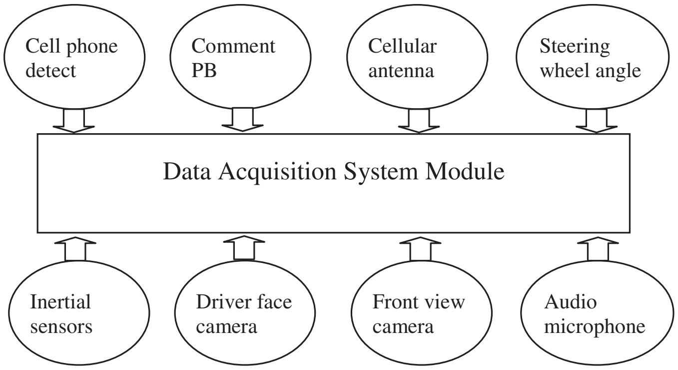


Figure 7.1. DAS used in Project 2 and Project 5.

shown in the left part of Figure 7.4. By filling the gaps, the real routes can be accurately located, as shown in the right part of Figure 7.4. Another example of data dropping out is illustrated in Figure 7.5, which depicts a situation in which data (speed and range) are missing for a few milliseconds while the valid target is being tracked. To impute missing data, a linear interpolation may be performed. Figure 7.5 shows a speed profile with missing values for a target (upper part) and the same profile after performing a linear interpolation (lower part of the figure).

Invalid readings and null values are other commonly occurring obstacles. For example, in examining the rain gauge data, invalid (unrealistic) measurements were found, as shown in Table 7.1 (196.088 mm). A rain gauge measurements algorithm was developed to identify such unusual rainfall values and replace them with the average of the before-and-after values in the case of the first problem. Additionally, Table 7.2 shows another example of unrealistic data in which the GPS time reports a null value for an entry. Algorithms were developed to replace invalid or missing data, as demonstrated in the right column of Table 7.2.

Details of challenges and potential problems regarding kinematic data, video data, reduced data, and other data sources are discussed in the following sections.

Kinematic Data

Sensor measurements in each study to collect kinematic data differ from one to another. Project 7 and Project 8 used a lane tracker developed by VTTI (i.e., Road Scout) to detect the vehicle's location within a lane. The maximum error was 6 in. for distance measuring and 1° for angular measuring. Similar equipment was used in Project 5 with a different accuracy

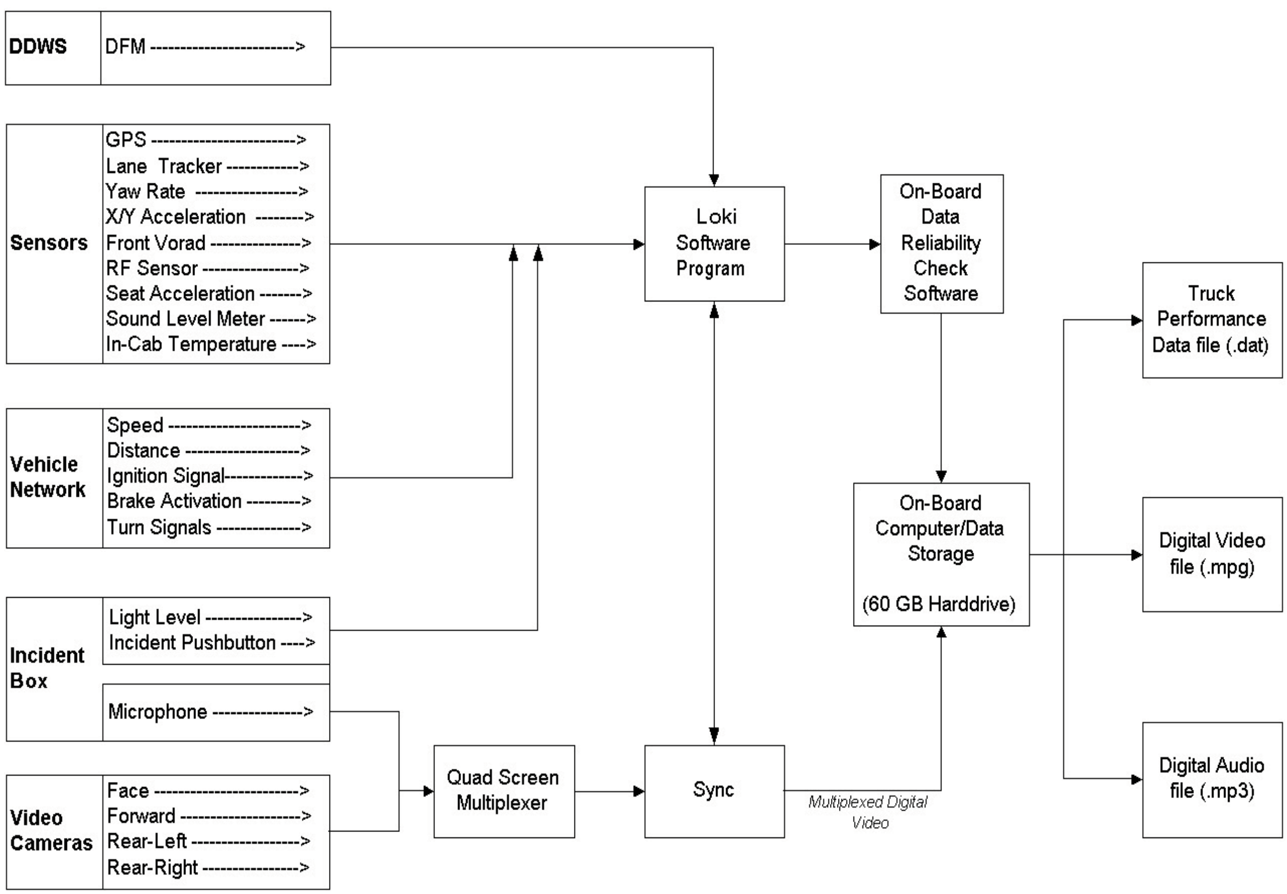


Figure 7.2. DAS used in Project 7.

Figure 7.3. Urban canyon blockage of GPS signals.

level. It used a monochrome charge-coupled device (CCD) camera that observed painted lane markers or other nonpainted visual features and could observe lane markers forward of the vehicle to approximately 30 m.

The radar systems used to detect surrounding objects and collect distance information for further derivation of data also vary from one project to another in numbers and configurations. The variations in settings resulted in varied accuracy levels. In Project 2, radar units were operating at 77 GHz frequency to track up to 15 targets. The sensing range was 100 m. This range is reduced on a winding road because of the radar's limited azimuth coverage. Project 5 included two forward-looking radar units configured at 20 Hz, two side-looking radar units configured at 50 Hz, and fields of view 120° wide. Radar systems used in VTTI studies were operating at 20 Hz. Project 6 had a radar range effective from 5 to 500 ft. Project 7 and Project 8 used radar systems with effective ranges from 5 to 600 ft. Radar units used in Project 11 increased the effective range to 700 ft.

When reviewing and reducing radar data, one must consider the type of radar used, the rate of data collection, how "noisy" the data are, and the assumptions used to mathematically smooth the data. One typically derived variable was the TTC between the subject vehicle and a leading vehicle. In Project 2, enhanced TTC was computed incorporating the accelerations of both vehicles, as well as the range to the vehicle ahead and the speeds of the vehicles. In VTTI studies, the TTC was derived from the range measured by either a radar-based VORAD forward object detection unit or a TRW sensor. Acceleration and deceleration of related vehicles were taken into consideration in all studies except Project 11.

Radar switching targets is another difficulty. When radar signal strength changes, the system might lose track of a target and then reacquire it. When this happens, the target will be assigned a different ID number, which can cause confusion. The manner in which the distance between the target and the equipped vehicle is recorded and organized in the data set may also generate errors. The data collection system simultaneously tracks up to seven targets that change over time. As shown in Table 7.3, VTTI has developed algorithms that identify and track the primary target of interest. To illustrate the importance of implementing such algorithms, consider the primary target of interest to be Target 231, as shown in Table 7.3. Before implementing the tracking algorithm

Figure 7.4. GPS gaps: (left) before links added; (right) after links added.

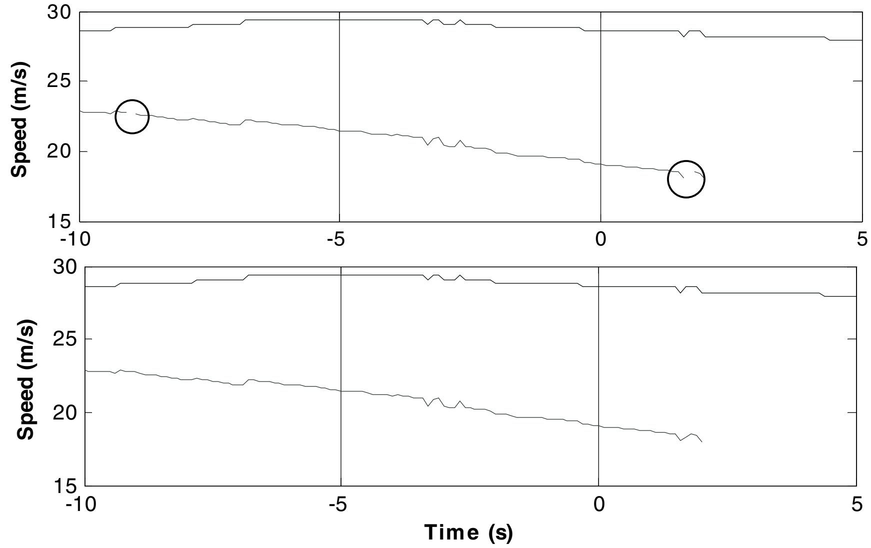


Figure 7.5. Speed profile: (top graph) with missing data; (bottom graph) after performing linear interpolation.

(i.e., using range values in the "VORAD1_Range_1" range calculations), erroneous variable computations would result, as shown in Figure 7.6. After applying the algorithm, correct target variables are identified.

A critical postprocessing data issue that one must address is the linking of the in-vehicle data with other data, including environmental, traffic, and weather data. Valid and accurate time and location variables should be available for researchers to complete the linkage. As described in earlier reports, most studies had GPS data recorded in their data sets. In Project 6, however, the local computer clocks without synchronization were used instead of GPS time, resulting in time errors. These errors deem synchronization infeasible.

Table 7.1. Rain Gauge Measurements

Rain Gauge Measurements (invalid values), *mm*	Rain Gauge Measurements (invalid values removed), *mm*
8.382	8.382
196.088	8.382
196.088	8.382
196.088	8.382
196.088	8.382
8.382	8.382
8.382	8.382

Table 7.2. GPS Imported Data (Null Values)

GPS Date Time (with null values)	GPS Date Time (corrected)
'2007-06-22 06:30:39.617'	'2007-06-22 06:30:39.617'
'2007-06-22 06:30:39.617'	'2007-06-22 06:30:39.617'
Null	'2007-06-22 06:30:39.617'
'2007-06-22 06:30:39.617'	'2007-06-22 06:30:39.617'
'2007-06-22 06:30:39.617'	'2007-06-22 06:30:39.617'
'2007-06-22 06:30:39.903'	'2007-06-22 06:30:39.903'

Video Data

During the extensive naturalistic driving data collection period, it was not unusual to have technical difficulties in cameras, including disconnection of cables and malfunction of cameras. Special supervision is needed to ensure smoother data collection and reduction.

Each study set up videos with varied views for its particular research goal. VTTI studies usually have four cameras capturing the front view, side views (left and right), and driver's

Table 7.3. Example Illustration of Radar Target Tracking

Time Step (s)	VORAD_ID			VORAD1_Range (Ft)		
	1	2	3	1	2	3
–0.6	231	248	247	50.1	370.1	212.2
–0.5	231	248	247	49.5	363.4	205.8
–0.4	248	231	247	354.5	49.6	193.1
–0.3	247	231	248	186.7	49.6	344.5
–0.2	247	248	231	173.9	331.8	49.5
–0.1	247	231	248	165	49.3	321.8

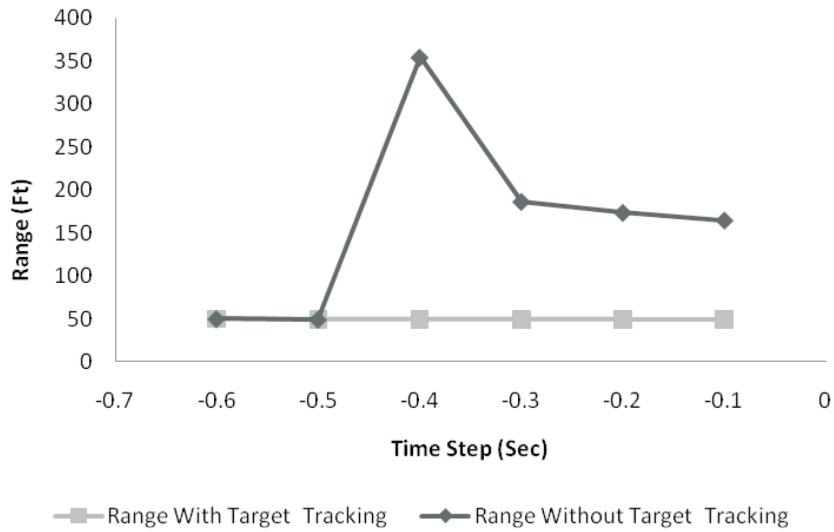


Figure 7.6. Range with and without target tracking.

face. Some projects added a fifth camera to capture the driver's hand and foot movements.

Figure 7.7 illustrates the camera views in Project 7 (*2*). In the UMTRI studies, only two cameras were used, providing a front view and a driver's face view, because less emphasis was put on observing driver behavior. Video data that does not include all four views have limited usage in this study, given that it is not possible to identify causes and behavior before crash or near-crash events.

In making decisions based on driver behavior, as is the case in this research effort, a prerequisite is satisfactory quality of video data. Incorrect brightness is a typical problem that prevents researchers from interpreting video data clearly.

Figure 7.8 shows some video captures in Project 5 (*3*). During some daytime driving, when the vehicle is heading directly into a setting sun, the camera images degraded because of sun glare. When some data from outside sources are unavailable, such as the weather data and level of congestion, the ability to acquire them through video data reduction becomes vital. If the video data are blurred, it is impossible to obtain such data. In the new DAS protocol of VTTI, the problem is solved by using Autobrite, as shown in Figure 7.9 (*4*). As can be seen, the quality of video data on the right of the figure is significantly improved.

Reduced Data

One challenge faced by researchers is data reduction in which raw data can be organized in a more functional format. Each study listed previously has a data reduction dictionary into which raw data were coded by reductionists, but the coding schemes of each dictionary are not identical. In Project 2, the

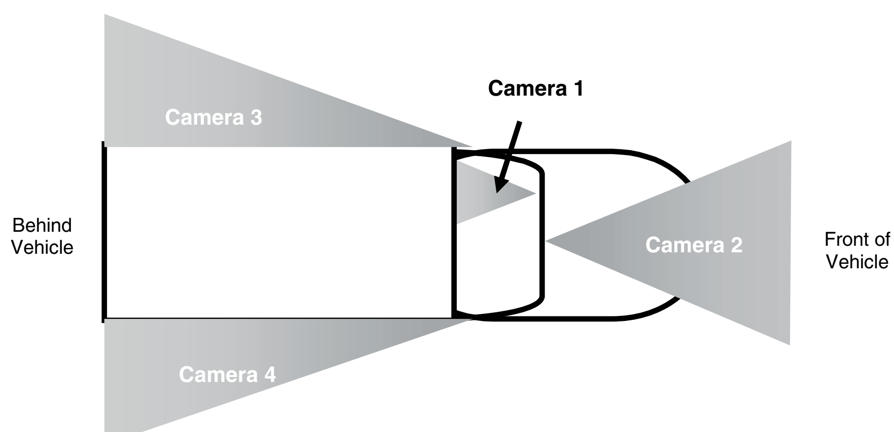


Figure 7.7. Camera directions and approximate fields of view in Project 7.

Figure 7.8. Simultaneous images from the cameras of an RDCWS vehicle heading into the sun.

variable "Time of Day" was coded in 0 or 1 for Day or Night, respectively. The "Location of eyes at time of the alert" was coded 0 through 9, representing "Looking forward at forward scene" at one extreme, to the more distracted statuses of "Head down, looking at center stack console area" or "Cannot accurately evaluate eye location." In contrast, data reduction at VTTI was more extensive. In a typical VTTI study (such as Project 7 and Project 8), "Date," "Day of Week," and "Time" were three independently coded variables to pinpoint the time of an event. "Light Condition" was a separate variable, in addition to time and day variables, coded from 01 to 05 to describe the light situation as "Daylight," "Dark," "Dark but lighted," "Dawn," or "Dusk." VTTI studies coded driver actions and distractions more elaborately. "Driver Potentially Distracting Driver Behavior" was a variable coded in 31 values describing situations including "bite nails," "remove/adjust jewelry," and even "comb/brush/fix hair." Besides behavior variables, some variables were designed to describe other statuses of drivers.

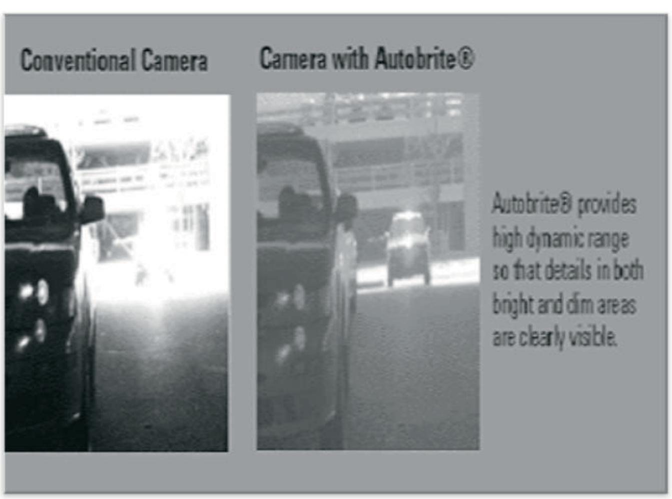


Figure 7.9. Prototype DAS camera under consideration.

For example, "Driver Actions/Factors/Behaviors Relating to Event" described drivers' emotions, coded in 60 values to represent "angry," "drowsy," "drunk," and others. When using data from multiple sources with different coding protocols, a proper unifying process is required to ensure the same standards in data analysis.

Reductionist capability is a dominant factor that affects the quality of reduced data. VTTI has a professional data reduction team composed of professional researchers and graduate research assistants and a separate data reduction laboratory led by laboratory managers. All the reductionists have extensive experience in data reduction and analysis. Before data reduction officially starts, reductionists are trained using a protocol written by the laboratory manager and project researchers. The laboratory manager works closely with the reductionists to assess their comprehension of the data reduction dictionary. UMTRI also has a professional data reduction team with researchers and graduate students. Students are responsible for relatively easy variable coding, such as weather condition, presence of passenger, and type of road. The research staff is responsible for more difficult variables that require judgments, such as degree of distraction and behavior coding.

Quality control in data reduction is critical in data post-processing and can be a decisive factor in the success of later analyses. A quality control procedure to support accurate and consistent coding was established at VTTI. For example, in Project 11 data reductionists performed 30 min of spot checks of their own or other reductionists' work each week. Besides the spot checks, inter- and intra-rater reliability tests were conducted every 3 months. Reliability tests were developed for which the reductionist was required to make validity judgments for 20 events. Three of the 20 events were also completely reduced; in other words, the reductionist recorded information for all reduction variables as opposed to simply marking the severity of the event. These three tests were repeated on the next test to obtain a measure of intra-rater reliability. At the same time, using the expert reductionist's evaluations of each epoch as a gold standard, the proportion of agreement between the expert and each rater was calculated for each test. This inter-rater test between expert and regular data reductionists was completed on the initial reduction for 6, 12, and 18 months of data reduction. The results indicated an intra-rater reliability of 99% for all three tests. The average inter-rater reliability score for this task was 92.1%. Discrepancies are mediated by a third, senior-level researcher (*2*). Similar quality control procedures were used by the UMTRI research team. Two researchers initially viewed and coded a small portion of the alerts independently. The coded data were compared to decide a percentage of agreement; each researcher then independently coded the remaining alerts. A third researcher examined their coding results and determined the degree of consistency in coding. The results showed a high level of agreement, which testified to the efficiency and consistency of the data reduction dictionary. The researchers then jointly viewed and recoded all video to modify factors that had not been agreed on. Each of these meticulous steps guarantees that the data reduction is under control.

Other Data Sources

Besides the possible risks and problems that exist in vehicle data, the availability and quality of environmental data—specifically, weather data, traffic count, crash, and work zone data—are worthy of attention. Accurate weather data are available at ASOS stations. Only vehicle data that were collected at locations close enough (e.g., within 15 mi) to ASOS stations can be associated with the weather data observed there. Another source of weather data is the Road and Weather Information System (RWIS). RWIS is a combination of technologies that uses historic and current climatological data to develop road and weather information. The information is then sent to road users and decision makers. RWIS usually includes an environmental sensor system, a model to develop forecasts, and a dissemination platform to publish data. More than 10 DOTs in the United States have started or plan to start RWISs. When vehicles were far from ASOS stations or RWIS, the only source of weather information was the weather variable coded in the reduced data by data reductionists.

Even at locations where weather data are available, risks that the data have errors in them still exist. In the example shown in Table 7.4, the variable Rain_Today Field, which records the cumulative rainfall since 12:00 a.m., is reset to zero. The controller times do not necessarily coincide with the local time, and thus resetting has to be done at 13:29, according to local time (shown as 17:29 as of Coordinated Universal Time [UTC]). To address this problem it was determined that the offset was a constant value that was location specific. An offset was allocated to each location in computing the precipitation rate to account for this error.

Table 7.4. Weather Station Input Data

Row Number	GPS	Minute of Day	Rain_ Today
57809	'2007-04-26 17:29:38.367'	1439	8.382
57810	'2007-04-26 17:29:38.367'	1439	8.382
57811	'2007-04-26 17:29:38.367'	0	0
57812	'2007-04-26 17:29:38.367'	0	0

For traffic count and traffic condition, crash, and work zone data, the quality and availability differ from state to state. As introduced in Chapter 6, some states have more continuous traffic count stations than others. For example, Virginia has maintained a traffic count database containing traffic data from more than 72,000 stations, among which 470 are continuous. West Virginia has only 60 continuous-count stations. Some states have more complete crash or work zone databases (e.g., New Jersey and Michigan) and others maintain only general crash records from police reports. When linking vehicle data to outside data sources, special attention should be paid to variations in data quality.

In summary, data elements designed to accomplish the objectives of the original research study may not be suitable for a study of driver behavior that causes nonrecurring congestion. As discussed in this chapter, if certain modifications can be feasibly executed, a candidate data set can be more valuable to serve the research goal of this study. Table 7.5 lists the possible modifications that can be made for each candidate data set to render them more compatible with this research goal.

As illustrated in the last row of Table 7.5, one more potential data set will be available in the near future. Being the most extensive naturalistic travel data collection effort, SHRP 2 Safety Project S07 will include high-quality video data and other in-vehicle data. With accurate location and time information, Project S07 data can be easily linked with other external environment data. Once integrated with other external data sets, the data set from Project S07 will be the most valuable candidate data set for studying nonrecurring congestion and its relationship to driver behavior. Details of Project S07 are discussed in Chapter 8 in the section on recommendations.

Table 7.5. Modification Needed for Each Data Set

	Feasibility (✓✓✓)	Modifications Needed	Cost for Modifications ($$$)
Project 2: ACAS FOT	✓	• Additional external data, such as work zone and crashes	$$$
		• Manual filtering of invalid triggers	
Project 5: RDCWS FOT	✓	• Additional external data, such as work zone and crashes	$$$
		• Manual filtering of invalid triggers	
Project 6: The 100-Car Study	✓✓	• More efficient identification of driver	$
Project 7: DDWS FOT	✓✓	• More comprehensive external data	$
Project 8: NTDS	✓✓	• More comprehensive external data	$
Project 11: NTNDS	✓✓	• More efficient identification of driver	
		• More comprehensive external data	$$
SHRP 2 Project S07	✓✓✓		$$$

References

1. University of Michigan Transportation Research Institute. *Automotive Collision Avoidance System Field Operational Test Report: Methodology and Results.* Report DOT HS 809 900. NHTSA, 2005.
2. Hickman, J. S., R. R. Knipling, R. L. Olson, M. C. Fumero, M. Blanco, and R. J. Hanowski. *Phase I—Preliminary Analysis of Data Collected in the Drowsy Driver Warning System Field Operational Test: Task 5, Preliminary Analysis of Drowsy Driver Warning System Field Operational Test Data.* NHTSA, 2005.
3. University of Michigan Transportation Research Institute. *Road Departure Crash Warning System Field Operational Test: Methodology and Results.* NHTSA, 2006.
4. Biding, T., and G. Lind. *Intelligent Speed Adaptation (ISA), Results of Large-Scale Trials in Borlänge, Lidköping, Lund and Umeå During the Period 1999–2002.* Swedish National Road Administration, 2002.

CHAPTER 8

Conclusions and Recommendations for Future Data Collection Efforts

The team reviewed reduced data from each candidate data set in Chapter 4. The original reduced data were not detailed or specific enough to recognize contributing factors to safety-related events. Consequently, additional data reduction was needed. The next section details the factors that contribute to crashes and near crashes, whether or not driver behavior can be adjusted, and corresponding countermeasures are recommended.

Contributing Factors and Correctable Driver Behaviors

For Project 5, RDCWS FOT, the original data reduction did not specifically pinpoint contributing factors to events. The variables coded for drivers and the related environment stated the situation only at the instant the events occurred. The team performed additional data reduction and identified contributing factors. As seen in Table 8.1, in most of the safety-related events, driver-related decision errors are the contributing factor. For both freeway and arterial safety-related events, more than 80% of the cases are caused by factors in this category. The next largest category on both types of roads is recognition errors in which drivers were distracted or failed to look.

For Project 6, the 100-Car Study, the factors that precipitated a safety-related event, contributed to the event, and were associated with the event were determined. These factors are grouped into pre-event maneuvers, precipitating factors, contributing factors, associated factors, and avoidance maneuvers. Of all the factors, contributing factors are the key in the study of driver behavior and were judged by trained data reductionists to have directly influenced the presence or severity of a crash or near crash. Three subcategories were further constructed for contributing factors to identify the causes of crashes: infrastructure factors and driving environment factors, such as road surface, traffic density, and weather; driver factors, such as driver inattention, drowsiness, and distraction; and vehicle factors, such as flat tires and vehicle breakdowns.

As revealed by the data, the factors that contributed to a crash usually were not caused by a sole factor but involved some form of factor interaction. For example, the driver could be distracted by both talking on a cell phone and adjusting the radio during an event, or the crash could have been caused by both inattention of the driver and a poorly designed roadway. Therefore, as shown in Tables 8.2 and 8.3, the resulting sum of percentages of contributing factors may add up to more than 100%. Not all 18 conflict categories are listed in these tables. Most of the categories of events have driver factors involved as a contributing factor, especially for the single- and lead-vehicle crashes, as well as for the single-, lead-, and following-vehicle near crashes. These categories have more than 100% driver factor–involved cases, demonstrating a high probability of human errors.

In some cases the contributing factors are descriptions of the status of the driver, vehicle, and environment at the moment the event happened. For example, if the driver was using a wireless device when the event happened, the driver factor would be "Secondary task" under the "Inattention to Forward Roadway" category. It is possible that these factors are not the direct causal factor leading to the events. To better distinguish preventable events from others, additional data reduction was performed by the team. Table 8.4 ascribes the crashes and near crashes to one mutually exclusive contributing category that can be considered as a dominant factor. The categories listed in this table better serve the purpose of studying the relationship between behavior and travel time reliability.

In summary, the data collected in the 100-Car Study were comprehensively and accurately reduced. The continuous video and other data are suitable for studying driving behavior and its impact on travel time reliability.

For Project 7, DDWS FOT, more than one vehicle was involved in multiple crashes. Because only the subject vehicle was equipped with data collection equipment, data reductionists could observe only scenarios related to that vehicle. Contributing factors were analyzed based on the observations

Table 8.1. Critical Contributing Factors for Project 5

Critical Factor	Freeway	Arterial	Total
Driver-related factor (critical nonperformance errors), including sleep; heart attack or other physical impairment of the ability to act; drowsiness, fatigue, or other reduced alertness; other critical nonperformance	0 (0.0%)	0 (0.0%)	0 (0.0%)
Driver-related factor (recognition errors), including inattention; internal distraction; external distraction; inadequate surveillance (e.g., failure to look); other or unknown recognition error	5 (6.5%)	0 (0.0%)	5 (5.0%)
Driver-related factor (decision errors), including too fast or too slow; misjudgment of gap; following too closely or unable to respond to unexpected actions; false assumption of other road user's actions; apparently intentional sign or signal violation; illegal U-turn or other illegal maneuver; failure to turn on headlamps; inadequate evasive action; aggressive driving; other or unknown decision error	66 (85.7%)	19 (82.6%)	85 (85.0%)
Driver-related factor (performance errors), including panic or freezing; overcompensation; poor directional control; other or unknown performance error	0 (0.0%)	0 (0.0%)	0 (0.0%)
Environment-related factor, including sign missing; view obstruction by roadway design; roadway geometry; sight distance; maintenance problems; slick roads; other highway-related conditions	0 (0.0%)	0 (0.0%)	0 (0.0%)
Environment-related factor, including glare, blowing debris, animal or object in roadway	0 (0.0%)	0 (0.0%)	0 (0.0%)
Crashes or near crashes caused by others	1 (1.3%)	1 (4.3%)	2 (2.0%)
Unknown reasons	5 (6.5%)	3 (13.0%)	8 (8/0%)
Total	**77 (100%)**	**23 (100%)**	**100 (100%)**

from equipped vehicles. The most frequent critical reason for crashes was "object in roadway," which constituted 57% of the total events. The next largest groups were driver-related factors (recognition errors) and driver-related factors (performance errors); each group had more than 14% of the cases. For tire strike cases, most were attributed to environment-related factors. For near crashes, driver-related factors (recognition errors and decision errors) constituted nearly half of all cases. Details of the critical factors are enumerated in Table 8.5.

In summary, the data collected in this study were comprehensive, and the data reduction was extensive. This study was conducted more recently than Project 6; consequently, the instrumentation used to collect the data was more accurate. Because only commercial trucks were studied, the data set has certain limitations with regard to the versatility of drivers and vehicles.

For Project 8, NTDS, a total of 320,011 triggers were visually inspected during data reduction. From those triggers, 2,899 safety-critical events were identified, including 13 crashes (eight of those were tire strikes), 61 near crashes, 1,594 crash-relevant conflicts, 1,215 unintentional lane deviations, and 16 illegal maneuvers. Additionally, a random sample of

Table 8.2. Contributing Factors for Crashes in Project 6

		Factor Category		
Crashes	Total Number	Driver	Environmental	Vehicle
Single vehicle	24	121%	38%	0%
Lead vehicle	15	127%	13%	0%
Following vehicle	12	83%	8%	0%
Object obstacle	9	144%	56%	0%
Parked vehicle	4	100%	50%	0%
Animal	2	0%	100%	0%
Turning across opposite direction	2	100%	50%	0%
Adjacent vehicle	1	100%	0%	0%

Table 8.3. Contributing Factors for Near Crashes in Project 6

Near Crashes	Total Number	Factor Category Driver	Environmental	Vehicle
Single vehicle	48	135%	29%	0%
Lead vehicle	380	119%	9%	0%
Following vehicle	70	110%	9%	0%
Object obstacle	6	83%	50%	0%
Parked vehicle	5	60%	0%	0%
Animal	10	70%	10%	0%
Turning across opposite direction	27	96%	30%	0%
Adjacent vehicle	115	90%	11%	0%
Merging vehicle	6	33%	67%	0%
Across path through intersection	27	89%	30%	0%
Oncoming	27	96%	30%	0%
Other	2	100%	50%	0%
Pedestrian	6	133%	50%	0%
Turning across in same direction	3	44%	11%	0%
Turning in same direction	28	54%	21%	0%
Unknown	1	200%	0%	0%

Table 8.4. Critical Factors Contributing to Crashes and Near Crashes

Critical Factor	Crashes	Near Crashes
Driver-related factor (critical nonperformance errors), including sleep; heart attack or other physical impairment of the ability to act; drowsiness, fatigue, or other reduced alertness; other critical nonperformance	8 (11.6%)	33 (4.3%)
Driver-related factor (recognition errors), including inattention; internal distraction; external distraction; inadequate surveillance (e.g., failure to look); other or unknown recognition error	22 (31.9%)	201 (26.4%)
Driver-related factor (decision errors), including too fast or too slow; misjudgment of gap; following too closely or unable to respond to unexpected actions; false assumption of other road user's actions; apparently intentional sign or signal violation; illegal U-turn or other illegal maneuver; failure to turn on headlamps; inadequate evasive action; aggressive driving; other or unknown decision error	19 (27.5%)	218 (28.6%)
Driver-related factor (performance errors), including panic or freezing; overcompensation; poor directional control; other or unknown performance error	1 (1.4%)	30 (3.9%)
Environment-related factor, including sign missing; view obstruction by roadway design; roadway geometry; sight distance; maintenance problems; slick roads; other highway-related conditions	2 (2.9%)	26 (3.4%)
Environment-related factor, including glare, blowing debris, animal or object in roadway	4 (5.8%)	29 (3.8%)
Crashes or near crashes caused by others	13 (18.8%)	224 (29.4%)
Total	**69 (100%)**	**761 (100%)**

Table 8.5. Categorized Critical Factors for Crashes and Near Crashes in Project 7

Critical Factor	Crashes		Crashes: Tire Strikes		Near Crashes	
Critical reason not coded to this vehicle.	1	7.1%	0	0%	29	29.6%
Driver-related factor (critical nonperformance errors), including sleep; heart attack or other physical impairment of the ability to act; drowsiness, fatigue, or other reduced alertness; other critical nonperformance	0	0%	0	0%	1	1.0%
Driver-related factor (recognition errors), including inattention; internal distraction; external distraction; inadequate surveillance (e.g., failure to look); other or unknown recognition error	2	14.3%	0	0%	30	30.6%
Driver-related factor (decision errors), including too fast or too slow; misjudgment of gap; following too closely or unable to respond to unexpected actions; false assumption of other road user's actions; apparently intentional sign or signal violation; illegal U-turn or other illegal maneuver; failure to turn on headlamps; inadequate evasive action; aggressive driving; other or unknown decision error	1	7.1%	2	14.3%	18	18.4%
Driver-related factor (performance errors), including panic or freezing; over-compensation; poor directional control; other or unknown performance error	2	14.3%	3	21.4%	7	7.1%
Environment-related factor, including sign missing; view obstruction by roadway design; roadway geometry; sight distance; maintenance problems; slick roads; other highway-related conditions	0	0%	9	64.3%	2	2.1%
Environment-related factor, including glare, blowing debris, animal or object in roadway	8	57.2%	0	0%	11	11.2%
Total	**14**	**100%**	**14**	**100%**	**98**	**100%**

456 baseline events, each 30 s long, was selected. Data reductionists used the data directory and coded a variety of variables from these 456 randomly selected baseline driving events or brief driving periods. One random baseline event was selected for each driver-week of data collection. Baseline events were described using many of the same variables used to describe safety-critical events. The goal of identifying baseline events was to provide a comparison between normal driving and driving during critical events. For example, the proportion of time spent driving under various conditions and the proportion of time drivers performed various behaviors (e.g., eating, drinking, talking on citizens band [CB] radio or cell phone) were compared across different situations (*1*).

Because only the subject vehicle was equipped with data collection units, only the behavior of the driver in that vehicle was coded and documented. As shown in Table 8.6, the most frequent critical factor for crashes was an object in the roadway, followed by driver-related factors associated with recognition errors, decision errors, and performance errors; each constituted 20% of the total cases. Not surprisingly, almost all (75%) the tire strikes involved some type of improper turn. The next two largest categories of contributing factor for crashes: tire strikes are driver performance error and driver decision error, respectively. For near crashes, the most frequent factor is driver-related recognition errors; more than 40% of near crashes were caused by inattention or distraction. Of these near crashes, almost one-quarter involved the subject driver not seeing the other vehicle during a lane change or merge.

Countermeasures

A close examination of the event causations reveals that a significant portion of the crashes or near crashes happened because of driver errors, such as inattention, distraction, or judgment errors. To prevent these events, the driver's response should be altered, his or her attention should be improved, or his or her driving habits should be corrected. The most frequently suggested functional countermeasures relating to modifying driver behavior include increasing driver recognition of specific highway crash threats (improving driver recognition of forward threats), increasing driver attention, improving driver situation awareness, and defensive driving. The team examined the video data from Project 5, and the results are listed in Table 8.7. Most events are preventable by modifying driver behavior or increasing the attention level. The percentage indicates the portion of crashes and near crashes that would have been avoided if the suggested countermeasures had been applied. It was not unusual that more than one countermeasure could be selected for an event when the contributing factor was a combination of factors. Therefore, the total may be more than 100%.

Because of the massive size of the data, the countermeasures identification for Project 6 was not as detailed as for projects

Table 8.6. Contributing Factors for Crashes in Project 8

Critical Factor	Crashes		Crashes: Tire Strikes		Near Crashes	
Critical reason not coded to this vehicle	0	0%	0	0%	16	26.2%
Driver-related factor (critical nonperformance errors), including sleep; heart attack or other physical impairment of the ability to act; drowsiness, fatigue, or other reduced alertness; other critical nonperformance	0	0%	0	0%	4	6.6%
Driver-related factor (recognition errors), including inattention; internal distraction; external distraction; inadequate surveillance (e.g., failure to look); other or unknown recognition error	1	20%	0	0%	27	44.3%
Driver-related factor (decision errors), including too fast or too slow; misjudgment of gap; following too closely or unable to respond to unexpected actions; false assumption of other road user's actions; apparently intentional sign or signal violation; illegal U-turn or other illegal maneuver; failure to turn on headlamps; inadequate evasive action; aggressive driving; other or unknown decision error	1	20%	1	12.5%	6	9.8%
Driver-related factor (performance errors), including panic or freezing; over-compensation; poor directional control; other or unknown performance error	1	20%	1	12.5%	5	8.2%
Environment-related factor, including sign missing; view obstruction by roadway design; roadway geometry; sight distance; maintenance problems; slick roads; other highway-related conditions	0	0%	6	75%	2	3.2%
Environment-related factor, including: glare, blowing debris, animal or object in roadway	2	40%	0	0%	1	1.6%
Total	**5**	**100%**	**8**	**100%**	**61**	**100%**

discussed earlier. The team did differentiate avoidable crashes from unavoidable or likely avoidable crashes and near crashes, as shown in Table 8.8. Almost 40% of crashes can or are likely to be prevented. More than 80% of near crashes can or are likely to be prevented, given reasonable countermeasures.

The detailed countermeasures to safety-critical events in Project 7 and Project 8 are illustrated in Tables 8.9 and 8.10 (*1; 2*), respectively. Because of multiple countermeasures applicable to one event, the total may be more than 100%.

Tables 8.9 and 8.10 list the functional countermeasures that describe an intervention into the driving situation. Specifically, to technologically modify drivers' behaviors, warning systems can be used to alert drivers so that they are not distracted or to correct driving habits to improve safety. Many of these systems, some of which are provided in the following list, have been tested in previous studies as technical countermeasures.

1. FCW System. An FCW system measures the headway between the subject vehicle and the leading vehicle. It issues a visual or audio warning when the equipped vehicle approaches the leading vehicle too rapidly. The system is effective in correcting drivers' performance errors and decision errors. It is even more effective in alerting an inattentive or a less-alert driver (*3*).
2. Lane Tracking System/LDWS. A lane-tracking system or an LDWS usually can measure lane-keeping behavior. For example, VTTI developed a lane tracker called Road Scout, which consists of a single analog black-and-white camera, a personal computer with a frame-grabber card, and an interface-to-vehicle network for obtaining ground speed. Distance from the center of the car to left and right lane markings, the angular offset between the car centerline and road centerline, the approximate road curvature, and marking characteristics can be measured and calculated to determine whether the car remains in the lane or is crossing lines. A similar system was developed by UMTRI. A lane-tracking system can be used mainly to alert a driver in circumstances of decreased alertness. It can also be used to correct drivers' recognition errors, decision errors, and performance errors (*1–3*).
3. ACC System. Instead of simply maintaining a preset target speed, as does a CCC system, an ACC system is primarily a speed and headway controller. It can modulate speed and use throttle and brakes simultaneously to manage headway to the leading vehicle. Usually an ACC system holds a maximum braking authority. In Project 2, this value is 0.3 *g*. When the headway decreases to the point at which the maximum braking response is required, the system will issue an alert or apply braking. This system will benefit drivers with recognition errors, a decreased alertness level, and some extent of decision errors and performance errors (*3*).
4. CSWS. A CSWS can usually help drivers slow down to a safe speed before entering an upcoming curve. It uses GPS and

Table 8.7. Countermeasures by Category for Project 5

	Crashes and Near Crashes				Total	
	Freeway		Arterial			
No countermeasure applicable	5	6.5%	6	26.0%	11	11.0%
1. Increase driver alertness (reduce drowsiness)	2	2.6%	0	0.0%	2	2.0%
2. Prevent drift lane departures	1	1.3%	0	0.0%	1	1.0%
3. Improve vehicle control on curves	0	0.0%	0	0.0%	0	0.0%
4. Improve vehicle control on slippery road surfaces	0	0.0%	0	0.0%	0	0.0%
5. Improve vehicle control during braking	0	0.0%	0	0.0%	0	0.0%
6. Improve vehicle control during evasive steering	0	0.0%	0	0.0%	0	0.0%
7. Increase driver attention to forward scene	15	19.5%	1	4.3%	16	16.0%
8. Improve driver use of mirrors or provide better information from mirrors	0	0.0%	0	0.0%	0	0.0%
9. Improve general driver situation awareness and defensive driving	1	1.3%	0	0.0%	1	1.0%
10. Reduce travel speed	1	1.3%	0	0.0%	1	1.0%
11. Reduce speed on downgrades	0	0.0%	0	0.0%	0	0.0%
12. Reduce speed on curves or turns	0	0.0%	0	0.0%	0	0.0%
13. Reduce speed at or on exits (including ramps)	0	0.0%	0	0.0%	0	0.0%
14. Limit top speed to 70 mph (except on downgrades)	0	0.0%	0	0.0%	0	0.0%
15. Increase driver recognition of specific highway crash threats: stopped vehicle(s) in lane ahead, traveling in same direction	14	18.2%	0	0.0%	14	14.0%
16. Increase driver recognition of specific highway crash threats: moving or decelerating vehicle(s) in lane ahead, traveling in same direction	55	71.4%	17	73.9%	72	72.0%
17. Increase driver recognition of specific highway crash threats: vehicle in left adjacent lane on highway	2	2.6%	0	0.0%	2	2.0%
18. Increase driver recognition of specific highway crash threats: vehicle in right adjacent lane on highway	1	1.3%	1	4.3%	2	2.0%
19. Increase driver recognition of specific highway crash threats: vehicle in left adjacent lane during merging maneuver	0	0.0%	0	0.0%	0	0.0%
20. Increase driver recognition of specific highway crash threats: vehicle in right adjacent lane during merging maneuver	0	0.0%	0	0.0%	0	0.0%
21. Increase driver recognition or gap judgment recrossing or oncoming traffic at intersections	0	0.0%	0	0.0%	0	0.0%
22. Improve driver response execution of crossing or turning maneuver at intersections (performance failure)	0	0.0%	0	0.0%	0	0.0%
23. Improve driver recognition or gap judgment response execution at intersection	0	0.0%	0	0.0%	0	0.0%
24. Improve driver compliance with intersection traffic signal controls (both intentional and unintentional intersection control violations)	0	0.0%	0	0.0%	0	0.0%
25. Improve driver compliance with intersection traffic sign controls	0	0.0%	0	0.0%	0	0.0%
26. Increase forward headway during vehicle following	9	11.7%	2	8.7%	11	11.0%
27. Improve driver night vision in the forward field	0	0.0%	0	0.0%	0	0.0%
28. Provide warning to prevent rear encroachment or tailgating by other vehicle	0	0.0%	0	0.0%	0	0.0%
29. Provide advisory to driver regarding reduced road-tire friction (i.e., associated with slippery roads)	0	0.0%	0	0.0%	0	0.0%
30. Prevent vehicle mechanical failure	0	0.0%	0	0.0%	0	0.0%

(continued on next page)

Table 8.7. Countermeasures by Category for Project 5 (continued)

	Crashes and Near Crashes				Total	
	Freeway		Arterial			
31. Prevent splash and spray from this vehicle affecting other vehicle(s)	0	0.0%	0	0.0%	0	0.0%
32. Improve driver recognition or gap judgment relating to oncoming vehicle during passing maneuver	0	0.0%	0	0.0%	0	0.0%
33. Prevent animals from crossing roadways	0	0.0%	0	0.0%	0	0.0%
34. Provide driver with navigation system	0	0.0%	0	0.0%	0	0.0%
35. Aid to vertical clearance estimation	0	0.0%	0	0.0%	0	0.0%
36. Prevent or reduce trailer off-tracking outside travel lane or path	0	0.0%	0	0.0%	0	0.0%
97. Provide advance warning of need to stop at traffic sign or signal	0	0.0%	0	0.0%	0	0.0%
98. Driver error or vehicle failure apparent but countermeasure unknown	0	0.0%	0	0.0%	0	0.0%
99. Unknown	0	0.0%	0	0.0%	0	0.0%
Total	***101***	***131.2%***	***21***	***91.3%***	***133***	***133.0%***
Events Total	**77**	**100%**	**23**	**100%**	**100**	**100.00%**

Table 8.8. Preventability of Crashes in Project 6

	Avoidable		Likely Avoidable		Unavoidable	
Amendable Factor	**Crashes**	**Near Crashes**	**Crashes**	**Near Crashes**	**Crashes**	**Near Crashes**
Correct driver-related factor, including sleep; drowsiness or other reduced alertness; other critical nonperformance	4 (5.8%)	20 (2.6%)	0 (0.0%)	1 (0.1%)	4 (5.8%)	12 (1.6%)
Correct driver-related factor, including inattention; internal distraction; external distraction; inadequate surveillance (e.g., failure to look); other or unknown recognition error	13 (18.8%)	109 (14.3%)	0 (0.0%)	29 (3.8%)	9 (13.0%)	63 (8.3%)
Correct driver-related factor, including too fast or too slow; misjudgment of gap; following too closely to respond to unexpected actions; false assumption of other road user's actions; apparently intentional sign or signal violation; illegal U-turn or other illegal maneuver; failure to turn on headlamps; inadequate evasive action; aggressive driving; other or unknown decision error	6 (8.7%)	134 (17.6%)	1 (1.4%)	60 (7.9%)	12 (17.4%)	24 (3.2%)
Correct driver-related factor, including poor directional control; other or unknown performance error	0 (0.0%)	5 (0.7%)	0 (0.0%)	20 (2.6%)	1 (1.4%)	5 (0.7%)
Correct environment-related factor, including sign missing; view obstruction by roadway design; roadway geometry; sight distance; maintenance problems; slick roads; other highway-related conditions	0 (0.0%)	12 (1.6%)	0 (0.0%)	14 (1.8%)	2 (2.9%)	0 (0.0%)
Correct environment-related factor, including glare, blowing debris, animal or object in roadway.	0 (0.0%)	9 (1.2%)	0 (0.0%)	20 (2.6%)	4 (5.8%)	0 (0.0%)
Not correctable: crashes or near crashes caused by others	3 (4.3%)	33 (4.3%)	0 (0.0%)	164 (21.6%)	10 (14.5%)	27 (3.5%)
Total	**26 (38%)**	**322 (42%)**	**1 (1%)**	**308 (40%)**	**42 (61%)**	**131 (17%)**

Table 8.9. Countermeasures by Category for Project 7

	Crashes		Crashes: Tire Strikes		Near Crashes	
No countermeasure applicable	1	7.1%	0	0.0%	18	18.4%
1. Increase driver alertness (reduce drowsiness)	3	21.4%	0	0.0%	7	7.1%
3. Prevent drift lane departures	0	0.0%	0	0.0%	4	4.1%
4. Improve vehicle control on curves	0	0.0%	0	0.0%	0	0.0%
5. Improve vehicle control on slippery road surfaces	0	0.0%	0	0.0%	0	0.0%
6. Improve vehicle control during braking	0	0.0%	0	0.0%	0	0.0%
7. Improve vehicle control during evasive steering	0	0.0%	0	0.0%	1	1.0%
8. Increase driver attention to forward scene	3	21.4%	0	0.0%	28	28.6%
9. Improve driver use of mirrors or provide better information from mirrors	0	0.0%	0	0.0%	3	3.1%
10. Improve general driver situation awareness and defensive driving	1	7.1%	0	0.0%	9	9.2%
12. Reduce travel speed	0	0.0%	0	0.0%	3	3.1%
13. Reduce speed on downgrades	0	0.0%	0	0.0%	0	0.0%
14. Reduce speed on curves or turns	0	0.0%	0	0.0%	1	1.0%
15. Reduce speed at or on exits (including ramps)	0	0.0%	0	0.0%	1	1.0%
16. Limit top speed to 70 mph (except on downgrades)	0	0.0%	0	0.0%	0	0.0%
17. Increase driver recognition of specific highway crash threats: stopped vehicle(s) in lane ahead, traveling in same direction	1	7.1%	0	0.0%	4	4.1%
18. Increase driver recognition of specific highway crash threats: moving or decelerating vehicle(s) in lane ahead, traveling in same direction	0	0.0%	0	0.0%	6	6.1%
19. Increase driver recognition of specific highway crash threats: vehicle in left adjacent lane on highway	0	0.0%	0	0.0%	3	3.1%
20. Increase driver recognition of specific highway crash threats: vehicle in right adjacent lane on highway	0	0.0%	0	0.0%	1	1.0%
21. Increase driver recognition of specific highway crash threats: vehicle in left adjacent lane during merging maneuver	0	0.0%	0	0.0%	0	0.0%
22. Increase driver recognition of specific highway crash threats: vehicle in right adjacent lane during merging maneuver	0	0.0%	0	0.0%	0	0.0%
23. Increase driver recognition or gap judgment regarding crossing or oncoming traffic at intersections	0	0.0%	0	0.0%	1	1.0%
25. Improve driver response execution of crossing or turning maneuver at intersections (performance failure)	2	14.3%	5	35.7%	2	2.0%
26. Improve driver recognition or gap judgment response execution at intersection	0	0.0%	0	0.0%	0	0.0%
27. Improve driver compliance with intersection traffic signal controls (both intentional and unintentional intersection control violations)	0	0.0%	0	0.0%	0	0.0%
28. Improve driver compliance with intersection traffic sign controls	0	0.0%	0	0.0%	0	0.0%
29. Increase forward headway during vehicle following	0	0.0%	0	0.0%	2	2.0%
30. Improve driver night vision in the forward field	0	0.0%	0	0.0%	5	5.1%
32. Provide warning to prevent rear encroachment or tailgating by other vehicle	0	0.0%	0	0.0%	0	0.0%
33. Provide advisory to driver regarding reduced road-tire friction (i.e., associated with slippery roads)	0	0.0%	0	0.0%	0	0.0%
34. Prevent vehicle mechanical failure	0	0.0%	0	0.0%	0	0.0%
36. Prevent splash and spray from this vehicle affecting other vehicle(s)	0	0.0%	0	0.0%	0	0.0%

(continued on next page)

Table 8.9. Countermeasures by Category for Project 7 (continued)

	Crashes		Crashes: Tire Strikes		Near Crashes	
37. Improve driver recognition or gap judgment relating to oncoming vehicle during passing maneuver	0	0.0%	0	0.0%	0	0.0%
38. Prevent animals from crossing roadways	4	28.6%	0	0.0%	9	9.2%
39. Provide driver with navigation system	2	14.3%	1	7.1%	1	1.0%
40. Aid to vertical clearance estimation	2	14.3%	0	0.0%	0	0.0%
41. Prevent or reduce trailer off-tracking outside travel lane or path	0	0.0%	7	50.0%	5	5.1%
42. Provide advance warning of need to stop at traffic sign or signal	0	0.0%	0	0.0%	3	3.1%
98. Driver error or vehicle failure apparent but countermeasure unknown	1	7.1%	2	14.3%	0	0.0%
99. Unknown	1	7.1%	0	0.0%	2	2.0%
Total	***21***	***150.0%***	***15***	***107.1%***	***119***	***121.4%***
Events Total	**14**	**100%**	**14**	**100%**	**98**	**100%**

Table 8.10. Countermeasures by Category for Project 8

	Crashes		Crashes: Tire Strikes		Near Crashes	
No countermeasure applicable	0	0%	0	0%	11	18%
1. Increase driver alertness (reduce drowsiness)	0	0%	1	13%	7	11%
2. Prevent drift lane departures	0	0%	0	0%	18	30%
3. Improve vehicle control on curves	0	0%	0	0%	1	2%
4. Improve vehicle control on slippery road surfaces	0	0%	0	0%	1	2%
5. Improve vehicle control during braking	0	0%	0	0%	0	0%
6. Improve vehicle control during evasive steering	0	0%	0	0%	0	0%
7. Increase driver attention to forward scene	0	0%	0	0%	16	26%
8. Improve driver use of mirrors or provide better information from mirrors	2	40%	0	0%	11	18%
9. Improve general driver situation awareness and defensive driving	3	60%	2	25%	23	38%
10. Reduce travel speed	0	0%	0	0%	2	3%
11. Reduce speed on downgrades	0	0%	0	0%	0	0%
12. Reduce speed on curves or turns	0	0%	0	0%	1	2%
13. Reduce speed at or on exits (including ramps)	0	0%	0	0%	0	0%
14. Limit top speed to 70 mph (except on downgrades)	0	0%	0	0%	0	0%
15. Increase driver recognition of specific highway crash threats: stopped vehicle(s) in lane ahead, traveling in same direction	0	0%	0	0%	0	0%
16. Increase driver recognition of specific highway crash threats: moving or decelerating vehicle(s) in lane ahead, traveling in same direction	0	0%	0	0%	3	5%
17. Increase driver recognition of specific highway crash threats: vehicle in left adjacent lane on highway	0	0%	0	0%	9	15%
18. Increase driver recognition of specific highway crash threats: vehicle in right adjacent lane on highway	1	20%	0	0%	6	10%
19. Increase driver recognition of specific highway crash threats: vehicle in left adjacent lane during merging maneuver	0	0%	0	0%	3	5%

(continued on next page)

Table 8.10. Countermeasures by Category for Project 8 (continued)

	Crashes		Crashes: Tire Strikes		Near Crashes	
20. Increase driver recognition of specific highway crash threats: vehicle in right adjacent lane during merging maneuver	0	0%	0	0%	1	2%
21. Increase driver recognition or gap judgment regarding crossing or oncoming traffic at intersections	0	0%	0	0%	0	0%
22. Improve driver response execution of crossing or turning maneuver at intersections (performance failure)	0	0%	1	13%	0	0%
23. Improve driver recognition or gap judgment response execution at intersection	0	0%	0	0%	0	0%
24. Improve driver compliance with intersection traffic signal controls (both intentional and unintentional intersection control violations)	0	0%	0	0%	1	2%
25. Improve driver compliance with intersection traffic sign controls	0	0%	0	0%	0	0%
26. Increase forward headway during vehicle following	0	0%	0	0%	2	3%
27. Improve driver night vision in the forward field	0	0%	0	0%	0	0%
28. Provide warning to prevent rear encroachment or tailgating by other vehicle	0	0%	0	0%	0	0%
29. Provide advisory to driver regarding reduced road-tire friction (i.e., associated with slippery roads)	0	0%	0	0%	0	0%
30. Prevent vehicle mechanical failure	0	0%	0	0%	0	0%
31. Prevent splash and spray from this vehicle affecting other vehicle(s)	0	0%	0	0%	0	0%
32. Improve driver recognition or gap judgment relating to oncoming vehicle during passing maneuver	0	0%	0	0%	0	0%
33. Prevent animals from crossing roadways	1	20%	0	0%	1	2%
34. Provide driver with navigation system	1	20%	1	13%	1	2%
35. Aid to vertical clearance estimation	0	0%	0	0%	0	0%
36. Prevent or reduce trailer off-tracking outside travel lane or path	0	0%	2	25%	0	0%
97. Provide advance warning of need to stop at traffic sign or signal	0	0%	5	63%	1	2%
98. Driver error or vehicle failure apparent but countermeasure unknown	0	0%	1	13%	0	0%
99. Unknown	0	0%	0	0%	0	0%
Total	***8***	***160%***	***13***	***163%***	***119***	***195%***
Event Total	**5**	**100%**	**8**	**100%**	**61**	**100%**

digital maps to anticipate curve locations and radiuses. Combined with recent driver control actions (turning signal and lateral acceleration), it determines if it is appropriate to issue a warning. The system is effective in correcting driver performance errors, recognition errors, and decision errors. It also helps to alert drivers of upcoming changes in roadway geometry (*4*).

5. Dilemma Zone Mitigation (DZM). Many crashes that occur at signalized intersections are associated with dilemma zones; for example, when faced with a yellow light, some drivers may decide to proceed through and others may decide to stop. Components of DZM usually include a carefully designed signal-timing cycle with an effective vehicle detection system that will identify the speed and size of vehicles, as well as the distance to the stopping line and provide additional safety by extending green time to allow safe passage through the intersection if necessary (*5*).
6. Lateral Vehicle Detection (LVD). LVD usually consists of lateral cameras, a lane change assistance system, and a lateral collision warning system. The main purpose of LVD is to aid drivers to detect movements of vehicles in adjacent lanes and conduct corresponding maneuvers. The system will issue a warning when it determines that a lateral vehicle is trying to cut in front of the subject vehicle in an unsafe way (*6*).
7. Intelligent Speed Adaption System. In this system, developed in Sweden, GPS was used to locate a car on a digital map. The speed limit on that roadway was retrieved from the database, and the real speed of the vehicle was compared with the speed limit. The system adopts interventions that are preprogrammed in the vehicle (*7*).

Besides these existing warning systems, potentially beneficial warning systems not yet tested might be effective in reducing safety-related events; for example, a system that is capable of detecting weather and road surface conditions (e.g., rainfall amount, snow amount, visibility, wet road surface) and proposing possible road friction parameter variations because of these conditions in order to issue corresponding warnings, and a customized warning system initiated by the user's individual car key, which can adjust warning-issuing threshold values according to different driving habits.

When making countermeasure recommendations, it should be recognized that emerging driver assistance systems may initiate some complexities and, therefore, the assessment of safety benefits is not straightforward. For example, when drivers rely on these safety systems, failure of such systems can be fatal. Incorporating some other countermeasures in a systematic approach will more than likely be beneficial. In conclusion, collision prevention should include a better design of roads, a more comprehensive recovery system, and a more coordinated safety management system. According to a report from the Organisation for Economic Co-operation and Development, some basic enforcement may be highly efficient. Seat belt usage, speed management, extra efforts to monitor high-risk drivers, and identification and monitoring of dangerous locations are all effective countermeasures that contribute to improvements in transportation system safety (*8*).

Conclusions

To determine the feasibility of using in-vehicle video data to make inferences about driver behavior that would allow investigation of the relationship between observable driver behavior and nonrecurring congestion to improve travel time reliability, the team explored the identified data sets to investigate the usefulness of video and other supplementary data, proposed models for the estimation of travel time reliability measures, and identified potential problems in current data sources. Based on the analysis of the six naturalistic data sources, this study demonstrates the following:

1. It is feasible to identify driver behavior before near crashes and crashes from video data collected in a naturalistic driving study and thus infer the causes of those events.
2. Recommendations can be made to change driver behavior and, therefore, prevent (or reduce) crashes and near crashes.
3. Naturalistic data are useful to identify impacts of crashes on traffic conditions. Given the small sample of crashes and the fact that the DAS does not gather data when the engine is off, it is not possible to study the impact of incidents on travel time reliability using in-vehicle data alone. When effectively integrated with external data sources, however, which is extremely feasible, given an accurate time and location stamp in the data set, naturalistic data can be highly efficient in recognizing the relationship between modifying driver behavior and nonrecurring congestion.
4. Increased coordination with weather and traffic volume data is required to determine when nonrecurring congestion exists and which driver actions result from these nonrecurring events.
5. It is possible to analyze naturalistic driving data to characterize typical levels of variability in travel times and develop measures for quantifying travel time reliability.

Although the team has successfully proved the feasibility of using video and other in-vehicle and external data to study driver behavior and related nonrecurring congestion, some limitations need to be enhanced when a full data analysis is conducted. These limitations are summarized as follows:

1. A limited number of safety-related events exist in the data sets the team examined because of the naturalistic nature of the data. This shortcoming can be improved by extending the time duration of data collection or increasing the number of participants. Both can be realized in the SHRP 2 Safety Project S07 study, in which a much larger data collection effort will be performed.
2. The external data sources in this study were examined only for availability and accuracy. Because of time constraints, no real connection was conducted to relate driver behavior to external driving environment.
3. Because of the limited size of travel time data in the naturalistic data sets, other data sets were also used to develop travel time reliability models. These models are general and apply regardless of the source of travel time data.

These limitations can be corrected if a larger video data set can be collected or the external data can be better linked with the in-vehicle data, which will be feasible in the next stage of the study.

Recommendations and Discussion

An important component of the next stage of SHRP 2 research is a large-scale naturalistic driving data collection project (Project S07). The field study data collection contractors will be responsible for advertising for participants with preprepared recruitment materials, scheduling participant drivers for installation and assessment, conducting driver intake testing and installing the DAS in the owner's vehicle, collecting data, addressing problems encountered during the study, investigating crashes, transmitting data, carrying out quality control procedures, and preparing periodic reports that document field study activities. The combined goal is to collect approximately 4,000 vehicle-years of data in a 30-month period. The following are the planned variables to be collected:

1. Antilock Brake System (ABS) Activation: Antilock brake activation indicator.
2. Acceleration, *x* axis: Vehicle acceleration in the longitudinal direction versus time.
3. Acceleration, *x* axis fast: Vehicle acceleration in the longitudinal direction versus time. Fast buffer (−9 s to +3 s) based on trigger (e.g., in crash or other high-acceleration event).
4. Acceleration, *y* axis: Vehicle acceleration in the lateral direction versus time.
5. Acceleration, *y* axis fast: Vehicle acceleration in the lateral direction versus time. Fast buffer (−9 s to +3 s) based on trigger (e.g., in crash or other high-acceleration event).
6. Acceleration, *z* axis: Vehicle acceleration vertically (up or down) versus time.
7. Acceleration, *z* axis fast: Vehicle acceleration vertically (up or down) versus time. Fast buffer (−9 s to +3 s) based on trigger (e.g., in crash or other high-acceleration event);
8. Airbag, Driver: Indicates deployment of the driver's airbag.
9. Alcohol: Presence of alcohol in the vehicle cabin.
10. Altitude, GPS: Altitude.
11. Audio: Audio recording for 30 s when incident button is pushed.
12. Average Fuel Economy after Fueling: Average fuel economy after fueling.
13. Cruise Control: Status of cruise control.
14. Date: UTC year, month, and day.
15. Distance: Distance of vehicle travel.
16. Driver Button Flag: Flag indicating that the driver has pressed the incident button.
17. Electronic Stability Control (ESC): ESC activation indicator.
18. Engine RPM: Instantaneous engine speed.
19. Face, Driver ID: Machine-vision–based identification of the driver within those observed in a specific vehicle. The system observes within a vehicle to identify drivers who drive that vehicle (i.e., not a unique identification across all drivers in the study).
20. Face, Gaze Zone: Estimation of the location of the driver's gaze categorized into zones in and around the vehicle.
21. Face, Gaze Zone Confidence: Confidence in the estimation of the zone at which the driver is looking.
22. Fuel Economy, Instantaneous: Instantaneous fuel economy.
23. Fuel Level: Fuel level.
24. Heading, GPS: Compass heading of vehicle from GPS.
25. Headlight Setting: State of headlamps.
26. Horn Status: Actuation of horn.
27. Illuminance, Ambient: Ambient exterior light.
28. LDWS: Status of original equipment manufacturer (OEM) lane departure warning system.
29. Lane Marking, Distance, Left: Distance from vehicle centerline to inside of left side lane marker based on vehicle-based machine vision.
30. Lane Marking, Distance, Right: Distance from vehicle centerline to inside of right side lane marker based on vehicle-based machine vision.
31. Lane Marking, Probability, Right: Probability that vehicle-based machine vision lane marking evaluation is providing correct data for the right side lane markings.
32. Lane Marking, Type, Left: Type of lane marking immediately to the left of vehicle using vehicle-based machine vision.
33. Lane Marking, Type, Right: Type of lane marking immediately to the right of vehicle using vehicle-based machine vision.
34. Lane Marking, Probability, Left: Probability that vehicle-based machine vision lane marking evaluation is providing correct data for the left side lane markings.
35. Lane Position Offset: Distance to the left or right of the center of the lane based on machine vision.
36. Lane Width: Distance between the inside edge of the innermost lane marking and the left and right of the vehicle.
37. Latitude: Vehicle position latitude.
38. Longitude: Vehicle position longitude.
39. Pedal, Accelerator Position: Position of the accelerator pedal collected from the vehicle network and normalized using manufacturer specifications.
40. Pedal, Brake: On or off press of brake pedal.
41. Pitch Rate, *y* axis: Vehicle angular velocity around the lateral axis.
42. Pitch Rate, *y* axis fast: Vehicle angular velocity around the lateral axis. Fast buffer (−9 s to +3 s) based on trigger (e.g., in crash or other high-acceleration event).
43. P-R-N-D-L: Gear position.
44. Radar, Azimuth Forward: Angular measure to target.
45. Radar, Range Rate Forward: Range rate to forward radar targets.
46. Radar, Range, Forward: Range to forward radar targets measured from the radar to the targets.
47. Radar, Target Identification: Numerical value used to differentiate one radar target from others.
48. Radius of Curvature, Machine Vision: Estimation of roadway curvature based on machine vision.
49. Roll Rate, *x* axis: Vehicle angular velocity around the longitudinal axis.
50. Roll Rate, *x* axis fast: Vehicle angular velocity around the longitudinal axis. Fast buffer (−9 s to +3 s) based on trigger (e.g., in crash or other high-acceleration event).

51. Satellites, Number of: Count of the number of satellites being used for GPS position fix.
52. Seat belt, Driver: Use of the seat belt by the driver.
53. Speed, GPS: Vehicle speed from GPS.
54. Speed, Vehicle Network: Vehicle speed indicated on speedometer collected from network.
55. Steering Wheel Position: Angular position and direction of the steering wheel from neutral position.
56. Sync: Integer used to identify one time sample of data when presenting rectangular data.
57. Temperature, Interior: Vehicle interior temperature.
58. Time: UTC Time. Local time offsets need to be applied.
59. Track Type: Classification of target based on radar.
60. Traction Control: Status of traction control system.
61. Turn Signal: State of illumination of turn signals.
62. Vehicle Angle Relative to Roadway: Vehicle angle relative to the roadway based on machine vision.
63. Video Frame: Frame number of video at point in time.
64. Video, Driver and Left Side View: Video capture of the driver and exterior area to the left of the vehicle.
65. Video, Forward Roadway: Video capture of forward roadway.
66. Video, Occupancy Snapshot: Occupancy snapshot.
67. Video, Rear View: Video capture to the rear of the vehicle.
68. Video, Right Side View: Video capture to the right of the vehicle.
69. Wiper Setting: Indicates setting of windshield wipers.
70. Yaw Rate, *z* axis: Vehicle angular velocity around the vertical axis.
71. Yaw Rate, *z* axis fast: Vehicle angular velocity around the vertical axis. Fast buffer (−9 s to +3 s) based on trigger (e.g., in crash or other high-acceleration event).

To ensure that data collected in the SHRP 2 Safety Project S07 study are versatile and comprehensive enough to be used to conduct full-scaled research to study nonrecurring congestion related to driver behavior, several recommendations have resulted from the findings of this study.

First, the procedure to recruit participants needs to be carefully designed. Ideally, a comprehensive population of drivers ranging evenly across every age category, income category, and occupation category should be included. When recruiting participants, it is crucial to make it clear to them that driver information is vital for the research. To better identify drivers, two methods can be used:

1. A formal statement needs to be included in the contract to make the signer the exclusive driver of the vehicle.
2. A touch-screen device can be installed on board to collect information before and after each trip. The touch-screen equipment can be designed so that a customized interface will be displayed to the driver to input trip-related information by selecting certain check boxes. The before-trip information-collecting interface may consist of a list of the first names of household members for the driver to select from, a list of trip purposes, weather conditions when the trip started, and any information about why the driver selected the time of departure. The after-trip information-collecting interface may include an "original trip purpose changed" option, a "route choice changed" option, and a "crash happened en route" option. Necessary hardware can be designed to connect the input touch-screen with the engine so that the driver can start the engine only after the information is input. To ensure safety while driving, the device should be disabled while the vehicle is in motion to prevent driver distraction. One major concern this type of device may impose on such studies is that it will remind drivers that they are being monitored and thus may reduce the naturalistic nature of the studies.

Second, to serve the research purpose, certain data are more important than others. The following four categories are imperative:

1. Basic onboard equipment should include devices that collect the following data: video; vehicle network information (speed, brake pedal, throttle, and turn signal); GPS data (latitude, longitude, and heading); *X, Y,* and *Z* acceleration; distances between the subject and surrounding objects; lane location information (*X, Y,* and *Z*); driver behavior (seat belt usage, lights on or off); and yaw rate.
2. Video cameras should shoot at least five views: front, back, right, left, and the driver. The resolution of the video camera should be high enough to identify ongoing traffic conditions, weather conditions, and the driver's hand movements and facial expressions. Correction of sun glare to improve video quality is available when needed.
3. The frequency setting should be high enough so that the video is continuous, the acceleration or deceleration of the vehicles should be clearly recorded, and the reaction times need to be recorded and measured. The recommended minimum frequency for GPS devices is 1 Hz and, for all other equipment, 10 Hz.
4. To improve the versatility of the data so that it can be used in other, related research, the vehicle performance parameters (e.g., engine speed, throttle position, and torque) should be recorded. Table 8.11 shows a sublist of variables that are vital to the next stage of this research and that will be collected in the Project S07 study. Units and minimum rates of data collection are suggested.

Third, the data collection system needs to run for an additional 10 min after the engine is turned off in case the vehicle is involved in an accident. During the additional data reduction,

Table 8.11. Recommended Variables for Collection

Variable Name	Units	Recommended Minimum Rate
1. Acceleration, *x* axis	*g*	10 Hz
2. Acceleration, *y* axis	*g*	10 Hz
3. Acceleration, *z* axis	*g*	10 Hz
4. Altitude, GPS	ft	1 Hz
5. Date	NA	NA
6. Distance	mi	NA
7. Engine RPM	rpm	NA
8. Face, Driver ID	NA	10 Hz
9. Face, Gaze Zone	NA	10 Hz
10. Fuel Economy, Instantaneous	mpg	NA
11. Heading, GPS	degree	1 Hz
12. LDWS	NA	NA
13. Lane Marking, Distance, Left	ft	10 Hz
14. Lane Marking, Distance, Right	ft	10 Hz
15. Lane Marking, Type, Left	NA	10 Hz
16. Lane Marking, Type, Right	NA	10 Hz
17. Lane Position Offset	ft	NA
18. Lane Width	ft	NA
19. Latitude	Ddd.sss	1 Hz
20. Longitude	Ddd.sss	1 Hz
21. Pedal, Accelerator Position	NA	NA
22. Pedal, Brake	NA	NA
23. Pitch Rate, *y* axis	degree/s	10 Hz
24. Radar, Azimuth Forward	rad	10 Hz
25. Radar, Range, Forward	ft	10 Hz
26. Radar, Target Identification	NA	10 Hz
27. Radius of Curvature, Machine Vision	NA	NA
28. Roll Rate, *x* axis	degree/s	10 Hz
29. Seat belt, Driver	NA	10 Hz
30. Speed, GPS	mph	1 Hz
31. Time	NA	NA
32. Track Type	NA	NA
33. Video Frame	NA	NA
34. Video, Driver and Left Side View	NA	10 Hz
35. Video, Forward Roadway	NA	10 Hz
36. Video, Occupancy Snapshot	NA	NA
37. Video, Rear View	NA	10 Hz
38. Video, Right Side View	NA	NA
39. Wiper Setting	NA	1 Hz
40. Yaw Rate, *z* axis	degree/s	10 Hz

data collection was usually found to stop the instant the driver stopped the vehicle. It is important, however, to observe the traffic conditions being affected by a safety-related event. In discussion with the SHRP 2 S06 contractor, a potential safety hazard was identified that may deem this recommendation infeasible. Specifically, continued data collection after an accident may result in a vehicle explosion if the vehicle gasoline tank is jeopardized.

Fourth, to improve the linking of vehicle data with external data, it is ideal to standardize the time and location data. For external data, the database in some states is built on the milepost system. The conversion of milepost locations to standard latitude and longitude coordinates should be conducted ahead of time. For vehicle data, the synchronized GPS clock should be used instead of the local computer time for better connection of the data with external traffic, crash, work zone, and weather data.

Fifth, because a limited number of crashes occurred in all the candidate data sets—especially severe crashes that affected traffic conditions—certain adjustments are needed to create a statistically significant database. A lengthier data collection effort or more drivers involved in the study would be ideal. For example, the 2,500-Car Study (SHRP 2 Safety Project S07), which will soon be conducted, is a quality candidate. Another solution is simulation, which can be used to compensate for data shortage.

Sixth, additional analysis of existing data is required to study typical levels of variability in driver departure times, typical levels of variability in trip travel times, and the level of variability in driver route choices. A characterization of this behavior is critical in attempting to quantify and develop travel time reliability measures because it identifies potential causes for travel time variability and thus can enhance travel time reliability models. These data may be augmented with tests on a driving simulator to study the impact of travel time reliability on driver route choice behavior.

Finally, although a number of studies have used video cameras to gather data, an ideal starting point is a compiled data source list that summarizes existing video-involved studies with specifications of data collected, limitations of data usage, and access issues. Such a list would help prevent redundancy in future investigation efforts.

This research can benefit from the data being collected under the IntelliDrive Program (IntelliDrive is a service mark of the U.S. Department of Transportation). The IntelliDrive Program is, as introduced on its website, "a multimodal initiative that aims to enable safe, interoperable networked wireless communications among vehicles, the infrastructure, and passenger's personal communications devices" (*9*). It will collect and disseminate data, including roadway, traffic condition, weather, crashes, and traffic control among vehicles. With the development of IntelliDrive, it is possible to use the data sets collected by the program to complement the scantiness of regular state-maintained traffic count and crash data.

References

1. Lerner, N., J. Jenness, J. Singer, S. G. Klauer, S. Lee, M. Donath, M. Manser, and M. Ward. *An Exploration of Vehicle-Based Monitoring of Novice Teen Drivers: Draft Report.* Virginia Tech Transportation Institute, Blacksburg, Va., 2008.
2. Hickman, J. S., R. R. Knipling, R. L. Olson, M. C. Fumero, M. Blanco, and R. J. Hanowski. *Phase I—Preliminary Analysis of Data Collected in the Drowsy Driver Warning System Field Operational Test: Task 5, Preliminary Analysis of Drowsy Driver Warning System Field Operational Test Data.* NHTSA, 2005.
3. University of Michigan Transportation Research Institute. *Automotive Collision Avoidance System Field Operational Test Report: Methodology and Results.* Report DOT HS 809 900 NHTSA, 2005.
4. University of Michigan Transportation Research Institute. *Road Departure Crash Warning System Field Operational Test: Methodology and Results.* NHTSA, 2006.
5. Dingus, T. SHRP 2 S05 Status Update and Current Design Plans. Presented at SHRP2 Safety Research Symposium, Washington, D.C., 2008. http://onlinepubs.trb.org/onlinepubs/shrp2/TomDingusSymposiumPresentation.pdf. Accessed May 17, 2011.
6. Blanco, M., J. S. Hickman, R. L. Olson, J. L. Bocanegra, R. J. Hanowski, A. Nakata, M. Greening, P. Madison, G. T. Holbrook, and D. Bowman. *Investigating Critical Incidents, Driver Restart Period, Sleep Quantity, and Crash Countermeasures in Commercial Vehicle Operations Using Naturalistic Data Collection.* FMCSA, 2008.
7. Fitch, G., H. Rakha, M. Arafeh, M. Blanco, S. Gupta, R. Zimmerman, and R. Hanowski. *Safety Benefit Evaluation of a Forward Collision Warning System: Final Report.* Report DOT HS 810 910. NHTSA, 2008.
8. International Transport Forum. *Towards Zero: Ambitious Road Safety Targets and the Safe System Approach.* Organisation for Economic Co-operation and Development, 2008. www.internationaltransportforum.org/Pub/pdf/08TowardsZeroE.pdf. Accessed May 17, 2011.
9. Research and Innovative Technology Administration, U.S. Department of Transportation. About Intelligent Transportation Systems. www.its.dot.gov/its_program/about_its.htm. Accessed May 17, 2011.

APPENDIX A

Project 2 Data Dictionary

Motivation and intent were important criteria. The motivations behind the driver's actions, when able to be clearly determined, weighed more heavily than actual behaviors. Events were coded according to judgment with respect to why the alert went off—as though one was looking through the eyes of the driver.

Time of Day

Dusk was difficult to classify as either day or night; this classification was subjective. Events were coded to reflect whichever the clip was most similar to, day or night.

- 0 = Day.
- 1 = Night.

Road Condition

Glare and reflection helped determine whether the road was dry or wet.

- 0 = Dry.
- 1 = Wet. (Any moisture on the road led to the classification as wet; there did not need to be standing water. The road was classified as wet if it was wet from snow but not snow covered.)
- 2 = Snow covered (Snow covered would have included ice covered if it was observed, but it was never observed. If any portion of the road, including turn lanes, was covered in snow, then the classification was snow covered.)

Precipitation

Spots on the windshield or wiper activity helped determine if there was in fact precipitation.

- 0 = None.
- 1 = Rain. (Light rain and drizzle were classified as rain, as were downpours.)
- 2 = Snow. (This category included sleet. Several cues helped indicate if the precipitation was in fact snow; snow tended to be larger and fall more slowly than rain, it looked like white flurries and was also present on the ground, reinforcing the classification as snow. Also, precipitation that occurred in December through February was assumed to be snow, not rain. Snow could be coded in other months, but the assumption that the precipitation was snow was not as strong.)

Location of Eyes at Time of the Alert

This category was coded at the actual time of the alert, when the radar display showed 100 for the alert level. Eye location was coded by what the reviewers could see of the driver's eyes at the time of the alert, even if they could not see the eyes preceding the alert. Reviewers coded the location of the driver's eyes even if they could see only one eye because it was assumed that the driver's eyes moved in parallel. Because of the absence of an eye-tracking camera and the limitations of the face camera, some ambiguity often existed about where the drivers were looking. Reviewers needed to be confident in the location of the driver's eyes in order to code as a specific location. In many instances, the reviewers were confident that the driver's eyes were not looking forward but could not tell specifically where the eyes were looking. These instances were coded as 8s. One such example is when the driver appeared to be looking at the camera. In this situation, it was difficult to determine if the driver was looking at the camera intentionally, glancing out the corner, or looking slightly out the left window; therefore, it was coded as an 8. The determination of whether glances were still forward or if they were away was also difficult and subjective. Reviewers agreed on an area or "box" that they considered to be looking forward; this allowed for slight glances but even many scans across the forward scene were considered glances away. This process defined "looking forward" narrowly and essentially meant straight forward.

Glances toward the right of the forward scene, the right area of the windshield, were glances away and were coded as 8s.

- 0 = Looking forward at forward scene. (Looking forward included looking at the head-up display [HUD].)
- 1 = Left outside mirror or window.
- 2 = Looking over left shoulder. (The driver's gaze needed to look over the driver's shoulder, although the driver's chin did not necessarily need to cross over the driver's shoulder.)
- 3 = Right outside mirror or window.
- 4 = Looking over right shoulder. (The driver's gaze needed to look over the driver's shoulder, although the driver's chin did not necessarily need to cross over the driver's shoulder.)
- 5 = Head down, looking at instrument panel or lap area. (Looking at the HUD was not considered part of the instrument panel.)
- 6 = Head down, looking at center stack console area. (Console means the area in which the stereo, thermostat, and clock are located.)
- 7 = Driver wearing sunglasses or glasses with glare. (Glare prohibited the ability to classify where the eyes were looking. In some instances, drivers were wearing sunglasses, but the reviewers thought that they could confidently identify the location of the drivers' eyes. In these instances, eye location was recorded.)
- 8 = Cannot accurately evaluate eye location. (An 8 was chosen when the reviewer was unsure of the eye position or classification within a reasonable level of confidence but not because of glasses. Typically, the reviewer could see the actual eye but could not determine where the gaze was directed. Eyes in transition were often coded as 8 because it was unclear where the driver's gaze was at that particular moment.)
- 9 = Other. (For example, the driver may clearly be looking at the passenger side floor. When a glance was coded as other, the location was noted in the notes section. The most common position recorded as other was the rearview mirror.)

Location of Eyes During the Last Nonforward Glance and Time from the Last Nonforward Glance

If the driver's eyes were on the forward scene at the moment of the alert but had looked away during some portion of the clip previous to the alert, this location was recorded. The reviewers also recorded the amount of time between when the driver's gaze began to return to the forward scene and the moment of the alert, according to the radar display showing alert level 100. We did not count the actual moment of the alert; the time represents the time between the change in gaze and the alert. Time was recorded in 10ths of seconds. If the driver was always looking forward, then the time from the last nonforward glance was left null because that category was not applicable. If the driver was looking away 0.1 s before the alert and then was looking forward at the time of the alert, the time from the last nonforward glance was recorded as 0. If the driver's eyes were not visible, typically because of glare, for any portion of the clip, the location was coded as a 7 because one could not be certain there was not a glance away. The only exception to this rule was when the reviewers could not see the driver's eyes and then the eyes became visible so that the reviewers could see the eyes and there was a glance away before the alert. This situation negates the fact that the reviewers could not see the eyes at the beginning of the clip, because there was a nonforward glance after the portion during which the eyes were unclassifiable. If the eyes were then unclassifiable again, before the alert but after the glance, the eyes were coded as a 7 because the reviewers could not be certain what happened during that portion of the clip. If the location of one eye could be determined but not that of the other eye, location was still coded. Reviewers were confident in coding eye position when only one eye could be seen because eyes normally move in parallel. If the driver's eyes glanced away before the alert and were in transition at the time of the alert, the last nonforward glance code reflected where they were looking at the time of the alert, not where they had previously been looking. For more details on eye location, see the information on eye location at the time of the alert. The criteria for classifying a glance as a specific location are the same as the criteria for eye location at the time of the alert.

- 0 = Always looking forward at the forward scene. (Looking forward includes looking at the HUD.)
- 1 = Left outside mirror or window.
- 2 = Looking over left shoulder.
- 3 = Right outside mirror or window.
- 4 = Looking over right shoulder.
- 5 = Head down, looking at instrument panel or lap area.
- 6 = Head down, looking at center stack console area. (Console means the area in which the stereo, thermostat, and clock are located.)
- 7 = Driver wearing sunglasses or glasses with glare. (Glare prohibited the ability to classify where the eyes are looking.)
- 8 = Cannot accurately evaluate eye location. (An 8 was chosen when the reviewer was unsure of the eye position or classification within a reasonable level of confidence but not because of glasses. Typically, the reviewer could see the actual eye but could not determine where the gaze was directed. Eyes in transition were often coded as 8 because it was unclear where the driver's gaze was at that particular moment.)
- 9 = Other. (For example, the driver might clearly be looking at the passenger side floor. When a glance was coded

as other, the location was noted in the notes section. The most common position recorded as other was the rearview mirror.)

Eyes on Task at Time of the Alert

- 0 = No. (The classification of no was used only when the reviewer could confidently determine that the driver's eyes were off the task of driving at the time of the alert [e.g., the driver was looking at a friend or the stereo system].)
- 1 = Yes. (The classification of yes does not mean looking forward; it means that the driver's eyes were on the task of driving.)
- 2 = Cannot determine. (For instance, the driver was wearing glasses with glare or reviewers could not see the driver's eyes for some other reason. This classification was also used when reviewers could not tell if the eye location was on task [e.g., the driver was looking out the window, but it was unclear whether the driver was looking at traffic or at a fancy building that was distracting the driver's attention]. In any case, reviewers did not know whether the driver was on task.)

Eyes in Transition

To classify the driver's eyes as in transition, they must have been in transition at the time of the alert and must have started the transition at least 0.1 s before the alert. The eyes could not be at the beginning of a transition or the end of one; they must have been in the transition at the time of the alert.

- 0 = No.
- 1 = Yes, toward forward scene.
- 2 = Yes, away from forward scene.
- 3 = Cannot tell. (Cannot tell was selected when the driver was wearing sunglasses or reviewers could not see the driver's eyes for some other reason; therefore, it was uncertain whether they were in transition.)

Visual Response to Alert and Time to Visual Response

Reviewers coded the time that it took the driver to initiate a visual response to the alert by filling in the number of 10ths of a second the response took. The time counted was the time between the alert and when the look was initiated, not including the moment of the alert or the moment of response. If the response was initiated within 1.0 s, then the driver was considered to have looked in response to the alert. The amount of time it took to look in response was always recorded for applicable situations, even if this was greater than 1.0 s. If the driver was already looking at the road and continued to look forward, the code was null (not applicable). If reviewers were not sure of the location of the driver's eyes, then the time to visual response was left as null. The time to visual response was recorded for Week 1, even though there was no alert to which to respond. The rationale for coding this was that a baseline would provide an idea of what a normal time to visual response was compared with the time to response with an alert.

- 0 = Looked in response. (The driver initiated a look in response to the alert within 1.0 s. Glances qualified as a look in response.)
- 1 = Did not look in response to alert. (The driver did not look within 1.0 s of the alert.)
- 2 = NA. (This option was always used for Week 1 because no alert occurred during that week; thus, this category could not be coded. This option was also selected when the driver was already looking forward at the time of the alert; this category was not applicable.)
- 3 = Cannot tell. (Because the driver was wearing sunglasses or other glasses with glare, reviewers could not tell where the driver's eyes were looking.)

Visual Occlusion

Occlusion was coded with regard to the driver as well as to the reviewers. For instance, heavy rain or bright sun might have occluded the scene for both parties, whereas blurry video occluded the scene only for reviewers. The occlusion did not necessarily have to impact the reviewers' ability to code the scene.

- 0 = None.
- 1 = Sun or headlight glare. (This classification includes when the scene was whitewashed from the sun. Only headlight glare was included in this section; taillight glare was coded as other.)
- 2 = Other, specified in notes section. (The most common entry was taillight glare.)

Startle Response

Coding the startle response was subjective, and the classification as such was often hotly debated. The driver had to be visibly rattled. The driver's startle was observed by body response, dialogue, or both.

Cursing was not sufficient to be coded as startle, because it might have resulted from anger or frustration, not startle. This category tried to capture startle either to the situation or to the alert.

- 0 = No.
- 1 = Yes.

Steering in Response

- 0 = No steering in response to alert. (Small, jerky reactions or slight wiggling in response to the alert or to the situation was classified as a 0 and not considered steering.)
- 1 = Driver steered partially or fully in response to the alert. (Steering, for review purposes, was an evasive maneuver in an attempt to prevent striking a vehicle; thus there must have been a significant amount of steering.)

Hand Location at Time of Alert

Because both hands were not often visible, reviewers coded what could confidently be inferred from the scene. At times playing the video farther helped determine what was ambiguous in a still frame at the time of the alert. For instance, at the time of the alert there may have been a small blur near the steering wheel. On continuation of the video the blur may have moved and come into view as a hand.

- 0 = Cannot see the position of either hand or cannot determine the position of either hand. (Reviewers coded 0 if a hand could be seen but they could not tell if it was on the wheel.)
- 1 = At least one hand on the steering wheel. (Reviewers used this code when the position of one hand could not be determined but they could see that at least one hand was on the steering wheel.)
- 2 = Both hands on the steering wheel.
- 3 = At least one hand off the steering wheel. (This code was used when the position of one hand could not be determined but at least one hand was clearly off the steering wheel.)
- 4 = One hand on, one hand off the steering wheel. (A 4 was classified when reviewers could clearly see both hands and one was on the wheel and the other was off.)
- 5 = Both hands off the steering wheel. (A 5 was classified when reviewers could clearly see both hands and both were off the wheel.)

Road Geometry

- 0 = Straight.
- 1 = Curve. (A curve must be of some substance to be considered a curve.)
- 2 = Approaching curve. (The classification of approaching a curve constituted situations in which the driver was almost in a curve, not when there was simply a curve in the distance.)
- 3 = Lane shift. (Road geometry was classified as a lane shift when there was a change in the lane structure, for instance, when a new lane was created. If the new lane was a turn lane, it was not classified as a lane shift, because the scenario should have covered the fact that there was a turn lane if that were relevant. Lane shifts were also considered shifts in traffic patterns, such as lane shifts in construction areas.)

Secondary Driving Behaviors

Audio was used to assist in coding whenever possible. For instance, reviewers may have heard the radio station change and seen the driver look at the console; this would indicate in-car system use. The default for nondriving behaviors was none, coded as 0.

Cell Phone

- 10 = Conversation, in use. (Conversation could be coded for listening, talking, or both while using the cell phone.)
- 11 = Reaching for phone. (This classification refers to when the driver reached for the handheld phone to speak on that phone. If the driver reached for the phone simply to answer the phone and talk on the headset the driver was wearing, the classification was other. Simply answering the phone involves far less physical activity by the driver than reaching for the phone and holding it during a conversation.)
- 12 = Dialing phone.

Headset, Hands-Free Phone

- 20 = Conversation. (This classification was selected when reviewers could tell that the driver was in a conversation.)
- 21 = Reaching for headset.
- 22 = Unsure of activity level. (The driver was wearing a headset, but it was not clear whether the headset was in use. The driver may have been listening to someone or wearing the headset in case there was an incoming call.)

Eating

- 30 = Highly involved. (High involvement includes such things as eating a burger or unwrapping food.)
- 31 = Low involvement. (Low involvement includes such things as eating candy or grabbing chips.)

Drinking

- 40 = Highly involved. (High involvement includes situations in which the driver was trying to open a straw or bottle or was blowing on a hot drink.)
- 41 = Low involvement. (Low involvement includes situations in which the driver was sipping a drink or drinking without looking.)
- 50 = Conversation. (The driver and someone in the car were carrying on a conversation. The driver can be listening during the clip, talking during the clip, or both.)

- 60 = In-car system use. (The driver was actively adjusting something. For example, the driver was not just listening to the stereo but also adjusting it. The car lighter was coded under the smoking section.)

Smoking

- 70 = Lighting. (This classification includes the in-car lighter.)
- 71 = Reaching for cigarettes or lighter. (This classification includes the in-car lighter.)
- 72 = Smoking.

Grooming

- 80 = Highly involved. (High involvement includes applying makeup or brushing hair.)
- 81 = Low involvement. (Low involvement includes scratching or running one's fingers through his or her hair.)
- 90 = Other/multiple behaviors, specified in notes section. (Such behaviors might include whistling or classifications that reviewers were unsure of [e.g., if the driver's lips were moving but there was no audio, the behavior might be singing or conversation].)

Alert Classifications

- 1 = False alarm. (Two reasons were given for the classification of an event as a false alarm. The first was that the target was out of the lane throughout the episode or it was off the roadway, including both vehicles and stationary objects. The second was that the kinematics of the scenario did not make sense [e.g., when there was a lead vehicle deceleration error resulting in an alert].)
- 2 = True/nuisance alert. (The target had to be a vehicle that was in the driver's path for at least a portion of the episode. The event may have been viewed by the driver as either a needed alert or a nuisance alert.)
- 3 = Instigated alert. (An alert was classified as instigated if the driver deliberately drove in such a way as to provoke an alert. This does not apply to when the driver was in adaptive cruise control [ACC] and was trying to see if the system would brake sufficiently; this was testing the ACC system rather than the forward crash warning [FCW] system.)

Target

A vehicle was considered in path if even a small portion—for example, a rear bumper—of the lead vehicle remained in the driver's lane at the time of the alert. Vehicles with both sport-utility vehicle (SUV) and car characteristics, a small SUV, were classified as a car. The classification as a car for these SUVs was because the body of the vehicle, in respect to how the driver may follow it, is more in line with that of a car than with that of an SUV.

- 0 = Vehicle in path—car.
- 1 = Vehicle in path—pickup, van, or SUV.
- 2 = Vehicle in path—other (e.g., motorcycle, semitrailer, commercial vehicle).
- 3 = Vehicle out of path—car.
- 4 = Vehicle out of path—pickup, van, or SUV.
- 5 = Vehicle out of path—other (e.g., motorcycle, semitrailer, commercial vehicle).
- 6 = Construction. (This includes all equipment associated with construction [e.g., barrels, cones, and construction vehicles].)
- 7 = Discrete roadside object. (This classification includes signposts, light poles, trees, fire hydrants, and mailboxes.)
- 8 = Overhead items. (This classification includes such items as overhead signs and bridges.)
- 9 = Bridge support.
- 10 = Guardrail/Jersey barrier.
- 11 = Other, to be specified in notes section.

Target Stationary or Moving

This category was coded using both visual cues and radar data. For vehicles that appeared to be either stopped or slowly moving, visual cues were used if the cues were clear—for example, lane markings—but otherwise, radar data was used to determine the classification.

- 1 = Stationary. (The target must have had a velocity of less than or equal to 1.3 m/s to be classified as stationary.)
- 2 = Moving.
- 3 = Stationary potential threat. (The classification of 3 was made when the target was a vehicle that was stopped [velocity of less than or equal to 1.3 m/s] but had the potential to move at any moment [e.g., stopped cars with drivers in them].)

Forward Conflict Scenario

The same scenarios are used for classifying supplementary scenarios as needed.

The supplementary scenario column also has the option of 0 = none. Supplementary scenarios are available in case there is another possible scenario or if the situation resembles another scenario that may be of interest to people working with that specific type of scenario.

Out of Host's Path

- 100 = False alarm. (Two reasons were given for the classification of an event as a false alarm. The first was that the target was out of the lane throughout the entire episode or it was off the roadway, including both vehicles and stationary objects. The second was that the kinematics of the scenario did not make sense [e.g., when there was a lead vehicle deceleration error resulting in an alert].)

In Host's Path

The following scenarios were initiated by the host vehicle. Note, however, that this and the following subcategories apply to the stereotypes of scenarios and may not apply to all cases. Nevertheless, these assumptions were used during some analyses.

- 200 = Host tailgating. (Tailgating was coded even if the lead vehicle was on an exit or entrance ramp. The criterion for tailgating was a headway of 0.8 s or less. If the host was using ACC and the headway matched the criterion for tailgating, it was not considered tailgating, because the system rather than the driver was controlling the headway.)
- 210 = Host approaches an accelerating vehicle. (This typically occurred when the host misjudged the lead vehicle's acceleration and the host accelerated too fast on a vehicle that was accelerating as well or if the host was approaching a vehicle as a traffic light turned green.)

The following scenarios played out naturally with neither the host nor the lead vehicle changing accelerations.

- 220 = Host approaches slower vehicle that is traveling at constant speed. (The criteria for this classification were as follows: no brake lights were visible; lead vehicle was always going slower than the host while the target was detected by ACAS; and the target did not appear to be decelerating, either visibly or using the radar data. Slight fluctuations in speed are normal, but the overall speed must have been fairly constant. Typically, the lead vehicle was in the distance at the beginning of the clip and then moved into view as the host gained on it.)
- 230 = Host approaches lead vehicle that is already stopped. (The lead vehicle must have been traveling less than 1.3 m/s during all portions of the clip in which it was acquired as a target. Please see the notes on Target Stationary or Moving for more details on how the determination of a target as stationary or moving was made.)

The following scenarios were initiated by the lead vehicle.

- 240 = Host follows a lead vehicle that decelerates to an unpredictable stop. (This scenario was classified as such when the traffic stopped but the host driver may have thought traffic would just slow and not stop; no stop sign or traffic signal was in view. The event was coded as 240 even if the lead vehicle had not stopped by the time of the alert, as long as the lead stopped within 12.5 s of the alert [the regular viewing time plus a 10-s continuation]. For coding purposes, stop means that the lead came close to 0 mph [less than or equal to 1.3 m/s].)
- 250 = Host follows a lead vehicle that decelerates to a predictable stop. (This scenario occurred when there was a traffic light, stop sign, visibly stopped traffic, or the car in front of the lead had its brake lights on. These cues made it so that the host could logically anticipate that the lead would stop. The event could be coded as 250 even if the lead had not stopped by the time of the alert, as long as it had stopped within 12.5 s of the alert [the regular viewing time plus a 10-s continuation]. For this scenario, a stop means coming close to 0 mph [less than or equal to 1.3 m/s].)
- 260 = Host follows a lead vehicle that decelerates in an unpredictable manner but does not stop. (For this scenario the brake lights on the lead vehicle had to be visible or the lead was noticeably slowing even if the brake lights could not be seen but the reason for slowing was not visible or not predictable [e.g., a cat crossing the street]. The classification as 260 was the default over a 270.)
- 270 = Host follows a lead vehicle that decelerates in a predictable manner but does not stop. (For this scenario the brake lights on the lead had to be visible or the lead was noticeably slowing even if the brake lights could not be seen. The code of 270 was selected only if it was clear from the cues available why the lead needed to decelerate [e.g., a slow moving piece of farm equipment was ahead or the car in front of the lead had its brake lights on].)

Transitional Host Path: One or Both Vehicles Change Lanes

The following scenarios were initiated by the host vehicle.

- 300 = Host cuts behind a lead vehicle. (The code 300 was used when the host cut into another lane closely behind the lead vehicle. This maneuver could not be part of a two-lane pass. See the description of a two-lane pass for more details.)
- 310 = Host performs two-lane pass in order to pass. (This scenario involves the host's cutting behind the lead in the process of making a two-lane pass maneuver: the host crossed the middle lane in order to enter two lanes over from the host's lane of origin. If two alerts occurred for the same smooth transition during a two-lane pass, reviewers coded the first alert as 310 and did not code the second alert in the series. Reviewers commented on the scenario in the notes section, labeling the second alert as such in the notes section. A two-lane pass does not require two alerts.)

- 315 = Host performs two-lane pass to exit or to make a turn that is carried out within the time of the clip. (The scenario was classified as such when the host cut behind a lead in the process of making a two-lane pass maneuver: the host crossed the middle lane in order to enter a lane two lanes over from the host's lane of origin. If two alerts occurred for the same smooth transition during a two-lane pass, reviewers coded the first alert as 315 and did not code the second alert in the series. Reviewers commented on the scenario in the notes section, labeling the second alert as such in the notes section. A two-lane pass does not require two alerts.)
- 320 = Host changes lanes to pass lead vehicle. (The target for this alert type was the vehicle in the host's original lane. For this event, a pass was coded even if the host had not passed at the time of the alert, as long as the host was acting like a pass was planned; for instance, the host was checking the mirrors and accelerating. If the pass was aborted and not carried out during the entirety of the extended clip [10 additional seconds], reviewers classified the event as a pass and marked the scenario as aborted. Reviewers added notes when necessary. If the host was passing a vehicle that was waiting to turn, reviewers coded whichever event seemed to be the reason for the alert and the other event could be chosen in the supplementary scenario section.)
- 330 = Host enters turn lane or other dedicated lane (e.g., exit lane) while approaching a lead vehicle in the new lane. (In this case, the target was the vehicle in the host's new lane.)
- 340 = Host enters a turn lane or other dedicated lane (e.g., exit lane) while approaching lead vehicle in original travel lane. (In this case, the target was the vehicle in the host's original lane.)
- 350 = Host weaves in order to avoid an obstacle but does not completely change lanes. (An event was coded as 350 if the host did not completely change lanes in the process of avoiding an obstacle. This was not a planned maneuver; there was no turn signal or other indications of a lane change until the last moment—an evasive reaction.)
- 355 = Host weaves in order to avoid an obstacle and completely changes lanes. (An event was coded as 355 if the host changed lanes completely in the process of avoiding an obstacle. The motivation for the lane change has to have been to avoid something in the original lane. This was not a planned maneuver; there was no turn signal or other indications of a lane change until the last moment—an evasive reaction.)

The following scenarios were initiated by the lead vehicle.

- 360 = Lead vehicle changes lanes and cuts in front of host. (The main precipitant of this scenario was the lead cutting in front of the host. This occurred when the lead was in a lane parallel to the host's and then cut in or merged close to the front of the host's vehicle.)
- 365 = Lead vehicle is forced to merge in front of host. (An event was classified as 365 when the lead needed to merge into the host's lane because an entrance ramp was ending or a lane was ending for another reason.)
- 370 = Lead executes a two-lane pass. (This scenario was coded when the lead passed from a lane parallel to the host's, across the host's path, and then over to the parallel lane on the other side of the host's car. The lead was only in the host's path momentarily.)
- 380 = Lead vehicle in adjacent lane weaves or encroaches unintentionally/unknowingly into host's lane. (The classification of an event as 380 refers to events in which the lead entered the host's lane unintentionally and momentarily. The brief entry into the host's lane by the lead vehicle caused the alert.)
- 390 = A vehicle crossing the host's roadway travels straight across host's path. (The scenario is characterized by a vehicle driving straight across the host's path. The target vehicle did not remain in the host's path; typically, the radar hit the side of the crossing vehicle, and the intersecting paths were perpendicular to each other.)
- 400 = A vehicle makes a left turn across the host's path in order to travel in a direction other than that of the host. (In this scenario the crossing vehicle turned left either from a perpendicular street or from a parallel lane traveling in the opposite direction of the host's lane. If the vehicle turned from a side street, it crosses the host's path and then continues into a parallel lane traveling in the opposite direction. If the vehicle turned left from a parallel lane, it crossed the host's path and continued onto a perpendicular street. The vehicle crossed the radar path with primarily a perpendicular angle, but the angle may be more steeply tilted than when the vehicle was simply crossing straight across the host's path in other scenarios.)
- 410 = A vehicle entering the host's roadway crosses host's path and moves into a lane parallel to the host's lane in the same direction.
- 420 = A vehicle pulls out from a side street or driveway and pulls in front of host and into the host's lane. (This scenario occurred when the lead started out perpendicular to the host car and turned into and in front of host's car.)
- 430 = Lead changes lanes out of host's lane. (This scenario developed when the lead departed the host's lane but the situation was not covered by another described scenario. One instance of this situation is when the host could logically anticipate that the lead vehicle would change lanes because of the lead's turn signal or another indication and, therefore, the host gained on the lead, sounding an alert.)

- 440 = Lead leaves host's lane to enter turn lane or other dedicated lane (e.g., exit ramps and turn lanes).
- 450 = Lead turns left from host's travel lane. (The target in this conflict is the turning car. The lead must have begun the turn, even slightly, for the scenario to be coded as 450. If the lead's turn signal was on but the turn had not yet been initiated, the event was coded as a predictable deceleration and given a supplementary scenario of 450.)
- 460 = Lead turns right from host's travel lane. (The target in this conflict is the turning car. The lead must have begun the turn, even slightly, for the scenario to be coded as 460. If the lead's turn signal was on but the turn had not yet been initiated, the event was coded as a predictable deceleration and given a supplementary scenario of 460.)

Scenario Completion

- 0 = Completed.
- 1 = Aborted.

Supplementary Scenario

This category allowed reviewers to select a second scenario to which the situation could be attributed. The supplementary scenario could also be a scenario that preceded or followed the imminent scenario but may have contributed to the development of the actual scenario. This category was designed to be a reference list for people interested in scenarios that could have been classified in other ways. The category also allowed reviewers to indicate two scenarios when the primary scenario was difficult to determine. If reviewers did not indicate a supplementary scenario, a 0 was entered for none.

Notes

A notes section recorded any unusual events or ambiguous situations not covered by categories for a particular question. This section also contains general notes on the clip if there was anything significant taking place that was not adequately covered by the coding process. See section Annex A-1 below for further details.

Annex A-1

The following are examples of items that are captured in the notes section, although other, unforeseen events are also noted.

Visual Occlusion

Rear taillights, glare from rain and wetness on the road, blurry video, dirty windshield, temporary incapacitation, sneezing, flying debris, faulty wiper or defroster, and an object in or over the eyes.

Nondriving Behaviors

Whistling, two or more behaviors, if there is no audio and the driver is clearly talking or singing but reviewers could not tell which, attempting to avoid an insect in the car, adjusting mirrors, reading a map, reading other materials, checking a watch, or yawning.

Target

Shadows and embankments.

APPENDIX B

Project 5 Data Dictionary

Coding Key for RDCWS Alerts

LDWS and CSWS alerts were coded using different criteria; driver behavior, however, was rated in the same way for each set of alerts. Those criteria are listed first below. Specific categories regarding scenario details were different for each system. Each of the scenario coding keys is described after the driver behavior key.

Driver Behaviors

Location of the driver's eyes during the last nonforward glance and time from the last nonforward glance.

If the driver's eyes were on the forward scene at the moment of the alert but had looked away during some portion of the clip before the alert, this location was recorded. Reviewers also recorded the amount of time between when the driver's gaze began to return to the forward scene and the moment of the alert, according to the driver-vehicle interface (DVI) display on the computer monitor. The actual moment of the alert was not counted; the time represents the time between the change in gaze and the alert. Time was recorded in 10ths of seconds. If the driver was always looking forward, the time from the last nonforward glance was left null because that category was not applicable. If the driver was looking away 0.1 s before the alert and then was looking forward at the time of the alert, the time from the last nonforward glance was recorded as 0. If the eyes were not visible, typically because of glare, for any portion of the clip, the location was coded as 9 because one could not be certain there was not a glance away. The only exception to this rule is when reviewers could not see the driver's eyes and then the eyes became visible so that reviewers could see the eyes and there was a glance away before the alert. This situation negates the fact that reviewers could not see the eyes at the beginning of the clip, because there was a nonforward glance after the portion during which the eyes were unclassifiable. If the eyes were unclassifiable again, before the alert but after the glance, the eyes were coded as 9 because reviewers could not be certain what happened during that portion of the clip. If one eye location could be determined and the other eye's location could not, location was still coded. Reviewers were confident in coding eye position when only one eye could be seen because normally eyes move in parallel. If the driver's eyes were away before the alert and in transition at the time of the alert, the last forward glance code reflected where they were looking at the time of the alert, not where they had previously been looking. For more details on eye location, see the information on Location of Eyes at Time of Alert. The criteria for classifying a glance as a specific location are the same as the criteria for eye location at the time of the alert.

- 0 = Always looking forward at the forward scene.
- 1 = Left outside mirror or window.
- 2 = Looking over left shoulder.
- 3 = Right outside mirror or window.
- 4 = Looking over right shoulder.
- 5 = Interior rearview mirror.
- 6 = Head down, looking at instrument panel or lap area.
- 7 = Head down, looking at center console area. (Console means the area where the stereo, thermostat, and clock are located.)
- 8 = Driver wearing sunglasses or glasses with glare. (Glare prohibited the ability to classify where the eyes were looking.)
- 9 = Cannot accurately evaluate eye location. (This was coded as 9 when reviewers were unsure of the eye position or classification within a reasonable level of confidence, although not because of glasses. Typically, reviewers could see the actual eye but could not determine where the gaze was directed. Eyes in transition were often coded as 9 because it was unclear where the driver's gaze was at that particular moment.)
- 10 = Other. (For example, the driver may clearly be looking at the passenger side floor. When a glance was coded as other, the location was noted in the notes section. The most common position recorded as other was the rearview mirror.)

Location of Eyes at Time of Alert

This category was coded at the actual time of the alert. Eye location was coded by what reviewers could see of the driver's eyes at the time of the alert, even if they could not see the eyes before the alert. Reviewers coded the location of the driver's eyes even if they could see only one eye because it was assumed that the driver's eyes moved in parallel. Because of the absence of an eye-tracking camera and the limitations of the face camera, there was often some ambiguity about where the drivers were looking. Reviewers needed to be confident in the location of the driver's eyes to code as a specific location. In many instances, reviewers were confident that the driver's eyes were not looking forward but could not tell specifically where the eyes were looking. These instances were coded as 9s. One such example is when the driver appeared to be looking at the camera. In this situation, it was difficult to determine if the driver was looking at the camera intentionally, glancing out the corner, or looking slightly out the left window; therefore, it was coded as 9. Another example is when the driver was looking toward the curve that elicited the alert. The exact location of the driver's eyes could not be determined in these instances, although a notation was made in the notes field. The determination of whether glances were still forward or if they were glances away was also difficult and subjective. Reviewers agreed on an area or box they considered to be looking forward; this allowed for slight glances but even many scans across the forward scene were considered glances away. This process defined looking forward narrowly and essentially as meaning straight forward. Glances toward the right of the forward scene, the right area of the windshield, were glances away and were coded as 9s.

- 0 = Looking forward at forward scene. (Looking forward included looking at the head-up display [HUD].)
- 1 = Left outside mirror or window.
- 2 = Looking over left shoulder. (The driver's gaze needed to look over the driver's shoulder but the driver's chin did not necessarily need to cross over the driver's shoulder.)
- 3 = Right outside mirror or window.
- 4 = Looking over right shoulder. (The driver's gaze needed to look over the driver's shoulder but the driver's chin did not necessarily need to cross over the driver's shoulder.)
- 5 = Interior rearview mirror.
- 6 = Head down, looking at instrument panel or lap area. (Looking at the HUD was not considered part of the instrument panel.)
- 7 = Head down, looking at center console area. (Console means the area where the stereo, thermostat, and clock are located.)
- 8 = Driver wearing sunglasses or glasses with glare. (The glare prohibited the ability to classify where the eyes were looking. In some instances, drivers were wearing sunglasses, but reviewers believed that they could confidently identify the location of the drivers' eyes. In these instances, eye location was recorded.)
- 9 = Cannot accurately evaluate eye location. (The code 9 was chosen when reviewers were unsure of the eye position or classification within a reasonable level of confidence but not because of glasses. Typically, reviewers could see the actual eye but could not determine where the gaze was directed. Eyes in transition were often coded as 9 because it was unclear where the driver's gaze was at that particular moment.)
- 10 = Other. (For example, the driver may clearly be looking at the passenger side floor. When a glance was coded as other, the location was noted in the notes section. The most common position recorded as other was the rearview mirror.)

Eyes on Task at Time of Alert

- 0 = No. (The classification of no was used only when reviewers could confidently determine that the driver's eyes were off the task of driving at the time of the alert [e.g., the driver was looking at a friend or the stereo system].)
- 1= Yes. (The classification of yes does not mean looking forward; it means that the driver's eyes were on the task of driving. Looking at the instrument panel, for example, was considered on task.)
- 2 = Cannot determine (For instance, the driver was wearing glasses with glare or reviewers could not see the driver's eyes for some other reason. This classification was also used when reviewers could not tell if the eye location was on task. For instance, the driver was looking out the window [e.g., toward a curve in the road], but it was unclear whether the driver was looking at the road and traffic or at a fancy building that was distracting the driver's attention. In any case, reviewers did not know whether the driver was on task.)

Eyes in Transition

To classify the eyes as in transition, the driver's eyes must have been in transition at the time of the alert and must have started the transition at least 0.1 s before the alert. The eyes could not be at the beginning of a transition or the end of one; they must have been in the transition at the time of the alert.

- 0 = No.
- 1 = Yes, toward forward scene.
- 2 = Yes, away from forward scene.
- 3 = Cannot tell. (Cannot tell was selected when the driver was wearing sunglasses or reviewers could not see the driver's eyes for some other reason; therefore, researchers were uncertain whether the eyes were in transition.)

Visual Response to Alert and Time to Visual Response

If the driver initiated a visual response to the alert, reviewers coded the time it took for the response by recording the number of 10ths of a second. The time counted was the time between the alert and when the look was initiated, not including the moment of the alert or the moment of response. If the response was initiated within 1.0 s, the driver was considered to have looked in response to the alert. The amount of time it took to look in response was always recorded for applicable situations, even if this was greater than 1.0 s. If the driver was already looking at the road and continued to look forward, the code was null (not applicable). If reviewers were not sure of the location of the driver's eyes, the time to visual response was left as null. The time to visual response was recorded for Week 1, even though there was no alert to which to respond. The rationale for coding this was that a baseline would provide an idea of what a normal time to visual response was compared with the time to response with an alert.

- 0 = Looked in response. (The driver initiated a look in response to the alert within 1.0 s. Glances qualified as a look in response.)
- 1 = Did not look in response to alert. (The driver did not look within 1.0 s of the alert.)
- 2 = NA. (This option was always used for Week 1 because there was no alert during Week 1; thus we could not code this category, although we still coded the time to visual response. This option was also selected when the driver was already looking forward at the time of the alert.)
- 3 = Cannot tell. (The driver was wearing sunglasses or other glasses with glare, and reviewers could not tell where the driver's eyes were.)

Visual Occlusion

Occlusion was coded with regard to the driver as well as to reviewers. For instance, heavy rain or bright sun might have occluded the scene for both parties, whereas blurry video occluded the scene only for the reviewer. The occlusion did not necessarily have to impact the reviewers' ability to code the scene.

- 0 = None.
- 1 = Sun or headlight glare. (This classification includes when the scene was whitewashed from the sun. Only headlight glare was included in this section; taillight glare was coded as other.)
- 2 = Other, specified in notes section. (The most common entry was taillight glare.)

Startle Response

This was subjective and the classification as such was often hotly debated. The driver had to be visibly rattled. The driver's startle was observed by body response or dialogue or both.

Cursing was not sufficient to be coded as startle, because it may have resulted from anger or frustration, not startle. This category tried to capture startle to either the situation or the alert.

- 0 = No.
- 1 = Yes.

Steering in Response

- 0 = No steering in response to alert. (Small, jerky reactions or slight wiggling in response to the alert or to the situation was classified as 0 and was not considered steering.)
- 1 = Driver steered partially or fully in response to the alert. (Steering, for review purposes, was an evasive maneuver in an attempt to avoid striking a vehicle; thus there must have been a significant amount of steering.)

Hand Location at Time of Alert

Both hands were not often visible, so reviewers coded what could confidently be inferred from the scene. At times, playing the video farther helped determine what was ambiguous in a still frame at the time of the alert. For instance, at the time of the alert there may have been a small blur near the steering wheel. On continuation of the video the blur may have moved and come into view as a hand.

- 0 = Cannot see the position of either hand or cannot determine the position of either hand. (Reviewers coded 0 if a hand could be seen but they could not tell if it was on the wheel.)
- 1 = At least one hand on steering wheel. (This was coded when the position of one hand could not be determined but reviewers could see that at least one hand was on the steering wheel.)
- 2 = Both hands on the steering wheel.
- 3 = At least one hand off the steering wheel. (This was coded when the position of one hand could not be determined but at least one hand was clearly off the steering wheel.)
- 4 = One hand on, one hand off the steering wheel. (The classification was 4 when reviewers could clearly see both hands and one was on the wheel but the other was off.)
- 5 = Both hands off the steering wheel. (This classification was used when reviewers could clearly see both hands and both were off the wheel.)

Secondary Driving Behaviors

Audio was used to assist in coding whenever possible. For instance, reviewers may have heard the radio station change and seen the driver look at the console; this would indicate in-car system use. The default for nondriving behaviors was none, coded as 0.

Cell Phone

- 10 = Conversation, in use. (Conversation could be coded for listening, talking, or both while using the cell phone.)
- 11 = Reaching for phone. (This classification was used when the driver reached for the handheld phone to speak on that phone. If the driver reached for the phone simply to answer the phone and talk on the headset the driver was wearing, the classification was other. Simply answering the phone involves far less physical activity by the driver than reaching for the phone and holding it during a conversation.)
- 12 = Dialing phone.

Headset, Hands-free Phone

- 20 = Conversation. (This was selected when reviewers could tell that the driver was in a conversation.)
- 21 = Reaching for headset.
- 22 = Unsure of activity level. (The driver was wearing a headset but it was not clear whether the headset was in use. The driver may have been listening to someone or wearing it in case of an incoming call.)

Eating

- 30 = High involvement. (High involvement includes such activities as eating a burger or unwrapping food.)
- 31 = Low involvement. (Low involvement includes such activities as eating candy or grabbing chips.)

Drinking

- 40 = High involvement. (High involvement includes situations in which the driver was trying to open a straw or bottle or was blowing on a hot drink.)
- 41 = Low involvement. (Low involvement includes situations in which the driver was sipping a drink or drinking without looking.)
- 50 = Conversation. (The driver and someone in the car were carrying on a conversation. The driver can be listening during the clip, talking during the clip, or doing both.)
- 60 = In-car system use. (The driver was actively adjusting something. For example, the driver was not just listening to the stereo but also adjusting the stereo. The car lighter was coded under the smoking section.)

Smoking

- 70 = Lighting. (This classification includes the in-car lighter.)
- 71 = Reaching for cigarettes or lighter. (This classification includes the in-car lighter.)
- 72 = Smoking.

Grooming

- 80 = High involvement. (High involvement includes applying makeup or brushing hair.)
- 81 = Low involvement. (Low involvement includes scratching or running one's fingers through his or her hair.)
- 90 = Other/multiple behaviors, specified in notes section. (Behaviors may include whistling or classifications that reviewers were unsure of [e.g., if the driver's lips were moving but there was no audio, the behavior might be singing or conversation].)

Seat Belt

- 0 = Yes.
- 1 = No.
- 2 = Cannot tell.

Curve Speed Warning System Scenario Elements

Road Type

- 0 = Freeway/interstate.
- 1 = Ramp. (A ramp was defined as an entrance or exit ramp from a freeway or any ramp that connected two arterial roads.)
- 2 = Ramp near merge point. (Near was defined as being within 10 s of the merge point or within 10 s of arriving at the straightening of the ramp leading to a merge.)
- 3 = Surface road.
- 4 = Other. (Enter in notes.)

Road Condition

Glare and reflection helped determine whether the road was dry or wet.

- 0 = Dry.
- 1 = Wet. (Any moisture on the road led to the classification as wet; there did not need to be standing water. The road was classified as wet if it was wet from snow but not snow covered.)

- 2 = Snow covered. (Snow covered included ice covered if it was observed, but it was never observed. If any portion of the road, including turn lanes, was covered in snow, the classification was snow covered.)

Precipitation

Spots on the windshield or wiper activity helped determine if there was precipitation.

- 0 = None.
- 1 = Rain. (Light rain and drizzle were classified as rain, as were downpours.)
- 2 = Snow. (This category included sleet. Several cues helped indicate that the precipitation was snow. Snow tended to be larger and fall more slowly than rain, it looked like white flurries, and was present on the ground, reinforcing the classification as snow. Precipitation that occurred in December through February was assumed to be snow rather than rain. Snow could be coded in other months, but the assumption that the precipitation was snow was not as strong.)

Number of Through Lanes

Turn lanes and dedicated exit lanes are not included in the count of the number of through lanes.

- 1 = 1.
- 2 = 2.
- 3 = 3.
- 4 = 4 or more.

Recent Lane Change

To be considered a recent lane change, the lane change had to occur no more than 5 s before the alert or the car had to be in the process of a lane change at the time of the alert.

- 0 = No.
- 1 = Yes, toward branch that triggered the alert.
- 2 = Yes, away from the branch that triggered the alert.
- 3 = Yes, but there was no branch triggering the alert or the branch triggering the alert is unknown.

Curve Confidence

This field was used to indicate when reviewers could not accurately determine which branch or curve triggered the alert. Most of the events categorized as confidence not high resulted from CSWS behavior that stems from artifacts of the map or CSWS implementation details.

- 0 = Confidence not high.
- 1 = Confidence high.

Nearby Overpass or Underpass

The criteria were that the driver had to pass an overpass or underpass 5 s before the alert or 10 s after the alert.

- 0 = No.
- 1 = Yes.

Change in Number of Through Lanes

- 0 = No.
- 1 = Yes.

Does the Vehicle Branch?

This addresses whether the vehicle is or will be taking a branch that triggers the CSWS alert.

- 0 = Not branching, and the alert is not triggered by a branch. (This can occur on a curvy rural road, for instance, or after the vehicle has exited onto a ramp and is approaching a curve.)
- 1 = Not branching, but passing branch that triggers alert.
- 2 = Branching onto segment that triggers alert. (This includes taking an exit or driving in a dedicated exit lane.)
- 3 = Branching but alert was triggered by curve on initial roadway.
- 9 = No confidence in identifying the curve.

Branch Type When Branch Is Triggering Alert

If the roadway is a ramp, the ramp being traveled is not considered a branch. For instance, if the vehicle has exited the freeway onto an exit ramp and the roadway classification is ramp, an alert triggered by a curve along that ramp would be coded as 0, no branch, because the vehicle is already on the ramp.

- 0 = A branch does not trigger the alert.
- 1 = Ramp.
- 2 = Turn lane.
- 3 = Michigan left.
- 4 = Intersection.
- 5 = Other.
- 9 = No confidence in identifying the curve.

Road Geometry

- 0 = Straight.
- 1 = Curve.

- 2 = Approaching curve. (This classification constituted situations in which the driver was approaching but not in a curve at the time of the alert. The driver had to be driving through the curve within 5 s after the alert in order to be classified as approaching curve.)

Notes

A notes section recorded any unusual events or ambiguous situations not covered by categories for a particular question. This section also contains general notes on the clip if there was anything significant taking place that was not adequately covered by the coding process. Examples of items captured in the notes section are described below, but other, unforeseen events are also noted.

Visual Occlusion

Rear taillights, glare from rain and wetness on the road, blurry video, dirty windshield, temporary incapacitation, sneezing, flying debris, faulty wiper or defroster, and object in or over the driver's eyes.

Nondriving Behaviors

Whistling, two or more behaviors, if there is no audio and the driver is clearly talking or singing but reviewers could not tell which, attempting to avoid insect in car, adjusting mirrors, reading map, reading other materials, checking watch, or yawning.

LDWS Scenario Elements

Road Type

- 0 = Freeway/interstate.
- 1 = Ramp.
- 2 = Ramp near merge point. (Near is defined as being within 10 s of the merge point or within 10 s of arriving at the straightening of the ramp leading to a merge.)
- 3 = Surface road.
- 4 = Other. (Enter in notes.)

Road Condition

Glare and reflection helped determine whether the road was dry or wet.

- 0 = Dry.
- 1 = Wet. (Any moisture on the road led to the classification as wet; there did not need to be standing water. The road was classified as wet if it was wet from snow but not snow covered.)
- 2 = Snow covered. (Snow covered included ice covered if it was observed, but it was never observed. If any portion of the road, including turn lanes, was covered in snow, the classification was snow covered.)

Precipitation

Spots on the windshield or wiper activity helped determine if there was precipitation.

- 0 = None.
- 1 = Rain. (Light rain and drizzle were classified as rain, as were downpours.)
- 2 = Snow. (This category included sleet. Several cues helped indicate that the precipitation was snow. Snow tended to be larger and fall more slowly than rain, it looked like white flurries and it was also present on the ground, reinforcing the classification as snow. Precipitation that occurred in December through February was assumed to be snow rather than rain. Snow could be coded in other months, but the assumption that the precipitation was snow was not as strong.)

Road Curvature

- 0 = Straight.
- 1 = Right-hand curve.
- 2 = Left-hand curve.

Lane Marking Change

- 0 = No.
- 1 = Yes.

Boundary Type

This field refers to which type of boundary was on the side of the alert. For example, for an imminent LDW to the left in which there was a solid lane boundary to the left, it would be coded as 0. Options 4 and 5 refer to double-boundary situations.

- 0 = Solid.
- 1 = Dashed.
- 2 = Double solid.
- 3 = No marking.
- 4 = Solid/dashed.
- 5 = Dashed/solid.
- 6 = Curb.
- 7 = Cannot tell.

Continuous Incidental Feature

This feature applies to continuous markings on the road that are not lane lines but may appear as lane lines to the LDWS—for example, tar markings, shadows, or tire marks on wet pavement.

- 0 = No.
- 1 = Yes.

Badly Placed Boundary

At times the LDWS's real or virtual boundary was not properly placed according to actual conditions on the roadway.

- 0 = No.
- 1 = Yes.

Boundary Interaction

Ultimately, the position of the vehicle's tires was used to determine its position in the lane. At the time of the alert, if the tires were on or over the lane line, the crossed/straddled line option was selected.

- 0 = Crossed/straddled line at alert.
- 1 = Lane change at alert.
- 2 = Centered/slightly off-center in lane.
- 3 = Drifted in lane.

Postboundary Maneuver

This field evaluates the first maneuver the vehicle makes after the alert. For example, if the vehicle was drifting in the lane at the time of the alert, then crossed the lane line, and finally returned to its original lane, only the eventually crossed option would be selected. The fact that the vehicle had ultimately returned to its original lane was addressed in the additional driving circumstances field, option corrected per the alert, which is detailed in the Additional Driving Circumstances section.

- 0 = Eventually crossed.
- 1 = Eventually returned to original lane.
- 2 = Stayed in lane.

Beyond the Boundary

The area within two-thirds of a lane width and outside the boundary in question was considered in this evaluation. Although the choices were not mutually exclusive, no attempt was made to quantify everything beyond the boundary. If the alert was propagated by the camera, the area directly to the right or left of the vehicle was evaluated. If, however, information from the radar produced the alert, every effort was made to discern which object(s) had provoked the alert based on available maneuvering room (AMR) bin information.

- 0 = Median/open space.
- 1 = Solid barrier.
- 2 = Turning lane.
- 3 = Empty lane.
- 4 = Adjacent same-direction vehicle.
- 5 = Fixed, discrete objects.
- 6 = Construction zone.
- 7 = Stalled/slow traffic in adjacent lane.
- 8 = Curb.
- 9 = Other/unknown.
- 10 = Adjacent opposing-direction vehicle.

Additional Driving Circumstances

These circumstances are intentional maneuvers by the driver that help explain why the vehicle crossed the boundary or, in the case of corrected per the alert, the action the driver took after the alert.

- 0 = None.
- 1 = Cut behind a car.
- 2 = Clear a temporary obstacle.
- 3 = Make room for a large truck.
- 4 = Corrected per the alert.
- 5 = Early or late exit/merge.

False Alert Comments

- 0 = None.
- 1 = Cannot identify target. (For a radar-induced alert.)
- 2 = Target seems far. (For a radar-induced alert, the target had to be within two-thirds of the lane width from the vehicle to be considered valid.)
- 3 = Appears too sensitive. (This classification is usually applied when it appeared that the driver was not drifting.)
- 4 = Other. (List in notes.)

Lighting Issues

- 0 = None.
- 1 = Possible road reflection.
- 2 = Recent change in road illumination.

Notes

A notes section recorded any unusual events or ambiguous situations not covered by categories for a particular question. This section also contains general notes on the clip if anything significant was taking place that was not adequately covered by the coding process.

APPENDIX C

Project 7 and Project 8 Event Data Dictionary

Event Variables

In the following variables, C-N-I-B is an abbreviation for crashes, near crashes, incidents, and baseline epochs.

Event Identifier (C-N-I-B)

Comment: Each event is assigned a file name that is automatically generated by the software.

Analyst Identifier (C-N-I-B)

Comment: Analysts/data reductionists are identified by their log-ins.

Trigger Type (C-N-I-B)

- 00 = Not applicable (baseline epoch).
- 01 = Lateral acceleration.
- 02 = Longitudinal acceleration.
- 03 = CI button.
- 04 = Lane deviation/bust.
- 05 = Normalized lane position.
- 06 = Forward time to collision (TTC).
- 07 = Forward range.
- 08 = Rear TTC.
- 09 = Rear range.
- 10 = Side object detection.
- 11 = Lane change cutoff.
- 12 = Yaw rate (swerve).
- 13 = Automatic collision notification (ACN).
- 14 = RF sensor.
- 15 = Glare event.
- 16 = Air bag.

Comment: These are taken from the 100-Car Study coding, although a number of 100-Car triggers are not being used in the current study. The total will be somewhat greater than the total event *N* because some events will have more than one trigger. This variable will be automatically generated by the software.

Trigger Quantitative Value (C-N-I)

Maximum or minimum value of relevant triggers. For TTC triggers, find the closest point at which the two vehicles are still in a path to collision and enter that number.

Event Classification (C-N-I-B)

- 00 = Invalid trigger. (In these events, sensor readings were spurious or otherwise not safety-relevant, because of a transient spike or some other anomaly.)
- 00a = No video. (One or more of the quadrants of video is out or not visible. It is not possible to obtain enough information to determine the event.)
- 01 = Baseline driving epoch (selected randomly). (These are 1-min periods that are randomly selected from the recorded data set. Baseline epochs are described using many of the same variables and data elements used to describe and classify crashes, near crashes, and incidents. Examples of such variables include ambient weather, roadway type, and driver behaviors. The creation of a baseline data set will enable the study to (1) describe and characterize normal driving for the study sample and (2) infer the increased or decreased risk associated with various conditions and driver behaviors by comparisons between the control [baseline] data set and the incident and near-crash data sets. For example, if 20% of incidents but only 10% of baseline epochs occurred during rain, one could infer that rain is associated with an increased incident rate and, therefore, increased risk.)
- 02 = Crash. (This includes any contact with an object [e.g., other vehicles, roadside barriers, objects on or off the roadway, pedestrians, cyclists, or animals], either moving or

fixed, at any speed in which kinetic energy is measurably transferred or dissipated.)

- 03 = Near crash (evasive maneuver). (This classification includes any circumstance that requires a rapid, evasive maneuver by the subject vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. A rapid, evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs that approaches the limits of vehicle capabilities. Any event in which the driver swerves off the side of the road and any part of the truck leaves the pavement is automatically coded a near crash.)
- 04 = Near crash (no evasive maneuver). (Any circumstance that results in extraordinary proximity of the subject vehicle to any other vehicle, pedestrian, cyclist, animal, or fixed object in which, because of apparent unawareness on the part of the driver(s), pedestrians, cyclists, or animals, there is no avoidance maneuver or response is coded in this manner. Extraordinary proximity is defined as a clear case in which the absence of an avoidance maneuver or response is inappropriate for the driving circumstances, including speed and sight distance. TTCs of less than 2.00 s are reviewed to assess whether they qualify as crash-relevant conflicts [or near crashes]; TTCs of less than 1.00 s are always coded as crash-relevant conflicts or near crashes.)
- 05 = Crash-relevant conflict (evasive maneuver). (This category includes any circumstance that requires a crash avoidance response on the part of the subject vehicle, any other vehicle, pedestrian, cyclist, or animal that is less severe than a rapid evasive maneuver [as defined above] but more severe than a normal maneuver. A crash avoidance response can include braking, steering, accelerating, or any combination of control inputs. A normal maneuver for the subject vehicle is defined as a control input that falls within the 99% confidence limit for control inputs for the initial study data sample. Examples of potential crash-relevant conflicts include hard braking by a driver because of a specific crash threat or proximity to other vehicles. Evasive maneuvers resulting in unsafe or illegal maneuvers or situations should be included in this category [or as near crashes if more severe]. Longitudinal decelerations of –0.35 *g* or greater are reviewed to assess whether they qualify as crash-relevant conflicts [or near crashes]; those with decelerations of –0.50 *g* or greater are always coded as crash-relevant conflicts or near crashes.)
- 06 = Crash-relevant conflict (no evasive maneuver). (Included in this classification is any circumstance that results in proximity of the subject vehicle to any other vehicle, pedestrian, cyclist, animal, or fixed object in which, because of apparent unawareness on the part of the driver(s), pedestrians, cyclists, or animals, there is no avoidance maneuver or response. Proximity is defined as a clear case in which the absence of an avoidance maneuver or response is inappropriate for the driving circumstances, including speed and sight distance.)
- 07 = Nonconflict. (This includes any incident that has an above-threshold trigger but does not result in a crash, near crash, or crash-relevant conflict as defined above. There is no abrupt evasive maneuver and no signs of any other unsafe condition, such as a lane break. Driver errors may be observed, but they do not result in a traffic conflict. Examples include hard braking by a driver in the absence of a specific crash threat and high lateral acceleration on curves not resulting in loss of control, lane departure, or proximity to other vehicles.)

Comment: Initial coding step. Invalid triggers and nonconflicts result in no further coding. Identification of two types of near crashes (i.e., evasive maneuver and proximity event) permits later disaggregation if desired. Definitions of each type of event are given above.

Date (C-N-I-B)

Comment: Raw data from vehicle.

Day of Week (C-N-I-B)

Comment: Raw data from vehicle.

Time (C-N-I-B)

Comment: Raw data from vehicle. For C-N-I events, the time of maximum or minimum trigger value is recorded. For baseline epochs, the end of the 30-s baseline period is recorded.

Format: Integer.

Vehicles or Nonmotorists Involved (C-N-I)

- 00 = Not applicable (baseline epoch).
- 01 = 1 vehicle (subject vehicle only).
- 02 = 2 vehicles.
- 03 = 3 vehicles.
- 04 = 4 or more vehicles.
- 05 = Subject vehicle + pedestrian.
- 06 = Subject vehicle + pedalcyclist.
- 07 = Subject vehicle + animal.
- 08 = Other.

Comment: Events that involve the subject vehicle and an object (i.e., struck or potentially struck) are coded 01. For some events (e.g., those that involve transient encroachment into an oncoming lane), it will be difficult to decide whether the event should be considered a one- or two-vehicle event. Consider the event a two-vehicle event if the crash resulting from the incident would probably have involved two vehicles and if either driver's maneuvers were influenced by the presence of the other vehicle (e.g., if Driver/Vehicle 1 [DV1] maneuvered to avoid Vehicle 2 [V2]). Consider the event a one-vehicle event if the presence of other vehicles presented no immediate threat and had no effect on Driver 1's maneuvers or behaviors.

Which Vehicle Is Considered to Be at Fault? (C-N-I)

- 00 = Not applicable (baseline epoch).
- 01 = Vehicle 1 (subject vehicle).
- 02 = Vehicle 2 (other vehicle, pedalcyclists, or animal).
- 09 = Unknown.

Comment: The at-fault vehicle is defined as the vehicle with the assigned critical reason.

Light Condition (C-N-I-B)

- 01 = Daylight.
- 02 = Dark.
- 03 = Dark but lighted.
- 04 = Dawn.
- 05 = Dusk.
- 09 = Unknown.

Comment: General estimate system (GES) A19.

Weather (Atmospheric Condition) (C-N-I-B)

- 01 = No adverse conditions.
- 02 = Rain.
- 03 = Sleet.
- 04 = Snow.
- 05 = Fog.
- 06 = Rain and fog.
- 07 = Sleet and fog.
- 08 = Other (smog, smoke, sand or dust, crosswind, hail).
- 09 = Unknown.

Comment: GES A20.

Roadway Surface Condition (C-N-I-B)

- 01 = Dry.
- 02 = Wet.
- 03 = Snow or slush.
- 04 = Ice.
- 05 = Sand, oil, dirt.
- 08 = Other.
- 09 = Unknown.

Comment: GES A15.

Relation to Junction (C-N-I-B)

- 00 = Nonjunction.
- 01 = Intersection.
- 02 = Intersection related.
- 03 = Driveway, alley access, etc.
- 03a = Parking lot.
- 04 = Entrance/exit ramp.
- 05 = Rail grade crossing.
- 06 = On a bridge.
- 07 = Crossover related.
- 08 = Other.
- 09 = Unknown.

Comment: GES variable A09. GES instructions for coding this variable will be reviewed to ensure consistency of coding approach with GES.

Construction Zone Related (C-N-I-B)

- 00 = Not construction zone related (or unknown).
- 01 = Construction zone (occurred in zone).
- 02 = Construction zone related (occurred in approach or otherwise related to zone).

Comment: Default code is 0. For the purposes of coding, consider any area with multiple traffic cones, barrels, and so forth to be a construction zone.

Traffic Density (C-N-I-B)

- 01 = LOS A: Free flow. (Individual users are virtually unaffected by the presence of others in the traffic stream. Freedom to select desired speeds and to maneuver within the traffic stream is extremely high. The general level of comfort and convenience provided to the motorist, passenger, or pedestrian is excellent.)
- 02 = LOS B: Flow with some restrictions. (In the range of stable traffic flow, but the presence of other users in the traffic stream begins to be noticeable. Freedom to select desired speeds is relatively unaffected, but there is a slight decline in the freedom to maneuver within the traffic stream from LOS A because the presence of others in the traffic stream begins to affect individual behavior.)
- 03 = LOS C: Stable flow; maneuverability and speed are more restricted. (Traffic flow is in the stable range but is beginning to enter the range of flow in which the opera-

tion of individual users becomes significantly affected by interactions with others in the traffic stream. The selection of speed is now affected by the presence of others, and maneuvering within the traffic stream requires substantial vigilance on the part of the user. The general level of comfort and convenience declines noticeably at this level.)
- 04 = LOS D: Unstable flow; temporary restrictions substantially slow driver. (This category represents high-density but stable traffic flow. Speed and freedom to maneuver are severely restricted, and the driver or pedestrian experiences a generally poor level of comfort and convenience. Small increases in traffic flow generally cause operational problems at this level.)
- 05 = LOS E: Flow is unstable; vehicles are unable to pass, temporary stoppages, etc. (Operating conditions are at or near the capacity level. All speeds are reduced to a low but relatively uniform value. Freedom to maneuver within the traffic stream is extremely difficult, and it is generally accomplished by forcing a vehicle or pedestrian to give way to accommodate such maneuvers. Comfort and convenience levels are extremely poor, and driver or pedestrian frustration is generally high. Operations at this level are usually unstable because small increases in flow or minor perturbations within the traffic stream will cause breakdowns.)
- 06 = LOS F: Forced traffic flow condition with low speeds and traffic volumes that are below capacity; queues form in particular locations. (This condition exists whenever the amount of traffic approaching a point exceeds the amount that can traverse the point. Queues form behind such locations. Operations within the queue are characterized by stop-and-go waves and are extremely unstable. Vehicles may progress at reasonable speeds for several hundred feet or more and then be required to stop in a cyclic manner. LOS F is used to describe operating conditions within the queue, as well as the point of the breakdown. In many cases, operating conditions of vehicles or pedestrians discharged from the queue may be quite good. Nevertheless, it is the point at which arrival flow exceeds discharge flow, which causes the queue to form. LOS F is an appropriate designation for such points.)
- 09 = Unknown/unable to determine.

Driver/Vehicle 1 Variables

Note: DV1 is always the study subject driver/vehicle (e.g., the truck driver or truck).

Subject Vehicle Number (C-N-I-B)

Format: Integer. Automatically generated.

Subject Driver Number (C-N-I-B)

Format: Integer. Automatically generated.

Trafficway Flow (C-N-I-B)

- 00 = Not physically divided (center two-way left turn lane).
- 01 = Not physically divided (two-way trafficway).
- 02 = Divided (median strip or barrier).
- 03 = One-way trafficway.
- 09 = Unknown.

Comment: GES variable V A11. Coded in relation to subject vehicle.

Number of Travel Lanes (C-N-I-B)

- 01 = 1.
- 02 = 2.
- 03 = 3.
- 04 = 4.
- 05 = 5.
- 06 = 6.
- 07 = 7+.
- 09 = Unknown.

Comment: GES V A12. Per GES, if road is divided, only lanes in travel direction are counted. If the road is undivided, all lanes are counted. Coded in relation to subject vehicle. Count all contiguous lanes at the time and location of the incident (e.g., include entrance and exit lanes if contiguous).

Truck Pre-Event Speed (C-N-I-B)

Format: Integer.
Comment: C-N-I events are coded for the period just before the occurrence of the critical event or just before any avoidance maneuver or both. For example, when braking is involved, the pre-event speed is the speed just before the beginning of braking. Baseline events are coded for the end of the 30-s baseline interval. Note that roadway speed limit cannot currently be determined because most speed limit signs are not legible on the videos. Future efforts (in particular, Phase 2) will consider automated ways to obtain this variable, such as the use of GPS and roadway geographic information systems.

Roadway Alignment (C-N-I-B)

- 01 = Straight.
- 02a = Curve right.
- 02b = Curve left.
- 09 = Unknown.

Comment: GES V A13, with expansion of curve choices. Coded in relation to subject vehicle.

Roadway Profile (C-N-I-B)

- 01 = Level (or unknown).
- 02a = Grade up.
- 02b = Grade down.
- 03 = Hillcrest.
- 04 = Sag.

Comment: GES V A14, with expansion of grade choices. Coded in relation to subject vehicle.

Driver Seat Belt Worn? (C-N-I-B)

- 01 = Yes.
- 02 = No.
- 09 = Unknown.

Comment: This issue is of current interest to FMCSA, and its capture would permit comparisons of driver behavior between drivers wearing and those not wearing seat belts. Judgment is based on whether a shoulder strap is visible; the lap belt typically cannot be seen.

Does the Driver Cover the Camera or Is the Camera Covered? (C-N-I-B)

- 00 = Yes.
- 01 = No/not observed.
- 02 = Attempts but fails.

Alcohol Use (C-N-I-B)

- 00 = None apparent.
- 01 = Suspected use observed in vehicle without overt effects on driving.
- 02 = Suspected use observed in vehicle with overt effects on driving.
- 03 = Reported by police (applicable only to crashes).
- 04 = Use not observed or reported but suspected based on driver behavior.
- 09 = Unknown.

Comment: Use indicated only if apparent from event review. Note: The remaining DV1 variables are precrash and event causation variables. Table C.1 lists these variables, indicates sources, and shows the corresponding variable for DV2.

Vehicle Pre-Event Movement (C-N-I-B)

- 00 = No driver present.
- 01 = Going straight.
- 02 = Decelerating in traffic lane.
- 03 = Accelerating in traffic lane.
- 04 = Starting in traffic lane.
- 05 = Stopped in traffic lane.
- 06 = Passing or overtaking another vehicle.
- 07 = Disabled or parked in travel lane.
- 08a = Leaving a parking position, moving forward.
- 08b = Leaving a parking position, backing.
- 09a = Entering a parking position, moving forward.
- 09b = Entering a parking position, backing.

Table C.1. Coded Precrash and Causation Variables

Variable Name	Principal Source(s) (e.g., other databases/studies)	Subject Vehicle (DV1) Variable No.	Other Vehicle (DV2) Variable No.
Vehicle pre-event movement	GES, LTCCS	27	44
Accident type (scenario role)	GES, LTCCS	28	45
Incident types	Two recent VTTI studies	29	46
Critical precrash event	LTCCS	30	47
Critical reason for the critical event	LTCCS	31	48[a]
Attempted avoidance maneuver	GES, LTCCS	32	49
Driver vision obscured by	GES	34	Not coded
Average PERCLOS value (1, 3, 5 minutes)	VTTI and other fatigue research	35–37	
Observer rating of drowsiness (1 minute)	Previous VTTI research	38	
Potentially distracting driver behaviors	GES	39	
Driver actions/factors relating to event	100-Car Study	40	50[a]
Applicable functional countermeasures	Various	41	51

[a]Abridged due to inability to observe specific DV2 behaviors and states.

- 10 = Turning right.
- 11 = Turning left.
- 12 = Making a U-turn.
- 13 = Backing up (other than parking).
- 14 = Negotiating a curve.
- 15 = Changing lanes.
- 16 = Merging.
- 17 = Successful avoidance maneuver to a previous critical event.
- 98 = Other.
- 99 = Unknown.

Comment: This is Large Truck Crash Causation Survey (LTCCS) Variable 4 with expanded choices for 8 and 9. For baseline epochs, the primary movement of the vehicle during the epoch is coded.

Accident Type (Scenario Role) (C-N-I)

- 00 = Not applicable (baseline epoch).
- Other codes (see Table C.2).

Comment: LTCCS Variable 10 and GES Variable V23. Because this variable includes intent, analysts should project probable scenario roles for incidents in which outcomes are not definite. In other words, if the trigger-related event had resulted in a crash, what would the crash scenario be? When specific scenarios cannot be projected, use the specific unknown choices (e.g., 5, 10, 16, and 33). Table C.2 illustrates accident types.

Additional clarifications:

- Drive off road codes (e.g., 01 and 06) are used when a vehicle has crossed or is projected to cross a roadside delineation such as a lane edge line (going onto the shoulder or median), a curb, or the edge of the pavement. This includes scenarios that involve parked vehicles and stationary objects if those objects are outside the roadway delineation (e.g., on an unpaved shoulder).
- Forward impact codes (e.g., 11 and 12) are used when the objects are in the travel lane or when there is no lane edge delineation as described above. Thus, a scenario involving a parked vehicle on the pavement where there is no lane edge delineation is coded 12.
- For left-side lane departures into the oncoming traffic lane, code 64 or 65 if the lateral encroachment is less than a few feet. Code 50 or 51 only if the lateral encroachment was sufficient to create a significant risk of a head-on crash.
- Hard-braking events at intersections in the absence of a specific crash or crash threat are coded 91 (intersecting straight paths, specifics unknown).

Incident Types (C-N-I)

- 00 = Not applicable (baseline epoch).
- 01/02 = Aborted lane change.
- 03/04 = Approaches traffic quickly (not used).
- 05/06/07/08 = Backing in roadway.
- 09/10 = Clear path for emergency vehicle.
- 11/12 = Conflict between merging and exiting traffic.
- 13/14 = Conflict with oncoming traffic.
- 15/16 = Exit then re-entrance onto roadway.
- 17/18 = Following too closely.
- 19/20 = Improper lane change.
- 21/22/23 = Improper passing.
- 24/25 = Improper U-turn.
- 26/27 = Lane change without sufficient gap.
- 28/29 = Lane drift.
- 30/31 = Late braking for stopped/stopping traffic.
- 32/33 = Lateral deviation of through vehicle.
- 34/35 = Left turn without clearance.
- 36/37 = Merge out of turn (before lead vehicle).
- 38/39/40 = Merge without sufficient gap.
- 41/42 = Obstruction in roadway.
- 43/44 = Proceeding through red traffic signal.
- 45/46 = Roadway entrance without clearance.
- 47/48 = Slow speed.
- 49/50 = Slow upon passing.
- 51/52/53 = Sudden braking in roadway.
- 54/55 = Through traffic does not allow lane change.
- 56/57/58 = Through traffic does not allow merge.
- 59/60 = Turn without sufficient warning.
- 61/62 = Turn/exit from incorrect lane.
- 63/64 = Wide turn into adjacent lane.
- 65 = Conflict with object/animal/pedalcyclist in roadway.
- 66 = Conflict with object/animal/pedalcyclist on side of road.
- 67 = Other single-vehicle event.
- 68/69 = Proximity to turning vehicle.
- 99 = Unknown.

Comment: This scenario classification has been used in Hanowski, Keisler, and Wierwille (*2*) and Hanowski, Olson, Hickman, and Dingus (*3*). Coding this variable will enable comparisons with that study. Diagrams of these scenarios are provided in Table C.3.

Critical Precrash Event for Vehicle 1 (C-N-I)

- 00 = Not applicable (baseline epoch).

Causes of Vehicle (V1) Loss of Control

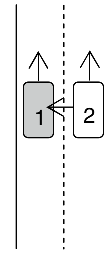
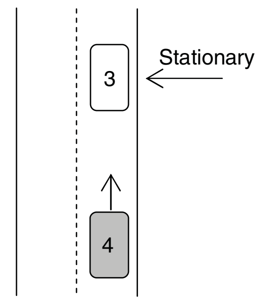
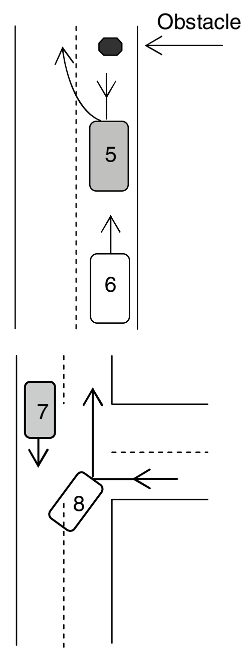
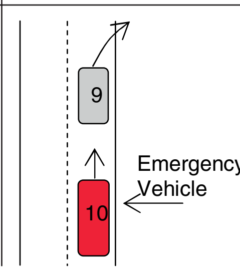
- 01 = Blow out or flat tire.
- 02 = Stalled engine.
- 03 = Disabling vehicle failure (e.g., wheel fell off).

(*text continues on page 107*)

Table C.2. Description of Accident Types (1)

Category	Configuration	ACCIDENT TYPES (Includes Intent)					
I. Single Driver	A. Right Roadside Departure	01 DRIVE OFF ROAD	02 CONTROL/ TRACTION LOSS	03 AVOID COLLISION WITH VEH., PED., ANIM.		04 SPECIFICS OTHER	05 SPECIFICS UNKNOWN
	B. Left Roadside Departure	06 DRIVE OFF ROAD	07 CONTROL/ TRACTION LOSS	08 AVOID COLLISION WITH VEH., PED., ANIM.		09 SPECIFICS OTHER	10 SPECIFICS UNKNOWN
	C. Forward Impact	11 PARKED VEHICLE	12 STATIONARY OBJECT	13 PEDESTRIAN/ ANIMAL	14 END DEPARTURE	15 SPECIFICS OTHER	16 SPECIFICS UNKNOWN
II. Same Trafficway Same Direction	D. Rear-End	20, 22, 21, 23 STOPPED 21, 22, 23	24, 26, 25, 27 SLOWER 25, 26, 27	28, 30, 29, 31 DECELERATING 29, 30, 31		(EACH - 32) SPECIFICS OTHER	(EACH - 33) SPECIFICS UNKNOWN
	E. Forward Impact	34, 35 CONTROL/ TRACTION LOSS	36, 37 CONTROL/ TRACTION LOSS	38, 39 AVOID COLLISION WITH VEHICLE	40, 41 AVOID COLLISION WITH OBJECT	(EACH - 42) SPECIFICS OTHER	(EACH - 43) SPECIFICS UNKNOWN
	F. Sideswipe Angle	44, 45	46, 45, 47			(EACH - 48) SPECIFICS OTHER	(EACH - 49) SPECIFICS UNKNOWN
III. Same Trafficway Opposite Direction	G. Head-On	50, 51 LATERAL MOVE				(EACH - 52) SPECIFICS OTHER	(EACH - 53) SPECIFICS UNKNOWN
	H. Forward Impact	54, 55 CONTROL/ TRACTION LOSS	56, 57 CONTROL/ TRACTION LOSS	58, 59 AVOID COLLISION WITH VEHICLE	60, 61 AVOID COLLISION WITH OBJECT	(EACH - 62) SPECIFICS OTHER	(EACH - 63) SPECIFICS UNKNOWN
	I. Sideswipe/ Angle	64, 65 LATERAL MOVE				(EACH - 66) SPECIFICS OTHER	(EACH - 67) SPECIFICS UNKNOWN
IV. Change Trafficway Vehicle Turning	J. Turn Across Path	68, 69 INITIAL OPPOSITE DIRECTIONS	70, 73, 71, 72 INITIAL SAME DIRECTION			(EACH - 74) SPECIFICS OTHER	(EACH - 75) SPECIFICS UNKNOWN
	K. Turn Into Path	77, 79, 76, 78 TURN INTO SAME DIRECTION	82, 80, 81, 83 TURN INTO OPPOSITE DIRECTIONS			(EACH - 84) SPECIFICS OTHER	(EACH - 85) SPECIFICS UNKNOWN
V. Intersecting Paths (Vehicle Damage)	L. Straight Paths	87, 86	88, 89			(EACH - 90) SPECIFICS OTHER	(EACH - 91) SPECIFICS UNKNOWN
VI. Miscellaneous	M. Backing Etc.	92 BACKING VEHICLE	93 OTHER VEHICLE OR OBJECT			98 OTHER ACCIDENT TYPE 99 UNKNOWN ACCIDENT TYPE 00 NO IMPACT	

Table C.3. Incident Type Descriptions

Incident Type	Description	Illustration
Aborted lane change	Driver tries to make a lane change into a lane where there is already a vehicle (driver doesn't see vehicle). The driver has to brake and move back into the original lane.	
Approaches traffic quickly (not used)	Driver approaches stopped or slowing traffic too quickly and has to brake hard or suddenly to avoid hitting the lead vehicle.	
Backing in roadway	Driver backs the vehicle while on a roadway in order to maneuver around an obstacle ahead on the roadway.	
Clear path for emergency vehicle	Driver is traveling ahead of an emergency vehicle (e.g., ambulance, fire truck) and has to move to the side of the road to let the emergency vehicle pass.	

(continued on next page)

Table C.3. Incident Type Descriptions (continued)

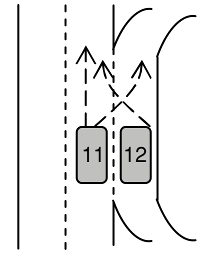
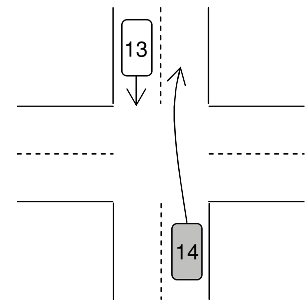
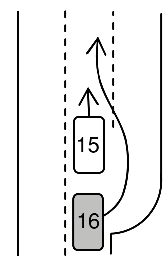
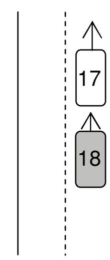
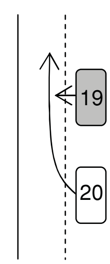
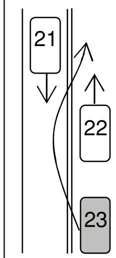
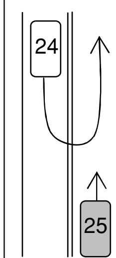
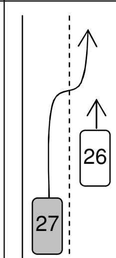
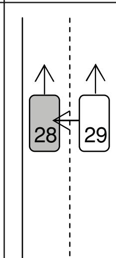
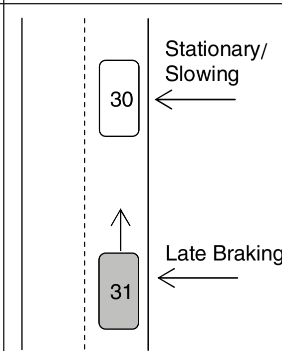
Incident Type	Description	Illustration
Conflict between merging or exiting traffic	Drivers entering or exiting a roadway, using a shared weaving section, conflict.	
Conflict with oncoming traffic	Driver is approaching oncoming traffic (e.g., through an intersection) and has to maneuver back into the correct lane to avoid an oncoming vehicle.	
Exit then re-entrance onto roadway	Driver exits a roadway and then crosses a solid white line to re-enter.	
Following too closely	Driver does not allow adequate spacing between his or her vehicle and the lead vehicle (e.g., tailgating).	
Improper lane change	Driver makes an improper lane change with regard to another vehicle (e.g., does not use blinker, changes lanes behind another vehicle and then does not let vehicle change lanes, changes lanes across multiple lanes).	

Table C.3. Incident Type Descriptions (continued)

Incident Type	Description	Illustration
Improper passing	Driver passes another vehicle when it is illegal or unsafe (e.g., passing across a double yellow line or without clearance from oncoming traffic).	
Improper U-turn	Driver makes a U-turn in the middle of the road (over the double yellow line) and blocks traffic in the opposite direction.	
Lane change without sufficient gap	Driver enters an adjacent lane without allowing adequate space between the driver's vehicle and the vehicle ahead or behind it.	
Lane drift	Driver drifts into an adjacent lane without intention to make a lane change.	
Late braking (or steering) for stopped or stopping traffic	Driver fails to slow in advance for stopped or stopping traffic and must brake or steer abruptly.	

(continued on next page)

Table C.3. Incident Type Descriptions (continued)

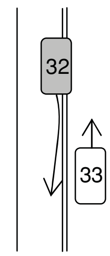
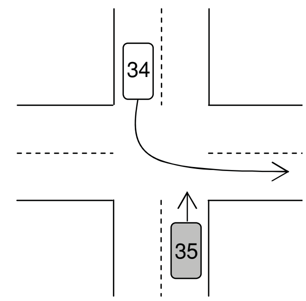
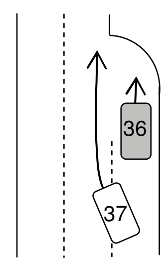
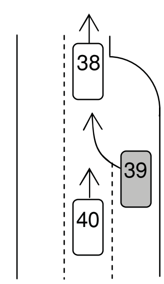
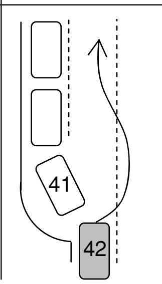
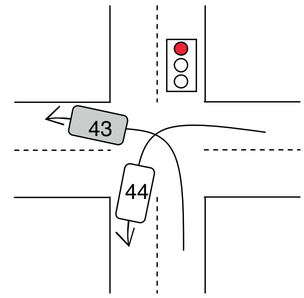
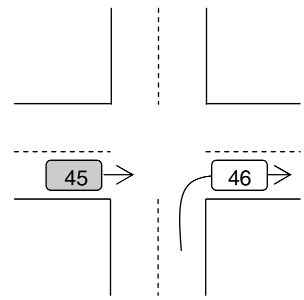
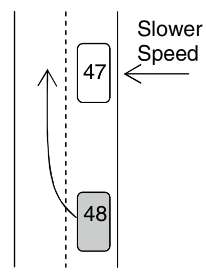
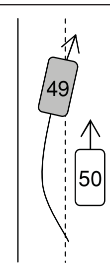
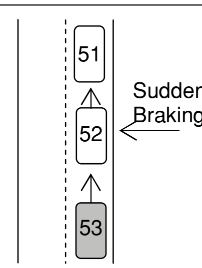
Incident Type	Description	Illustration
Lateral deviation of through vehicle	Driver has substantial lateral deviation of a through vehicle. Vehicle may or may not deviate from the lane.	
Left turn without clearance	Driver turns left without adequate clearance from either oncoming through traffic or cross traffic from the left. Driver crosses another driver's path while entering an intersecting roadway.	
Merge out of turn (before lead vehicle)	Driver merges onto a roadway before the lead vehicle. The lead vehicle must wait for the merged vehicle to pass before it is safe to enter the main highway.	
Merge without sufficient gap	Driver merges into traffic without a sufficient gap to either the front or the back of one or more vehicles.	
Obstruction in roadway	Stationary object blocks through traffic, such as traffic that is backed up or an animal in the roadway.	

Table C.3. Incident Type Descriptions (continued)

Incident Type	Description	Illustration
Proceeding through red traffic signal	Driver fails to respond to a red traffic signal, conflicting with a vehicle proceeding through the intersection legally.	
Roadway entrance without clearance	Driver turns onto a roadway without adequate clearance from through traffic.	
Slow speed	Driver is traveling at a much slower speed than the rest of the traffic, causing following traffic to pass the slow vehicle to avoid a conflict.	
Slow on passing	Driver moves in front of another vehicle and then slows, causing the second (passed) vehicle to slow as well or go around the first vehicle.	
Sudden braking in roadway	Driver is traveling ahead of another vehicle and brakes suddenly and improperly in the roadway for traffic or a traffic light, causing the following vehicle to come close to the braking vehicle or to also brake suddenly.	

(continued on next page)

Table C.3. Incident Type Descriptions (continued)

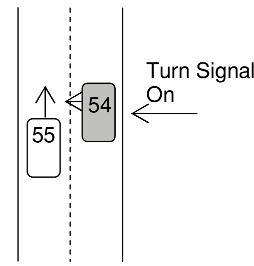
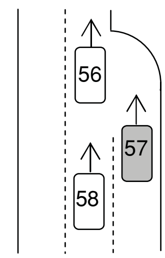
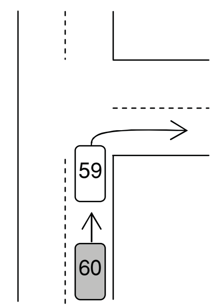
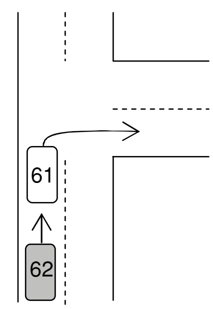
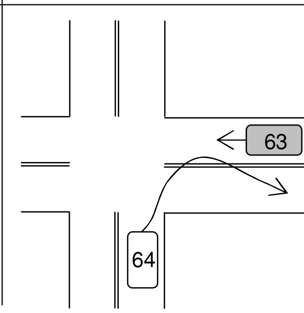
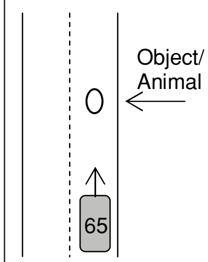
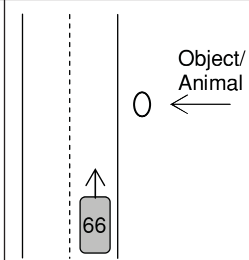
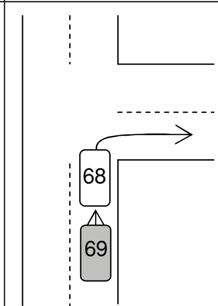
Incident Type	Description	Illustration
Through traffic does not allow lane change	Driver is trying to make a lane change (with the turn signal on) but traffic in the adjacent lane will not allow the lane change to be completed.	
Through traffic does not allow merge	Through traffic obstructs a driver from entering the roadway.	
Turn without sufficient warning	Driver slows and turns without using a turn signal or without using a turn signal in advance.	
Turn or exit from incorrect lane	Driver turns onto a side road from the incorrect lane (e.g., driver makes a right turn from the left lane instead of the right lane).	
Wide turn into adjacent lane	Vehicle partially enters an adjacent lane when turning. Traffic in the adjacent lane may be moving in the same or opposite direction.	

Table C.3. Incident Type Descriptions (continued)

Incident Type	Description	Illustration
Conflict with object/animal/ pedalcyclist in roadway	Vehicle approaches an object, animal, or pedalcyclist in the roadway and either makes contact with it or performs an evasive maneuver to avoid it.	
Conflict with object/animal/ pedalcyclist on side of roadway	Vehicle approaches an object, animal, or pedalcyclist on the side of the road and either makes contact with it or performs an evasive maneuver to avoid it.	
Proximity to turning vehicle	Lead vehicle is making a right or left turn or changing lanes to the right or left and the following vehicle comes close to the rear of the lead vehicle as they pass.	
Other single-vehicle event	Vehicle is involved in a single-vehicle event; for example, runs off the side of the road without a threat of hitting a fixed object.	
Unable to determine	It is not possible to determine which vehicle is at fault; therefore, it is not possible to assign an incident type to the event.	

(*continued from page 99*)

- 04 = Nondisabling vehicle problem (e.g., hood flew up).
- 05 = Poor road conditions (e.g., wet road, puddle, pothole, ice).
- 06 = Traveling too fast for conditions.
- 07 = Jackknife event.
- 08 = Cargo shift.
- 09 = Braking.
- 10 = Steering.
- 18 = Other cause of control loss.
- 19 = Unknown cause of control loss.

Travel of Vehicle (V1)

- 20 = Toward or over the lane line on left side of travel lane.
- 21 = Toward or over the lane line on right side of travel lane.
- 22 = Toward or off the edge of the road on the left side.
- 23 = Toward or off the edge of the road on the right side.
- 24 = End departure.
- 25 = Turning left at intersection.
- 26 = Turning right at intersection.
- 27 = Crossing over (passing through) intersection.
- 28 = This vehicle decelerating.
- 29 = Unknown travel direction.

Other Motor Vehicle (V2) in Lane

- 50 = Other vehicle stopped.
- 51 = Traveling in same direction with lower steady speed.
- 52 = Traveling in same direction while decelerating.
- 53 = Traveling in same direction with higher speed.
- 54 = Traveling in opposite direction.
- 55 = In crossover.

- 56 = Backing
- 59 = Unknown travel direction of other motor vehicle in lane.

Other Motor Vehicle (V2) Encroaching into Lane

- 60 = From adjacent lane (same direction), toward or over left lane line.
- 61 = From adjacent lane (same direction), toward or over right lane line.
- 62 = From opposite direction, toward or over left lane line.
- 63 = From opposite direction, toward or over right lane line.
- 64 = From parking lane.
- 65 = From crossing street, turning into same direction.
- 66 = From crossing street, across path.
- 67 = From crossing street, turning into opposite direction.
- 68 = From crossing street, intended path not known.
- 70 = From driveway, turning into same direction.
- 71 = From driveway, across path.
- 72 = From driveway, turning into opposite direction.
- 73 = From driveway, intended path not known.
- 74 = From entrance to limited-access highway.
- 78 = Encroachment by other vehicle, details unknown.

Pedestrian, Pedalcyclist, or Other Nonmotorist

- 80 = Pedestrian in roadway.
- 81 = Pedestrian approaching roadway.
- 82 = Pedestrian, unknown location.
- 83 = Pedalcyclist or other nonmotorist in roadway.
- 84 = Pedalcyclist or other nonmotorist approaching roadway.
- 85 = Pedalcyclist or other nonmotorist, unknown location.

Object or Animal

- 87 = Animal in roadway.
- 88 = Animal approaching roadway.
- 89 = Animal, unknown location.
- 90 = Object in roadway.
- 91 = Object approaching roadway.
- 92 = Object, unknown location.

Other

- 93 = This vehicle not involved in first harmful event.
- 98 = Other critical precrash event.
- 99 = Unknown.

Comment: This is LTCCS Variable 5. It is coded for both vehicles in a two-vehicle incident. However, the critical reason (see below) is coded for only one vehicle. For consistency with the accident type variable (28), lane edges between travel lanes and nontravel lanes (e.g., shoulders) are considered road edges; for example, events involving V1's crossing of these edges are coded 22 or 23. Unlike the accident type variable, the analyst should code the actual precipitating event and should not project or extrapolate the event. In the list above, note the addition of 09 = loss of control caused by braking and 10 = steering.

DV1 Critical Reason for the Critical Event (C-N-I)

- 000a = Not applicable (baseline epoch).
- 000b = Critical reason not coded to this vehicle.

Driver-Related Factor: Critical Nonperformance Errors

- 100 = Sleep (i.e., actually asleep).
- 101 = Heart attack or other physical impairment of the ability to act.
- 107 = Drowsiness, fatigue, or other reduced alertness (not asleep).
- 108 = Other critical nonperformance.
- 109 = Unknown critical nonperformance.

Driver-Related Factor: Recognition Errors

- 110 = Inattention (i.e., daydreaming).
- 111 = Internal distraction.
- 112 = External distraction.
- 113 = Inadequate surveillance (e.g., failed to look, looked but did not see).
- 118 = Other recognition error.
- 119 = Unknown recognition error.

Driver-Related Factor: Decision Errors

- 120 = Too fast for conditions (e.g., for safe vehicle control or to be able to respond to unexpected actions of other road users).
- 121 = Too slow for traffic stream.
- 122 = Misjudgment of gap or other road user's speed.
- 123 = Following too closely to respond to unexpected actions (proximity for 2 or more seconds).
- 124 = False assumption of other road user's actions.
- 125 = Illegal maneuver.
- 125a = Apparently intentional sign/signal violation.
- 125b = Illegal U-turn.

- 125c = Other illegal maneuver.
- 126 = Failure to turn on headlamps.
- 127 = Inadequate evasive action (e.g., braking only instead of braking and steering; releasing accelerator only instead of braking).
- 128a = Aggressive driving behavior: intimidation. (Any behavior emitted by a driver while driving that is intended to cause physical or psychological harm to another person.)
- 128b = Aggressive driving behavior: wanton, neglectful, or reckless behavior. (Excessive risky driving behaviors performed without intent to harm others, such as weaving through traffic, maneuvering without signaling, running red lights, frequent lane changing, and tailgating.)
- 138 = Other decision error.
- 139 = Unknown decision error.
- 140 = Apparent recognition or decision error (unknown which).

Driver-Related Factor: Performance Errors

- 141 = Panic/freezing.
- 142 = Overcompensation.
- 143 = Poor directional control (e.g., failing to control vehicle with skill ordinarily expected).
- 148 = Other performance error.
- 149 = Unknown performance error.
- 199 = Type of driver error unknown.

Vehicle-Related Factor

- 200 = Tires/wheels failed.
- 201 = Brakes failed.
- 202 = Steering failed.
- 203 = Cargo shifted.
- 204 = Trailer attachment failed.
- 205 = Suspension failed.
- 206 = Lights failed.
- 207 = Vehicle-related vision obstructions.
- 208 = Body, doors, hood failed.
- 209 = Jackknifed.
- 298 = Other vehicle failure.
- 299 = Unknown vehicle failure.

Environment-Related Factor: Highway Related

- 500 = Signs/signals missing.
- 501 = Signs/signals erroneous/defective.
- 502 = Signs/signals inadequate.
- 503 = View obstructed by roadway design.
- 504 = View obstructed by other vehicles in crash circumstance.
- 505 = Road design, roadway geometry (e.g., ramp curvature).
- 506 = Road design, sight distance.
- 507 = Road design, other.
- 508 = Maintenance problems (e.g., potholes, deteriorated road edges).
- 509 = Slick roads (low-friction road surface caused by ice, loose debris, any other cause).
- 518 = Other highway-related condition.

Environment-Related Factor: Weather Related

- 521 = Rain, snow (Note: code loss-of-control as 509).
- 522 = Fog.
- 523 = Wind gust.
- 528 = Other weather-related condition.

Environment-Related Factor: Other

- 530 = Glare.
- 531 = Blowing debris.
- 532 = Animal in roadway (no driver error).
- 533 = Pedestrian or pedalcyclist in roadway (no driver error).
- 538 = Other sudden change in ambience.
- 999 = Unknown reason for critical event.

Comment: LTCCS Variable 6 with revisions. "This vehicle" is always used for the vehicle being coded. Note that vehicle-related factors are rarely apparent to data reductionists.

Vehicle 1 Attempted Avoidance Maneuver (C-N-I)

- 00 = No driver present.
- 0a = Not applicable (baseline epoch).
- 01 = No avoidance maneuver.
- 02 = Braking (no lockup or lockup unknown).
- 03 = Braking (lockup).
- 04 = Braking (lockup unknown).
- 05 = Releasing brakes.
- 06 = Steered to left.
- 07 = Steered to right.
- 08 = Braked and steered to left.
- 08a = Braked and steered to left (no lockup or lockup unknown).
- 08b = Braked and steered to left (lockup).
- 09 = Braked and steered to right.
- 09a = Braked and steered to right (no lockup or lockup unknown).
- 09b = Braked and steered to right (lockup).
- 10 = Accelerated.
- 11 = Accelerated and steered to left.
- 12 = Accelerated and steered to right.
- 13 = Released gas pedal without braking.

- 14 = Released gas pedal (without braking) and steered to left.
- 15 = Released gas pedal (without braking) and steered to right.
- 98 = Other actions.
- 99 = Unknown if driver attempted any corrective action.

Comment: LTCCS Variable 7 and GES V27, corrective action attempted. Released gas pedal elements added because this evasive maneuver by subject drivers is sometimes observed.

Relevant Object (C-N-I)

Analyst chooses the most relevant object; that is, one that was struck in a crash or that constituted a crash threat for near crashes and crash-relevant conflicts.

- 00a = Not applicable (baseline epoch).
- 00b = Not applicable (single-vehicle event but no critical object [e.g., shoulder only]).
- 00c = Not applicable (two-vehicle event, pedestrian, animal, etc.).
- 01 = Parked motor vehicle.

Fixed Objects

- 02 = Building.
- 03 = Impact attenuator/crash cushion.
- 04 = Bridge structure (e.g., abutment).
- 05 = Guardrail.
- 06 = Concrete traffic barrier or other longitudinal barrier (e.g., Jersey barrier).
- 07 = Post, pole, or support (e.g., sign or light).
- 08 = Culvert or ditch.
- 09 = Curb.
- 10 = Embankment.
- 11 = Fence.
- 12 = Wall.
- 13 = Fire hydrant.
- 14 = Shrubbery or bush.
- 15 = Tree (not overhang; see below).
- 16 = Boulder.
- *17 = Loading dock.*
- *18 = Loading equipment (e.g., forklift or pallets).*
- *19 = Cargo.*

Overhanging Objects

The following were coded only if struck or potentially struck by the top of a truck trailer.

- *20 = Tree branch.*
- 21 = *Overhanging part of sign or post.*
- 22 = *Bridge/overpass.*
- 23 = *Building.*
- *24 = Telephone wires.*

Nonfixed Objects

- *25 = Vehicle parts, including tire parts.*
- *26 = Spilled cargo.*
- *27 = Dead animal in roadway.*
- *28 = Broken tree limbs or other tree/shrub parts.*
- *29 = Trash/debris.*
- *30 = Construction barrel.*
- *31 = Construction cone.*
- *98 = Other.*
- *99 = Unknown object hit.*

Comment: Most objects are the same as those used in GES A06, first harmful event. Those in italics are not A06 codes.

Driver 1 Vision Obscured (C-N-I)

- 00 = No obstruction.
- 01 = Rain, snow, fog, smoke, sand, dust.
- 02 = Reflected glare, sunlight, headlights.
- 03 = Curve or hill.
- 04 = Building, billboard, or other design features (includes signs and embankment).
- 05 = Trees, crops, vegetation.
- 06 = Moving vehicle (including load).
- 07 = Parked vehicle.
- 08 = Splash or spray of passing vehicle (any other vehicle).
- 09 = Inadequate defrost or defog system.
- 10 = Inadequate lighting system (includes vehicle or object in dark area).
- 11 = Obstruction interior to vehicle.
- 12 = Mirrors.
- 13 = Head restraints.
- 14 = Broken or improperly cleaned windshield.
- 15 = Fog.
- 16 = Other vehicle or object in blind spot.
- 50 = Hit-and-run vehicle.
- 95 = No driver present.
- 96 = Not reported.
- 97 = Vision obscured, no details.
- 98 = Other obstruction.
- 99 = Unknown whether vision was obstructed.

Comment: GES Variable D4. Element 16 added because of relevance to large trucks. Elements 50, 95, and 96 are not applicable.

Driver Fatigue Monitor Operating Mode (C-N-I-B)

- 01 = Auto-manual.
- 02 = Manual.
- 03 = Auto (if operating mode = auto, driver fatigue monitor [DFM] is automatically nonoperative).

DFM Sensitivity Level (C-N-I-B)

- 01 = Low.
- 02 = Medium.
- 03 = High.

Rules to Follow When Trying to Determine If DFM Is in Standby

- When speed is below 30 mph (48.28 kph) and ambient brightness is above 100, the DFM is in standby.
- When the speed is above 35 mph (56.32 kph) and ambient brightness is less than 50, the DFM is active.
- Ambient brightness (0 = dark; 255 = bright).

Special note: Sometimes when the DFM should be functioning according to the rules above but often during dawn and dusk, it still does not operate correctly. If it looks light in the video but the ambient brightness values are within the correct range, a judgment call might be needed to determine whether it is working. Please ask if there are any questions.

Average PERCLOS over 1 Min (C-N-I-B)

Comment: Recorded parameter from DFM, averaged over a 1-min period before the initiating event. Coded when available for time epoch.

Format: Percent; 999 = DFM not operative.

Average PERCLOS over 3 Min (C-N-I-B)

Comment: Recorded parameter from DFM, averaged over a 3-min period before the initiating event. Coded when available for time epoch.

Format: Percent; 999 = DFM not operative.

Average PERCLOS over 5 Min (C-N-I-B)

Comment: Recorded parameter from DFM, averaged over a 5-min period before the initiating event. Coded when available for time epoch.

Format: Percent; 999 = DFM not operative.

Observer Rating of Driver Drowsiness (C-N-I-B)

Note: Analysts use a 100-point scale to code observer rating of drowsiness (ORD). They can choose any value (e.g., 35, 62, or 87) on the scale in Figure C.1. The five given points are to be used as guidelines.

If ORD is 25 or greater, mark drowsy, sleepy, asleep, fatigued, other reduced alertness under Driver 1 behaviors.

- 999 = Driver is wearing sunglasses or eyes are otherwise blocked from view.
- 00 = Not drowsy. (No signs of being drowsy.)
- 25 = Slightly drowsy. (Driver shows minor signs of being drowsy [e.g., single yawn, single stretch, or droopy eyes for a short period]; quickly recovers; does not have any apparent impact on vehicle control.)
- 50 = Moderately drowsy. (Driver shows signs of being drowsy [e.g., yawns, stretches, moves around in seat, droopy eyes for a slightly longer period, or minor blinking]; takes slightly longer to recover; does not have any apparent impact on vehicle control.)
- 75 = Very drowsy. (Driver shows signs of being drowsy [e.g., yawns often, has very heavy or droopy eyes, or blinks frequently]; duration lasts much longer; does not have any apparent impact on vehicle control.)
- 100 = Extremely drowsy. (Driver shows extreme signs of being drowsy [e.g., yawns often, has very heavy or droopy eyes, has trouble keeping eyes open, or blinks very frequently]; duration lasts much longer; has apparent impact on vehicle control.)

Comment: An ORD is assigned for 1 min before the event based on review of driver videos. Three-, 6-, and 20-min ORDs are not obtained, because of the labor required and difficulties in averaging reliably over these periods.

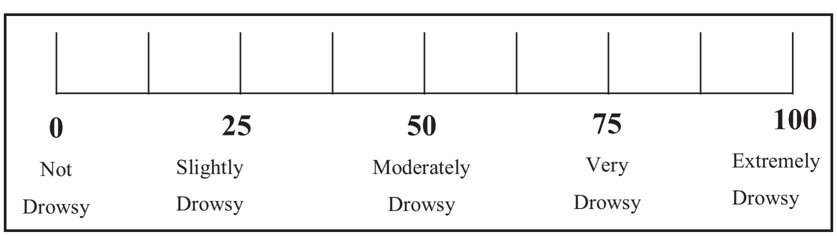


Figure C.1. Driver drowsiness scale.

Driver 1 Potentially Distracting Driver Behaviors (C-N-I-B)

Analysts code up to four behaviors observed during 10 s before the maximum or minimum trigger value or during final 10 s of a 30-s baseline epoch. Code observed behaviors regardless of their apparent relevance to the incident. This is similar to GES but significantly modified. If there are more than four behaviors, select the ones that occur closest in time to the trigger.

- 00 = None observed.
- 01 = Looked but did not see (e.g., driver looked in direction of crash threat but apparently did not recognize threat). (Not applicable to baseline epochs.)
- 02a = Interact with or look at other occupant(s).
- 02b = Interact with or look at pet in vehicle.
- 03a = Look at or for object in vehicle.
- 03b = Reach for object in vehicle (e.g., handheld cell phone, hands-free cell phone, PDA, CB microphone or other communications device, or other object).
- 04a = Talk or listen to handheld phone.
- 04b = Talk or listen to hands-free phone.
- 04c = Talk or listen to CB microphone or other communications device.
- 05a = Dial handheld phone.
- 05b = Dial hands-free phone.
- 05c = Operate PDA (inputting or reading).
- 06 = Adjust instrument panel (e.g., climate control, radio, cassette or CD).
- 07a = Look at left-side mirror/out left-side window.
- 07b = Look at right-side mirror/out right-side window.
- 07c = Look back in sleeper berth.
- 07d = Shift gears.
- 07e = Look down (e.g., at lap or at something on the floor).
- 08 = Use or reach for other devices.
- 09 = Appear drowsy, sleepy, asleep, fatigued.
- 10a = Look at previous crash or highway incident.
- 10b = Look at construction zone signs, barriers, flagperson, etc.
- 10c = Look at outside person.
- 10d = Look at outside animal, object, store, etc.
- 10e = Look at undetermined outside event, person, or object.
- 11a = Eat with utensil.
- 11b = Eat without utensil (includes chewing other than gum [e.g., toothpick]).
- 11c = Drink from covered container (e.g., with straw).
- 11d = Drink from open container.
- 11e = Chew gum.
- 12a = Smoking-related behavior—reaching, lighting, or extinguishing.
- 12b = Smoking-related behavior—other (e.g., cigarette in hand or mouth).
- 13a = Read book, newspaper, etc.
- 13b = Read or look at map.
- 14 = Talk/sing/"dance" with no indication of passenger.
- 15a = Handle or interact with dispatching, electronic recording, or navigational device.
- 15b = Read or look at dispatching, electronic recording, or navigational device.
- 16a = Comb/brush/fix hair.
- 16b = Apply makeup.
- 16c = Shave.
- 16d = Brush/floss teeth.
- 16e = Bite nails/cuticles.
- 16f = Remove/adjust jewelry.
- 16g = Remove/insert contact lenses.
- 16h = Other personal hygiene.
- 17 = Look at or handle driver fatigue monitor (DFM).
- 18 = Look at or handle data acquisition system (DAS) (e.g., in-vehicle camera).
- 19 = Appears inattentive or lost in thought.
- 20 = Other potentially distracting behavior.

Comment: Similar to GES Variable D7 (driver distracted by), with expansions of many elements to capture direct observations. All observed behaviors or conditions that occur within 10 s before the maximum trigger are coded without regard to apparent relevance to the conflict. Baseline epochs are coded only for activities that occur within the last 10 s of the 30-s baseline epoch. Handheld and hands-free phone data are coded separately to permit comparisons.

Driver 1 Actions, Factors, or Behaviors Related to the Event (C-N-I)

Note: Analysts code up to four factors believed to have relevance to the occurrence of the incident (e.g., contributing factors). If there are more than four factors, the four most important are selected.

- 00a = Not applicable (baseline epoch).
- 00b = None coded.
- 01 = Apparent excessive speed for conditions or location (regardless of speed limit; does not include tailgating unless above speed limit).
- 02 = Drowsy, sleepy, asleep, fatigued, other reduced alertness.
- 03 = Angry.
- 04 = Other emotional state.
- 05 = Inattentive or distracted.
- 06 = Apparent impairment (e.g., drowsy, drunk, distracted)—specific type unknown.
- 07 = Driving slowly (below speed limit or in relation to other traffic).
- 08 = Illegal passing (e.g., across double line).

- 09 = Passing on right.
- 10 = Other improper or unsafe passing.
- 11a = Cutting in, too close in front of other vehicle.
- 11b = Cutting in at safe distance but then decelerating, causing conflict.
- 12 = Cutting in, too close behind other vehicle.
- 13 = Making turn from wrong lane (e.g., across lanes).
- 14 = Did not see other vehicle during lane change or merge.
- 15 = Driving in other vehicle's blind zone.
- 16 = Aggressive driving, specific, directed menacing actions.
- 17 = Aggressive driving, other (reckless driving without directed menacing actions).
- 18 = Wrong side of road, not overtaking (includes partial or full drift into oncoming lane).
- 19 = Following too closely.
- 19a = Inadequate evasive action.
- 20 = Failed to signal, or improper signal.
- 21 = Improper turn, wide right turn.
- 22 = Improper turn, cut corner on left turn.
- 23 = Other improper turning.
- 24 = Improper backing, did not see.
- 25 = Improper backing, other.
- 26 = Improper start from parked position.
- 27 = Disregarded officer or watchman.
- 28 = Signal violation, apparently did not see signal.
- 29 = Signal violation, intentionally ran red light.
- 30 = Signal violation, tried to beat signal change.
- 31 = Stop sign violation, apparently did not see stop sign.
- 32 = Stop sign violation, intentionally ran stop sign at speed.
- 33 = Stop sign violation, rolling stop.
- 34 = Other sign (e.g., yield) violation, apparently did not see sign.
- 35 = Other sign (e.g., yield) violation, intentionally disregarded.
- 36 = Other sign violation.
- 37 = Nonsigned crossing violation (e.g., driveway entering roadway).
- 38 = Right-of-way error in relation to other vehicle or person, apparent recognition failure (e.g., did not see other vehicle).
- 39 = Right-of-way error in relation to other vehicle or person, apparent decision failure (e.g., did see other vehicle before action but misjudged gap).
- 40 = Right-of-way error in relation to other vehicle or person, other or unknown cause.
- 41 = Sudden or improper stopping on roadway.
- 42 = Parking in improper or dangerous location (e.g., shoulder of interstate).
- 43 = Speeding or other unsafe actions in work zone.
- 44 = Failure to dim headlights.
- 45 = Driving without lights or insufficient lights.
- 46 = Avoiding pedestrian.
- 47 = Avoiding other vehicle.
- 48 = Avoiding animal.
- 48a = Avoiding object.
- 49 = Apparent unfamiliarity with roadway.
- 50 = Apparent unfamiliarity with vehicle (e.g., displays and controls).
- 51 = Use of conventional cruise control (CCC) contributed to late braking (does not imply malfunction of CCC).
- 52 = Excessive braking/deceleration creating potential hazard.
- 53 = Loss of control on slippery road surface.
- 54 = Loss of control on dry (or unknown) surface.
- 55 = Apparent vehicle failure (e.g., brakes).
- 56 = Other.

Comment: This variable was used in the 100-Car Naturalistic Driving Study, although some new elements have been added. Also, the coding rule is different; in the 100-Car Study, analysts coded up to three factors for each driver in descending order of judged importance. In the current study, analysts code all that apply in no order of importance. Thus, the data from the two studies are not directly comparable. Note that Element 6 is not relevant to Driver 1, because analysts are not able to identify impairment type.

Applicable Countermeasures for DV1 (C-N-I)

On the basis of the variables above that relate to the event scenario, pre-event actions and states, and event causation, a senior analyst identifies applicable functional countermeasures. For crashes, an applicable DV1 functional countermeasure is one that would probably have prevented the crash, either by preventing the genesis of the unsafe condition or by improving driver response to the unsafe condition. Near crashes and crash-relevant conflicts are analyzed as if a crash had occurred. Table C.4 shows functional countermeasures and coding rules for them. The coding of functional countermeasures is based on both algorithmic determination from previous coded variables and analyst judgment. In many cases, particular accident type, critical reason, or other causation-related codes algorithmically determine applicable functional countermeasures. Some countermeasure choices, however, are coded based on senior analyst judgment.

Driver/Vehicle 2 Variables

Vehicle/Person 2 Type (C-N-I)

- 00a = Not applicable (baseline epoch).
- 00b = Not applicable (single-vehicle event; includes single vehicle + object).
- 01 = Automobile.

(*text continues on page 117*)

Table C.4. Applicable Countermeasures for Vehicle 1

No.	Functional Countermeasure	Scenario/Driver Error Source(s)	Code DV2?	Comments
0a	Not applicable (baseline epoch)	NA	Yes	
0b	No countermeasure applicable to this driver/vehicle (no driver error and/or coded to other vehicle only)	NA	Yes	
0c	No obvious or plausible countermeasure applicable to this driver/vehicle (e.g., insufficient information because of random occurrence)	NA	Yes	
0d	Not applicable: single-vehicle event	Veh/Nonmotorists involved = 01, 05–07	Yes	Never coded for V1
1	Increase driver alertness (reduce drowsiness)	CR = 100 or 107 *Or* analyst judgment considering PERCLOS, ORD, driver behavior	No	
2	Improve commercial driver hours-of-service (HOS) compliance (i.e., reflective of alertness-related incident during HOS violation period)		No	Not coded during Phase I; potential for Phase II.
3	Prevent drift lane departures (e.g., caused by fatigue, inattention, or misjudgment of lines)	AT = 01 or 06	Yes	No evidence of intention (e.g., lane change).
4	Improve vehicle control/stability on curves	Trigger Type = 1 *And* PEM = 14 *And* AT = 02, 07, 46, 47, or 50	Yes	Assumes potential rollover or other LOC event; no triggers for V2.
5	Improve vehicle control/stability on slippery road surfaces	Road surface = 2–5 *And* CPE = 05	Yes	
6	Improve vehicle control/stability during braking	CPE = 09 *Or* Avoidance maneuver = 3	Yes	
7	Improve vehicle control/stability during evasive steering	CPE = 10 *Or* Avoidance maneuver = 6–9 with LOC	Yes	
8	Increase driver attention to forward visual scene (e.g., eyes on road)	Analyst judgment, considering potential distractions coded (V39) and CR (e.g., 110–119, 140)	No	
9	Increase/improve driver use of mirrors or provide better information from mirrors (or from other indirect visibility systems)	AT = 46, 47, 70, 73, 76, 78, or others TBD *and* Vision Obscured = 12 or 16	No	
10	Improve general driver situation awareness and/or proactive/defensive driving	Analyst judgment	No	Not coded if 1 and/or 8 are coded.

Table C.4. Applicable Countermeasures for Vehicle 1 (continued)

No.	Functional Countermeasure	Scenario/Driver Error Source(s)	Code DV2?	Comments
12	Reduce road/highway travel speed	CR = 120 *Or* Driver behavior = 1, 43	Yes	Includes all road configurations and thus is inclusive of 14–16 but does not include all speeds above speed limit; must be significant factor.
13	Reduce speed on downgrades	CR = 120 *and* Profile = 2b *or* Driver B = 1, 43 *and* Profile = 2b	No	
14	Reduce speed on curves or turns	CR = 120 *and* Alignment = 2a, 2b *Or* Driver B = 1, 43 *and* Alignment = 2a, 2b	No	
15	Reduce speed at or on exits (including ramps)	CR = 120 *and* Profile = 2b *or* Driver B = 1, 43 *and* Profile = 2b	No	
16	Limit top speed to 70 mph (except on downgrades)	Prevented speed greater than 70 mph; analyst judgment Evidence: CR = 120; Driver A/F/B = 1	No	
17	Increase driver recognition/appreciation of specific highway crash threats: stopped vehicle(s) in lane ahead traveling in same direction	AT = 11, 20 *And* CR = 107–119	Yes	
18	Increase driver recognition/appreciation of specific highway crash threats: moving/decelerating vehicle(s) in lane ahead traveling in same direction	AT = 24, 28 *And* CR = 107–119	Yes	
19	Increase driver recognition/appreciation of specific highway crash threats: vehicle in left adjacent lane on highway	AT = 47 *And* CR = 107–119	Yes	
20	Increase driver recognition/appreciation of specific highway crash threats: vehicle in right adjacent lane on highway	AT = 46 *And* CR = 107–114	Yes	
21	Increase driver recognition/appreciation of specific highway crash threats: vehicle in left adjacent lane during merging maneuver	AT = 47, 78 *And* PEM = 16 *And* CR = 107–119	Yes	
22	Increase driver recognition/appreciation of specific highway crash threats: vehicle in right adjacent lane during merging maneuver	AT = 46, 76 *And* PEM = 16 *And* CR = 107–119	Yes	
23	Increase driver recognition of crossing or oncoming traffic at intersections	AT = 76, 78, 80, 82–91 *And* CR = 107–119	Yes	
24	Improve driver gap judgment relating to crossing or oncoming traffic at intersections	AT = 76, 78, 80, 82–91 *And* CR = 122	Yes	

(continued on next page)

Table C.4. Applicable Countermeasures for Vehicle 1 (continued)

No.	Functional Countermeasure	Scenario/Driver Error Source(s)	Code DV2?	Comments
25	Improve driver response execution of crossing or turning maneuver at intersections (performance failure)	AT = 76, 78, 80, 82–91 *And* CR = 141–199	Yes	
26	Improve driver recognition, gap judgment, or response execution at intersection (specific cause not determined)	AT = 76, 78, 80, 82–91 *And* CR = 140 or 199	Yes	
27	Improve driver compliance with intersection traffic signal (e.g., red light) controls (includes both intentional and unintentional intersection control violations)	Driver A/F/B = 28–30	Yes	
28	Improve driver compliance with intersection traffic sign (e.g., stop or yield sign) controls (includes both intentional and unintentional intersection control violations)	Driver A/F/B = 31–33	Yes	
29	Increase forward headway during vehicle following	AT = 24, 28 *And* CR = 123	Yes	Applies to tailgating scenarios, not rapid closing scenarios.
30	Improve driver night vision in forward field	Light = 2, 3 *And* AT = 1–14, 20, 34, 36, 38, 40 *And* analyst judgment	Yes	CM would provide earlier driver recognition of distant object (e.g., pedestrian walking in roadway)
32	Provide warning to prevent rear encroachment or tailgating by other vehicle (i.e., this vehicle is lead vehicle, other vehicle is following)	AT = 21, 22, 23, 25, 26, 27, 29, 30, 31	Yes	Reciprocal relation between 17 and 18 and 32 (i.e., if one vehicle is coded 17 or 18, other vehicle is coded 32).
33	Provide advisory to driver regarding reduced road-tire friction (i.e., associated with slippery roads)	Roadway surface condition = 2–5 *And* LOC *And* analyst judgment	No	
34	Prevent vehicle mechanical failure (e.g., brakes, steering, or tire blowout)	CR = 200–209, 298, 299	Yes	Probably undercounted in instrumented vehicle studies.
35	Other, specify	Analyst judgment	Yes	When possible, analyst will specify associated precrash/causation algorithm and add to list of CMs.
36	Prevent splash and spray from this vehicle affecting other vehicle(s)	AT = 25, 26, 35–41, 45–47 *And* analyst judgment *And* Roadway surface condition = 2, 3	Yes	
37	Improve driver recognition/gap judgment relating to oncoming vehicle during passing maneuver	PEM = 06 *And* AT = 50 or 64 *And* CR = 110–119, 120–122, or 128–140	Yes	
38	Prevent animals from crossing roadways	Vehicle/Person 2 Type = 13 or 14	No	Applicable to all animal-related events.

Table C.4. Applicable Countermeasures for Vehicle 1 (continued)

No.	Functional Countermeasure	Scenario/Driver Error Source(s)	Code DV2?	Comments
39	Navigation system/routing aid	Driver A/F/B = 49	No	
40	Aid to vertical clearance estimation	Object = overhanging object	No	Used when truck hits or has the potential to hit overhanging object (e.g., tree limb).
98	Driver error and/or vehicle failure apparent for this vehicle but countermeasure(s) to address it unknown	Vehicle has CR but no other CM specified	Yes	Not coded if other CMs coded.
99	Unknown		Yes	Not coded if other CMs coded.

Key: AT = accident type; CR = critical reason; CM = countermeasure; PEM = pre-event movement; CPE = critical precrash event; A = actions; B = behaviors; F = factors; TBD = to be determined; LOC = loss of control.

(*continued from page 113*)

- 02 = Van (minivan or standard van).
- 03 = Pickup truck.
- 03a = SUV (includes Jeep).
- 04 = Bus (transit or motor coach).
- 05 = School bus.
- 06 = Single-unit straight truck (includes panel truck, U-Haul truck).
- 07 = Tractor-trailer.
- 08 = Motorcycle or moped.
- 09 = Emergency vehicle (police, fire, EMS; in service)
- 10 = Vehicle pulling trailer (other than tractor-trailer).
- 11 = Other vehicle type.
- 12 = Pedestrian.
- 13 = Pedalcyclist.
- 14 = Deer.
- 15 = Other animal.
- 99 = Unknown vehicle type.

Comment: Highly abridged version of GES V5, body type. If Driver/Vehicle 2 is a pedestrian, cyclist, animal, or object, most other DV 1 file variables are coded not applicable.

Vehicle 2 Position (in Relation to V1) (C-N-I)

- 00 = Not applicable (baseline epoch).
- 00a = Not applicable (single-vehicle event).
- K = Top of vehicle.

Comment: The vehicle in Figure C.2 represents the DV1 (the truck). The relative position of Vehicle 2 (in relation to Vehicle 1) is coded for the time in which the critical event occurs; that is, the event creating the crash risk. Vehicles in the adjacent left lane are coded J, I, H, or G, depending on their position. Vehicles in the adjacent right lane are coded B, C, D, or E, depending on their position. Baseline epochs are coded 0.

Vehicle 2 Pre-Event Movement (C-N-I)

- 00 = No driver present.
- 00a = Not applicable (single-vehicle event).
- 01 = Going straight.
- 02 = Decelerating in traffic lane.
- 03 = Accelerating in traffic lane.
- 04 = Starting in traffic lane.
- 05 = Stopped in traffic lane.
- 06 = Passing or overtaking another vehicle.
- 07 = Disabled or parked in travel lane.
- 08a = Leaving a parking position, moving forward.
- 08b = Leaving a parking position, backing.
- 09a = Entering a parking position, moving forward.
- 09b = Entering a parking position, backing.
- 10 = Turning right.
- 11 = Turning left.
- 12 = Making a U-turn.
- 13 = Backing up (other than parking).
- 14 = Negotiating a curve.
- 15 = Changing lanes.
- 16 = Merging.
- 17 = Successful avoidance maneuver to a previous critical event.
- 98 = Other.
- 99 = Unknown.

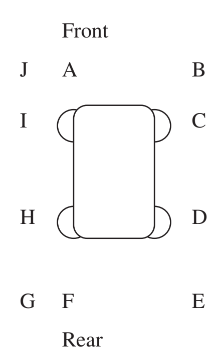


Figure C.2. Vehicle diagram.

Comment: This is LTCCS Variable 4 with expanded choices for 8 and 9. For baseline epochs, the primary movement of the vehicle during the epoch is coded.

Vehicle 2 Accident Type (Scenario Role) (C-N-I)

- 00 = Not applicable (baseline epoch).
- 00a = Not applicable (single-vehicle event).

Other Codes: See diagram shown earlier for Variable 28.

Vehicle 2 Incident Type (C-N-I)

- 00a = Not applicable (baseline epoch).
- 00b = Not applicable (single-vehicle event; includes those with pedestrian or animal).
- 01/02 = Aborted lane change.
- 03/04 = Approaches traffic quickly.
- 05/06/07/08 = Backing in roadway.
- 09/10 = Clear path for emergency vehicle.
- 11/12 = Conflict between merging and existing traffic.
- 13/14 = Conflict with oncoming traffic.
- 15/16 = Exit then re-entrance onto roadway.
- 17/18 = Following too closely.
- 19/20/21 = Improper lane change.
- 22/23 = Improper passing.
- 24/25 = Improper U-turn.
- 26/27 = Lane change without sufficient gap.
- 28/29 = Lane drift.
- 30/31 = Late braking for stopped/stopping traffic.
- 32/33 = Lateral deviation of through vehicle.
- 34/35 = Left turn without clearance.
- 36/37 = Merge out of turn (before lead vehicle).
- 38/39/40 = Merge without sufficient gap.
- 41/42 = Obstruction in roadway.
- 43/44 = Proceeding through red traffic signal.
- 45/46 = Roadway entrance without clearance.
- 47/48 = Slow speed.
- 49/50 = Slow upon passing.
- 51/52/53 = Sudden braking in roadway.
- 54/55 = Through traffic does not allow lane change.
- 56/57/58 = Through traffic does not allow merge.
- 59/60 = Turn without sufficient warning.
- 61/62 = Turn/exit from incorrect lane.
- 63/64 = Wide turn into adjacent lane.
- 68/69 = Proximity to turning vehicle.
- 99 = Unknown.

Comment: This scenario classification has been used in Hanowski, Keisler, and Wierwille (*2*) and Hanowski, Olson, Hickman, and Dingus (*3*). Coding this variable enables comparisons with that study. See Variable 29 for diagrams of these scenarios.

Vehicle 2 Critical Precrash Event (C-N-I)

- 00 = Not applicable (baseline epoch).
- 00a = Not applicable (single-vehicle event).

This vehicle (V2) loss of control because of

- 01 = Blow out or flat tire.
- 02 = Stalled engine.
- 03 = Disabling vehicle failure (e.g., wheel fell off).
- 04 = Nondisabling vehicle problem (e.g., hood flew up).
- 05 = Poor road conditions (wet road, puddle, pothole, ice, etc.).
- 06 = Traveling too fast for conditions.
- 07 = Jackknife event.
- 08 = Cargo shift.
- 09 = Braking.
- 10 = Steering.
- 18 = Other cause of control loss.
- 19 = Unknown cause of control loss.

This vehicle (V1) is traveling

- 20 = Toward or over the lane line on left side of travel lane.
- 21 = Toward or over the lane line on right side of travel lane.
- 22 = Toward or off the edge of the road on the left side.
- 23 = Toward or off the edge of the road on the right side.
- 24 = End departure.
- 25 = Turning left at intersection.
- 26 = Turning right at intersection.
- 27 = Crossing over (passing through) intersection.
- 28 = This vehicle decelerating.
- 29 = Unknown travel direction.

Other motor vehicle (V2) in lane

- 50 = Other vehicle stopped.
- 51 = Traveling in same direction with lower steady speed.
- 52 = Traveling in same direction while decelerating.
- 53 = Traveling in same direction with higher speed.
- 54 = Traveling in opposite direction.
- 55 = In crossover.
- 56 = Backing.
- 59 = Unknown travel direction of other motor vehicle in lane.

Other motor vehicle (V2) encroaching into lane

- 60 = From adjacent lane (same direction), toward or over left lane line.

- 61 = From adjacent lane (same direction), toward or over right lane line.
- 62 = From opposite direction, toward or over left lane line.
- 63 = From opposite direction, toward or over right lane line.
- 64 = From parking lane.
- 65 = From crossing street, turning into same direction.
- 66 = From crossing street, across path.
- 67 = From crossing street, turning into opposite direction.
- 68 = From crossing street, intended path not known.
- 70 = From driveway, turning into same direction.
- 71 = From driveway, across path.
- 72 = From driveway, turning into opposite direction.
- 73 = From driveway, intended path not known.
- 74 = From entrance to limited-access highway.
- 78 = Encroachment by other vehicle, details unknown.

Pedestrian, pedalcyclist, or other nonmotorist

- 80 = Pedestrian in roadway.
- 81 = Pedestrian approaching roadway.
- 82 = Pedestrian, unknown location.
- 83 = Pedalcyclist or other nonmotorist in roadway.
- 84 = Pedalcyclist or other nonmotorist approaching roadway.
- 85 = Pedalcyclist or other nonmotorist, unknown location.

Object or animal

- 87 = Animal in roadway.
- 88 = Animal approaching roadway.
- 89 = Animal, unknown location.
- 90 = Object in roadway.
- 91 = Object approaching roadway.
- 92 = Object, unknown location.

Other

- 93 = This vehicle not involved in first harmful event.
- 98 = Other critical precrash event.
- 99 = Unknown.

Comment: This is LTCCS Variable 5. Per discussion with Ralph Craft of FMCSA, this variable is coded for both vehicles in a two-vehicle incident. However, the critical reason (see below) is coded for only one vehicle. In the list above, note addition of 09 = loss of control caused by braking and 10 = loss of control caused by steering.

DV2 Critical Reason for Critical Event (C-N-I)

- 000a = Not applicable (baseline epoch).
- 000b = Not applicable (single-vehicle event).
- 000c = Critical reason not coded to this vehicle.

Driver-Related Factor: Critical Nonperformance Errors

- 100 = Sleep (i.e., actually asleep).
- 101 = Heart attack or other physical impairment of the ability to act.
- 107 = Drowsiness, fatigue, or other reduced alertness (not asleep).
- 108 = Other critical nonperformance.
- 109 = Apparent critical nonperformance (includes any apparent driver impairment).

Driver-Related Factor: Recognition Errors

- 110 = Inattention (i.e., daydreaming).
- 111 = Internal distraction.
- 112 = External distraction.
- 113 = Inadequate surveillance (e.g., failed to look, looked but did not see).
- 118 = Other recognition error.
- 119 = Apparent recognition error.

Driver-Related Factor: Decision Errors

- 120 = Too fast for conditions (e.g., for safe vehicle control or to be able to respond to unexpected actions of other road users).
- 121 = Too slow for traffic stream.
- 122 = Misjudgment of gap or other's speed.
- 123 = Following too closely to respond to unexpected actions (proximity for 2 or more seconds).
- 124 = False assumption of other road user's actions.
- 125 = Illegal maneuver.
- 125a = Apparently intentional sign/signal violation.
- 125b = Illegal U-turn.
- 125c = Other illegal maneuver.
- 126 = Failure to turn on headlamps.
- 127 = Inadequate evasive action (e.g., braking only instead of braking and steering; releasing accelerator only instead of braking).
- 128a = Aggressive driving behavior: intimidation. (Any behavior emitted by a driver while driving that is intended to cause physical or psychological harm to another person.)
- 128b = Aggressive driving behavior: wanton, neglectful, or reckless behavior. (Excessive risky driving behaviors performed without intent to harm others, such as weaving through traffic, maneuvering without signaling, running red lights, frequent lane changing, and tailgating.)
- 138 = Other decision error.
- 139 = Apparent, unknown decision error.
- 140 = Apparent recognition or decision error (unknown which).

Driver-Related Factor: Performance Errors

- 141 = Panic/freezing.
- 142 = Overcompensation.
- 143 = Poor directional control (e.g., failure to control vehicle with skill ordinarily expected).
- 148 = Other performance error.
- 149 = Apparent performance error.
- 199 = Type of driver error unknown.

Vehicle-Related Factors

- 200 = Tires/wheels failed.
- 201 = Brakes failed.
- 202 = Steering failed.
- 203 = Cargo shifted.
- 204 = Trailer attachment failed.
- 205 = Suspension failed.
- 206 = Lights failed.
- 207 = Vehicle-related vision obstructions.
- 208 = Body, doors, hood failed.
- 209 = Jackknifed.
- 298 = Apparent other vehicle failure.
- 299 = Unknown vehicle failure.

Environment-Related Factor: Highway

- 500 = Signs/signals missing.
- 501 = Signs/signals erroneous/defective.
- 502 = Signs/signals inadequate.
- 503 = View obstructed by roadway design.
- 504 = View obstructed by other vehicles in crash circumstance.
- 505 = Road design, roadway geometry (e.g., ramp curvature).
- 506 = Road design, sight distance.
- 507 = Road design, other.
- 508 = Maintenance problems (potholes, deteriorated road edges, etc.).
- 509 = Slick roads (low-friction road surface caused by ice, loose debris, any other cause).
- 518 = Other highway-related condition.

Environment-Related Factor: Weather

- 521 = Rain, snow. (Note: Code loss-of-control as 509.)
- 522 = Fog.
- 523 = Wind gust.
- 528 = Other weather-related condition.

Environment-Related Factor: Other

- 530 = Glare.
- 531 = Blowing debris.
- 532 = Animal in roadway (no driver error).
- 538 = Other sudden change in ambience.
- 999 = Unknown reason for critical event.

Comment: LTCCS Variable 6, with revisions that reflect lack of information about Driver 2. Many critical reason elements available for DV1 are not allowed for DV2, because they require observation of precrash driver behavior. The remaining elements for DV2 are either maneuvers or conditions visible from outside the vehicle (e.g., most of the decision error choices) or reasonable general inferences (e.g., codes 109, 119, 139, 140, 149).

Attempted Avoidance Maneuver (C-N-I)

- 00 = No driver present.
- 00a = Not applicable (baseline epoch).
- 00b = Not applicable (single-vehicle event).
- 01 = No avoidance maneuver.
- 02 = Braking (no lockup or lockup unknown).
- 03 = Braking (lockup).
- 04 = Braking (lockup unknown).
- 05 = Releasing brakes.
- 06 = Steered to left.
- 07 = Steered to right.
- 08 = Braked and steered to left.
- 08a = Braked and steered to left (no lockup or lockup unknown).
- 08b = Braked and steered to left (lockup).
- 09 = Braked and steered to right.
- 09a = Braked and steered to right (no lockup or lockup unknown).
- 09b = Braked and steered to right (lockup).
- 10 = Accelerated.
- 11 = Accelerated and steered to left.
- 12 = Accelerated and steered to right.
- 98 = Other actions.
- 99 = Unknown if driver attempted any corrective action.

Comment: LTCCS Variable 7 and GES V27, corrective action attempted. The released gas pedal elements available for DV1 are not available for DV2, because they would not be observable from outside the vehicle.

Driver Behavior: Driver 2 Actions or Factors Relating to Event (C-N-I)

Note: Analysts code up to four factors believed to have relevance to the occurrence of the incident (e.g., contributing factors). If there are more than four, the four most important are selected.

- 00a = Not applicable (baseline epoch).
- 00b = Not applicable (single-vehicle event).
- 00 = None coded.

- 01 = Apparent excessive speed for conditions or location (regardless of speed limit; does not include tailgating unless above speed limit).
- 02 = Drowsy, sleepy, asleep, fatigued, other reduced alertness.
- 03 = Angry.
- 04 = Other emotional state.
- 05 = Alert but inattentive or distracted.
- 06a = Vehicle drift or slow weave consistent with possible drowsy/distracted driving.
- 06b = Erratic steering, weaving, lane break, or other vehicle motion consistent with possible alcohol-impaired driving.
- 07 = Driving slowly (below speed limit or in relation to other traffic).
- 08 = Illegal passing (e.g., across double line).
- 09 = Passing on right.
- 10 = Other improper or unsafe passing.
- 11a = Cutting in, too close in front of other vehicle.
- 11b = Cutting in at safe distance but then decelerating, causing conflict.
- 12 = Cutting in, too close behind other vehicle.
- 13 = Making turn from wrong lane (e.g., across lanes).
- 14 = Did not see other vehicle during lane change or merge.
- 15 = Driving in other vehicle's blind zone.
- 16 = Aggressive driving, specific, directed menacing actions.
- 17 = Aggressive driving, other (reckless driving without directed menacing actions).
- 18 = Wrong side of road, not overtaking (includes partial or full drift into oncoming lane).
- 19 = Following too closely.
- 19a = Inadequate evasive action.
- 20 = Failed to signal, or improper signal.
- 21 = Improper turn, wide right turn.
- 22 = Improper turn, cut corner on left turn.
- 23 = Other improper turning.
- 24 = Improper backing, (apparently) did not see.
- 25 = Improper backing, other.
- 26 = Improper start from parked position.
- 27 = Disregarded officer or watchman.
- 28 = Signal violation.
- 29 = Not used.
- 30 = Signal violation, tried to beat signal change.
- 31 = Stop sign violation.
- 32 = Not used.
- 33 = Stop sign violation, rolling stop.
- 34 = Other sign (e.g., yield) violation.
- 35 = Not used.
- 36 = Other sign violation.
- 37 = Nonsigned crossing violation (e.g., driveway entering roadway).
- 38 = Right-of-way error in relation to other vehicle or person.
- 39 = Not used.
- 40 = Not used.
- 41 = Sudden or improper stopping on roadway.
- 42 = Parking in improper or dangerous location (e.g., shoulder of interstate).
- 43 = Speeding or other unsafe actions in work zone.
- 44 = Failure to dim headlights.
- 45 = Driving without lights or insufficient lights.
- 46 = Avoiding pedestrian.
- 47 = Avoiding other vehicle.
- 48 = Avoiding animal.
- 48a = Avoiding object.
- 49 = Apparent unfamiliarity with roadway.
- 50 = Apparent unfamiliarity with vehicle (e.g., displays and controls).
- 51 = Use of cruise control contributed to late braking.
- 52 = Excessive braking/deceleration, creating potential hazard.
- 53 = Loss of control on slippery road surface.
- 54 = Loss of control on dry (or unknown) surface.
- 55 = Apparent vehicle failure (e.g., brakes).
- 56 = Other.
- 57 = Unknown.

Comment: Parallel variable to 40 (see Table C.1). Note, however, that a number of element choices relating to specific driver behaviors or impairments are disallowed because these are not observable for Driver 2. Also, for signal, sign, and right-of-way violations, analysts code the violation but do not attempt to ascertain whether the violation was intentional or caused by recognition failure. Thus, several elements are not used. As noted under 40, this variable was used in the 100-Car Naturalistic Driving Study, although some new elements have been added. Also, the coding rule is different; in the 100-Car Study, analysts coded up to three factors for each driver in descending order of judged importance. In the current study, analysts code all that apply in no order of importance. Thus, the data from the two studies are not directly comparable.

Applicable Functional Countermeasures for DV2 (C-N-I)

On the basis of the variables above that relate to the event scenario, pre-event actions and states, and event causation, senior analysts identify applicable functional countermeasures. For crashes, an applicable DV2 functional countermeasure is one that would probably have prevented the crash by either preventing the genesis of the unsafe condition or improving driver response to the unsafe condition. Near crashes and crash-relevant conflicts are analyzed as if a crash had occurred. Variable 41 (see Table C.1) provides a table of functional countermeasures and shows coding rules for them. The coding of functional countermeasures is based on both algorithmic determination from previous coded variables and analyst judgment. In many cases, particular accident type, critical reason, or other causation-related codes algorithmically determine applicable functional countermeasures. Some countermeasure choices, however, are coded based on senior analyst judgment. Most potential functional countermeasures are coded for DV2, but some are not, because little information is available to analysts on the specific Driver 2 behaviors and states.

General Variables

Event Comments (C-N-I-B)

Comment: This text variable will permit analysts to provide any comments on the event, including information not captured by data variables, assumptions made about the event affecting coding, and coding issues that arose. Ordinarily this will not contain information that is captured by the coded variables.

References

1. Thiriez, K., G. Radja, and G. Toth. *Large Truck Crash Causation Study: Interim Report.* Report DOT HS 809 527. NHTSA, 2002.
2. R. J. Hanowski, A. S. Keisler, and W. W. Wierwille. *Light Vehicle-Heavy Vehicle Interactions: A Preliminary Assessment Using Critical Incident Analysis.* Report FMCSA-RT-04-004. FMCSA, 2004.
3. R. J. Hanowski, R. L. Olson, J. S. Hickman, and T. A. Dingus. *The 100-Car Naturalistic Driving Study: A Descriptive Analysis of Light Vehicle-Heavy Vehicle Interactions from Light Vehicle Driver's Perspective.* Virginia Tech Transportation Institute, Blacksburg, Va., 2004.

APPENDIX D

Project 7 and Project 8 Environmental Data Dictionary

Event Variables

Driver ID

Enter the driver ID (from Slacker Tool) in the following format 00##.

Date

Enter the date on which the event occurred in the following format: mm/dd/yyyy.

Time

Enter the GMT time at the point of the trigger in the following format: hh:mm AM/PM. If the time is obviously wrong (e.g., it says 10:00 PM but it is daylight), write "time is wrong" in the text box.

Day of Week

Using a calendar, enter the day of week on which the event occurred.

Vehicles/Nonmotorists Involved

- 00 = Not applicable (baseline epoch).
- 01 = 1 vehicle (subject vehicle only or subject vehicle + object).
- 02 = 2 vehicles.
- 03 = 3 vehicles.
- 04 = 4 or more vehicles.
- 05 = Subject vehicle + pedestrian.
- 06 = Subject vehicle + pedalcyclist.
- 07 = Subject vehicle + animal.
- 08 = Other.

Note: For some events (e.g., those involving transient encroachment into an oncoming lane), it is difficult to decide whether the event should be considered a one- or two-vehicle event. Consider the event a two-vehicle event if the crash resulting from the incident would probably have involved two vehicles or if either driver's maneuvers were influenced by the presence of the other vehicle (e.g., if DV1 maneuvered to avoid V2). Consider the event a one-vehicle event if the presence of other vehicles presented no immediate threat and had no effect on DV1's maneuvers or behaviors.

Vehicle/Nonmotorist 2 Type

- 00a = Not applicable (baseline epoch).
- 00b = Not applicable (single-vehicle event, no object).
- 01 = Automobile.
- 02 = Van (minivan or standard van).
- 03 = Pickup truck.
- 04 = SUV (includes Jeep).
- 05 = School bus.
- 06 = Transit bus.
- 07 = Greyhound bus.
- 08 = Conversion bus.
- 09 = Single-unit straight truck: multistop/step van.
- 10 = Single-unit straight truck: box.
- 11 = Single-unit straight truck: dump.
- 12 = Single-unit straight truck: garbage/recycling.
- 13 = Single-unit straight truck: concrete mixer.
- 14 = Single-unit straight truck: beverage.
- 15 = Single-unit straight truck: flatbed.
- 16 = Single-unit straight truck: tow truck.
- 17 = Single-unit straight truck: other.
- 18 = Single-unit straight truck: unknown.
- 19 = Straight truck + trailer.
- 20 = Tractor-trailer: cab only.
- 21 = Tractor-trailer: cab + trailer.
- 22 = Tractor-trailer: flatbed.
- 23 = Tractor-trailer: tank.
- 24 = Tractor-trailer: car carrier.
- 25 = Tractor-trailer: livestock.

- 26 = Tractor-trailer: lowboy trailer.
- 27 = Tractor-trailer: dump trailer.
- 28 = Tractor-trailer: multiple trailers.
- 29 = Tractor-trailer: multiple trailers, grain.
- 30 = Tractor-trailer: other.
- 31 = Other large construction equipment.
- 32 = Ambulance.
- 33 = Fire truck.
- 34 = Motorcycle or moped.
- 35 = Police car.
- 36 = Vehicle pulling trailer (other than tractor-trailer).
- 37 = Other vehicle type.
- 38 = Pedestrian.
- 39 = Pedalcyclist.
- 40 = Deer.
- 41 = Other animal.
- 42 = Object (single-vehicle event with relevant object).
- 43 = Unknown.

Note: Highly abridged version of GES V5, body type; codes above do not match GES codes.

Relevant Object

Choose the most relevant object (i.e., one that was struck in a crash or that constituted a crash threat) for near crashes and crash-relevant conflicts.

- 00a = Not applicable (baseline epoch).
- 00b = Not applicable (single-vehicle event, no relevant object; e.g., shoulder only).
- 00c = Not applicable (multivehicle event, pedestrian, animal, etc.).
- 01 = Parked motor vehicle.

Fixed Objects

- 02 = Building.
- 03 = Impact attenuator/crash cushion.
- 04 = Bridge structure (e.g., abutment).
- 05 = Guardrail.
- 06 = Concrete traffic barrier or other longitudinal barrier (e.g., Jersey barrier).
- 07 = Post, pole, or support (e.g., sign or light).
- 07a = Mailbox.
- 08 = Culvert/ditch/edge of road.
- 09 = Curb.
- 10 = Embankment.
- 11 = Fence.
- 12 = Wall.
- 13 = Fire hydrant.
- 14 = Shrubbery or bush.
- 15 = Tree (not overhang; see below).
- 16 = Boulder.
- *17 = Loading dock.*
- *18 = Loading equipment (e.g., forklift or pallets).*
- *19 = Cargo.*

Overhanging Objects (Only If Struck or Potentially Struck by Top of Truck or Trailer)

- *20 = Tree branch.*
- *21 = Overhanging part of sign or post.*
- *22 = Bridge/overpass.*
- *23 = Building.*
- *24 = Telephone wires.*

Nonfixed Objects

- *25 = Vehicle parts, including tire parts.*
- *26 = Spilled cargo.*
- *27 = Dead animal in roadway.*
- *28 = Broken tree limbs or other tree/shrub parts.*
- *29 = Trash/debris.*
- *30 = Construction barrel.*
- *31 = Construction cone.*

Other

- *98 = Other.*
- *99 = Unknown object hit.*

Note: GES A06, first harmful event. Options in italics are not A06 codes.

Vehicle/Nonmotorist 2 Position (in Relation to V1)

The vehicle in Figure D.1 represents the subject vehicle (V1, the truck). The relative position of Vehicle 2 (in relation to V1) is coded for the time in which the critical event occurs (i.e., the event creating the crash risk). Vehicles in the adjacent left lane are coded J, I, H, or G, depending on their position. Vehicles in the adjacent right lane are coded B, C, D, or

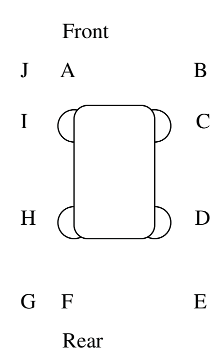


Figure D.1. Vehicle diagram.

E, depending on their position. Also code the position of animals, pedestrians, pedalcyclists, and objects.

- 00a = Not applicable (baseline epoch).
- 00b = Not applicable (single-vehicle event, no object).
- K = Top of vehicle.

Driver/Vehicle 1 variables
Driver/Vehicle 1 (DV1) is always the study subject driver/vehicle (i.e., the truck or truck driver).

Driver Seat Belt Worn?

- 00 = No.
- 01 = Yes.
- 02 = Unknown.

Possible to Do Observer Rating of Drowsiness?

- 00 = Yes.
- 01 = No, wearing sunglasses.
- 02 = No, not enough video.
- 03 = No, cannot see driver's eyes.

Driver 1 Vision Obscured

- 00 = No obstruction.
- 01 = Rain, snow, smoke, sand, dust.
- 02 = Reflected glare, sunlight, headlights.
- 03 = Curve or hill.
- 04 = Building, billboard, or other design features (includes signs, embankment).
- 05 = Trees, crops, vegetation
- 06 = Moving vehicle (including load).
- 07 = Parked vehicle.
- 08 = Splash or spray of passing vehicle or any other vehicle.
- 09 = Inadequate defrost or defog system.
- 10 = Inadequate lighting system (includes vehicle or object in dark area).
- 11 = Obstruction interior to vehicle.
- 12 = Mirrors.
- 13 = Head restraints.
- 14 = Broken or improperly cleaned windshield.
- 15 = Fog.
- 16 = Other vehicle or object in blind spot.
- 97 = Vision obscured, no details.
- 98 = Other obstruction.
- 99 = Unknown whether vision was obstructed.

Note: GES Variable D4. Element 16 added because of relevance to large trucks.

Environmental Variables

Environmental variables are coded at the time of the trigger.

Light Condition

- 01 = Daylight.
- 02 = Dark.
- 03 = Dark but lighted.
- 04 = Dawn.
- 05 = Dusk.
- 09 = Unknown.

Note: GES A19.

Weather

- 01 = No adverse conditions.
- 02 = Rain.
- 03 = Sleet.
- 04 = Snow.
- 05 = Fog.
- 06 = Rain and fog.
- 07 = Sleet and fog.
- 08 = Other (smog, smoke, sand/dust, crosswind, hail).
- 09 = Unknown.

Comment: GES A20.

Roadway Surface Condition

- 01 = Dry.
- 02 = Wet.
- 03 = Snow or slush.
- 04 = Ice.
- 05 = Sand, oil, dirt.
- 08 = Other.
- 09 = Unknown.

Comment: GES A15.

Relationship to Junction

- 00 = Nonjunction.
- 01 = Intersection.
- 02 = Intersection related.
- 03 = Driveway, alley access, etc.
- 03a = Parking lot.
- 04 = Entrance/exit ramp.
- 05 = Rail grade crossing.
- 06 = On a bridge.
- 07 = Crossover related.

- 08 = Other.
- 09 = Unknown.

Comment: GES variable A09. Baseline epoch coded at time of trigger.

Trafficway Flow

- 00 = Not physically divided (center two-way left-turn lane).
- 01 = Not physically divided (two-way trafficway).
- 02 = Divided (median strip or barrier).
- 03 = One-way trafficway.
- 09 = Unknown.

Note: GES variable V A11. Coded in relation to subject vehicle; baseline epoch coded at time of trigger.

Number of Travel Lanes

- 01 = 1.
- 02 = 2.
- 03 = 3.
- 04 = 4.
- 05 = 5.
- 06 = 6.
- 07 = 7+.
- 09 = Unknown.

Note: GES V A12. Per GES, if road is divided, only lanes in travel direction are counted. If undivided, all lanes are counted. Coded in relation to subject vehicle; baseline epoch coded at time of trigger. Count all contiguous lanes at the time and location of the incident; for example, include entrance or exit lanes if contiguous. Do not include lanes if blocked by cones or barrels.

Roadway Alignment

- 01 = Straight.
- 02a = Curve right.
- 02b = Curve left.
- 09 = Unknown.

Note: GES V A13, with expansion of curve choices. Coded in relation to subject vehicle; baseline epoch coded at time of trigger.

Roadway Profile

- 01 = Level (or unknown).
- 02a = Grade up.
- 02b = Grade down.
- 03 = Hillcrest.
- 04 = Sag.

Note: GES V A14, with expansion of grade choices. Coded in relation to subject vehicle; baseline epoch coded at time of trigger.

Traffic Density

Code the traffic density for the time before the precrash event.

- 01 = LOS A: Free flow. (Individual users are virtually unaffected by the presence of others in the traffic stream. Freedom to select desired speeds and to maneuver within the traffic stream is extremely high. The general level of comfort and convenience provided to the motorist, passenger, or pedestrian is excellent.)
- 02 = LOS B: Flow with some restrictions. (In the range of stable traffic flow, but the presence of other users in the traffic stream begins to be noticeable. Freedom to select desired speeds is relatively unaffected, but there is a slight decline in the freedom to maneuver within the traffic stream from LOS A because the presence of others in the traffic stream begins to affect individual behavior.)
- 03 = LOS C: Stable flow; maneuverability and speed are more restricted. (Traffic flow is in the stable range but is beginning to enter the range of flow in which the operation of individual users becomes significantly affected by interactions with others in the traffic stream. The selection of speed is now affected by the presence of others, and maneuvering within the traffic stream requires substantial vigilance on the part of the user. The general level of comfort and convenience declines noticeably at this level.)
- 04 = LOS D: Unstable flow; temporary restrictions substantially slow the driver. (This category represents high-density but stable traffic flow. Speed and freedom to maneuver are severely restricted, and the driver or pedestrian experiences a generally poor level of comfort and convenience. Small increases in traffic flow generally cause operational problems at this level.)
- 05 = LOS E: Flow is unstable; vehicles are unable to pass, temporary stoppages, etc. (Operating conditions are at or near capacity level. All speeds are reduced to a low but relatively uniform value. Freedom to maneuver within the traffic stream is extremely difficult and is generally accomplished by forcing a vehicle or pedestrian to give way to accommodate such maneuvers. Comfort and convenience levels are extremely poor, and driver or pedestrian frustration is generally high. Operations at this level are usually unstable because small increases in flow or minor perturbations within the traffic stream will cause breakdowns.)

- 06 = LOS F: Forced traffic flow condition with low speeds and traffic volumes that are below capacity; queues form in particular locations. (This condition exists whenever the amount of traffic approaching a point exceeds the amount that can traverse the point. Queues form behind such locations. Operations within the queue are characterized by stop-and-go waves and are extremely unstable. Vehicles may progress at reasonable speeds for several hundred feet or more and then be required to stop in a cyclic manner. LOS F is used to describe operating conditions within the queue, as well as the point of the breakdown. In many cases, operating conditions of vehicles or pedestrians discharged from the queue may be quite good. Nevertheless, it is the point at which arrival flow exceeds discharge flow, which causes the queue to form. LOS F is an appropriate designation for such points.)
- 09 = Unknown/unable to determine.

Construction Zone Related

- 00 = Not construction zone related (or unknown).
- 01 = Construction zone (occurred in zone).
- 02 = Construction zone related (occurred in approach or otherwise related to zone).

Note: Any area with one or more traffic cones, barrels, and so forth is considered to be a construction zone.

Truck Pre-Event Speed

Note: The pre-event speed is coded for the period just before the occurrence of the critical event or just before any avoidance maneuver or both. For example, when braking is involved, the pre-event speed is the speed just before the beginning of braking. If there is no avoidance maneuver, enter the speed at the time of the trigger.

- 999 = Unknown.

General Variables

Event Comments

Note: This text variable permits analysts to provide any comments on the event, including information not captured by data variables, assumptions made about the event affecting coding, and coding issues that arose. Ordinarily this will not contain information that is captured by the coded variables.

Related SHRP 2 Research

Roadway Information Database Development and Technical Coordination and Quality Assurance of the Mobile Data Collection Project (S04A)

Design of the In-Vehicle Driving Behavior and Crash Risk Study (S05)

In-Vehicle Driving Behavior Field Study (S07)

W W W . T R B . O R G / S H R P 2

ISBN 978-0-309-12898-8

90000

9 780309 128988