DESCRIPTION OF THE SHRP 2 NATURALISTIC DATABASE AND THE CRASH, NEAR-CRASH, AND BASELINE DATA SETS

TASK REPORT

Prepared for
The Strategic Highway Research Program 2
Transportation Research Board
of
The National Academies

Jonathan M. Hankey
Miguel A. Perez
Julie A. McClafferty
Virginia Tech Transportation Institute
Blacksburg, VA
April 2016

ACKNOWLEDGEMENT OF SPONSORSHIP

This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials, and was conducted in the Strategic Highway Research Program, which is administered by the Transportation Research Board of the National Academies.

DISCLAIMER

DESCRIPTION OF THE SHRP 2 NATURALISTIC DATABASE AND THE CRASH, NEAR-CRASH, AND BASELINE DATA SETS

TASK REPORT

Prepared for
The Strategic Highway Research Program 2
Transportation Research Board
of
The National Academies

Jonathan M. Hankey, Virginia Tech Transportation Institute,
Blacksburg, VA
Julie A. McClafferty, Virginia Tech Transportation Institute,
Blacksburg, VA
Miguel A. Perez, Virginia Tech Transportation Institute,
Blacksburg, VA

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

AUTHOR ACKNOWLEDGEMENTS

The work reported here required a substantial team effort. The authors wish to acknowledge the team of data reductionists at the Virginia Tech Transportation Institute (VTTI) who reviewed the thousands of possible crashes and near-crashes that were flagged for review and validation. They coded 75 variables for each identified crash and near-crash, as well as many of the same variables across more than 32,000 baselines. The following Data Reduction Coordinators at VTTI were instrumental in coordinating the data reduction work performed: Megan Moore, Nelson Gunter, Brunilda Swannell, Ben Boucher, Stuart Walker, Robert White, and Matt Marchese. The authors also want to thank the Information Technology team at VTTI who kept the database running smoothly despite the many constraints in place. Feng Guo, a statistician at VTTI, was responsible for developing the initial baseline sample. Finally, the authors would like to thank Ken Campbell and Jim Hedlund for their help and guidance throughout this project.

ABSTRACT

The focus of this project was to identify and prepare crash, near-crash, and baseline data sets extracted from the Second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS) trip files, then to make that information available to researchers for use in their analysis projects. A dozen trigger algorithms were executed on 5,512,900 trip files in the SHRP 2 NDS, and a manual validation of these algorithms identified 1,549 crashes and 2,705 near-crashes. A longitudinal deceleration-based algorithm produced the highest percentage of valid crashes and near-crashes. Baselines were selected via a random sample stratified by participant and proportion of time driven. Triggered epochs and the resulting crashes and near-crashes were reviewed and analyzed by a large team of data reductionists and quality control coordinators following a rigorous training, testing, and monitoring protocol. As a result, 20,000 baselines, including all drivers in the SHRP 2 NDS, were prepared and are recommended for researchers using a case-cohort design. An additional 12,586 baselines are also available for researchers who may require more power in their analyses but are able to forego a fully proportional representation of all drivers in the study. Researchers using this data set are encouraged to review the data dictionaries on the InSight website prior to doing analysis and to be particularly careful in selecting the best subset of crashes, near-crashes, and baselines that informs their research questions.

CHAPTER 1: BACKGROUND

The Second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS), the largest study of naturalistic driving behaviors to date, monitored approximately 3,400 participant drivers and produced over 4,300 years of naturalistic driving data between 2010 and 2013. Data were collected from six sites around the United States. The largest collection sites were in Seattle, Washington; Tampa, Florida; and Buffalo, New York. Each of these sites collected over 20% of the data. Durham, North Carolina, accounted for approximately 15% of the data, while State College, Pennsylvania, and Bloomington, Indiana, each accounted for over 5% of the data. Over 3,300 participant vehicles were instrumented with a data acquisition system (DAS) that collected four video views (driver's face, driver's hands, forward roadway, rear roadway), vehicle network information (e.g., speed, brake, accelerator position), and information from additional sensors included with the DAS (e.g., forward radar, accelerometers). A SHRP 2 report, "Technical Coordination and Quality Control" (Dingus et al., 2014), provides further description of the DAS used in the study; another report, "Comparing the SHRP 2 NDS Sample with National Data" (Antin et al., 2014), provides more information about the data collection sites and how they compare to national data.

Description of SHRP 2 Database

Collecting and archiving this data was a massive undertaking that involved hundreds of people. In total, 6,559,367 files were collected during the SHRP 2 NDS. A trip file usually encompassed a whole trip from approximately 30 seconds after the ignition was turned on until the ignition was turned off. However, long trips (e.g., trips over an hour) could be split into more than one trip file. An additional safety mechanism built into the DAS could also shut down the system when the battery voltage was low in order to protect the vehicle. In these cases, a trip file could contain less than the entire trip, and multiple trip files could be associated with the same trip (e.g., if the battery voltage improved as the trip continued). Consequently, the data ingestion process was designed to keep as much of the usable data as possible, even if the trip was short or some of the data were missing.

Approximately 1.15% (75,370) of these trip files were excluded from the database; over 70% of these excluded files had missing or unusable video (54,468). It is likely that a number of these excluded trip files with missing video were DAS bench tests, since the data collection sites typically tested each system prior to DAS installation in, and upon DAS removal from, a vehicle. Installers also stated that they would conduct multiple bench tests during slow periods of work as they were familiarizing themselves with the DAS. However, the actual number of these bench tests is indeterminate. The remaining approximately 21,000 trip files were excluded because they were test vehicles, bench tests including video, or data from drivers that could not be correctly identified as consented participants. Additionally, at the beginning of the study when the data collection sites came online, some problems arose with vehicle and driver identification. While quickly rectified, these problems caused the loss of some trip files. In general, the proportion of trip files lost seems exceptionally low given the size of the project, the use of a newly designed DAS, and the fact that the six data collection sites had minimal or no experience in large-scale naturalistic driving data collection.

Only participants who signed an informed consent form agreeing to be in the SHRP 2 NDS could actually be considered in the study, which required that each collected trip file be manually reviewed to exclude data from non-consented drivers. Overall, data reductionists reviewed nearly 99% (6,483,997) of the trip files. During this review, the correct participant identification number was assigned to each trip file, thus making it easier for future researchers to access the drivers of interest for their specific research questions. A SHRP 2 report, "Identification of Consented Driver Trips in the SHRP 2 Naturalistic Driving Study Data Set," (McClafferty et al., 2015) describes this driver identification task in detail and includes the number of trip files driven by each participant. In total, approximately 85% of the collected trip files were included in the SHRP 2 NDS database (see Table 1.1). The largest category of excluded trip files consisted of unconsented drivers, accounting for approximately 10.5% of the manually reviewed trip files. Only nine of the 3,358 vehicles included in this study had no unconsented drivers. The largest number of unconsented trip files in one vehicle was 4,927. In that case, consented drivers drove approximately 60% of the time. On average, each vehicle contributed 204 unconsented trip files to the database. As shown in Figure 1.1 below, a total of 1,870 vehicles, over 50% of all vehicles in the study, had 5% or fewer unconsented trip files. Seventy-three percent of the SHRP 2 vehicles (i.e., 2,451) had 10% or fewer unconsented trip files.

Table 1.1 Number of SHRP 2 Trip Files in Each Driver Category

Driver Category	Number of Trip Files	Percent of Files
Consented Driver	5,512,900	85.02%
Unknown (likely unconsented)	684,733	10.56%
Trip took place prior to consent	39,936	0.62%
No Driver	221,051	3.41%
Data Collection Site Technician	12,829	0.20%
Multiple Drivers	12,548	0.19%
Total	*6,483,997*	*100.00%*

Trip files also had to be excluded if they occurred before the driver gave consent, which accounted for approximately 0.6% of omitted trip files. Note that, in most cases, these exclusions occurred for secondary drivers (i.e., sporadic drivers of the vehicle who were inducted into the study after the primary driver had enrolled and the vehicle had been instrumented). Of the 3,358 vehicles, 821 had at least one trip file that had to be excluded for this reason; however, over 80% of these 821 vehicles had 10 trip files or fewer lost due to trips being collected before driver consent. Twenty-five vehicles had over 500 trip files excluded due to trips being taken prior to consent. On average, these 25 vehicles had a third of their trip files excluded, with the largest number of excluded trip files being 2,667.

Trip files with no driver present (e.g., vehicle was empty and parked with the motor running) accounted for 3.4% of the excluded trip files. In about 0.2% of the trip files, the data collection site technician was recognized and coded as such. Finally in another 0.2% of the trip files, multiple drivers drove during the same trip. All trips with each of these types of trip files were also excluded from the data set.

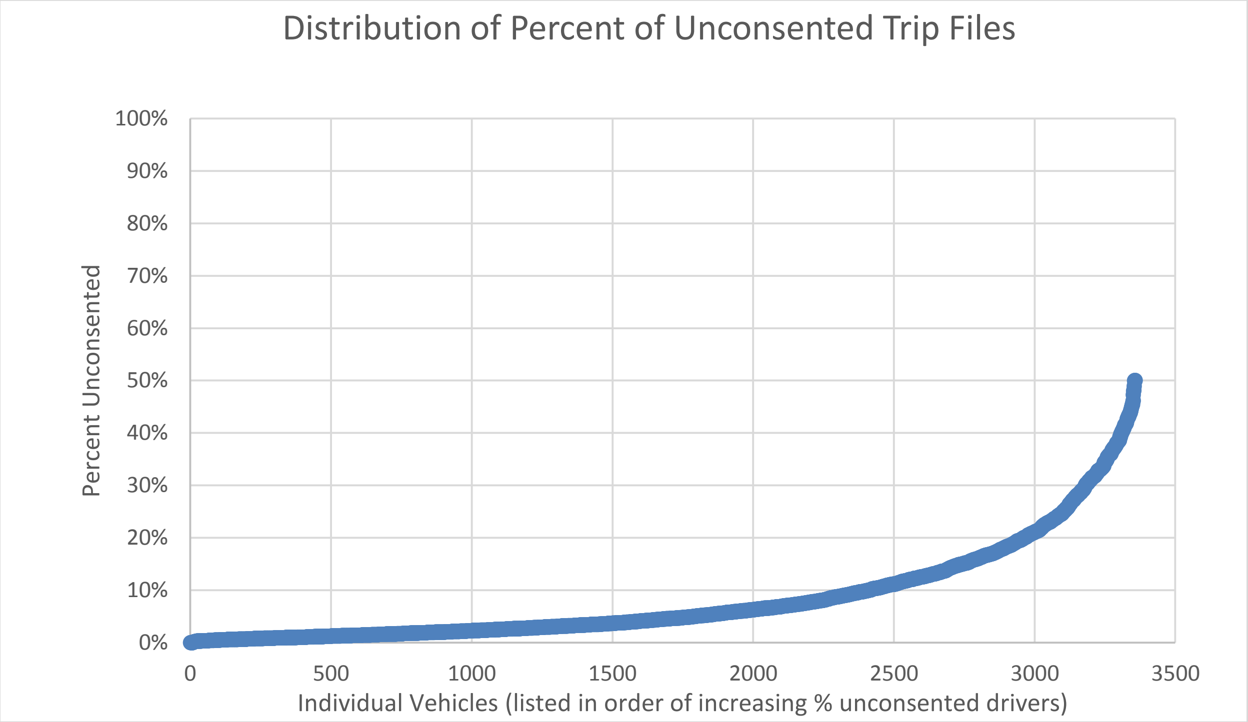


Figure 1.1 Percent of each vehicle's trip files that were unconsented drivers

A total of 5,512,900 consented driver trip files exist in the SHRP 2 NDS. The largest number of trip files driven in a vehicle was 20,939, while the next highest was approximately half that many at 10,336. Figure 1.2 plots the number of trip files from each of the other vehicles, with the vehicles arranged in order of increasing number of trip files. Approximately 32% of the vehicles had 1,000 or fewer trip files. Approximately 70% of the vehicles had 2,000 or fewer trip files. Approximately 11% had greater than 3,000 trip files (Figure 1.2).

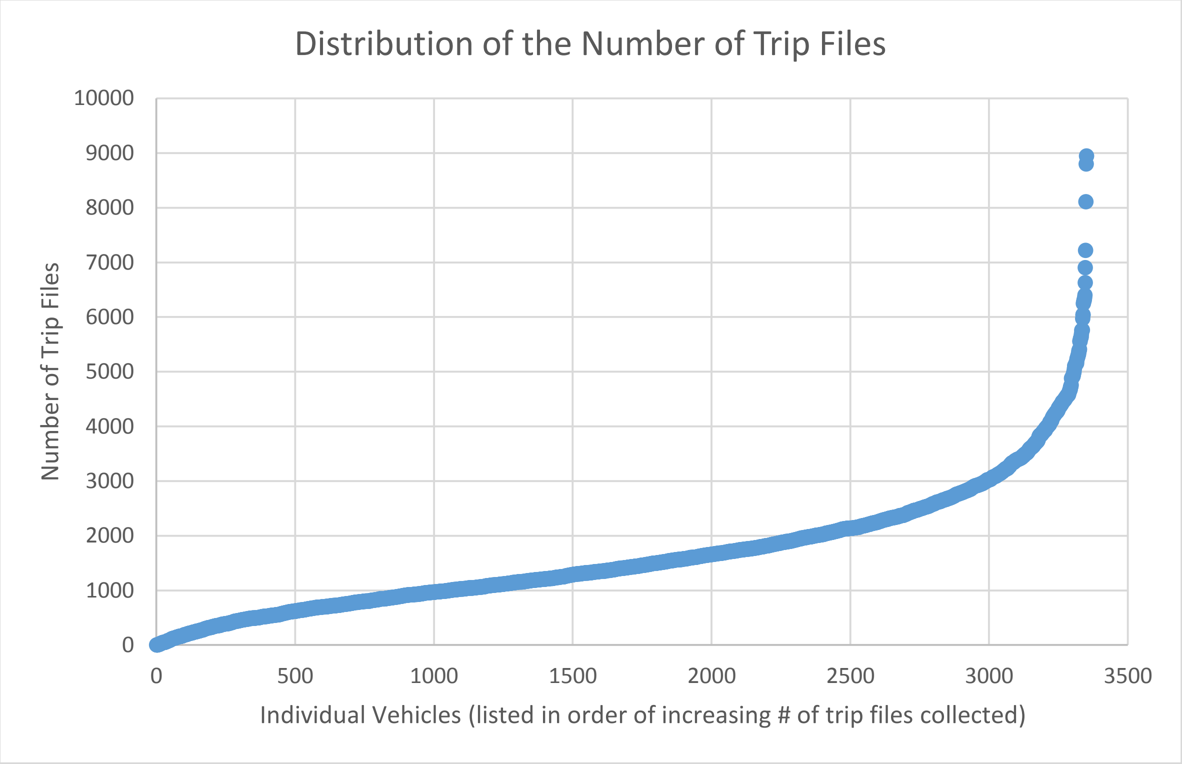


Figure 1.2.Number of trip files from each vehicle in the SHRP 2 NDS.

InSight Website

The SHRP 2 NDS collected approximately two petabytes of data, which can be categorized as shown in Figure 1.3. Given the volume of data, several usability tasks were accomplished to make the data more accessible and usable for researchers, including the "face" of the usability effort, the InSight website (https://insight.shrp2nds.us/).

The InSight website was developed to facilitate use of this enormous and useful data set within the transportation research community and beyond. The website was designed to allow some research questions to be answered directly, as well as to provide the information necessary for planning how to answer other research questions requiring more in-depth exploration of the SHRP 2 NDS. InSight also includes thorough data and variable dictionaries (e.g., SHRP 2 Researcher Dictionary for Video Reduction Data) to assist researchers attempting to interpret variables. Special care was taken to remove all personally identifying information from the website in order to maximize the number of potential researchers who could access the data and to limit the extent of any access restrictions. Although the initial InSight website is now complete, future enhancements are also expected. Figure 1.4 shows the website's query page.

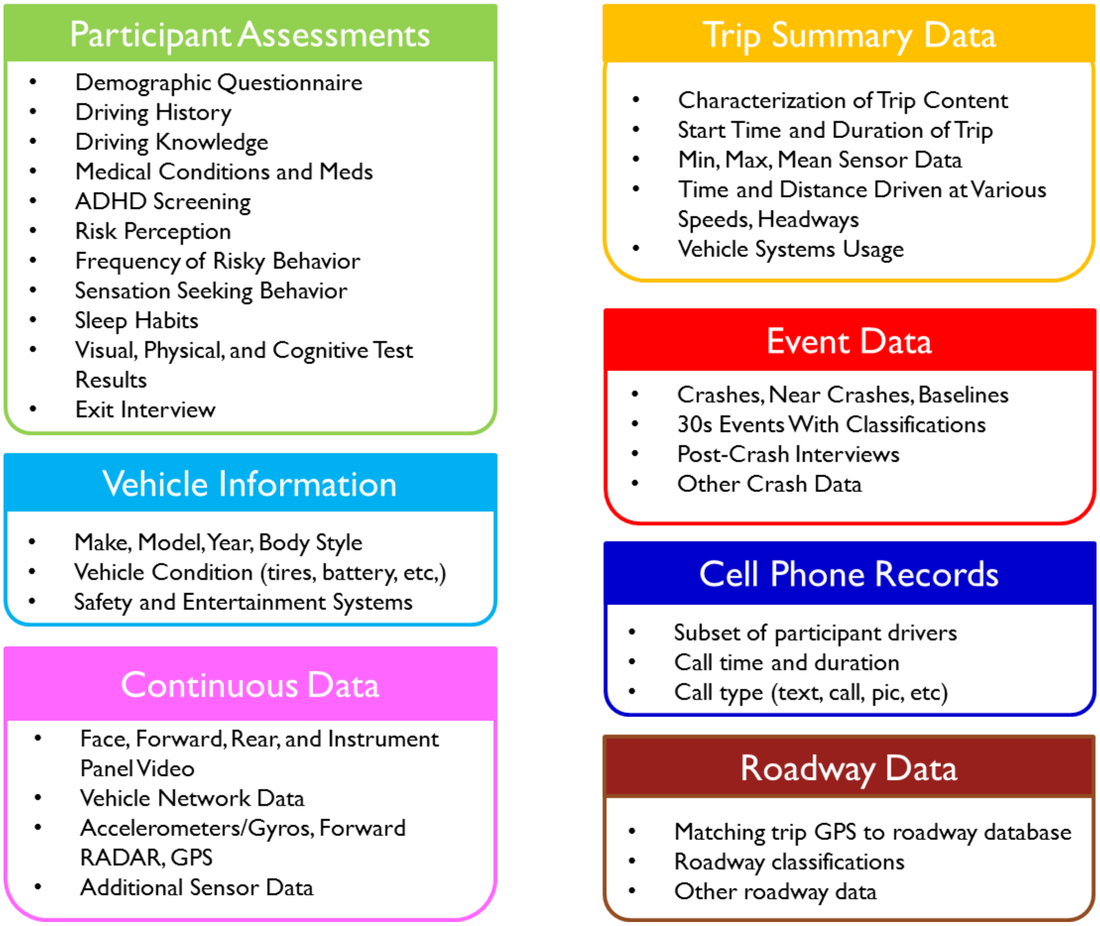


Figure 1.3. Representation of the data categories collected in the SHRP 2 project.

The InSight website is expected to be used heavily to work with data from the following sources, and researchers will interact with this data primarily via the query page shown in Figure 1.4.

- Detailed participant assessments, including demographic questionnaires, health and driving questionnaires, and vision tests, completed on over 3,100 drivers. (A SHRP 2 report titled “Enhancing Usability of Select Driver Assessment Data” explains some enhancements that were performed on the driver assessment data [Antin et al., forthcoming].)

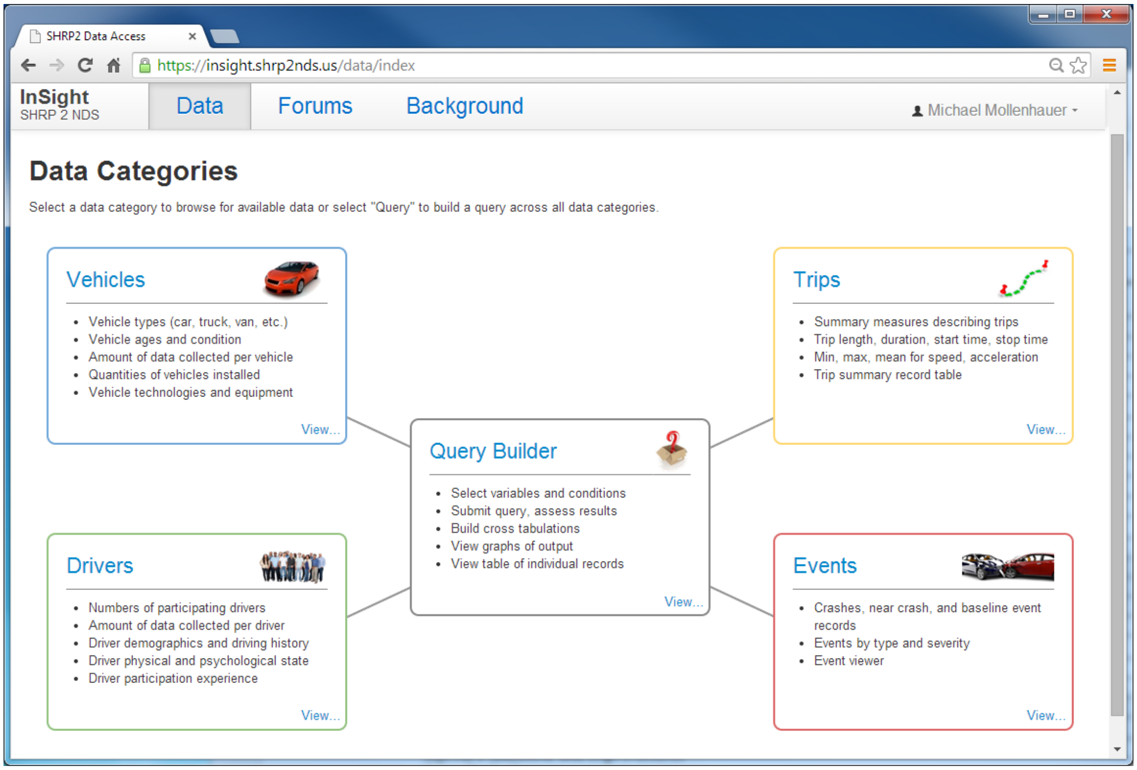
- Vehicle information (e.g., safety and entertainment options) for all 3,358 vehicles.

- Summary variables on over 5.5 million trip files (e.g., the maximum speed reached during a trip file, the maximum deceleration achieved, trip file duration).

- An interactive “heat map” detailing the roads driven by drivers in the study and how many times they were driven. (The matching process is described in a SHRP 2 report titled “Linking the Study Data to the Roadway Information Database” [McLaughlin and Hankey, 2015]).

- Event data from the crashes, near-crashes, and baselines that were identified and selected.

The event data mentioned in the final bullet point above are the focus of the remainder of this report.



Data Categories

Select a data category to browse for available data or select "Query" to build a query across all data categories.

Vehicles
- Vehicle types (car, truck, van, etc.)
- Vehicle ages and condition
- Amount of data collected per vehicle
- Quantities of vehicles installed
- Vehicle technologies and equipment

View...

Trips
- Summary measures describing trips
- Trip length, duration, start time, stop time
- Min, max, mean for speed, acceleration
- Trip summary record table

View...

Query Builder
- Select variables and conditions
- Submit query, assess results
- Build cross tabulations
- View graphs of output
- View table of individual records

View...

Drivers
- Numbers of participating drivers
- Amount of data collected per driver
- Driver demographics and driving history
- Driver physical and psychological state
- Driver participation experience

View...

Events
- Crashes, near crash, and baseline event records
- Events by type and severity
- Event viewer

View...

Figure 1.4. Data query page from InSight website.

CHAPTER 2: RESEARCH APPROACH

A notable aspect of the SHRP 2 NDS is the inclusion of crash, near-crash, and baseline data. As evidenced by the analyses for the 100-Car NDS database and subsequent similar efforts (Dingus et al., 2006), detailed observation and coding of these events can lead to critical findings related to traffic safety. Estimates of prevalence and risk from different driver behaviors, environmental conditions, and roadway characteristics can be derived when crashes and near-crashes are observed and catalogued in a naturalistic data set. In addition, baselines are necessary for comparisons with crashes and near-crashes to, for example, calculate risk estimates. Thus, a major component of the InSight website is the data from crashes, near-crashes, and baselines, which are the three categories designated as "events."

This chapter will describe:

- What crashes, near-crashes, and baselines (i.e., events) are;
- How events were identified and selected from the SHRP 2 NDS;
- Which events were selected;
- How additional information was extracted from the video and other data sources and used to annotate these events;
- Additional guidance and suggestions for researchers using these events to answer research questions.

Crash, Near-Crash, and Baseline Definition and Identification

In this context, it is important to operationally define what a crash and a near-crash represent. Following are the research definitions used at the Virginia Tech Transportation Institute (VTTI) for these events.

- **Crash**: Any contact that the subject vehicle has with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated is considered a crash. This also includes non-premeditated departures of the roadway where at least one tire leaves the paved or intended travel surface of the road, as well as instances where the subject vehicle strikes another vehicle, roadside barrier, pedestrian, cyclist, animal, or object on or off the roadway.

- **Near-Crash**: Any circumstance that requires a rapid evasive maneuver by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal, to avoid a crash is considered a near-crash. A rapid evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs.

Throughout this report, the combination of crashes and near-crashes will be referred to as safety-critical events (SCEs). The SHRP 2 NDS used multiple methods to identify SCEs, including the following.

- **Data Collection Site Report**: During the data collection, a participant reported they were involved in a crash, or the data collection site staff noticed previously unrecorded damage to a vehicle while it was active in the study. The information provided from the sites was used to comb through the data to find the potential event reported by the participant or suspected by the staff. Video verification was then used to determine if an SCE occurred and was collected by the DAS.

- **Automatic Crash Notification (ACN)**: The vehicle's DAS, via the execution of internal algorithms, detected a kinematic signature that could indicate a crash and uploaded a snippet of data around the potential event for analyst review. Video verification was then used to determine if an SCE occurred.

- **Critical Incident (CI) Button**: The vehicle's DAS included a critical incident button that participants could press in order to identify potential events of interest. Some of these potential events were actual SCEs; others were drivers completing their initial introduction as participants for driver identification. When the button was pressed, it created a flag in the corresponding trip file that was used to identify when a potential SCE occurred. Video verification was then used to determine if an SCE occurred.

- **Analyst Identified**: The event was identified by manually reviewing video footage for other purposes, such as while reducing a baseline or during driver ID. Sometimes events that were not identified in other ways were caught this way.

- **"Trigger" Execution**: The most systematic approach to finding an SCE was post hoc processing of incoming or resident data via custom algorithms called "triggers." These algorithms used kinematic and behavioral signatures that had a high probability of being present during particular types of SCEs. Different thresholds were used based on project resources and the importance of identifying the majority of crashes and near-crashes. Video verification then was used to determine if an SCE occurred.

Crashes and near-crashes are not the only essential elements in determining the risk of being involved in an SCE or the only essential elements for many other analyses. Although it is important to quantify this data, collections of crash and near-crash information alone do not provide any true indication of how prevalent certain activities, environmental conditions, or roadway attributes are. Without knowledge of that prevalence, or exposure, it is possible to incorrectly infer that potential risk factors modify risk only because they occur frequently in crashes. For example, consider mirror-scanning behavior, which occurs frequently while driving. Because of this frequency, it is likely to appear in situations when crashes or near-crashes occur. Without the knowledge that this behavior is also frequent in normal driving, it would be easy to incorrectly infer that crashes and near-crashes are partly due to mirror-scanning behaviors.

Exposure metrics can be obtained using different approaches. The most common of these approaches in NDS data sets is to select samples of "normal" driving and code them in a fashion similar to that used for crashes and near-crashes. These "normal" driving events are typically referred to as baseline events. The sampling approach used to select these baseline events varies based on the type of analysis being conducted. For the purpose of this project, a baseline sample was selected for potential use in the exposure metrics, a process that is further detailed later in this chapter. It is fully expected that as future researchers use the SHRP 2 NDS, additional baseline samples will be specified and made available.

The flow chart in Figure 2.1 provides an overview of the SCE and baseline identification process.

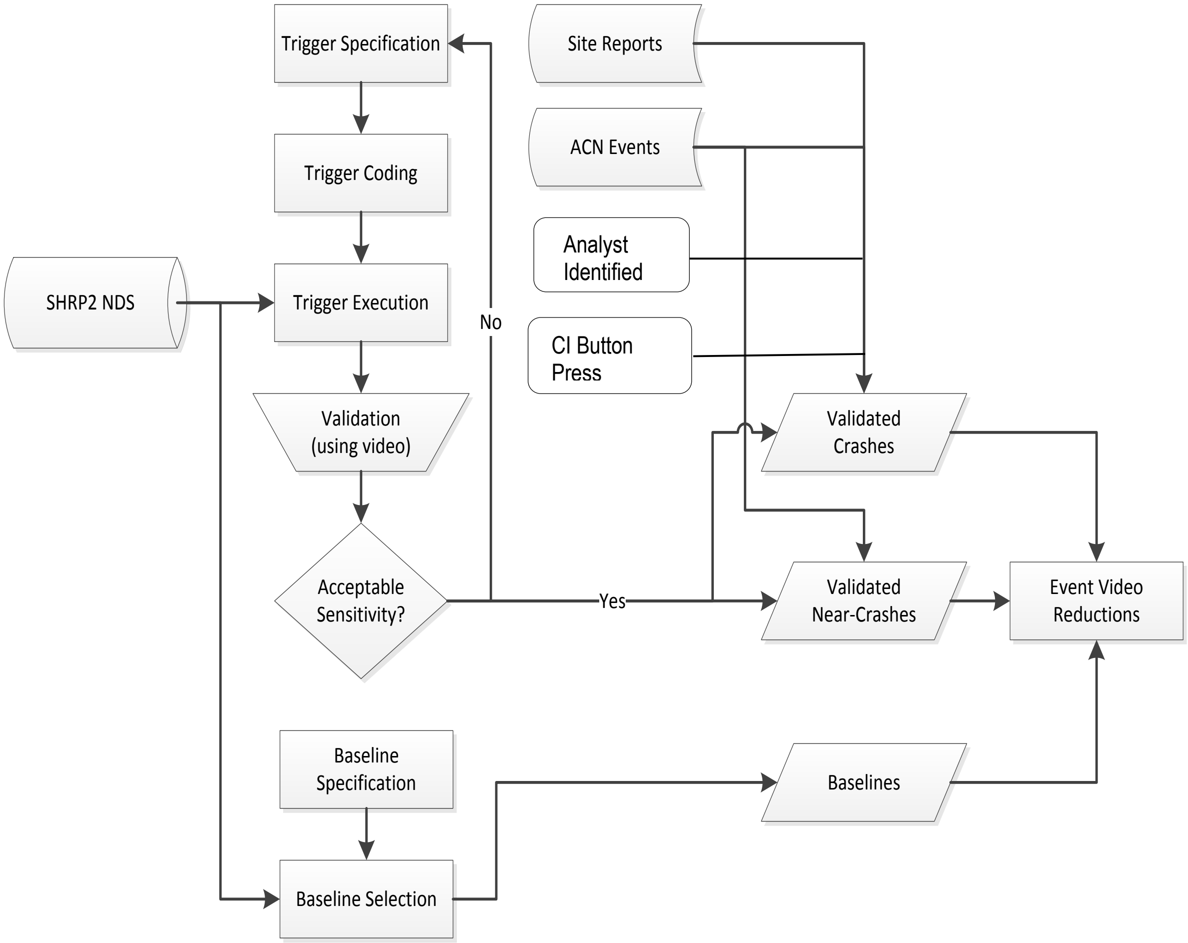


Figure 2.1. Flowchart overview of SCE and baseline process.

Trigger Specification and Validation

Table 2.1 shows a set of initial trigger specifications that were coded for the SHRP 2 NDS data. These were based on previous triggers conducted on other VTTI naturalistic data and new triggers that made use of some unique aspects of the SHRP 2 NDS. For example, the freeway deceleration trigger utilized the matching of trip files to the roadway (McLaughlin and Hankey, 2015) and excluded all non-freeway areas. Since data ingestion was conducted concurrently with some of the trigger coding, some of the coding was done based on efficient use of the available resources instead of the most effective trigger. As a practical matter, the trigger coding was performed so as to allow the entire project (i.e., collection, database testing, and analysis, etc.) to continue running smoothly.

As part of this process, trigger performance was continuously monitored, tracked, and used to determine the number of potential SCEs identified by each trigger. Also, errors in the data files found via the trigger algorithm and video reviews were flagged for further quality improvements (See Table 2.1.)

Table 2.1 Description of Initial Trigger Specifications

Trigger Type	Description
Longitudinal Deceleration	The level of longitudinal acceleration is less than or equal to -0.65 g. The threshold is exceeded for at least one timestamp. Multiple triggers within a 2-s window are joined into the same potential event.
Longitudinal Acceleration	The level of longitudinal acceleration is greater than or equal to 0.50 g. The threshold is exceeded for at least one timestamp. Multiple triggers within a 2-s window are joined into the same potential event.
Freeway Deceleration	The level of longitudinal acceleration is less than or equal to -0.3 g when the vehicle travels on a freeway segment, as defined by the base map used in the Roadway Information Database (RID). This level of deceleration or higher lasts for at least one timestamp. Multiple triggers within a 2-s window are joined into the same potential event.
Lateral Acceleration	The lateral acceleration is greater than or equal to 0.75 g or less than or equal to -0.75 g. The threshold is exceeded for at least 0.2 s. Multiple triggers within a 2-s window are joined into the same potential event.
Swerve	The derivative of yaw rate is monitored to find cases where the signal defines one complete cycle of a sine waveform whose minimum and maximum exceed ±15 deg/s/s within 2 s. The minimum speed is 5 m/s (~11 mph).
Yaw Rate	Vehicle swerves from ±8 degrees per second to ±8 degrees per second (in the opposite direction) within a window of 0.75 s. The minimum activation speed is 13.4 m/s (30 mph). Multiple triggers within a 2-s window are joined into the same potential event. As a noise-reduction measure, the algorithm looks at the first three autocorrelation lag values for the yaw rate of each event. If any of the autocorrelation values is negative, meaning there is a lot of noise, the potential event is discarded. This works well for 10 Hz; a different number of autocorrelation lag values may be necessary at other data collection frequencies.

Trigger Type	Description
Advanced Safety System Activation	Some vehicles provided the ability to monitor for activation of advanced safety systems such as anti-lock braking, traction control, airbag deployment, and electronic stability control systems. For these vehicles, different algorithms searched the appropriate network variables to look for system activations.
Longitudinal Jerk	The derivative of longitudinal acceleration was less than -1.0 g/s for 1 s while the vehicle travels at 5 m/s (~11 mph) or higher speeds.
Steering Evasive Maneuver	The absolute value of the derivative of lateral acceleration is greater than 1.0 g/s for 0.8 s while the vehicle travels at 5 m/s (~11 mph) or higher speeds.

While trigger specifications were developed to be selective and to reduce the likelihood that an important SCE was missed, this selectivity came at a cost: more “false alarm” events appeared than would have been produced under more stringent requirements for the trigger criteria. For this reason, every SCE was manually validated using the available video and parametric data. This validation process also provided information about the approximate severity of the SCE and the approximate timing of key events. SCE severity, which is included in the database, was assigned during the subsequent event reduction.

Each of the events reviewed during the validation was assigned one of the following designations (or severities). Of note, only crashes and near-crashes identified per the following definitions were subject to further analysis. These definitions can also be found on the InSight website.

1) **Crash**: This refers to any contact that the subject vehicle has with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated. It also includes non-premeditated departures of the roadway where at least one tire leaves the paved or intended travel surface of the road, as well as the subject vehicle striking other vehicles, roadside barriers, animals, pedestrians, cyclists, or objects on or off of the roadway. During the event reduction (discussed in the section entitled “Event Reduction and Annotation”), events identified as crashes were further classified into the following Crash Severity categories.

    a) *Level 1 Severe Crash*: Any crash that includes an airbag deployment; any known injury of driver, pedal cyclist, or pedestrian (one sufficient to warrant a doctor’s visit, including those self-reported and those apparent from video); a vehicle rollover; a high Delta-V; or vehicle damage requiring towing. A high Delta-V is defined as a change in speed of the subject vehicle in any direction during impact greater than 20 mph (excluding curb strikes) or (more commonly) acceleration on any axis greater than +/-2 g (excluding curb strikes).

    b) *Level 2 Crash Moderate Severity*: Not a level 1 crash; minimum of approximately $1,500 worth of damage as estimated from video. It also includes crashes that reach acceleration on any axis greater than +/-1.3 g (excluding curb strikes). Examples are most large animal and sign strikes.

c) *Level 3 Crash Minor Severity*: Not a level 1 or 2 crash; the vehicle makes physical contact with another object or departs the road but sustains only minimal or no damage. This includes most road departures (unless criteria for a more severe crash are met), small animal strikes, all curb and tire strikes potentially in conflict with oncoming traffic, and other curb strikes with an increased risk element (i.e., the crash may have been worse if the curb had not been there).

d) *Level 4 Crash Tire Strike, Low Risk*: Not a level 1, 2, or 3 crash; the tire is struck with little or no risk element (e.g., clipping a curb during a tight turn).

2) **Near-Crash**: This refers to any circumstance requiring a rapid evasive maneuver by the subject vehicle or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. Near-crashes must meet the following four criteria.

   a) *Not a Crash*: The vehicle must not make contact with any object, moving or fixed, and the maneuver must not result in a road departure.

   b) *Not Premeditated*: The maneuver performed by the subject must not be premeditated. This criterion does not rule out near-crashes caused by unexpected events experienced during a premeditated maneuver (e.g., a premeditated aggressive lane change resulting in a conflict with an unseen vehicle in the adjacent lane that requires a rapid evasive maneuver by one of the vehicles).

   c) *Evasion Required*: An evasive maneuver is performed or required by the subject or another vehicle, pedestrian, cyclist, or animal to avoid a crash. This is defined as steering, braking, accelerating, or combination of control inputs performed to avoid a potential crash.

   d) *Rapidity Required*: Rapidity refers to the swiftness of the response given the amount of time from the beginning of the participant’s reaction and the potential time of impact.

3) **Crash-Relevant Conflict**: This refers to any circumstance that requires an evasive maneuver on the part of the participant vehicle or any other vehicle, pedestrian, cyclist, or animal that is less urgent than a rapid evasive maneuver (as defined above in near-crash) but greater in urgency than a “normal maneuver” to avoid a crash. A crash-avoidance response can include braking, steering, accelerating, or any combination of control inputs. Crash-relevant conflicts must meet the following four criteria.

   a) *Not a Crash*: The vehicle must not make contact with any object, moving or fixed, and the maneuver must not result in a road departure.

b) *Not Premeditated*: The maneuver performed by the subject must not be premeditated. This criterion does not rule out crash-relevant conflicts caused by unexpected events experienced during a premeditated maneuver (e.g., a premeditated aggressive lane change resulting in a conflict with an unseen vehicle in the adjacent lane that requires a non-rapid evasive maneuver by one of the vehicles).

c) *Evasion Required*: An evasive maneuver to avoid a crash was required by the subject or another vehicle, pedestrian, cyclist, or animal. An evasive maneuver is defined as steering, braking, accelerating, or combination of control inputs performed to avoid a potential crash.

d) *Rapidity Not Required*: The evasive maneuver is not required to be rapid.

4) **Non-Participant Conflict**: This refers to any event captured on video—crash-relevant, near-crash, or crash—that does not involve the participant driver.

5) **Non-Conflict**: This refers to any incident or maneuver within the bounds of "normal driving" behaviors and scenarios that is accurately represented by the time series data that created the flagged event. The driver may react to situational conditions and events, but the reaction is not evasive and the situation does not place the subject or others at elevated risk.

Baseline Specification and Selection

The selection of baselines is one of the key aspects of most analyses involving naturalistic driving data, especially those related to the estimation of risk. Many sampling approaches are possible, all with the goal of minimizing the potential for biases into the analysis sample. The goal of the baselines is to provide an estimate of what constitutes "normal driving" and "typical driver behavior" across the sample; tradeoffs in data selection have to be made to avoid overrepresentation of participants, driving environments, and many other potential confounding factors.

To select the baselines for this project, VTTI developed a proposed baseline sampling method that could be used to answer multiple research questions. This method was presented to an Expert Technical Group (ETG) with experience conducting reviews of naturalistic driving studies in order to solicit feedback. Feedback was also obtained from key members of the SHRP 2 NDS Technical Coordinating Committee (TCC) who had relevant knowledge.

The objective of the baseline sample was to provide the necessary information to answer a variety of research questions in the following categories.

- **Exposure**: the prevalence of factors under normal driving conditions (i.e., analyzing the baseline file by itself).

- **Risk Evaluation**: the base for evaluating the relative risk of factors (i.e., comparing crash and/or near-crash events to the baselines).

Both a case-cohort type sample and case-crossover type sample were considered. A case-cohort baseline sample relies on a random sample of the population in which the cases were found. A case-crossover type sample uses a matched sampling scheme in which each case has several matched samples that control for certain factors (e.g., time of day, weather, roadway infrastructure, etc.; Guo and Hankey, 2009). There are trade-offs for both types of baselines. The case-crossover design baseline cannot be used for estimating factor prevalence, one of the key goals of this SHRP 2 baseline sample. Conversely, given a very specific research question, a case-cohort design may have less power and be more prone to bias than a well-designed case-crossover baseline sample. Table 2.2 summarizes the pros and cons of the two approaches that were considered.

A case-cohort design was ultimately selected primarily because, given the funding available to complete this task, it could be used to answer a wider range of research questions. The design also provides the most flexibility when considering multiple research questions.

Table 2.2. Case-Cohort and Case-Crossover Comparison

	Pros	Cons
Case-Cohort	• Can be used to answer a wide range of research questions • Can be used to estimate prevalence • Easy to implement and cost-benefit efficient • Flexible for future expansion	• Might be more prone to bias in risk assessment
Case-Crossover	• Less prone to bias when applied properly	• Difficult to reach consensus on matching factors • Hard to sample • Hard to analyze • Cannot be used for estimating prevalence

A random sampling scheme stratified by participant and proportion of time driven was used. All participants were included in the sample regardless of whether they were involved in a crash or near-crash. A minimum of one baseline was included for each driver in the study. Time driving was operationally defined to include only driving speeds above 5 mph, which was done to eliminate the effect of long stopping time and focus on periods of time when the vehicle was at risk of an at-fault crash.

As a practical matter, the baselines had to be selected prior to the end of data collection for every driver in order to complete the baseline selection and reduction as near to the completion of the overall SHRP 2 NDS as possible. To accommodate this and reduce the chance for bias, samples for individual drivers were drawn once their files had been fully ingested into the SHRP 2 NDS database, allowing for baseline selection and annotation to start before the database was fully populated. Then, the baseline sample was rebalanced across all drivers once all data was fully ingested. This resulted in the availability of both “balanced-sample” baselines and “additional” baselines that are available for analyses where the balancing requirement is less important than having as large a baseline sample as possible.

Event Reduction and Annotation

The purpose of manually coding event characteristics is to evaluate the sequence of actions in the seconds prior to crashes and near-crashes and in baseline driving epochs, to document observed driver behaviors and mannerisms, and to record environmental/roadway variables that are not automatically recorded. One goal of data reduction is to define a relatively complete set of data elements, while not precluding further data reduction in the future, and to make the database more directly useful for search and analysis. The specific purpose of this effort was to extract particular information about the SCEs that would interest researchers who wanted to understand what factors led to and were associated with the SCE of interest.

The key to extracting this information is an extensive data dictionary and thorough operational definitions. One of the primary sources of the data dictionary was the General Estimates System (GES) database compiled by the National Highway Traffic Safety Administration (NHTSA; NHTSA, 2014). The GES is designed to be a nationally representative sample of police-reported motor vehicle crashes ranging from minor to fatal. The data from the GES are used to answer many motor vehicle safety research questions, and the GES itself is designed to extract and code information from police agency reports for this purpose. Therefore, it provided a great starting point for developing a video-based event reduction dictionary. However, in this case, the variables are not extracted from police crash reports but from review of event video from in-vehicle cameras and other data collected by the DAS.

Changes to the GES variable definitions were necessary, in large part because of the differences between the GES intent and perspective and the intent and perspective of general naturalistic driving data analysis. The former is based on crash information only, and is gleaned from police accident reports based on post-event site visits, crash reconstruction, and interviews with involved parties. The latter is based on crash and near-crash information, and is gleaned from on-scene, real-time video analysis, from the perspective of only one driver. Thus, the definitions included in a video-based event reduction dictionary were modified from the GES model accordingly, although maintaining the GES variable references within the modified dictionary continues to be useful as general guide rather than as a direct instruction.

The first video-based event reduction dictionary developed by VTTI was for use in the *The 100 Car Naturalistic Driving Study* (Neale at al., 2002; Dingus et al., 2006), sponsored by NHTSA. This data was collected and reduced between 2002 and 2005. This dictionary was later employed in *The Naturalistic Teenage Driving Study* (Simons-Morton et al., 2011), sponsored by the National Institute of Health (NIH), which took place between 2006 and 2008, and other VTTI studies over time. Since its original development, the dictionary has evolved and been further developed to better describe critical event scenarios. The SHRP 2 SCEs and baselines were coded according to (or brought up to the standard of) Version 3.4 of the SHRP 2 Researcher Dictionary for Video Reduction, published in February 2015. Examples of some of the definitions in this version of the dictionary and their associations with related GES definitions are shown in Appendix A. The full dictionary is available for reference and download on the InSight website.

Each crash and near-crash identified during the trigger validation, as well as each sampled baseline, was manually analyzed and annotated using the data dictionary in preparation for statistical and qualitative analysis. All available video views, along with the parametric collected data, were available as information sources for this annotation. There were 75 reduced/annotated variables, as indicated in Table 2.3. All 75 variables were coded for SCEs. An asterisk after the variable indicates that it was also coded for baselines (37 variables). The specific detailed information that was added and the associated operational definitions can be found on the InSight website. These variables can be placed into one of three categories, described below, according to the type of information provided and the aspect of the event or environmental condition assessed.

- **Event Variables***:* Variables used to establish the scenario and sequence of events prior to and throughout a critical event. These variables include event severity, event nature (e.g., conflict with lead vs. crossing vehicle), vehicle configurations, pre-incident maneuver, precipitating event, driver reaction, post-maneuver control, information about other drivers/vehicles/objects involved (e.g., type, position, maneuvers), and fault assignment.

- **Driver Variables***:* Variables used to systematically describe the subject driver prior to and during the critical event. These include driver ID, driver behavior (e.g., speeding, aggressive driving), driver impairments (e.g., drowsiness, anger, substance abuse), secondary task presence and duration (e.g., cell phone use), placement of hands on the wheel, visual obstructions, and seatbelt use.

- **Environmental Variables***:* Variables used to describe environmental and/or roadway conditions consistent with GES and other crash databases. These include roadway surface condition, traffic flow, number of travel lanes, traffic density, traffic control device at event onset, relation to junction of vehicle at event onset, roadway alignment (e.g., curve, grade), locality type (e.g., residential, interstate), ambient lighting, and weather.

Before researchers start to use this data, it is strongly recommended that they review the latest version of the dictionary on the InSight website. Users should read and have a thorough understanding of all variable definitions and should be familiar with the categories within each variable. It is especially important to have a clear concept of the Precipitating Event variable (variable #8), because the proper coding of many other variables depends upon an accurate identification of this variable and the Event Start variable (variable #2), which is the point in time at which the event begins (see Table 2.3). The Event Start is the first variable coded following the general classification of the SCE that is to be analyzed. This is the anchor point for the SCE, and all variables are populated related to that anchor point. For many variables, such as driver distraction, the reductionist starts coding five seconds before the precipitating event (e.g., lead vehicle braking) and continues coding until the conflict ends.

While naturalistic driving data contains a large set of data that otherwise would be unknown in standard crash analysis, such as the timing of events and subject drivers' behaviors, it is also important to recognize the limitations of naturalistic driving data. There are some incident factors that often cannot be determined due to the limited perspective of the naturalistic driving data (one vehicle only, referred to as the participant vehicle) and video data (from the perspective of a camera inside the participant vehicle looking out). Usually, very little is known about the behaviors, secondary tasks, reactions, impairments, etc., of drivers of other vehicles involved in the incident, or about certain non-visual SCE elements not sensed directly by the DAS.

Baseline analysis utilizes a subset of 37 variables included in this dictionary (denoted with an asterisk in Table 2.3). Precipitating Event or Event Start do not apply to baseline driving; instead, the baseline anchor point is defined as occurring 1 second prior to the end (last timestamp) of the baseline event. Accordingly, in baseline analyses, variables that reference the time of the Precipitating Event or Event Start for crash and near-crash analyses instead reference the endpoint of the baseline event minus 1,000 timestamps (milliseconds) (See Table 2.3.).

Table 2.3 Variables Used in Video Reduction

Variable Number	Variable Name
1*	Subject Number
2	Event Start
3	Subject Reaction Start
4	Impact or Proximity Time
5	Event End
6*	Pre-Incident Maneuver
7*	Maneuver Judgment
8	Precipitating Event
9	Vehicle 1 (Subject) Configuration
10	Vehicle 2 Configuration
11	Vehicle 3 Configuration
12	Event Nature 1
13	Incident Type 1
14*	Event Severity 1
15	Crash Severity 1
16	V1 Evasive Maneuver 1
17	V1 Post-Maneuver Control 1
18	Event Nature 2
19	Incident Type 2
20	Event Severity 2
21	Crash Severity 2
22	V1 Evasive Maneuver 2
23	V1 Post-Maneuver Control 2
24	Airbag Deployment
25	Vehicle Rollover
26*	Driver Behavior 1
27*	Driver Behavior 2
28*	Driver Behavior 3
29*	Driver Impairments
30*	Front Seat Passengers
31*	Rear Seat Passengers
32*	Secondary Task 1
33*	Secondary Task 1 Start Time
34*	Secondary Task 1 End Time
35	Secondary Task 1 Outcome
36*	Secondary Task 2
37*	Secondary Task 2 Start Time
38*	Secondary Task 2 End Time
39	Secondary Task 2 Outcome
40*	Secondary Task 3
41*	Secondary Task 3 Start Time
42*	Secondary Task 3 End Time
43	Secondary Task 3 Outcome
44*	Hands on the Wheel
45*	Driver Seatbelt Use
46	Vehicle Contributing Factors
47	Infrastructure
48	Visual Obstructions
49*	Lighting
50*	Weather
51*	Surface Condition
52*	Traffic Flow
53*	Contiguous Travel Lanes
54*	Through Travel Lanes
55*	V1 Lane Occupied
56*	Traffic Density
57*	Traffic Control
58*	Relation to Junction
59*	Intersection Influence
60*	Alignment
61*	Grade
62*	Locality
63*	Construction Zone
64	Number of Other Motorists/Non-Motorists
65	Number of Objects/Animals
66	Fault
67	Motorist/Non-Motorist/Animal/Object 2 Location
68	Motorist/Non-Motorist/Animal/Object 2 Type
69	Motorist/Non-Motorist 2 Pre-Incident Maneuver
70	Motorist/Non-Motorist 2 Evasive Maneuver
71	Motorist/Non-Motorist/Animal/Object 3 Location
72	Motorist/Non-Motorist/Animal/Object 3 Type
73	Motorist/Non-Motorist 3 Pre-Incident Maneuver
74	Motorist/Non-Motorist 3 Evasive Maneuver
75*	Final Narrative/Additional Notes

Data Reduction Process and Quality Control

The review of triggers along with the identification, reduction, and annotation of SCEs and baselines was a significant effort that required over 100 trained video data reductionists. All reductionists were subject to similar recruitment, training, and quality control processes as discussed in the following sections.

Recruitment and Reduction Lab Policies

Data reductionists at VTTI are part-time employees, recruited largely from the pool of Virginia Tech students, as well as from other nearby colleges and universities and the surrounding community. VTTI's reductionists often have backgrounds in related fields (e.g., psychology or human factors engineering), but other less technical fields have also yielded very successful data reductionists. Characteristics that contribute to quality data reductionists include a solid work ethic, a strong attention to detail, good critical thinking skills, and a ready willingness to ask questions when protocols are unclear. Because these qualities often cannot be judged from a resume or written application, interviews, reference checks, trial periods and proficiency tests for new hires are an integral part of the recruitment and retention process.

Minimum qualifications to be considered for a VTTI data reductionist position include a high school diploma, being of legal age to sign required non-disclosure statements, some college coursework (strongly preferred), proficiency in the English language to facilitate understanding of protocols and clear communication, possession of a valid U.S. driver's license to ensure experience with the driving scenarios to be analyzed, computer proficiency in common applications, demonstrated ability to pay close attention to detail for extended periods of time, strong critical thinking skills, prior data entry or detail-oriented work experience, and the ability to work required hours. Reductionists-in-training are required to pass a proficiency test before beginning to code new data, and all data reduced by newly hired or newly trained staff are closely reviewed for accuracy and consistency.

The following security policies are enforced in all of VTTI's data reduction labs.

- **Data Access**: All computers used for data reduction are restricted from internet access and all methods for externally saving data are disabled (e.g., USB ports, CD writers, etc.). All data are stored on secure servers accessed through a secure firewall, and access is granted only by secure log-in.

- **Human Subjects Protection and Non-Disclosure Agreements**: All reductionists receive training in Human Subjects Protection, including certification by an Institutional Research Board (IRB), and research confidentiality. All reductionists sign a Confidentiality/Non-Disclosure Agreement before working with any data or protocols.

- **Data Reduction Lab Access**: Access to the data reduction labs is restricted to authorized personnel only. Data reduction labs have a closed-door policy with access granted by key or numeric combination.

- **Personal Belongings**: Most personal belongings are disallowed at data reduction workstations. This includes any electronic devices with either wireless connection capabilities or audio or video recording capabilities, backpacks, handbags, and any non-work-related books or papers.

Additional policies are in effect for quality assurance and productivity purposes:

- **Hours Limitations**: Due to the focus-intensive nature of data reduction, new data reductionists are limited to a maximum of 4 hours per day of work. With several months of experience and demonstrated accuracy, reductionists may be approved for a maximum of 6 hours per day. Data reductionists only work when senior staff (lab proctors and supervisors) are on duty to ensure policy adherence and guarantee a quick response to reduction-related questions.

- **Regular Breaks**: All data reductionists take mandatory 10-minute breaks after each hour of work to prevent fatigue-related errors.

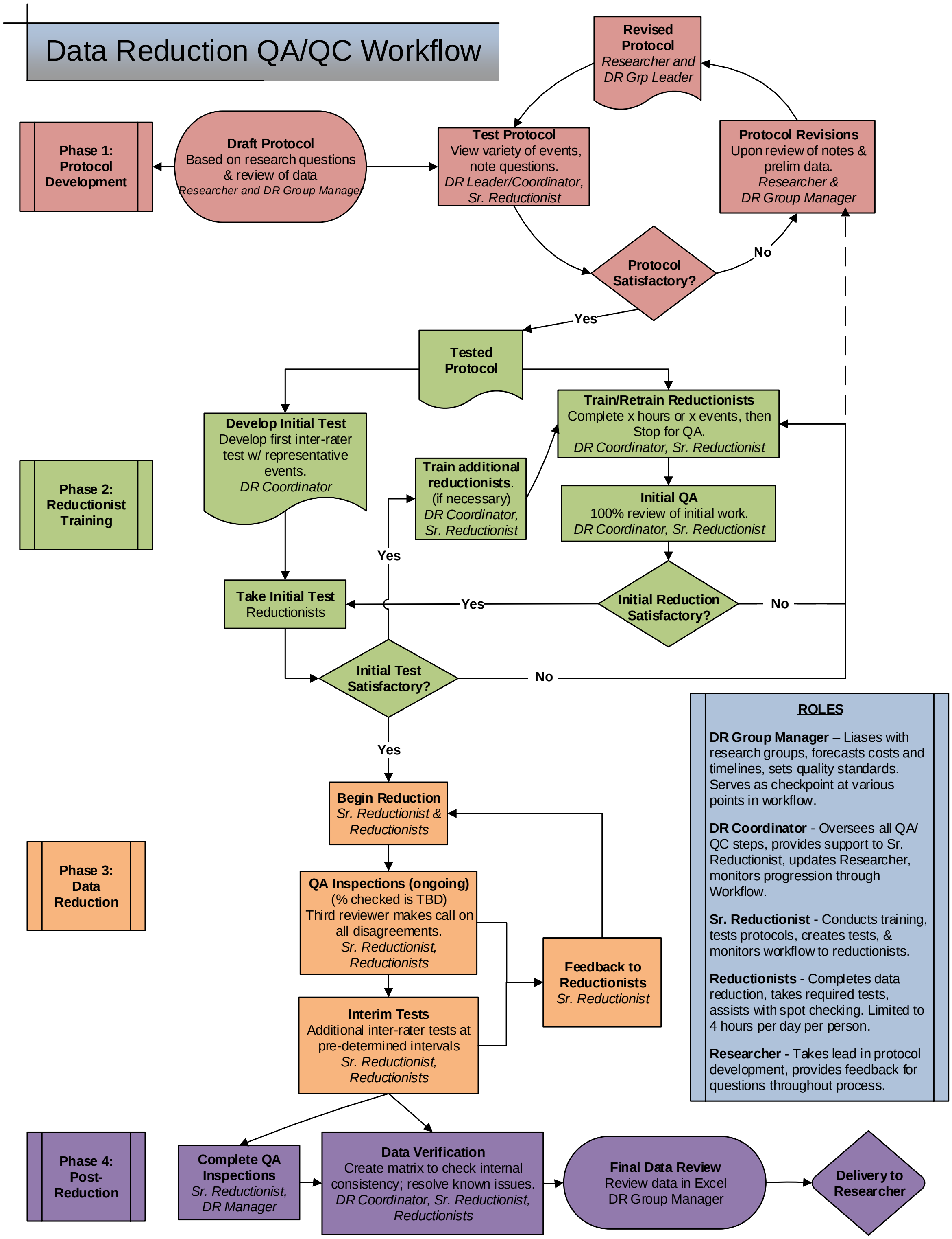
- **Protocol Clarifications**: Reductionists are instructed to ask for clarification any time that an event being analyzed doesn't meet the specific criteria in the protocol. This sometimes results in either re-training the individual or revising the protocol.

Reductionist Training and Quality Control Practices

Most data reduction at VTTI, including reduction performed for the project described in this report, follows a standard quality assurance/quality control (QA/QC) workflow (Figure 2.2). This workflow has four phases, all of which are equally critical to the quality of manually reduced data. These phases include Protocol Development, Reductionist Training, Data Reduction, and Post-Reduction, with tasks assigned at each level to one of four roles. Each role and phase is discussed below and also illustrated in the diagram.

Five roles are critical to the data reduction process.

1. **The Researcher or Research Group** may be internal or external to VTTI and plays a major role in protocol development, providing feedback in response to questions throughout the reduction process.
2. **The Data Reduction Group Manager** oversees the data reduction project from protocol development through reduction, analysis, and reporting. The manager liaises with research groups, forecasts costs and timelines, sets quality standards, and serves as a checkpoint at various points in the workflow.
3. **The Data Reduction Coordinator** oversees all QA/QC steps; hires, tests, and trains data reductionists; and monitors progression through the workflow. Most questions from reductionists can be fielded by the Coordinator; those that cannot are taken to the Group Manager.
4. **Senior Data Reductionists** (or lab proctors) are experienced, high quality data reductionists who assist the Data Reduction Coordinator with reductionist training and QA/QC, test new protocols before reduction work begins, and monitor the reductionists' workflow.
5. **Data Reductionists**, of course, perform the bulk of the data reduction. They also participate in the QA/QC process by completing required proficiency tests, receiving feedback, and assisting with spot checks.



ROLES

DR Group Manager – Liases with research groups, forecasts costs and timelines, sets quality standards. Serves as checkpoint at various points in workflow.

DR Coordinator - Oversees all QA/QC steps, provides support to Sr. Reductionist, updates Researcher, monitors progression through Workflow.

Sr. Reductionist - Conducts training, tests protocols, creates tests, & monitors workflow to reductionists.

Reductionists - Completes data reduction, takes required tests, assists with spot checking. Limited to 4 hours per day per person.

Researcher - Takes lead in protocol development, provides feedback for questions throughout process.

Figure 2.2. The Data Reduction Quality Assurance/Quality Control Workflow at VTTI.

Each role has specific tasks and checkpoints in one or more phases of the Data Reduction QA/QC Workflow:

1. **Phase 1: Protocol Development.** QA/QC of manual data reduction begins before reductionists ever see a protocol. Phase I engages a protocol writing, testing, and revision loop that requires intense collaboration between the Researcher, Data Reduction Group Manager, and Coordinator. Due to the reduction-intensive nature of the Data Reduction Group Manager's and Coordinator's work, they often have more experience than the Researcher in finding potential ambiguities, knowing when categories may be missing from certain variables, and adapting new protocols to be consistent (if possible) with previous protocols for later cross-analysis. The Coordinator and Senior Reductionist take the lead in testing the protocol and provide feedback to the Group Manager. The goal is to assess how well the protocol performs in answering the research question, whether the data was coded as intended and will provide the information required, and what problems were encountered by the testers. If revisions are significant, a second round of testing is performed. Once the protocol is satisfactory, it enters the second phase, Reductionist Training. However, a protocol can always be returned to the first phase if new problems are identified in subsequent phases. The nature of data reduction protocols is one of continuous improvement, with all specific project data coded to date brought up to the current standard.

2. **Phase 2: Reductionist Training**. Phase 2 consists of two concurrent processes—training and testing—involving small cohorts of Data Reductionists, ideally three to four at a time to keep the initial quality control manageable. Once a cohort moves into Phase 3, additional reductionists can enter the training loop in similar groups. Depending on the complexity of the protocol and the reductionists' experience, the training period may require only a day or two or up to multiple weeks. In this study, for reductionists with previous reduction experience, the training period for reducing baseline events usually lasted about 5 days. For similarly experienced reductionists, the training period for reducing SCEs was closer to 10 days.

    a. Reductionist training during Phase 2 is a collaborative process between the Coordinator and Senior Reductionists. Training for this study began by thoroughly reviewing the data dictionary with the Reductionists and providing both paper and electronic copies of the dictionary to encourage frequent referencing. Then, several reduced example events were reviewed alongside the video, and any questions were answered. Once the formal training session was completed (usually about 2 hours), reductionists reviewed a set of approximately 40 pre-coded example events on their own. These "review" events were selected to expose trainees to a variety of common event scenarios and further familiarize them with the reduction process. Once this self-paced event review was completed (usually within 2–4 days), reductionists began reducing new events under close supervision. This initial reduction was limited to a few hours or a few events (no more than 1 work shift), and then 100% of each Reductionists' work was reviewed by a Senior Reductionist or Coordinator. Detailed feedback was provided to the original Reductionists, who then made any necessary corrections to their own work. If the initial review was unsatisfactory, then reductionists were re-trained and asked to code another small set of events. Throughout this process, both reductionist and protocol performance was evaluated. If a certain aspect of the protocol was consistently problematic, this was usually a sign that either more in-depth training on a particular variable needed to be conducted, or that the protocol needed to be modified or clarified to increase reliability.

   b. A major part of training is proficiency testing. Prior to the first cohort's training session, the Reduction Coordinator develops the first test and corresponding gold standard responses in coordination with the Group Manager to include a sample of events that approximately represents the range and frequency of conditions expected to be present in the data set. For this study, the initial proficiency test for SCEs comprised 10 events, and the initial test for baselines comprised 20 events. After completing the initial training and reviews (as described in Phase 2a above), reductionists take this test. A score of 90% is considered passing (90% of questions answered correctly). All trainees are provided with feedback on their performance and specific errors. Reductionists who do not achieve a score of 90% or who exhibit systematic errors are retrained and administered a second test with a different set of events. If scores on either test are satisfactory and no systematic errors are observed, then reductionists move into Phase 3. If scores remain unsatisfactory, retraining or additional protocol revisions are considered, or reductionists are removed from the project.

3. **Phase 3: Data Reduction**. Three tools for ongoing QA/QC are used during the data reduction phase: Spot Checks (which start immediately upon entering Phase 3) and periodic expert- and intra-rater tests. All methods are continued until all events have been reduced and Phase 4 begins.

   a. A spot check is an experienced reductionist review of completed data reduction work. Feedback to confirm and/or improve accuracy is provided to the original reductionist, who then makes (and learns from) any necessary corrections. Spot checks are a very useful tool for assessing reductionists' understanding of protocols, revealing potential ambiguities that may need to be addressed in the protocol or data dictionary, and monitoring overall data quality. In this project, the SCE reduction task underwent a 100% spot check policy, which was maintained throughout the reduction task due to the protocol complexity. Thus, all SCEs events were subject to review by at least two people. For baselines, the 100% rate was dropped to 50% for individual reductionists as accuracy improved with experience, and sometimes eventually reduced to 25% if exceptionally high accuracy was attained. If consistent or increasing errors were found during later spot checks, then the check rate was again increased for that individual.

   b. Expert-rater tests are administered periodically to ensure consistency with gold standard coding. Six such tests were administered to SCE and baseline reductionists during the SHRP 2 study. Similar to the initial proficiency test conducted in Phase 2, this test included a sample of events with a range of conditions present in the data. The tests were designed to be completed in a typical 4-hour work shift. Each of the six tests included a different set of events, and the goal of each test was to receive responses that matched a gold standard from all Reductionists. This gold standard was coded collaboratively between Reduction Coordinators and the Group Manager, and reductionist responses were compared to that gold standard. Unsatisfactory scores signaled a need for retraining.

   c. Intra-rater tests were conducted as part of the expert-rater process by including a small number of events in each test that were repeated from test to test. The intra-rater test measured the consistency with which individual Reductionists coded data over time. The goal in this test was for each reductionist to code these events in the same way they coded them during the previous test(s).

4. **Phase 4: Data Delivery**. In Phase 4, the data reduction team works to prepare the data set for delivery back to the Researcher (for SHRP 2, this included release on InSight) so that statistical analysis can begin. First, any remaining spots checks (Phase 3a) are completed and any remaining discrepancies between original reductionist and reviewer are resolved. Then, based on the spot check review, and a pragmatic review of the protocol, all potential logical errors or inconsistencies are reviewed in the resulting dataset. This is the Data Verification step. This type of review looks for missing data, outliers, and potential inconsistencies between variables that should be internally consistent with each other (e.g., if Locality is coded as Interstate, then the Relation to Junction should not be "Intersection" but may be "Interchange"). The Data Reduction Group Manager and Coordinator work together to identify these potential issues and locate them (if present) in the dataset, and then work with Reductionists to resolve them. As a final review before data delivery in this study, all the data was pulled into a spreadsheet and checked again for internal consistency, completeness, and outliers. Any questionable events were again flagged for review. For example, a review would have been requested for events with unexpected or rarely used categories or responses outside the expected range. Generally, thorough post-reduction processing for SCEs and baselines can take two or more weeks. If performed periodically during Phase 3 as well, these final data cleaning steps can usually be completed more quickly.

CHAPTER 3: FINDINGS AND APPLICATIONS

In all, the data reduction process to date has resulted in a database that includes 4,254 SCEs (1,549 crashes; 2,705 near crashes) and 32,586 baselines (20,000 of which are balanced on driver exposure). As with any naturalistic dataset, this database is considered dynamic; updates will be made periodically to add and/or update events.

This chapter will describe:

- The performance and output of the trigger algorithms;
- The number and distribution of reduced SCEs and baselines;
- The results of expert- and intra-rater testing;
- Important considerations for selecting samples of SCEs and/or baselines for further analysis.

Trigger Algorithm Performance

In total, a dozen triggers were used to attempt to identify the SCEs. As stated previously, the thresholds selected for each trigger were designed to identify as many crashes and near-crashes as possible. Possible SCEs were rank-ordered for review based on the severity of the kinematic signature, and those with more severe kinematic triggers were reviewed first. This was done not only to have a higher potential hit versus miss rate for the algorithm but also to ensure that more severe crashes were identified earlier. Table 3.1 illustrates the number of triggers executed and their relative success at identifying crashes, near-crashes, and crash-relevant conflicts (CRC). Note that unlike crashes and near-crashes, CRCs were not reduced for further analysis.

Table 3.1 Trigger Algorithm Performance

Trigger Type	Number of Valid Triggers			Percent Valid
	Crash	Near-Crash	Crash-Relevant Conflict	
Longitudinal Deceleration	626	3,950	4,680	22%
Lateral Acceleration	218	71	48	2%
Yaw Rates	25	61	48	3%
Freeway Deceleration	28	1,500	2,746	8%
ABS	456	1,462	2,549	4%
Airbag	1	0	0	<1%
ESC	17	6	39	1%
Traction Control	96	22	101	1%

Steering Evasive Maneuver	384	56	42	1%
Swerve	32	17	28	<1%
Longitudinal Acceleration	241	41	25	2%
CI Button	2,062			15%*

*Includes SCEs and CRCs that the participant witnessed but had tertiary or no involvement in.

In Table 3.1, it is important to note that each trigger algorithm is not independent of the others. For example, high longitudinal deceleration could flag a possible SCE, and the anti-lock brake system (ABS) could flag the same SCE. Therefore, a valid crash would be counted in both the longitudinal deceleration and the ABS rows. The longitudinal deceleration trigger performed best, with 22% of the potential SCEs being confirmed as an SCE or CRC. Participants pressing the critical incident button that opened a 30-second audio recording window to describe something of interest produced a valid SCE or CRC 15% of the time. It is important to note that the critical incident button press included some events that unfolded in front of the driver but in which their vehicle was not involved. The freeway deceleration trigger performed the next best with 8% of events being identified as valid. Some of the other triggers did not perform as well (lower valid rates) or had thresholds that were too liberal and produced a low number of potential SCEs. Trained data reductionists ultimately reviewed portions of approximately 6.7% of the trip files in the SHRP 2 NDS while validating these SCEs.

The next section describes the crashes and near-crashes identified through this effort. In the future, it is recommended that the processed data from the forward radar also be used to run additional trigger algorithms. An InSight report titled "Task 1.6: Radar Post-Processing" (Gorman et al. 2015) describes this process in further detail.

Safety-Critical Events

The results of the SCE verification process are shown in Table 3.2. Overall, 1,549 crashes (including vehicle roll-over events that do not appear on InSight) and 2,705 near-crashes have been identified. Out of the crashes, it is likely that 271 of them would have been reportable to the police (Level II or higher), although the actual number of crashes reported to the police in this project is unknown. Note that there are additional crashes and near-crashes that are still undetected in the SHRP 2 NDS. The focus of the project was identifying crashes, so it is expected that there are substantially more undetected near-crashes than crashes. In the case of crashes, it is also expected that the vast majority of remaining undetected crashes are less severe. This is a result of SCEs with greater kinematic signatures being reviewed first. For near-crashes, this trend is also expected, but the chance of an undetected near-crash of greater severity is much more likely. The kinematic signature of a severe near-crash is not as highly correlated with severity as a crash. Ultimately, as more researchers work with the database and more funding is dedicated to finding crashes, it is expected that both the number of crashes and near-crashes will increase. (See Table 3.2.)

Table 3.2 Crash and Near-Crash Frequency in InSight Database

Crash Severity	Category Name	Count
I	Airbag, Injury, Roll Over, High Delta-V Crash	112
II	Police Reportable Crash	159
III	Physical Contact with Another Object	633
IV	Tire Strike Low Risk	637
Near-Crash	Near Crash	2,705
	Total	**4,246**

Baselines

The result of the baseline sample was a random sample stratified by vehicle and proportion of time driven over 5 mph. This sample contains 20,000 balanced-sample baselines and includes all drivers in the SHRP 2 NDS. As part of developing this balanced sample, an additional 12,586 baseline samples were selected and reduced for some drivers. These are also available for researchers to potentially include in their studies in order to increase statistical power, though this will be at the expense of a balanced stratified sample.

Intra-Rater and Expert-Rater Reliability

Six rounds of tests were conducted at roughly two-month intervals during the reduction of SCE and baseline events. Note that the first test did not include an intra-rater, as it was used to establish a reference point for future intra-rater tests; therefore, only five intra-rater tests were administered. Each test was administered to all reductionists assigned to the project at the time the test was assigned. Individual reductionists took between one and six expert-rater tests and between one and five intra-rater tests.

Each SCE test included six new (expert-rater) events and one repeated (intra-rater) event for a total of seven events on each test. Each baseline test included 17 new (expert-rater) events and three repeated (intra-rater) events for a total of 20 events on each test. Per the data reduction dictionary, SCE tests required responses to 75 variables; baseline tests required responses to a subset of 37 variables. Both the SCE test and the baseline test required, on average, 3.5–4 hours to complete. When scoring the results, a 200 millisecond “window” of acceptance was applied to variables that required video timestamps as responses (Event Start, Event End, Driver Reaction, Impact Proximity, and Secondary Task Start/End 1,2,3).

Five SCE intra-rater tests were administered, with results displayed below in Table 3.3. Ten raters in all were assigned to the SCE reduction task over the course of the project, six of whom took at least four of the five intra-rater tests. (Raters needed be assigned to the reduction task during two consecutive tests to have an intra-rater to score.) Each test score represents the average proportion of responses that were in agreement with the previous test taken by the same rater across all baseline questions in each test. Average intra-rater scores ranged from 73% to 94%, with an overall average of 91%.

Traffic Density was the categorical variable with the lowest intra-rater reliability at 65%, but when an allowance was made to allow for one-degree of variability (e.g., Level of Service [LOS] B vs. LOS C), this variable’s intra-rater reliability increased to 100%. This scoring trend was consistent regardless of whether SCEs or baselines were the focus, and regardless of whether the test was expert- or intra-rater (all of which are described in more detail below). Researchers are therefore encouraged to collapse the seven Traffic Density categories in their analyses into fewer levels (e.g., LOS A-B, LOS C-D, LOS E-F or LOS A-B-C, LOS D-E-F). This will allow for a higher level of confidence when interpreting results."

Other variables with SCE intra-rater scores below 80% include Precipitating Event, V1 Evasive Maneuver, Crash Severity 2, Intersection Junction, and V2 Pre-incident Maneuver. As illustrated by Table 3.3, scores on tests 3–5 tended to be higher than on tests 1 and 2, indicating a typical learning curve and providing evidence of higher consistency with more experience. Lower individual scores were represented by Raters 8, 9, and 10, who were assigned to the project only a very short time, as evidenced by the fact that only 1 intra-rater test was taken by each.

Table 3.3 SCE Intra-Rater Scores

Rater #	# Tests Taken	Test 1	Test 2	Test 3	Test 4	Test 5	Average Score
Rater 1	5	92%	78%	91%	95%	97%	91%
Rater 2	5	86%	89%	90%	93%	94%	90%
Rater 3	5	90%	91%	97%	98%	90%	93%
Rater 4	4	NT	90%	94%	99%	97%	95%
Rater 5	5	82%	93%	93%	97%	95%	92%
Rater 6	5	77%	89%	98%	97%	99%	92%
Rater 7	1	NT*	NT	NT	NT	94%	94%
Rater 8	1	NT	NT	NT	NT	82%	82%
Rater 9	1	NT	83%	NT	NT	NT	83%
Rater 10	1	NT	73%	NT	NT	NT	73%
Average	33	85%	86%	94%	96%	93%	91%

*NT: Test not taken.

Five baseline intra-rater tests were administered, with overall results displayed below in Table 3.4. Each score represents the average proportion of responses that were in agreement with the previous test taken by the same rater across all baseline questions in each test. Twenty-six raters in all were assigned to the baseline reduction task over the course of the project, many of whom took multiple tests. Average intra-rater rater scores ranged from 95% to 100%. Because the scores were so high, sequential individual test scores are not reported here, as they do not provide any additional information. This is evidence of the shorter learning curve for baseline reduction as compared to SCE reduction. Again, Traffic Density was the categorical variable with the lowest baseline intra-rater reliability at 85%, but when an allowance was made to allow for one-degree of variability (e.g., LOS B vs. LOS C), this variable's intra-rater reliability increased to 99%. Locality, was the only other variable (out of 37 variables) with an average intra-rater score below 90%.

Table 3.4 Baseline Intra-Rater Scores

Rater #	# Tests Taken	Score
Rater 1	3	97%
Rater 2	5	96%
Rater 3	4	98%
Rater 4	2	97%
Rater 5	5	99%
Rater 6	5	98%
Rater 7	5	98%
Rater 8	2	99%
Rater 9	2	97%
Rater 10	3	96%
Rater 11	2	97%
Rater 12	1	100%
Rater 13	4	97%

Rater #	# Tests Taken	Score
Rater 14	3	96%
Rater 15	1	99%
Rater 16	2	97%
Rater 17	5	97%
Rater 18	5	97%
Rater 19	5	98%
Rater 20	3	98%
Rater 21	3	95%
Rater 22	2	97%
Rater 23	4	97%
Rater 24	1	97%
Rater 25	5	98%
Rater 26	5	97%
Overall	**87**	**97%**

The SCE expert-rater test was administered six times with between 6 and 11 raters each time. Each test score represents the proportion of raters who met the gold standard across all SCE questions on each test. Average test scores (see Table 3.5) ranged from 85% to 92%, which was deemed satisfactory for a task as complex as SCE reduction. Traffic Density was the only categorical variable below 70% (at 63%), but when an allowance was made to allow for one-degree of variability (e.g., LOS B vs. LOS C), this variable's intra-rater reliability increased to 92%. Other categorical variables with expert-rater scores of 70-80% include V1 Pre-incident Maneuver, V1 Evasive Maneuver, Driver Behavior 1, Relation to Junction, V2 Pre-incident Maneuver, and V2 Reaction. All other variables yielded scores above 80% with 32 variables yielding scores above 90%.

Table 3.5 SCE Average Expert-Rater Scores

Test	# Raters	Score
Test 1	9	88%
Test 2	11	86%
Test 3	6	85%
Test 4	6	92%
Test 5	8	90%
Test 6	8	90%
Overall	48	88%

The baseline expert-rater test was also administered six times with between 14 and 32 raters each time. Each test score represents the proportion of raters who agreed with a gold standard across all baseline questions in each test. Average test scores were all between 92% and 93% (Table 3.6), also quite satisfactory. Traffic Density was again the only categorical variable below 70% (at 68%) but when an allowance was made to allow for one-degree of variability (e.g., LOS B vs. LOS C), this variable's intra-rater reliability increased to 98%. Other categorical variables with expert-rater scores of 70-80% include Pre-incident Maneuver, Relation to Junction, and Locality. All other variables yielded scores above 80% with 33 of the 37 applicable variables yielding scores above 90%.

Table 3.6 Baseline Expert-Rater Scores

Test	# Raters	Score
Test 1	19	93%
Test 2	14	92%
Test 3	21	92%
Test 4	24	92%
Test 5	20	93%
Test 6	32	92%
Overall	**130**	**92%**

The expert-rater scores reported above for both the SCE and baseline reductions suggest a high agreement between individual raters and a gold standard. This indicates that, at any given time, reductionists were generally analyzing events the same way that an expert would, and the consistency in scores over time (successive tests) indicates that this level of accuracy was carried through over time.

Considerations for Data Selection

It is anticipated that the SHRP 2 NDS database will continue to be a powerful, valuable tool for researchers for decades to come. This data can be used to answer thousands of research questions posited by hundreds if not thousands of researchers. However, it is critical to view and understand this as a "living" database. Researchers may encounter quality issues that must be fixed or at least noted for future researchers; their subsequent analysis and annotation may also yield additional information. Consequently, the specific findings described in this report are expected to be updated as time progresses. Thus, it will be the responsibility of researchers to ensure they have the most current representation of the database.

The SHRP 2 NDS database described in this report is expected to be heavily used by researchers. As of this writing, over 1,000 people have registered as qualified researchers on the InSight website, giving them access to this SHRP 2 data, which many are planning to use in analyses. It is fully expected that competing analyses and contradictory results will be reported. Arguably, these contradictions can lead to improvements in the overall science; therefore, it is imperative that researchers reference in their reports which data set was used and what data were excluded or included. This information can be used in any examinations of conflicting results and yield potential updates or method improvements.

VTTI has implemented several tools and provided suggestions in order to assist researchers in selecting the correct data from the database when pursuing their research questions. The following subsections detail the important devices and information that must be reviewed and considered whenever a researcher begins to use the database.

Referencing Data Sets

VTTI has created a mechanism that retains the original form of the data sets as they were built, as well as a referencing scheme to ensure that researchers understand what data were used. Because researchers often further refine a data set to meet the requirements of their specific research question, documenting exactly what information was included and excluded facilitates both the reproduction of the research and the understanding of what the research is based on.

Data Dictionary

VTTI also produced a series of data dictionaries that are located on the InSight website. These dictionaries encompass every variable in the database, including the driver assessment data, time series data, and vehicle-related data. The dictionaries are also "living" documents that will be updated as more variables are added to the SHRP 2 NDS and refinements are made to existing variables. When a change is made from one version of the dictionary to the next, the information will be highlighted on a revision page that notes what was changed and how it was changed.

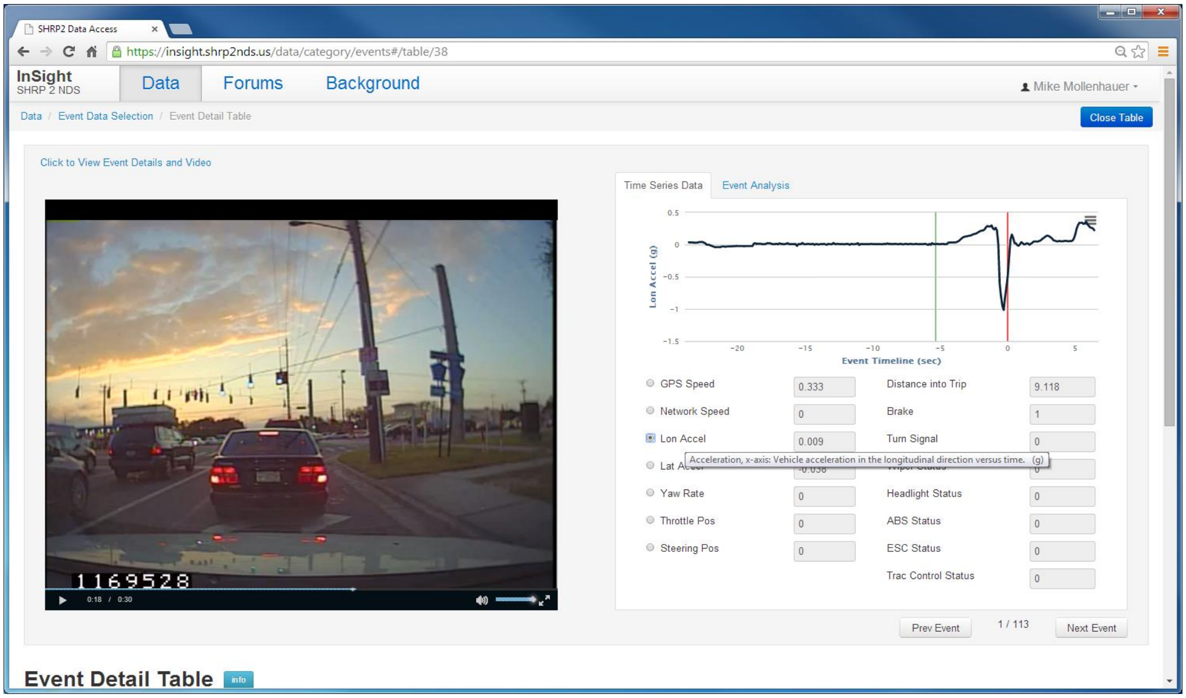
Researchers must review these dictionaries before doing any analysis to make sure the operational definitions meet the requirements necessary to answer their questions. Based on the name of a variable alone, they could easily misinterpret the data. For example, the SHRP 2 database contains a driver impairment variable, which encompasses the following categories:

- None apparent
- Drowsy, sleepy, asleep, fatigued
- Ill, blackout
- Angry
- Other emotional state
- Drugs, medication
- Drugs, alcohol
- Other illicit drugs
- Restricted to wheelchair
- Impaired due to previous injury
- Deaf
- Other
- Unknown

If a researcher wanted to conduct an analysis on driving under the influence of alcohol or drugs only, some of these classifications should be excluded. Without reviewing the data dictionaries, a researcher may incorrectly assume certain information is included or excluded.

SCE Selection

The selection of which SCEs to study is vital to correctly answering a research question. Researchers need to determine exactly what research question they want to answer and then ensure that the sample they select correctly addresses this question. In most cases, SCEs are included and excluded based on the research question being posed. On the website, the InSight viewing tool (Figure 3.1) allows researchers to review each SCE to confirm that it truly belongs in the analysis based on the parameters of the research. The viewing tool allows researchers to watch an SCE from the forward-view camera perspective along with the associated parametric data to judge for themselves whether an event should be included.



○ GPS Speed 0.333
○ Network Speed 0
● Lon Accel 0.009
○ Lat Accel -0.038
○ Yaw Rate 0
○ Throttle Pos 0
○ Steering Pos 0

Distance into Trip 9.118
Brake 1
Turn Signal 0
Wiper Status 0
Headlight Status 0
ABS Status 0
ESC Status 0
Trac Control Status 0

Acceleration, x-axis: Vehicle acceleration in the longitudinal direction versus time. (g)

Figure 3.1. InSight event viewing tool.

The SCE variable includes both crashes and near-crashes; crashes are then further categorized into four categories from the most to least severe. Depending on the research question, some researchers may choose to remove all Level IV crashes from their analysis. Other researchers may want to review run-off-road Level IV events and decide which should be included and which should be excluded. Similarly, some researchers may want to include near-crashes in their analysis whereas others may want to exclude them.

Although a large number of SCEs have been identified, it is expected that others still reside undetected in the SHRP 2 NDS. For example, participants could press an incident button to describe a near-crash that occurred. If that near-crash was a “T-bone” incident (i.e., another vehicle almost crashed into the side of the participant vehicle), it is possible that no other trigger could have detected it. Also, crashes or near-crashes that did not show up on any of the triggers could still be identified in the future by an analyst who is reviewing the data for other reasons. Finally, the thresholds chosen for the algorithm may not have found a crash that was self-reported by a participant. Sometimes how a crash or near-crash is identified or could have been identified is important to consider in the research question. In these cases, a table has been developed to indicate both how the crash or near-crash was identified, or could have been identified. Researchers can apply this data to include or exclude specific SCEs for their analysis.

Baseline Selection

Baselines in this context should provide information on exposure that researchers use in their analysis. As of the writing of this report, there are 32,586 randomly selected, reduced baselines available. Roughly two-thirds of this sample was balanced to a 20,000 baseline sample stratified by participant and proportion of time driven over 5 mph. Researchers should consider this balanced sample for their case-cohort design since it provides a balanced representation of normal driving for all the drivers in the SHRP 2 NDS. However, the baseline selected is ultimately the decision of the researcher. It is likely that some researchers will use the entire 32,586 sample, whereas others, depending on their research question or sampling of drivers in the SHRP 2 NDS, will use a subset of the 20,000 baseline sample. Documenting what baseline sample was used is highly recommended. Additionally, researchers may select brand new baselines that are necessary to support their analysis technique. Ideally, researchers will make this new baseline available to others, thereby opening the door to more baseline samples that can be considered in the future.

CHAPTER 4: SUMMARY

The focus of this project was to identify and prepare crash, near-crash, and baseline data sets for researchers to use in their analysis projects. The overall SHRP 2 NDS includes 5,512,900 valid trip files driven; approximately 1% of the trip files originally collected were excluded, primarily due to missing face video. Manual reviews verified that 15% of the original trip files had to be excluded mostly due to unconsented drivers. A dozen trigger algorithms executed on these 5,512,900 trip files identified 1,549 crashes and 2,705 near-crashes. A longitudinal deceleration-based trigger algorithm performed the best at identifying events, producing the highest percentage (22%) of valid events. All the potential SCEs were validated or rejected through manual review. Each SCE was prepared and annotated with an additional 75 variables, including a final narrative that explains exactly how the SCE occurred. Expert- and intra-rater testing revealed a high (80-90%) level of reliability in the SCE data reduction effort.

A random sample stratified by participant and proportion of time driven produced 20,000 baselines. This sampling scheme includes all the drivers in the SHRP 2 NDS and is recommended for researchers using a case-cohort design. An additional 12,586 baselines are available for researchers who want to use an increased sample size at the expense of the balancing and stratification. All 32,586 baselines were coded using many of the same variables in the SCE reduction. Expert- and intra-rater testing revealed a high (90-98%) level of reliability in the baseline data reduction effort.

Researchers using this data set are encouraged to perform the following steps:

- Review the data dictionaries on the InSight website prior to doing analysis. Understanding the full definitions of these variables is important to avoid misinterpretation.

- Determine which crashes and near-crashes should be included to answer their research question.

- Understand what the precipitating event is in a crash and near-crash and where it is located. Much of the information extracted from the video reduction uses this as an anchor point.

- Reference data sets used and indicate what events were included and excluded. Beyond allowing repeatability of findings, it is likely that with the number of researchers using this data, contradictory results will occur and will have to be defended and/or explained.

SUGGESTED READING

As a final comment, the use of data collected by other organizations is becoming increasingly common and more accepted; the primary benefit of this is cost savings. The SHRP 2 NDS cost millions of dollars to collect and is available at little to no cost to researchers. However, it is the responsibility of researchers to understand the analyzed data whether they collected and/or paid for it or not. The following references include discussion related to and recommended for such secondary analyses:

- Akerstrom, M., K. Jacobsson, and D. Wasterfors. (2004). Reanalysis of Previously Collected Material. In *Qualitative Research Practice* (Clive Seale, Giampietro Gobo, Jaber Gubrium, and David Silverman, eds.), Sage Publications, Thousand Oaks, CA.

- Corti, L. and P. Thompson. (2004). Secondary Analysis of Archived Data. In C *Qualitative Research Practice* (Clive Seale, Giampietro Gobo, Jaber Gubrium, and David Silverman, eds.), Sage Publications, Thousand Oaks, CA.

- Dale, A., S. Arber, and M. Procter. (1988). *Doing Secondary Analysis,* Unwin Hyman Ltd., London.

- Hyman, H. (1972). *Secondary Analysis of Sample Surveys,* Wesleyan University Press, Middletown, Connecticut.

- Kiecolt, K. and L. Nathan. (1985). *Secondary Analysis of Survey Data – Sage University Paper Series on Quantitative Applications in the Social Sciences*, 53. Sage Publications, Beverly Hills, CA.

REFERENCES

Antin, J. A., K. Stulce, L. Eichelberger, and J. Hankey. (2015). *Comparing the SHRP 2 NDS Sample with National Data.* Transportation Research Board of the National Academies, Washington, D.C. Retrieved from http://www.trb.org/Publications/Blurbs/172183.aspx.

Antin, J. A., K. Stulce, L. Eichelberger, and J. Hankey. (forthcoming). *Enhancing Usability of Select Driver Assessment Data*. Transportation Research Board of the National Academies, Washington, D.C. To be published at http://www.trb.org/Publications/PubsSHRP2ResearchReports.aspx.

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. (2006). *The 100-Car Naturalistic Driving Study, Phase II – Results of the 100-Car Field Experiment.* DOT HS 810 593. National Highway Traffic Safety Administration, U.S. Department of Transportation.

Dingus, T. A., J. Hankey, J. F. Antin, S. E. Lee, L. Eichelberger, K. Stulce, D. McGraw, M. Perez, and L. Stowe. (2014). *SHRP 2 Report S2-S06-RW-1: Naturalistic Driving Study: Technical Coordination and Quality Control*. Transportation Research Board of the National Academies, Washington, D.C.

Gorman, T., L. Stowe, and J. Hankey. (2015). *S31: NDS Data Dissemination Activities, Task 1.6: Radar Post-Processing*. Transportation Research Board of the National Academies, Washington, D.C. Retrieved from https://insight.shrp2nds.us/projectBackground/index.

Guo, F., and J. Hankey. (2009). *Modeling 100-Car Safety Events: A Case-Based Approach for Analyzing Naturalistic Driving Data*. National Surface Transportation Safety Center for Excellence, Blacksburg, VA.

McClafferty, J. A., M. A. Perez, and J. M. Hankey. (2016). *Identification of Consented Driver Trips in the SHRP 2 Naturalistic Driving Study Data Set*. Transportation Research Board of the National Academies, Washington, D.C. Retrieved from https://vtechworks.lib.vt.edu/handle/10919/70848.

McLaughlin, S., and J. M. Hankey. (2015). S*HRP 2 Report S2-S31-RW-3: Naturalistic Driving Study: Linking the Study Data to the Roadway Information Database*. Transportation Research Board of the National Academies, Washington, D.C. Retrieved from http://www.trb.org/Main/Blurbs/172174.aspx.

National Highway Traffic Safety Administration. (2014). *National Automotive Sampling System (NASS) General Estimates System (GES) Analytical User's Manual 1988-2012*. National Highway Traffic Safety Administration.

Neale, V. L., S. G. Klauer, R. R. Knipling, T.A. Dingus, G.T. Holbrook, and A. Petersen. (2002). *The 100 Car Naturalistic Driving Study, Phase I-Experimental Design*. DOT HS-809

536. National Highway Traffic Safety Administration, U.S. Department of Transportation.

Simons-Morton, B. G., M.C. Ouimet, Z. Zhang, S.E. Klauer, S.E. Lee, J. Wang, ... and T.A. Dingus. (2011). The effect of passengers and risk-taking friends on risky driving and crashes/near crashes among novice teenagers. *Journal of Adolescent Health, 49*(6), 587-593.

ACRONYMS

ACN	automatic crash notification
ABS	anti-lock brake system
CI	critical incident
CRC	crash-relevant conflict
ETG	Expert Technical Group
DAS	data acquisition system
IRB	Institutional Review Board
GES	General Estimates System
LOS	Level of Service
NDS	naturalistic driving study
NHTSA	National Highway Traffic Safety Administration
NIH	National Institute of Health
QA/QC	Quality Assurance/Quality Control
RID	Roadway Information Database
SCE	safety-critical event
SHRP 2	Second Strategic Highway Research Program
TCC	Technical Coordinating Committee
VTTI	Virginia Tech Transportation Institute

APPENDIX A. EXCERPT FROM SHRP 2 RESEARCHER DICTIONARY FOR VIDEO REDUCTION DATA

(FULL DICTIONARY AVAILABLE ON INSIGHT WEBSITE)

*Variable # (*Baseline)*	*Variable Name*	*Variable Definition*	*GES Related Variable(s) (modified from GES)*
1*	**Subject Number**	All consented drivers (primary and secondary) are assigned a unique numeric ID number, which can be used for cross-referencing demographic information, etc. For SHRP2, subject numbers are between 1 and 7 digits.	
2	**Event Start**	The point in the video when the sequence of events defining the occurrence of the incident, near-crash, or crash begins. Defined as the point at which the Precipitating Event begins (see Precipitating Event [V8]). Value is a timestamp, in milliseconds after the start of the file. NOTE: for cases in which the origin of the Precipitating Event is not visible in the video (e.g., "Other vehicle ahead - stopped on roadway more than 2 seconds" or "Pedestrian in roadway"), the start point for the Precipitating Event would be when the event is first visible in the forward view of the subject vehicle. ALSO NOTE: For baseline events, the Event Start is defined as 1 second (1,000 timestamps) prior to the end of the baseline event.	
3	**Subject Reaction Start**	The timestamp, in milliseconds after the start of the file, when the driver is first seen to recognize and begin to react to the safety critical incidents occurring. Defined as the first change in facial expression to one of alarm or surprise or the first movement of a body part in a way that indicates awareness and/or the start of an evasive maneuver, whichever occurs first. Reaction time can be coded after the time of impact for low-risk tire strikes if the driver is acting to prevent a worse collision and for certain rear-end, struck collisions if the driver is acting to prevent a second (rear-end, striking) incident.	
4	**Impact or Proximity Time**	The timestamp, in milliseconds after the start of the file, when the subject vehicle and other object of conflict first make impact. In the case of a near-crash, the timestamp when the subject vehicle and other object of conflict are at their closest distance to each other. If more than one incident type occurs, this is coded for the most severe (crash or near-crash) or the first incident type if both are the same severity.	
5	**Event End**	The timestamp in the video, milliseconds from the start of the file, when the sequence of events defining the occurrence of the incident, near-crash, or crash ends. Defined as the point at which final evasive maneuvers have been completed and all vehicles, objects, pedestrians, animals, etc., involved have either stopped or returned to normal patterns of road use, whichever occurs first.	

Variable # (*Baseline)	***Variable Name***	***Variable Definition***	***GES Related Variable(s) (modified from GES)***
6*	**Pre-Incident Maneuver**	This represents the last type of action or driving maneuver that the subject vehicle driver engaged in or was engaged in just prior to or at the time of the Precipitating Event, beginning anywhere from about 2 to 6 seconds before the Precipitating Event (V8). This variable is independent of the driver's engagement in secondary tasks and the Precipitating Event, but should be determined after the precipitating event is defined. It is a vehicle kinematic measure--based on what the vehicle does (movement and position of the vehicle), not on what the driver is doing inside the vehicle. For baselines, this is the action or driving maneuver that the subject is engaged in for the last 2-6 seconds of the baseline event prior to the baseline anchor point (Event Start, V2), which occurs 1 second before the end of the baseline event.	V21 (Vehicle Maneuver/Movement Prior to Critical Event (Precrash 1))
7*	**Maneuver Judgment**	Judgment of the safety and legality of the Pre-Incident Maneuver (V6). This is a vehicle kinematic measure-based on what the vehicle does, independent of the driver's engagement in secondary tasks and the Precipitating Event (V8). (e.g., Driving while texting on a cell phone may not be safe or legal, but it is not a consideration in this variable.) Although the determination of whether the maneuver is safe or unsafe is situation-dependent, the position of the vehicle itself is the main determinant of this factor, and a maneuver may or may not be safe, depending on the vehicle position.	
8	**Precipitating Event**	The state of environment or action that began the event sequence under analysis. What environmental state or what action by the subject vehicle, another vehicle, person, animal, or non-fixed object was critical to this vehicle becoming involved in the crash or near-crash? This is a vehicle kinematic measure (based on what the vehicle does--an action, not a driver behavior). It does not include factors such as driver distraction, fatigue, or disciplining a child. This is the critical event which made the crash or near-crash possible. It may help to use the "but for" test; "but for this action, would the crash or near-crash have occurred?" This is independent of fault. For example, Vehicle A is speeding when Vehicle B crosses Vehicle A's path causing a crash, the Precipitating Event would be Vehicle B crossing Vehicle A's path. If two possible Precipitating Events occur simultaneously, choose the event that imparted the greatest effect on the crash or near-crash. If more than one sequential event contributed to the crash or near-crash, determination of which is the Precipitating Event depends upon whether the driver had enough time or vehicular control to avoid the latter event. If the driver avoids one event and immediately encounters another potentially harmful event (with no time or ability to avoid the latter), then the Precipitating Event is the first obstacle or event that was successfully avoided (this is where the critical envelope begins, and is the reference point for the other variables). If the driver had ample time or vehicular control to avoid the latter event, then that latter event would be coded as the Precipitating Event (the critical envelope would begin here, and all other variables would be coded based on this event). Note that a parking lot is considered a roadway--thus a barrier or light pole in the parking lot would be considered an object in the roadway.	V26 (Critical Event-Precrash 2 (Event))