
# Design of the In-Vehicle Driving Behavior and Crash Risk Study

## DETAILS

32 pages | | PAPERBACK
ISBN 978-0-309-12895-7 | DOI 10.17226/14494

## AUTHORS

Jon Antin; Suzie Lee; Jon Hankey; Tom Dingus; Transportation Research Board

The Second
STRATEGIC HIGHWAY RESEARCH PROGRAM

SHRP 2 REPORT S2-S05-RR-1

# Design of the In-Vehicle Driving Behavior and Crash Risk Study

## In Support of the SHRP 2 Naturalistic Driving Study

JON ANTIN, SUZIE LEE, JON HANKEY, AND TOM DINGUS
Virginia Tech Transportation Institute
Blacksburg, Virginia

TRANSPORTATION RESEARCH BOARD
WASHINGTON, D.C.
2011
www.TRB.org

## Subscriber Categories

Highways
Safety and Human Factors

## The Second Strategic Highway Research Program

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

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

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

SHRP 2 Report S2-S05-RR-1

ISBN: 978-0-309-12895-7



## Copyright Information



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

## Notice



### SHRP 2 Reports

Available by subscription and through the TRB online bookstore:

www.TRB.org/bookstore

Contact the TRB Business Office:
202-334-3213

More information about SHRP 2:

www.TRB.org/SHRP2

# THE NATIONAL ACADEMIES

***Advisers to the Nation on Science, Engineering, and Medicine***

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

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

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

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

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

**www.national-academies.org**

## SHRP 2 STAFF



## ACKNOWLEDGMENTS

This work was sponsored by the Federal Highway Administration (FHWA) in cooperation with the American Association of State Highway and Transportation Officials (AASHTO). It was conducted in the second Strategic Highway Research Program (SHRP 2), which is administered by the Transportation Research Board of the National Academies. The project was managed by Kenneth Campbell, Chief Program Officer for SHRP 2 Safety.

The research reported in this document was performed under SHRP 2 Safety Project S05, Design of the In-Vehicle Driving Behavior and Crash Risk Study, by the Virginia Tech Transportation Institute (VTTI) at Virginia Tech. VTTI was the primary contractor for this study and was supported by subcontracts through the University of Michigan Transportation Research Institute (UMTRI) and Battelle Memorial Institute (Battelle).

Dr. Jon Antin is the Leader of the Light Vehicle Safety Group and driver testing specialist at VTTI. The other authors of this report are Dr. Suzanne (Suzie) Lee, Research Scientist and human subject specialist at VTTI; Dr. Jonathan M. Hankey, Director of the Center for Automotive Safety Research at VTTI, and Associate Principal Investigator for this study; and Dr. Thomas Dingus, Director of VTTI and the Principal Investigator for this study.

Contributions were provided throughout the course of the S05 study design project from a variety of authors at VTTI, UMTRI, and Battelle.

# FOREWORD

Kenneth L. Campbell, PhD, *SHRP 2 Chief Program Officer, Safety*

This report describes the study design for the SHRP 2 Naturalistic Driving Study (NDS). Using a sophisticated recording package installed in vehicles, it will collect information on the day-to-day driving of about 3,100 volunteer drivers for up to 2 years. Participants will be recruited in six sites throughout the United States. In this report, potential users of the SHRP 2 NDS data or its findings will find information about the participant recruitment plan, informed consent and data protection procedures, driver assessment tests, the capabilities of the data acquisition system (DAS), quality control, and project management.

---

The objective of the SHRP 2 NDS is to reduce traffic injuries and fatalities by finding ways to prevent collisions and reduce their severity. Every 1% reduction in crashes will prevent 330 deaths and about $2 billion annually in medical expenses and other losses from these crashes. Moreover, crashes are a leading cause of nonrecurring congestion. Collision prevention has added benefits in terms of reduced delay, fuel consumption, and emissions. The focus of the NDS is to provide objective information on the role of driver behavior and performance in traffic collisions and the interrelationship of the driver with vehicle, roadway, and environmental factors.

Planning activities began with an enumeration of high-priority safety research questions. More than 400 candidate research questions were obtained from various committee and other outreach activities. The high-priority safety topics identified include road departure and intersection collisions. Data requirements to address the research questions drove the specification of the DAS capabilities. The final DAS, designed to maximize the utility of the data within budget and technical constraints, records continuously whenever the vehicle is running.

Participants will be recruited through a national call center and traditional solicitation activities, including print media, message boards, and posters at each NDS site. Six site contractors have been selected to obtain the consent of participants and install the DAS. Data recorded by the DAS will be retrieved every 4 to 6 months and uploaded to the NDS data storage facility at Virginia Tech. Roadway characteristics are also being measured in the NDS sites. GPS locations recorded by the DAS will be used to link roadway data with the information collected on the vehicle. In subsequent projects, analysts will access the data to address the high-priority safety questions.

# CONTENTS

# Executive Summary

This final report provides a summary of the key aspects of the planning effort supporting the second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS), which focuses on Safety Project S06 (Technical Coordination and Independent Quality Assurance for Field Study) and Safety Project S07 (In-Vehicle Driving Behavior Field Study). The integration of these projects with Safety Project S04A (Roadway Information Database Development and Technical Coordination and Quality Assurance of the Mobile Data Collection Project) and Safety Project S04B (Mobile Data Collection) is also discussed.

The objective of the SHRP 2 Safety Project S05: Design of the In-Vehicle Driving Behavior and Crash Risk Study (Study Design) was to design the SHRP 2 NDS, which will collect data—on the order of 1 petabyte (1,000 terabytes)—on “naturalistic,” or real-world, driving behavior over a 2-year period beginning in fall 2010.

During the SHRP 2 NDS, contracts will be awarded for data collection at six sites chosen from among 11 prequalified sites. These sites were selected in an attempt to provide as nationally diverse a source of data as possible, with a fairly wide range of geography, weather, state laws, road types, and road usage. The field study data collection contractors (S07) at each site will be responsible for implementing the plan devised in the planning project for each of the following: installing and uninstalling the data acquisition system (DAS) units; participant assessment; data collection; addressing participant management problems encountered during the study; investigating crashes and data transmission; carrying out quality control procedures; and preparing periodic reports documenting the field study activities. The S06 contractor—the Virginia Tech Transportation Institute (VTTI)—will serve in the oversight and integration role for all activities that will be conducted under procedures defined by the study design (S05) contractor.

Participant recruitment and screening for all sites will be conducted by a single national call center. The study design seeks to control the distribution of drivers in the study by age and gender according to the experimental design. To provide a more representative group of participants, the call center will first seek to recruit individuals who have been randomly selected from listed samples of households and cell phones located in the vicinity of the study sites. As recruitment activities progress, they may have to be targeted towards age/gender cells that prove more difficult to fill. Finally, traditional recruitment activities (such as advertisements and flyers using prepared recruitment materials provided by VTTI) may also be conducted by the S07 site contractors, as needed.

Individuals who volunteer to participate will be compensated as follows: The data collection system will be installed in their personal vehicle for either 1 or 2 years. Compensation for 1 year of study participation will total a maximum of $300. For 2-year participants, the compensation will total a maximum of $600. If a participant withdraws from the study for any reason prior to the scheduled end date, compensation will be prorated at a rate of $25 per month of study participation.

In addition to the primary volunteers, it is expected that family members or others may also regularly drive study vehicles. Such individuals are considered secondary drivers. These secondary drivers will also be asked to provide informed consent so that data collected while they are driving may be analyzed as appropriate. No additional compensation will be provided to or for any secondary driver.

Individuals who drive the vehicle less frequently or who remain unconsented for whatever reason will not be a part of the study, and any personally identifiable data collected (such as facial images and Global Positioning System [GPS] coordinates) while such individuals are driving will not be used. All human subject protocols are subject to review by the appropriate institutional review boards (IRBs).

The SHRP 2 NDS will track approximately 1,950 DAS units on vehicles for 2 years, for a total of 3,900 data years. The DAS is designed to be easily mounted on a wide range of vehicles, including passenger cars, vans, sport utility vehicles (SUVs), and pickup trucks, but not all makes and models in these categories will be eligible for inclusion in the study. Some DAS units will be installed and maintained for 2 full years while others will be removed at the end of the 1st year, after which new participants will be recruited for the 2nd study year. Recruiting new drivers after the 1st year increases the number of primary drivers in the study to 3,102.

The DAS comprises several key components. The head unit assembly, mounted to the right of a participant's rearview mirror, will feature three video cameras to capture moving images of the driver's face, the forward roadway, and a dashboard and center-stack view. In addition, another camera housed in the head unit will periodically capture a still image of the cabin. This still image will be used to determine the presence of passengers, and it will be irretrievably blurred to protect the anonymity of unconsented passengers. A fifth video camera will be mounted in the rear deck to capture the rear and right-side views. Additionally, an alcohol sensor, an incident button, an illuminance sensor, and inertial acceleration and gyroscopic sensors will be incorporated into the head unit assembly.

The main unit houses the processing, recording, and communication functions of the system that will be active during any vehicle operation. The main unit features the ability to transfer limited data quantities via cellular interface and features other standard communication interfaces such as a controller area network (CAN) and universal serial bus (USB). Vehicle network information available from the CAN will include the activation status of variables, such as the antilock brake system (ABS), electronic stability control (ESC), brake assist, and traction control. Lastly, a radar unit will be mounted to the front of the vehicle. Radar data will be transmitted via Bluetooth technology to the main unit to minimize installation time and costs. The radar will capture data concerning the relative position and speed of nearby objects, such as lead vehicles.

In addition to the continuous data gathered on the vehicles, functional assessments will be conducted for all primary participants along dimensions related to cognitive, visual-cognitive, perceptual, physical, psychological predilection, and overall health. Crash investigation will also be performed on selected significant crashes of interest (e.g., police-reported or reportable crashes). Data from any or all sources may include highly personal or sensitive data (e.g., health information or crash records). As such, data security will be strictly maintained throughout the process. This includes fulfilling all IRB practices during data collection as well as during all subsequent analyses for the life of the data set.

A suite of quality assurance, quality control, and project management activities will be applied throughout the study to ensure that the highest-quality data are collected within budget and schedule constraints.

As previously summarized, the S05 study design project has developed a complete study plan for the SHRP 2 NDS. The resulting data, expected to exceed 1 petabyte in size—about the size of a million 1-gigabyte USB flash drives—will provide a wealth of information regarding driving behavior, lane departures, and intersection activities, which is anticipated to be of interest to transportation safety researchers and others for at least 20 years. It is important to note that all privacy protections promised regarding participants and their data are to be afforded in perpetuity. All future research efforts that seek to use the data collected in the SHRP 2 NDS will require IRB approval. In addition, researchers must establish a data-sharing agreement that guarantees privacy at least to the extent specified in the Informed Consent document. SHRP 2 anticipates the release of multiple projects under Safety Project S08, Analysis of In-Vehicle Field Study Data and Countermeasure Implications. The objective of S08 is to quantify the contribution of relevant driver, roadway, vehicle, and environmental factors to the research questions selected and assess the countermeasure implications of the findings. The knowledge gleaned from the SHRP 2 analyses, as well as the many additional analyses anticipated to be performed by other researchers, will support public policy, rulemaking, infrastructure improvements, and other—as yet unknown—activities targeted at reducing the fatalities on our nation's roadways.

CHAPTER 1

# Introduction

The second Strategic Highway Research Program (SHRP 2) is administered by the Transportation Research Board (TRB) of the National Academies and has the goal of addressing the challenges of moving people and goods efficiently and safely on the nation's highways. There are four areas of emphasis associated with SHRP 2: Safety, Reliability, Renewal, and Capacity. Each of these areas in turn encompasses several research projects. The SHRP 2 Safety research program evaluates the role of driving behavior in traffic safety. While driving behavior has largely been acknowledged as a primary factor in most collisions, the complex interrelationship of driver performance with factors such as roadway design, environmental conditions, and vehicle features—and how those factors influence the risk of collisions and casualties—remains largely unknown. Crash reports and investigations can capture only limited information on driver activities before the collision and cannot provide comparable information on driving activities in non-collision situations. By using innovative research methods combined with advanced technologies, SHRP 2 sees opportunities to improve traffic safety via the research project described below. Specific goals of the SHRP 2 Safety program include the following:

- Understanding how drivers interact with and adapt to various factors such as
  - The vehicle;
  - Traffic and traffic control devices;
  - The environment; and
  - Roadway characteristics.
- Assessing the changes in collision risk associated with each of these factors and their interactions.

In response to the TRB SHRP 2 solicitation S05: Design of the In-Vehicle Driving Behavior and Crash Risk Study (Study Design), the Virginia Tech Transportation Institute (VTTI) joined with the University of Michigan Transportation Research Institute (UMTRI) and Battelle Memorial Institute (Battelle) to lay the foundation for the most significant program of highway safety research in more than 30 years. The successful completion of this research effort will produce an unprecedented database of near-crash, precrash, crash, driving behavior, driving performance, roadway information, and vehicle kinematics data that will allow substantial progress to be made in both the crash-causation and crash-countermeasure domains. By working to improve highway safety through a more comprehensive understanding of driving behavior, the SHRP 2 Safety analysis projects, which will be conducted as a series of studies under Safety Project S08, will advance the directives of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) legislation passed by Congress in 2005 and of the SHRP 2 Safety plan.

This research agenda comes at a critical time in the history of our nation's surface transportation safety. The project is part of an effort to make driving safer in the United States. Vehicular crashes, because they do not discriminate by age, constitute a leading killer in terms of years of lost life. For example, newly licensed teen drivers are three times more likely to be involved in a fatal crash than their older counterparts. They are also more likely to die in a car crash than from all other sources of disease and accidental injury *combined.*

This project's research is an ambitious undertaking. The required data collection, archiving, reduction, and analysis efforts are of far greater magnitude and scope than any driving safety-related studies that have been attempted or accomplished to date anywhere in the world. In fact, advances in sensors, computing, data storage, video compression, and data-mining technology have only recently made this project feasible. In addition, the analytical methodologies, hierarchies and models, and the generation of hypotheses to be tested will need to be taken to a new level of quality and utility. Since achieving a successful outcome is a difficult task, the design of the project is a critical element to the success of the SHRP 2 NDS.

This final report provides a high-level summary of the key areas of the planning study that will be used in supporting the NDS, which includes Safety Project S06 (Technical Coordination and Independent Quality Assurance for Field Study), Safety Project S07 (In-Vehicle Driving Behavior Field Study), and Safety Project S04 (Roadway Data Collection). Figure 1.1 illustrates the general relationships among the SHRP 2 Safety projects as well as their approximate time frames. In brief, the planning accomplished in S05 Study Design supports the procurement of NDS data collection hardware (S12) as well as the NDS data collection activities to be accomplished in S06 and S07. Similarly, the evaluations performed in Safety Project S03 support the collection of the roadway data in the S04 projects. The analysis projects to be released in Safety Project S08 will be guided by the development of analysis methods and plans in Safety Projects S01 and S02. Finally, all data collected in the S04, S06, and S07 efforts will be integrated for S08 analysis projects, as well as for other analyses going forward.

Safety Project S06 is intended to provide technical coordination and independent quality assurance for the in-vehicle driving behavior field studies (S07) and technical coordination with the roadway data collection projects (S04A and S04B). This implementation project, as designed in Safety Project S05, includes the following:

- Coordinating the recruitment and protection of human subjects;
- Managing the data acquisition system (DAS) procurement, testing, and warranties;
- Integrating data collection systems;
- Providing support to the S07 site contractors;
- Controlling quality aspects and managing oversight activities; and
- Managing and reporting all aspects related to the study.

Safety Project S07 is tasked with conducting the in-vehicle driving behavior field study that will be performed at six locations throughout the continental United States. The key responsibilities include the following:

- Participant intake, including consenting, assessing, and initiating participant compensation;
- Installing DAS units on participant vehicles;
- Retrieving hard drives on which the data have been collected and uploading the data to the S07 staging computer (which will automatically transmit the data to the S06 servers);
- Deinstalling DAS units when participants exit the study; and
- Troubleshooting problems encountered, with the assistance of the S06 contractor.

Safety Project S04 relates to the large field study of driver behavior and will provide and manage road information to be used in safety analysis. It has two parts: A and B. Safety Project S04A represents the effort to provide quality assurance and quality control for the S04B Safety Project. Safety Project S04A will acquire existing roadway data, provide technical

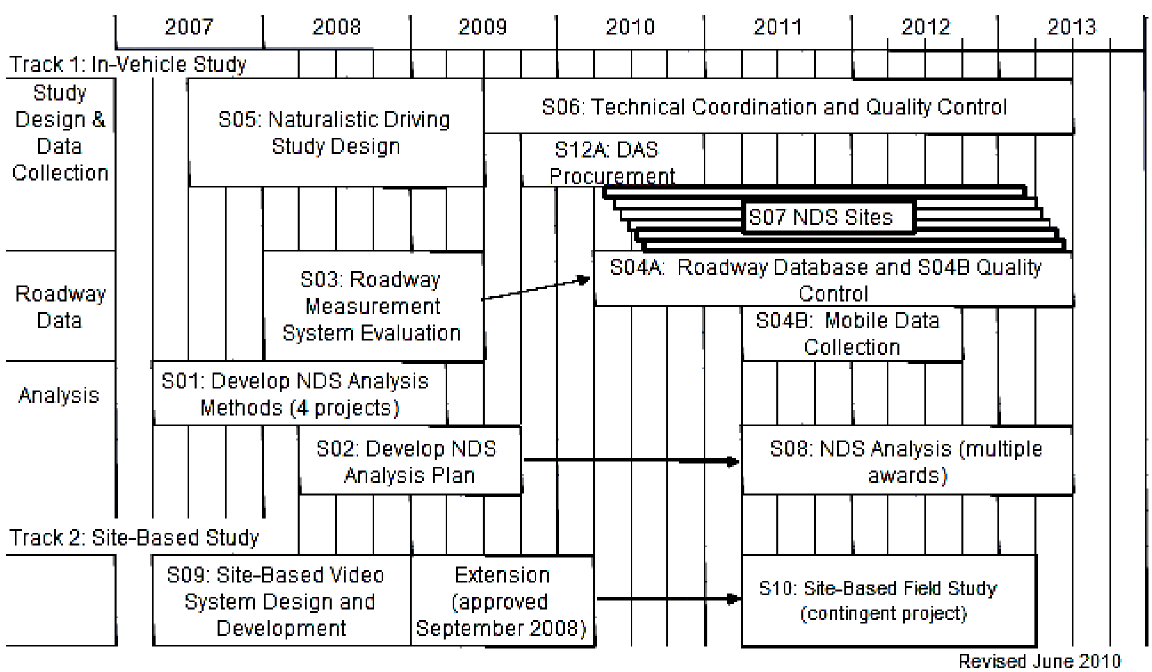


***Figure 1.1. SHRP 2 safety projects.***

coordination for collection of roadway data under Safety Project S04B, and produce the SHRP 2 Roadway Information Database. Safety Project S04B is focused on collecting roadway characteristic data on selected roads within the in-vehicle driving behavior field study areas. This work will be facilitated by the technical coordination contract under S06 to help the S04 contractor(s) determine the most salient roadways to cover in each study area. Ultimately, this will lead to the production of a roadway database in a geographic information system (GIS) linked via an appropriate linear reference system, and merged with existing roadway data from other sources. The study will incorporate variables using a roadway measurement system qualified in the SHRP 2 Roadway Measurement System.

CHAPTER 2

# Study Design

The study design elements have been defined from a variety of perspectives to try to make certain that there is sufficient statistical power to detect statistically significant effects in as many cases as possible. In this way, the data produced by the study will be able to answer as many of a comprehensive set of categorized research questions as feasible, while staying within the bounds of known project constraints such as funding and time. These constraints serve to limit the project in several ways, such as the duration of the data collection and the number of data collection sites that can be established and managed.

The set of research questions has guided decision-making processes throughout the design of the SHRP 2 NDS. Additionally, sampling statisticians and other experts were consulted in the development of the sample design and contractor site-selection process. Each of these elements, together with the process for participant recruitment, is discussed below.

## Specific Research Questions

On the basis of numerous source documents, including *The 100-Car Naturalistic Driving Study, Phase II* (Dingus et al. 2006), the *Tri-Level Study of the Causes of Traffic Accidents Final Report* (Treat et al. 1979), and *Pre-Crash Scenario Typology for Crash Avoidance Research* (Najm et al. 2007), as well as expertise from researchers in the transportation safety field, 392 research questions were identified for consideration in the NDS design. These questions address different types of crashes and other non–crash-specific safety areas of interest. They also provided the basis for making a variety of decisions regarding the number and types of DASs used to collect data during the study, the types and priority of additional sensors and DAS capabilities, and the sampling and analysis trade-offs.

The questions were subcategorized into the following perspectives (the list also includes examples of study design decisions intended to address the questions, as applicable):

- **Traffic-, roadway-, and environment-based questions.** Roadway data will be collected in the Roadway Data Collection effort (S04). Front-mounted radar combined with the images from several synched video cameras will provide traffic information. Environment-based questions can be addressed in several ways. Information will be available about time of day, ambient illuminance levels, windshield wiper status, and, for select crashes, the relevant details of the crash site will be documented by a trained crash site investigator.
- **Vehicle-based questions.** Vehicle details, including make, model, and year, as well as all onboard safety features, will be documented for each vehicle in the study.
- **Driver- or driver-error-based questions.** Each primary driver in the study will be characterized in terms of his or her demographic details. In addition, each primary driver will undergo a fairly extensive suite of assessments designed to characterize him or her along various driving-relevant dimensions of ability (e.g., visual perception). In addition, continuous video will be collected from four cameras. These can be reviewed for any event of interest to address error-based questions.
- **Passenger-based questions.** A fifth camera will be aimed to capture an irrevocably blurred still image of the cabin every 10 minutes. This is done to capture basic information on the number of passengers as well as an approximation of their ages and genders for each trip.
- **Infotainment-system-based and nomadic-device-based questions.** Information on Infotainment system status will be available from the controller area network (CAN) bus for some of the vehicle fleet. Nomadic device (e.g., handheld technology, such as a cell phone) usage can be determined using appropriate sampling schemes, though hands-free usage will be more difficult to reliably detect.
- **Aggressive-driving-based questions.** Aggressive driving patterns can be seen in the exceedance of preset (but adjustable) levels of various sensors or combinations thereof

as expressed in pretested algorithms (e.g., longitudinal deceleration < –0.5 *g and* forward time-to-collision < 3.0 s).

- **Vision-, attention-, and distraction-based questions.** Continuous video of the participant's face permits the scoring of glance direction and duration which can be analyzed to address these areas of relevance. This can be accomplished via manual frame-by-frame video data reduction, but machine-vision-based techniques are being developed at VTTI to enable some of the glance-related scoring to be performed in an automated fashion.
- **Speed- and speeding-based questions.** A set of redundant sensors provides continuous vehicle speed information. In addition, onboard Global Positioning System (GPS) sensor data can be combined with GIS-based information on speed limits to permit researchers to look at speeding-related behaviors (e.g., actual speed relative to the posted speed limit) and safety-related events.
- **Crash-countermeasure-based questions.** Crash countermeasures can be evaluated by examining actual crash-related events and superimposing different countermeasure algorithms to evaluate their relative effectiveness. Similarly, these same algorithms can be applied to non-event baseline epochs to determine their relative propensity for creating false alarms.
- **Passing-maneuver-based questions.** Passing maneuvers can be detected in the data stream (i.e., via the yaw sensor) and differentiated from swerves on the basis of turn signal status and visual validation of the passing maneuver.
- **Multifactor/Multivariate questions.** All of the above types of questions (and many more not listed) can be looked at in any combination desired by a researcher. For instance, a researcher may be interested in looking at a speeding-related countermeasure to evaluate whether it may have been possible to prevent any speeding-related crash events observed.

The research questions were grouped by the SHRP 2 Safety Project S02 contractor into the following high-level categories:

- How does driver distraction influence crash likelihood?
- How does driver fatigue affect crash likelihood?
- How do aggressive driving behaviors influence crash likelihood?
- What is the influence of driver impairment (e.g., alcohol) on crash likelihood?
- How do driver interactions with roadway features influence the likelihood of lane departure crashes?
- How do driver interactions with intersection features (configuration and operations) influence crash likelihood?
- How do advanced driver support systems influence crash likelihood?
- What variables or pre-event factors are the most effective crash surrogate measures? What explanatory factors are associated with crashes or crash surrogates? And what analytical models can be developed to predict crash or crash surrogates?

The list of research questions will continue to expand as the full implications of the data are analyzed and understood. Experience with other naturalistic studies, and the expectation for the SHRP 2 NDS, suggests that the data will be useful for addressing many questions for years after the data are collected. The specific research questions are discussed in greater detail in Appendix A of the S02 Phase 1 report (Boyle et al. forthcoming).

## Sample Design Plan

There are two broad aspects to the NDS sample design: (1) contractor-site selection and (2) participant-vehicle sample design. The first refers to the process of determining the number and location of the data-gathering sites and the contractors who will manage each; the second refers to the methods and factors used to recruit and select participants. These two aspects ultimately work toward the definition of a unified plan that specifies the quantity and characteristics of participants to be recruited at and across all site locations.

From the standpoint of robust experimental design, it is desirable to ensure that there is good representation in the participant sample of the basic driver–vehicle variables of relevance and interest. In this study, the primary variables of interest are age, gender, and vehicle type.

It has been well documented that age is a strong indicator of driving risk, with the youngest and oldest drivers having an elevated crash rate per mile driven compared with other age groups. Exposure details were carefully considered in an effort to understand the implications of driver age and gender in terms of crash rates per mile and per numbers of licensed drivers. Since the target data collection budget was already established, the following are some of the major trade-offs that were considered:

- **Representative versus risk-prone sample.** Do we strive for a more representative sample, which enhances generalizability of the data, or do we strive for a sample we believe to be more crash-prone (i.e., one emphasizing the extremes of the age range) to observe more safety-related incidents of interest?
- **Overall costs of participant pay versus the ability to enhance participant attraction and retention via meaningful compensation amounts.** It is reasonable to expect that the greater the compensation, the greater the recruitment uptake and retention rates would be. However, this is constrained by the project funding available for this purpose. It was ultimately determined that participants would

be compensated by the nominal amount of $25 per month of participation.

- **Total number of sites versus number of DAS units managed per site.** It was ultimately determined that the maximum number of DAS units that could be supported was 1,950. Then, the issue becomes how many data collection sites must be established to manage those DAS units. If too many sites are established, then the costs would be prohibitive. If too few sites are established, then each may be asked to manage more DAS units than resources at a particular site may allow.
- **Explicitly including stratification variables in the experimental design versus the difficulty of filling each cell.** The more variables formally included in the experimental design, the greater the risk of ending up with experimental cells with too few participants—or perhaps none at all. However, if factors of interest are not formally included, then the actual sample may not include sufficient numbers for analysis, and they are likely to be unevenly spread across the other factors of interest.
- **Total number of primary participants versus months of data collection per participant.** It is desirable to have as many participants as possible in the study, yet recruitment is expensive, and it is costly to move a DAS from one vehicle to another.
- **Total cost of data collection versus cost per data-year.** Just as it is desirable to have as many participants in the study as possible, so too is it desirable to have as many years of data or data-years as possible (i.e., where a data-year is equivalent to the amount of data generated by a single participant over the course of a single year). Of course, there is a substantial cost for each study year. However, the cost per data-year tends to diminish as the study period is extended, thus making the study simultaneously more expensive yet more cost-efficient.
- **Site recruitment size versus a contractor's ability to manage the square mileage.** The larger the size of a site's recruitment area, the greater the probability of finding a sufficient number of participants. However, if a site is too spread out geographically, then not only will this cause the contractor difficulty in managing all the DAS and participant issues, but it will also make it more difficult for those participants at the most distant points from the installation site.

Passenger cars and light trucks will be the focus of this study because these types of vehicles accounted for almost 95% of all vehicles on the road in 2007, as reported by Ward's Automotive Group (2010), and 94% of all motor vehicle crashes, as reported by the National Highway Traffic Safety Administration (2007). Vehicles that will be instrumented for the SHRP 2 NDS include the following: passenger cars (sedans, coupes, hatchbacks, and station wagons), pickup trucks, sport utility vehicles (SUVs; including crossover vehicles), and minivans. It is expected that only vehicle model-years later than 2002 will be targeted for recruitment to help ensure that only mechanically sound vehicles with access to vehicle network data, such as speed and accelerator position, will be included in the sample. The goal is for the DAS to support installation in about half of the light-vehicle population on the road.

On the basis of U.S. light-vehicle sales from model years 2000–2007, a list of the top 50 most popular vehicles were identified for possible inclusion in the study. These top sellers are represented by the members of either the Alliance of Automobile Manufacturers (AAM) or the Association of International Automobile Manufacturers (AIAM); these organizations represent a cumulative total of 89% of all U.S. light-vehicle sales for the selected time period. Relationships are being pursued with original equipment manufacturers (OEMs) belonging to AAM and/or AIAM with the objective of having CAN Parameter IDs (PIDs) available to obtain and interpret additional data from the onboard vehicle network for high-volume models; the types of data of interest include speed, wiper usage, brake actuation, accelerator position, and turn signal usage, as well as steering data. Within the advanced technology group of participants, additional data will be collected regarding the usage of in-vehicle communication systems and advanced infotainment systems. Additionally, information about driver monitoring, feedback, and collision warning systems will be available from some participating manufacturers.

Using the research questions to guide the process, the research team conducted an analysis to estimate the statistical power afforded by the experimental design to detect a statistically significant effect associated with the various age and gender groupings. The analysis generally indicated that the study was sufficiently powered to address the age by gender questions. This analysis was conducted on a single variable, speed variability, where previous data could be used to estimate that variable's mean and standard deviation (both required for power estimation). Since its standard deviation was fairly high relative to its mean, this variable represents a relatively conservative estimate of the statistical power that can be expected with various analyses. Still, it must be recognized that the power associated with each analysis will be dependent upon the particular means and standard deviations of the measures used and the actual magnitude of the differences observed.

The experimental design shown in Table 2.1 is based on the preceding investigations, and revisions were made throughout the course of the S05 study design project. Note that there are more data years than participants because some participants will participate for the full two years of data collection instead of just one. Also note that the sample design emphasizes the extremes of the age spectrum more than the middle-aged

*Table 2.1. Sample Design (with Target Cell Values)*

| Gender and Age Range | Age Range Description | One Year | Two Years | DAS Units | Primary Participants | Data-Years |
|---|---|---|---|---|---|---|
| M 16–17 | Minor teen | 72 | 28 | 100 | 172 | 200 |
| M 18–20 | Adult teen | 72 | 28 | 100 | 172 | 200 |
| M 21–25 | Young adult | 72 | 28 | 100 | 172 | 200 |
| M 26–35 | Adult | 72 | 28 | 100 | 172 | 200 |
| M 36–50 | Middle adult | 72 | 28 | 100 | 172 | 200 |
| M 51–65 | Mature adult | 72 | 28 | 100 | 172 | 200 |
| M 66–75 | Younger older driver | 72 | 28 | 100 | 172 | 200 |
| M 76+ | Older older driver | 72 | 28 | 100 | 172 | 200 |
| F 16–17 | Minor teen | 72 | 28 | 100 | 172 | 200 |
| F 18–20 | Adult teen | 72 | 28 | 100 | 172 | 200 |
| F 21–25 | Young adult | 72 | 28 | 100 | 172 | 200 |
| F 26–35 | Adult | 72 | 28 | 100 | 172 | 200 |
| F 36–50 | Middle adult | 72 | 28 | 100 | 172 | 200 |
| F 51–65 | Mature adult | 72 | 28 | 100 | 172 | 200 |
| F 66–75 | Younger older driver | 72 | 28 | 100 | 172 | 200 |
| F 76+ | Older older driver | 72 | 28 | 100 | 172 | 200 |
| Any | Advanced vehicle technology | 0 | 350 | 350 | 350 | 700 |
| **Total** | | **1,152** | **798** | **1,950** | **3,102** | **3,900** |

groups, as the extremes are where the age-related problems tend to manifest.

## Data Collection Sites

The objective in selecting the suite of sites for the SHRP 2 NDS was to represent to the extent feasible the wide range of driving and geographic conditions and other relevant characteristics found across the United States. Of course, the sites could not be selected independent of the site contractors: for each site selected, there had to be a suitably experienced contractor who responded to the Request for Proposals (RFP) and proposed to manage a site in a particular location. The site contractors must be technically qualified for all aspects of the study as well as have (or be able to secure) the necessary facilities to carry out the data collection plan. Note that each site selected could only represent some portion of the total environmental diversity (i.e., in terms of geographic location, predominant terrain, land use, or weather patterns, to name a few key factors). It is important for the NDS to have a broad range of environmental conditions represented to understand these factors in combination with other driver, vehicle, and roadway factors.

Potential contractors responded to two rounds of Request for Qualifications (RFQ) that addressed the qualifications of the site contractor, the characteristics of the proposed site, and the availability and quality of existing state or other local data (e.g., driver licensing, roadway inventories, driver history, and crash data). Of those respondents, 11 contractors were identified as potential SHRP 2 NDS study sites. Nine of these prequalified contractors responded to an RFP issued in March 2009. Proposals specified the details regarding how proposers would go about managing one or more data collection sites and how the sites would contribute to the collective balance of the entire suite of sites in terms of the key factors such as geography, weather, state law, road types, and land usage (e.g., urban versus rural). Proposals were evaluated on the basis of defined contractor qualifications, including human participant research experience; well-trained, permanent, full-time technical staff; and sufficient and suitable office, assessment, and garage space. To handle a project of this magnitude, in which a typical site will be managing up to three times as many primary participant-vehicles as the *entire 100-Car Study* (Dingus et al. 2006), requires substantial experience in working with participants and managing unanticipated circumstances. Site contractors were selected on the basis of their individual merits, as well as the collective representation they bring to the group of selected S07 sites. Figure 2.1 shows the six sites chosen to serve as the SHRP 2 NDS data collection sites. The sites are also listed below, along with the contractors who will manage each site.


*Figure 2.1. Sites for SHRP 2 Naturalistic Driving Study, indicating the number of DAS units managed per year at each site.*

- Erie County, N.Y.: Calspan–University of Buffalo Research Center (CUBRC);
- Seattle, Wash.: Battelle Memorial Institute;
- Central Pennsylvania centered on State College: Pennsylvania State University;
- Central Indiana centered on Bloomington: Indiana University;
- Tampa Bay, Fla.: Calspan–University of Buffalo Research Center (CUBRC); and
- Durham, N.C.: Westat.

## Participant Management

This study will require the participation of more than 3,000 volunteers in total, spread across the study's six data collection sites. Managing these participants from recruitment through the end of their participation will require several key activities as described below.

### Participant Recruitment

Previous naturalistic driving studies have typically recruited volunteers through so-called traditional means, including media advertisements and posted flyers. However, a probability-based sampling approach can provide a less biased, more representative sample from which generalizations can more accurately be drawn. To begin to address the questions of cost, efficiency, and measures of scope associated with a probability-based sampling approach for a study of this size, a pilot study was conducted by Battelle Memorial Institute to identify and assess potential challenges related to effectiveness, key differences, and relative costs of using phone recruiting.

### Recruitment Pilot Test

A work plan was developed and thoroughly tested and validated by staff at Battelle with the objective of testing several major aspects of the participant recruiting and screening process. The key objectives were as follows:

1. Determine the extent to which a random selection approach can be used and how this approach compares with more traditional approaches;
2. Determine the incentives that will be required to get people to participate; and
3. Assess the participant-screening forms and participant-assessment protocols.

From the outcomes of this pilot study, it is expected that an approach that combines both the call-out and traditional recruiting methods will provide a reasonable balance of randomized selection with efficient and reliable methods for obtaining the participants needed in the SHRP 2 NDS within the narrow time frame allotted.

A decision was made to use a centralized call center to conduct phone-based recruiting of volunteers to participate in the NDS. Current expectations are that, initially, the phone-based recruiting approach will be used exclusively. It may then need to be supplemented with more traditional recruiting methods centered around each site (e.g., newspaper ads and flyers placed on vehicles) to target the harder-to-fill cells in the experimental design. For consistency, the call center will also field incoming inquiries generated by the traditional recruiting ads. By centralizing as much of the recruitment process as is practical, the recruitment process can be kept much more consistent across sites, and substantial unnecessary duplication of activities across sites (such as development of recruiting protocols, materials, and data entry software) can be avoided.

## Consent and Privacy

It is important for the informed consent process, and the privacy it guarantees to participants and their data, to be addressed in close coordination with nearly all of the other S05 tasks and across the S07 sites. To this end, none of the participants' directly identifying information (e.g., name, address, social security number) will ever be associated with any of their data for any level of analysis. However, some participant data may inherently incorporate identifying elements (e.g., face video, voice recordings, some questionnaire responses, vehicle inventory, crash time and location, and some GPS information). To maintain participant privacy in light of such factors, each research project that proposes to use the SHRP 2 NDS data set must adhere to data access procedures, including attaining IRB approval, that will be developed as appropriate. Then, only the level of data required for the specified analyses will be made available.

The main drivers of study vehicles are targeted as primary participants in the study. Up to three other individuals per primary participant, typically a participant's family members—who can be expected to also regularly drive the study vehicle—are considered secondary drivers. Study personnel will attempt to get informed consent from up to three of these secondary drivers and ask them to complete select questionnaires. However, for secondary drivers who opt out or for others who may drive the vehicle less frequently, their data that include inherently identifying information (such as facial images and GPS coordinates) will be de-identified. Efforts related to the protection of human subjects will be discussed in greater detail in Chapter 5. All human subject protocols are subject to review by the appropriate IRBs.

## Participant Compensation

Primary drivers who volunteer to participate will be compensated for their participation in the SHRP 2 NDS. The data collection system will be installed in participants' personal vehicles for either 1 or 2 years. Compensation for one year of study participation will total $300, maximum; for two years of study participation, the compensation will total $600, maximum. If a participant withdraws from the study prior to the scheduled time, compensation for that participant will be prorated at a rate of $25 per month of participation at the time of exit. No compensation will be provided to secondary (or other) drivers of the study vehicle, whether consented or not.

CHAPTER 3

# Data to Be Collected

## Driver Demographics and Vehicle Inventory

Basic demographic data will be gathered from each primary driver. The list currently includes the following:

- Gender;
- Date of birth;
- Ethnicity;
- Race;
- Country of birth;
- Education level;
- Marital status;
- Household makeup;
- Household ownership;
- Working status;
- Vocation;
- Household income;
- Household population (i.e., number of people living at participant's residence);
- Household age categories;
- Number of vehicles (i.e., number of vehicles residing at the participant's residence);
- Zip code;
- Years of residence (i.e., number of years of residence in current neighborhood);
- Vehicular travel (i.e., estimated number of miles driven last year by the participant);
- Business use (i.e., is the instrumented vehicle used for business purposes?);
- Business purpose (i.e., if the instrumented vehicle is used for business purposes, what are those business purposes?); and
- Licensure age (i.e., at what age did the participant receive his or her first driver's license?).

In addition, the instrumented vehicle will be inventoried to record its basic facts as well as additional options and features. This inventory will include at least the following:

- VIN (vehicle identification number, nonpersonally identifying digits only);
- Make;
- Model;
- Year;
- Style/trim level (e.g., LX, EX, EX-L, or Touring for the 2010 Honda Odyssey);
- Body style (e.g., coupe, sedan, or wagon);
- Color;
- Safety features (e.g., antilock brake system [ABS]; electronic stability control [ESC]; front, side, curtain airbags; forward collision warning [FCW]; lane departure warning [LDW]; adaptive cruise control [ACC]); and
- Infotainment features (e.g., integrated navigation or infotainment systems).

## Driver Assessment

In a study such as this, it is desirable to measure certain driving-related functional capabilities and limitations of the participants. These, singly or combined in some fashion, may help to explain some of the variability in the driving or crash-related data observed. If so, such factors may lead to the development of countermeasures targeting such individual differences or impairments.

As part of the development of the driver testing suite, a review of the relevant literature was conducted to help devise the assessment plan. In addition, a blue-ribbon panel of experts thoroughly reviewed all aspects of the plan and provided invaluable, iterative feedback. All of their concerns were addressed in some fashion, and all members of the panel ultimately indicated their approval of the plan that is presented herein.

On the basis of the exercise described, an approach was defined wherein each participant will be assessed on several dimensions thought to be relevant to driving and potentially predictive of some portion of driving behaviors, problems, or crash-related events (e.g., crashes and near crashes). A systematic approach was undertaken to identify the most relevant dimensions and essential criteria in the development of the resulting driver assessment testing plan. The sleep questionnaire was constructed in close coordination with the Harvard Medical School Division of Sleep Medicine.

The selected relevant dimensions were executive function and cognition; visual perception; various visual–cognitive, physical, and psychomotor abilities; personality factors; sleep-related factors; medicines and medical conditions; driving knowledge; and history. Executive function (EF) broadly encompasses a set of cognitive skills that are responsible for the planning, initiation, sequencing, and monitoring of complex goal-directed behavior (Royall et al. 2002). The criteria used in selecting the testing instruments to measure each of these dimensions were comprehensiveness, evidence of predictive value, feasibility of administration, uniqueness, persistence, and feasibility of replication.

The driver assessments are intended to be administered in a 2- to 3-hour period that will, for the most part, occur simultaneously with vehicle instrumentation. A pilot test of the suite of driver assessments was conducted and, on the basis of the outcome of that pilot test, it is believed that they can be administered within that time frame. The current driver testing assessments are listed below.

Some instruments will be administered by trained personnel; other instruments include online questionnaires that can be filled in at the assessment site or later (e.g., within a week of installation). The first participant compensation payment will be made after all driver assessment activities, including filling in all questionnaires, have been completed.

The assessments to be implemented are listed in Tables 3.1 to 3.7.

***Table 3.1. Visual Perception Assessments***

| Measurement Construct | Description | In Person/ Online | Estimated Time (min) |
|---|---|---|---|
| High-light contrast sensitivity | Optec 6500P (Figure 3.1) | In person | 5 |
| Low-light contrast sensitivity | | | 5 |
| Near static acuity | | | 5 |
| Far static acuity | | | 5 |
| Depth perception | | | 5 |
| Color vision | | | 5 |
| Peripheral vision | | | 5 |

***Table 3.2. Visual-Cognitive Assessments***

| Measurement Construct | Description | In Person/ Online | Estimated Time (min) | Comments |
|---|---|---|---|---|
| Spatial relationships | Motor-Free Visual Perception Test (MVPT) Visual Closure Subtest | In person | 3 | Included in DrivingHealth Inventory software |
| ▪ Central vision and processing speed<br>▪ Divided/selective attention | Useful Field of View (UFOV)—Part 2 only | In person | 5 | |
| Divided attention | Trail-Making Test (A & B) | In person | 6 | |
| | Rapid Pace Walk (discussed in Table 3.4 below but also included here, since it is recorded within the DrivingHealth Inventory software) | In person | | |

***Table 3.3. Cognitive and Psychomotor Assessments***

| Measurement Construct | Description | In Person/ Online | Estimated Time (min) | Comments |
|---|---|---|---|---|
| ▪ Executive function/working memory<br>▪ Reaction time | Connors' Continuous Performance Test II (CPT II), Ver. 5.1 | In person | 15 | |
| Dementia | Clock-Drawing Test:<br>▪ Draw a clock<br>▪ Put in all the numbers<br>▪ Set the hands at 10 past 11 | In person | 2 | Administered to all participants |

*Table 3.4. Physical Assessments*

| Measurement Construct | Description | In Person/ Online | Estimated Time (min) |
|---|---|---|---|
| Lower limb strength/ mobility | Rapid Pace Walk | In person | 1 |
| Upper body strength | Jamar Hand Dynamometer | In person | 2 |

## Data Acquisition System Variables

The DAS comprises three primary units: the head unit, the main unit, and the front radar assembly as illustrated in Figure 3.2. Cabling will provide hard-wired connections between the head unit and main unit as well as a power source. Additionally, cabling will connect to the onboard diagnostic port (OBD II) so that specific vehicle network data can be obtained. Turn signals are collected from network data as available, or a cable connection will be established to capture turn-signal usage data. Note that the head unit is actually composed of two subassemblies: the rear camera and cellular antenna (General Packet Radio Service, GPRS), both located on the rear package shelf (or similar location in SUVs and minivans).

The head unit (Figure 3.3) assembly holds three cameras that will capture video images as illustrated in Figure 3.4. These image components include the forward roadway (upper left), driver face (upper right), and the pedals and instrument cluster interactions (lower left). The area behind and to the right of the subject vehicle (lower right) will be captured by a separate camera mounted on the rear of the cabin. Note that the face image is oriented at a 90-degree angle from the other images to maximize the efficiency of pixel allocation among all four images. That image and all images will be correctly oriented before any type of video analysis. Also, another camera, located in the head unit, will periodically take a still

*Table 3.5. Health-Related Assessments*

| Measurement Construct | Description | In Person/ Online | Estimated Time (min) |
|---|---|---|---|
| Sleep quality | Harvard sleep questionnaire | Online | 10 |
| Potentially driver-impairing medical conditions and medications | Custom comprehensive questionnaire based on material from the American Medical Association (AMA) and National Highway Traffic Safety Administration (NHTSA) | Online | 20 |

*Table 3.6. Psychological Assessments*

| Measurement Construct | Description | In Person/ Online | Estimated Time (min) |
|---|---|---|---|
| Risk-taking behavior | Cox Assessment of Risk Driving Scale (CARDS) combined with DeJoy Risk Perception Questionnaire | Online | 2 |
| Risk perception | CARDS items combined with DeJoy Risk Perception Questionnaire | Online | 2 |
| Attention deficit hyperactivity disorder (ADHD) | Barkley's Adult ADHD Quick Screen | Online | 2 |
| Driver style/behavior | Manchester Driver Behavior Questionnaire (DBQ) | Online | 5 |
| Thrill/adventure seeking | Sensation Seeking Scale (SSS) | Online | 5 |

*Table 3.7. Driving Knowledge/History Assessments*

| Measurement Construct | Description | In Person/Online | Estimated Time (min) |
|---|---|---|---|
| Driving knowledge | Custom questionnaire based on several state licensing practice tests | Online | 10 |
| Driving history | Custom questionnaire | Online | 5 |

*Figure 3.1. Optec 6500P all-in-one vision test apparatus.*

image of the cabin. These still images will be used to determine the presence of passengers, and they will be irretrievably blurred to protect the anonymity of unconsented passengers. This assembly will be mounted to the right and rear of a participant's rearview mirror. The forward camera will capture a color image, and the remaining cameras will be low-cost, small-form-factor monochromatic cameras that are sensitive to infrared (IR) illumination (for low-light video capture). The forward camera will collect color video to provide more comprehensive information about the forward or driver's view. Furthermore, for those researchers interested in traffic signal state detection, the color camera will provide the ability to conduct such research through a post hoc machine application or manual data analysis. Additionally, the alcohol sensor, incident button, illuminance sensor, inertial acceleration, and gyroscopic sensors will be incorporated into the head unit assembly.

In determining which sensors could be the most valuable, an extensive review of the research questions was conducted. The outcome of that prioritization led to the chosen sensors. The selected sensors reflect the best combination of desirable data that can be obtained within the cost and engineering constraints of the project and that meet the highest-priority research questions.

The main unit (Figure 3.5) is host to the computer functions of coordinating sensor nodes, communications (internally and via the cellular capability), and data storage on the hard drive (HD). It will have the capability of storing data on an onboard solid state HD for 4 to 6 months of typical driving. It will feature continuous asynchronous data collection. That is, each sensor or variable will be recorded at its native frequency and time-stamped with a real-time clock, without regard to the frequency at which any other data are being recorded. This is the most efficient approach to dealing with storage limitations. Vehicle network information is expected to be obtained from most of the vehicles for driver control interactions. For some makes and models, additional advanced technology variables are expected to be captured, such as the ABS, ESC, brake assist, traction control, etc. This feature requires OEM cooperation as described previously. Accelerometers will nominally collect X, Y, Z acceleration at 10 Hz, but these data will be continuously buffered at 500 Hz and saved at this higher rate under certain circumstances to provide higher-resolution acceleration information during a crash event. Yaw rate gyro will also be collected continuously at 10 Hz.

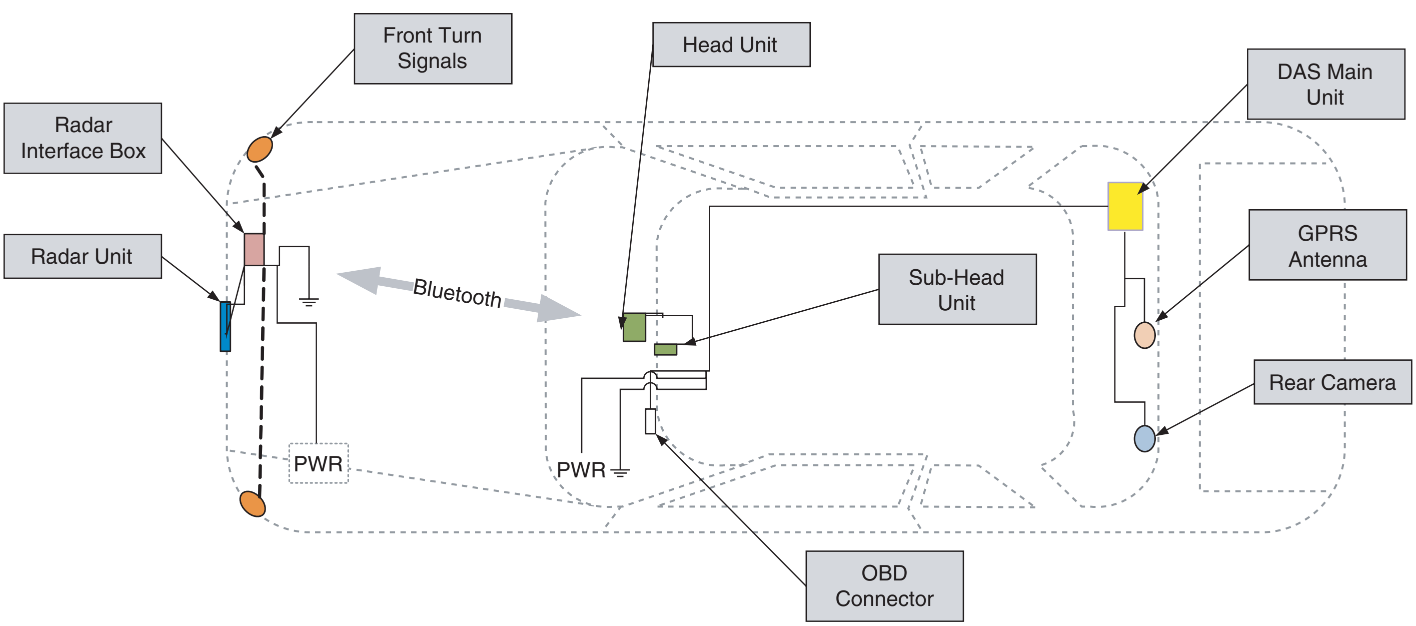


*Figure 3.2. DAS schematic.*

Figure 3.3. Head unit assembly prototype.

The GPRS cellular antennae will be mounted in the rear of the cabin. Wide field of view (FOV) forward radar capable of assessing oncoming traffic will be small, lightweight, and mounted to the front license plate holder (Figure 3.6). Radar data will be transmitted via Bluetooth wireless technology to the main unit and will capture data concerning the relative position and speed of multiple objects. Use of Bluetooth for this purpose alleviates the need of running cables through the vehicle's firewall to the main unit, thereby decreasing risk to the participant and saving valuable installation time and real costs associated with permanently altering a participant's vehicle. This is important because each S07 site must maintain an installation throughput rate of two vehicles per bay per day (i.e., where the largest sites have three installation bays and the smallest have a single installation bay).

DAS sensors and associated variables are listed in Table 3.8.


Figure 3.4. Composite snapshot of four continuous video camera views.

Figure 3.5. Main unit prototype.

## Machine Vision Applications

Several machine–vision-based applications will be incorporated into the DAS, including a lane tracker, head position monitor, and driver identification software. There are criteria for what to include as resident software on the DAS, as opposed to that which can be applied to stored data post hoc. The video data are being stored in a compressed format, because there are insufficient resources to store it in its native resolution. However, several of the machine–vision-based applications require the full resolution video to perform as expected, so these must operate on the native video coming directly from the cameras before the information is stored on the DAS hard drive. On the other hand, every such application running in real time consumes limited storage and pro-

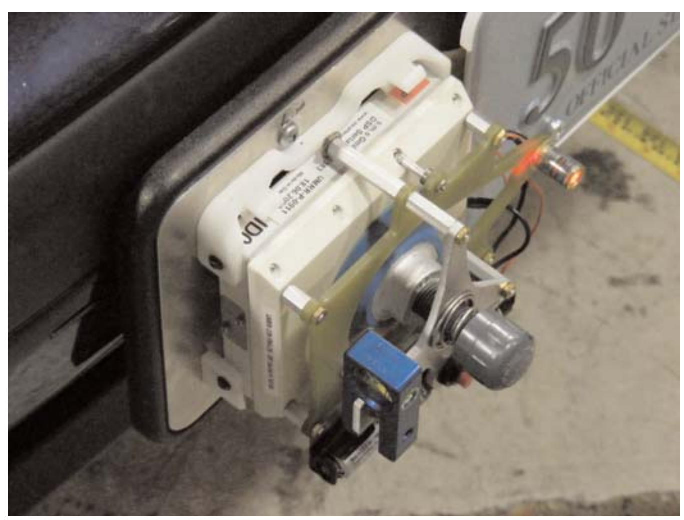

Figure 3.6. Forward radar assembly prototype.

*Table 3.8. Data Acquisition System Variables*

| Location | Sensor | Data to Be Collected |
|---|---|---|
| Head unit | Multiple cameras/video | Video images of the forward view, center stack view, rear and passenger side view, and driver face view, information for machine vision (MV) processes (including lane tracking and eyes forward data), as well as periodic, irrevocably blurred still photographs of the cabin interior to capture passenger presence |
| Main unit | Accelerometer data | In 3 axes:<br>Forward/reverse (x)<br>Right/left (y)<br>Down/up (z) |
| Main unit | Rate sensors | Yaw rate |
| Head unit | GPS | Latitude, longitude, elevation, time, velocity |
| Radar unit | Forward radar | Object ID, range, and range rate |
| Main unit | Cell phone module | Health checks, location notification, collision notification, and remote software upgrades |
| Head unit | Illuminance sensor | Ambient lighting levels |
| Head unit | Passive alcohol sensor | Presence of alcohol within the vehicle cabin |
| Head unit | Incident push button | In the event of an unusual or interesting traffic safety-related event, allows participant to open an audio recording channel for 30 seconds; also “flags” the data stream for ease of location during data analysis |
| Head unit | Audio | Available only in concert with the incident push button as noted above |
| | Turn signals (from vehicle network data or directly from the signals themselves) | Turn signal actuation, which distinguishes between left and right indicated turns |
| Main unit | Vehicle network data | Where available, the use of the accelerator, brakes, ABS, gear position, steering wheel angle, speed, horn, seat belt information, airbag deployment, and many other such variables |

cessing resources. Therefore the design must balance the need that certain applications have for real-time, onboard processing with the storage and processing limitations of the DAS.

The VTTI Road Scout (VRS) application is designed to track lane markings in real time on the DAS. Having this application on the DAS provides the advantage of operating on uncompressed video. The VRS has the ability to determine location of lane lines, horizontal curvature of the road, and angular offset of the vehicle within the lane. The VRS can determine if the lane lines are single, double, solid, or dashed. Testing has shown that the lane-tracking algorithm has a high degree of accuracy when lane markings are clearly present and visible; however, the lane tracking functionality is unable to gather sufficient roadway data in conditions where snow or other occlusions are present in the roadway. The data available from the MV lane tracker will be useful in answering many questions about road departure, because it is anticipated that many road departures and unintentional lane departures will be captured during the SHRP 2 NDS.

The VTTI Mask Head Tracker also operates on the DAS so that it can operate on uncompressed video (Figure 3.7). It is capable of identifying and distinguishing between a few general glance locations (e.g., forward roadway, mirrors, center stack) using software designed to find and determine characteristics of a person’s face in an image, and track those characteristics through subsequent images collected with the face view camera. The mask generates a three-dimensional representation of a person’s face using triangular surfaces to define the shape. The software and system operation will function in real time on the DAS using the raw video before compression.

A face recognition software solution is being sought that will permit researchers to automatically determine if a driver is a consented participant. Systems of this type are too processing-intensive to operate in real time, so this type of processing and analysis will be done on a post hoc basis. Driver identification will rely on a biometric software application to provide automated face recognition of drivers on the basis of their unique facial characteristics. Face recognition software would substan-


*Figure 3.7. VTTI Mask Head Tracker: Calibrated eyes forward (left); calibrated eyes on speedometer (right).*

tially reduce reliance on a human data reductionist to open each trip file to perform visual verification of the participant.

## Crash Investigation

Data from crash investigations provide an important complement and extension of the naturalistic driving data that will be collected as part of the NDS. While the data captured by the instrumented vehicle will be extensive, it is expected that they will not be a complete record of every detail of a crash, so the methodology recommended here is designed to collect the additional data needed in a way that is both feasible and effective. Data from the crash investigations should significantly enhance understanding of the actions, conditions, and behaviors that led to the crash. In addition, the comparison of crash investigation data, police accident reports (PARs), and DAS data will provide interesting insights into the reliability of these different sources of crash information.

Given that more than 1,000 crashes of all severity levels (as well as perhaps an order of magnitude more of number of near crashes) are expected during this study (see Table 3.9), not all crashes recorded during the study are expected to be investigated because of several constraints, including cost and, since many crashes are expected to be minor, the lack of crash site information. Crash investigations will only be carried out for crashes that meet certain criteria of interest. Such criteria may

***Table 3.9. Crashes and Near Crashes in the SHRP 2 Naturalistic Driving Study Estimated by Three Methods (Based on 1,950 DASs for 2 Years)***

| Crash/Incident Severity | Based on the *100-Car Study* | *100-Car Study,* Modified by Fatality Rate | Based on Crash Rates from GES[a] and NHTS[b] |
|---|---|---|---|
| Police reported | 624 | 363 | 230 |
| Nonreported, reportable | 975 | 566 | 360 |
| Nonpolice reported: low-*g* contact or tire strike | 1,599 | 929 | 590 |
| **Total crashes** | **3,198** | **1,859** | **1,180** |
| Near crashes | 29,679 | 17,247 | 10,952 |

[a] General Estimates System.
[b] National Household Travel Survey, 2001 (FHWA 2010).

be altered as the study proceeds, but they could include such factors as the following:

- Severity (e.g., airbag(s) deployed or injuries sustained);
- Crash type (e.g., intersection-related or lane change/ran-off-road);
- Driver age (e.g., teen driver or older driver);
- Land use (e.g., rural versus urban); and
- Advanced technology vehicle (e.g., equipped with crash warning/avoidance technology).

The rates based on the *100-Car Study* (Dingus et al. 2006) (see the first data column in Table 3.9) simply and directly extrapolate the crash rates observed in the *100-Car Study* to the size and scope of the current study. These estimates were then modified on the basis of the ratio of the relatively high fatality rates for Washington, D.C. (the site of the *100-Car Study*) and that of the United States overall (see the second data column in Table 3.9). Fatal crashes were selected, since General Estimates System (GES) crash estimates are not available for a particular locality, such as the Washington, D.C., and northern Virginia area. According to Fatality Analysis Reporting System data, Washington, D.C., had a fatality rate of 29.15 per 100,000 registered vehicles in 2003 compared with the latest available national rate of 16.05 per 100,000 registered vehicles in 2006 (NHTSA 2010; NHTSA 2007). This analysis assumes that the relative traffic fatality rates between Washington, D.C., and that of the United States overall can also be used to roughly approximate the relative crash and near-crash rates. The numbers in the third data column in Table 3.9 are derived from GES and 2001 National Household Travel Survey (FHWA 2010) estimates.

Near crashes have been operationally defined by VTTI researchers (Dingus et al. 2006) as any circumstance that requires a rapid, evasive maneuver by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. A rapid, evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle's capabilities. However, there is nothing prohibiting other researchers from defining the concept as they see fit.

In most cases, S06 personnel will be notified of a crash by the DAS via cellular communication channels. Included in the communications process will be key details about the crash, including a snippet of video covering the time just before and immediately after the crash. S06 personnel will then assess those data against the crash severity criteria bulleted above to determine if the crash event warrants further investigation.

If the crash is selected for further investigation, the S07 contractor will make an attempt to gather or retrieve the following information:

- DAS data.
- PAR (as available from participant or public records).
- Participant interview (using an instrument provided by the S06 contractor)—as soon as feasible postcrash to determine the following:
  - Predrive factors:
    - Recent sleep patterns;
    - Fatigue levels;
    - Emotional states; and
    - Stress levels.
  - Driving factors:
    - Weather;
    - Traffic; and
    - Obstructions.
  - Crash factors.
- Aerial view (as available, for example, from Google Earth or Google Street View).
- Vehicle photos (can be taken by data retrieval technicians)—front, back, sides.
- 37 Crashes categorization (Najm et al. 2007).

Considering data privacy and IRB issues, the only people to be interviewed in all crash investigations will be the consented drivers (and possibly other consented passengers) participating in the SHRP 2 NDS. Note that all the crash data noted above can be gathered remotely or with already-existing personnel.

A detailed list of data elements to be collected has been selected to be consistent with common data elements in the National Center for Statistics and Analysis national crash data, either by adopting the structure or by structuring individual data elements so that they can be mapped into the national data. Investigations will include documentation and data collection as related to precrash driver assessment, interviews, the crash site, and vehicle examination (as available).

The type of precrash and postcrash assessment information to be collected will be similar to the recent National Motor Vehicle Crash Causation Survey (NMVCCS) conducted by NHTSA (2008).

On the basis of the criteria of severity, type (e.g., run-off-road or intersection related), or some condition of uniqueness, the S06 contractor will notify the S07 contractor to dispatch an experienced investigator to visit the site within 48 hours (but not while police/emergency medical service [EMS] personnel are actively working the crash scene) to produce or retrieve the following additional information:

- A detailed description of the crash etiology;
- Crash site documentation and description (using software such as Easy Street Draw);
- Crash-site photographs showing the approach to point of impact for each involved vehicle and looking back from the point of impact for each vehicle; and
- Photos of physical evidence such as skid marks, gouges in the roadway or median, and impact points.

Crash-site investigation is expected to be conducted by experienced crash-site/scene investigators.

CHAPTER 4

# Quality Processes

## Quality Goals and Objectives

The overall goal of SHRP 2 NDS data quality processes is to work within budgetary and scheduling constraints to obtain the most complete and highest-quality data set possible. More detailed quality objectives include gathering data from a high-quality sample, gathering the full complement of data expected, gathering intended data efficiently, storing intended data securely, and maintaining data privacy and security.

The specific data quality processes have a threefold purpose, as follows:

1. To identify and preclude or reduce errors at their source;
2. To identify DAS system failures to initiate prompt maintenance actions during data collection; and
3. To mitigate the effects of data collection errors by assuring that the data entering the database are as high-quality as possible.

To implement the deceptively simple quality policy noted above, it will take the efforts of many individuals representing at least the following entities: S06 contractor's personnel, S07 contractors' personnel, DAS manufacturer(s), and the participants.

## Quality Activities: DAS Installation and Deinstallation Processes

It is important to ensure the required throughput (an average of two installations per day per bay) and quality in DAS installations. Training is the quality assurance process that will be used to ensure that DASs are installed in a timely and high-quality manner. All DAS installers will be trained at VTTI facilities by experienced personnel to make sure they have the ability to maintain such a schedule. DAS units must be installed without permanent destruction or defacing of the participant's vehicle in any way. Once the DAS is fully installed, its functionality will be verified with VTTI-provided custom software on the installer's laptop computer (also provided and imaged by VTTI), and the installation is not considered complete until this verification has been accomplished.

There are two other key aspects for ensuring a high-quality installation: the design of the DAS and the installation panel and lasers. The DAS is designed to minimize the chances of misinstallation because it can only be installed in the correct way. In addition, VTTI engineers have devised sophisticated yet easy-to-implement alignment protocols and laser hardware to ensure correct DAS installation (see Figure 4.1).

## Health Check

The DAS is equipped with cellular communications technology facilitating the automatic health check. This is a process whereby the DAS will automatically send a small batch of key data back to S06 servers. Automated algorithms will be applied to this batch of data to detect potential problems with DAS functionality.

The health checks will be generated and sent to the communications server at VTTI on a weekly basis starting on the day of installation of the DAS unit, unless some subsystem failure warrants immediate action. As a message is received by the VTTI server, it will be added to the database of stored health checks. A formula will be applied to each of the subsystems to determine whether the system should be considered functional (good) or whether it requires further attention. For all health checks classified as "bad" by the server process looking at communications, an issue will be immediately and automatically generated in the issue tracking software, Request Tracker (RT). The issue will then make its way through the triage process until it is fully resolved.

## Training

Training is a key responsibility of the S06 contractor to facilitate quality and consistency in S07 processes.

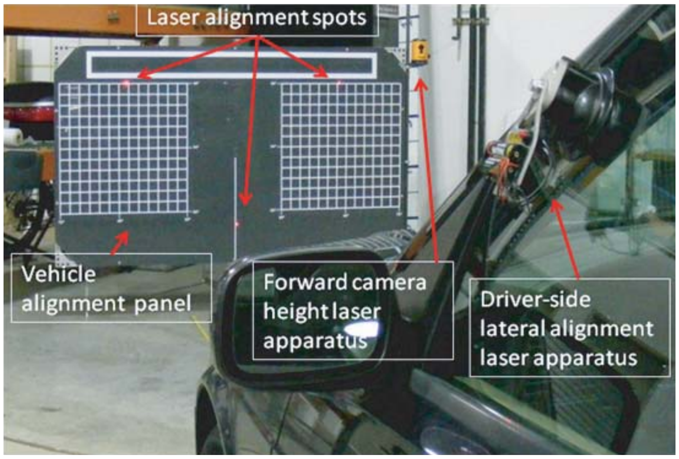


***Figure 4.1. Alignment panel and driver-side lateral alignment laser apparatus with left-right-center laser spots.***

## Training for Site Contractors

Training programs and materials will be administered to the S07 contractors. Training sessions must be conducted for DAS installers, driver assessment administrators, and others handling participant data. Also, all designated project personnel must be IRB certified. As the project is likely to experience attrition, a train-the-trainer approach will be employed. The initial set of trainees will travel to the S06 facility for a training session. DAS installers will undergo a 3-day training session, and assessment researchers will undergo a single day of training. Data handling will require a relatively small amount of training and will be incorporated within the DAS installation training sessions. Subsequently, SHRP 2 and S06 personnel will travel to the S07 sites to make certain that their facilities are suitable and procedures are consistently being applied. For each element of the training program, training criteria will be established for the individuals involved and the S07 site as a whole to determine whether training has been successful and what remediation is prescribed if the training has not met the criteria of success.

### *Participant Processing and Assessment Training*

Participant processing and assessment will be done consistently across S07 sites. Participant processing starts from the time the participant arrives at the S07 site and continues until the time that the instrumented vehicle is returned to him/her. It covers such requirements as the initial stages of the consent process (including reviewing and signing the consent form and other documents) and setting up the participant payment schedule described in Chapter 2. Driver assessment data must be collected using a consistent set of tools and protocols to avoid introducing additional variance into the data. Specific protocols regarding how to properly administer each of the driver assessments (and the entire experimental flow from one assessment to another, including suggested breaks and data-recording techniques) will be included in the training program. This training program will be primarily classroom based, with some hands-on use of the test equipment.

### *DAS Installation Training*

In terms of installation throughput, each S07 contractor must maintain installation throughput in accordance with the predetermined schedule (e.g., two vehicle instrumentations per work day per 150 DASs managed). DAS data quality relies on a correct installation. Also, installation must be accomplished in a manner that leaves the vehicle in a state where the participant feels comfortable with the aesthetics, and where it can eventually be returned to its preinstall condition. Training materials and a generalized training program instructing technicians how to install the DAS into a vehicle will be developed by the S06 contractor. The training program will include classroom, as well as hands-on, training in the garage. In addition, specific wiring diagrams and installation procedures will be prepared for some vehicles in the fleet. A wiki site will also be available for the sharing of lessons learned to facilitate ease and consistency of installations.

### *Data-Handling Training*

Although the data-handling process will be automated to the greatest degree feasible, there are specific steps that the S07 contractors must accomplish. S07 technicians will be trained on how to remove the hard drive from the DAS, replace the DAS with a new/refurbished hard drive, test the DAS functioning in the field, and use the hard drive bays to upload data to the staging server that will then automatically upload data to the S06 contractor servers. The S07 technicians will be trained on how to insert the hard drive into the reader, how to know when it is finished being processed, and how to determine when the hard drive can be put back into rotation for future use in the field. The training program will include classroom as well as hands-on training.

### *Institutional Review Board Training*

Each individual who interacts with a human participant, or may come into contact with personal information or items belonging to a participant, must provide evidence of successful completion of standard IRB training. This would apply to experimenters performing driver assessments, managers, and DAS installers as well as crash investigators who may have contact with participants. The S06 contractor will not provide this training but will provide recommended, low- or no-cost sources for this type of training. However, some specific instruction regarding IRB issues relating specifically to naturalistic data collection will be developed and given to both the S07 DAS technicians and the S07 participant processors.

## Site Contractor Inspections

During the study, with a heavy emphasis on the early stages for establishment of conformity and integrity, the S06 contractor will visit each S07 facility to assess the following areas:

- Observation of full DAS installation to verify installation/calibration integrity.
- Observation of participant assessments, where such would not affect or disrupt the assessments.
- Observation of S07 contractors' activities to
  - Note any S07 contractor-specific anomalies;
  - Determine protocol compliance;
  - Determine ability to continue operation per protocol requirements;
  - Identify issues or problem areas that require corrective actions;
  - Provide feedback to S07 contractors regarding performance; and
  - Seek input regarding any areas that the S07 contractors feel the S06 contractor could improve to enhance operational quality and efficiency.

The S06 contractor will periodically produce a report summarizing the results of these observations.

## Continual Project Assessment

Mission Control Software (MCS) will allow S06 and S07 managers to track study progress in a variety of ways (e.g., DAS inventory management, participant/vehicle information, and recruitment summaries). The workflow processes of participant intake and DAS tracking will provide the data reflected in the MCS application. More detailed information will be viewable for each individual participant or logically aggregated group (e.g., by S07 site). This additional information includes at least the following:

- Where each DAS is located (e.g., in vehicle A, or at site B, or en route to site C).
- How many of those participants with a DAS installed in their vehicles have completed assessments.
- Data gathered to date in terms of miles, hours, aggregate months installed, and so forth.
- Participant progress by age group—can be viewed per S07 site or overall. Shows the number of participants at each stage in the process.
- Data collection progress by age group—can be viewed per S07 site or overall.
- Data collection progress compared to plan.
- DAS kit inventory types by S07 site.
- HD status.
- Variable costs compared to planned expenditures.
- Participants installed compared with goals by
  - Age group;
  - Gender;
  - 12 or 24 months; and
  - Advanced vehicle.

## Intraproject Communication

### *Standing Meetings*

The management plan will entail, at a minimum, routinely scheduled standing 30- to 60-minute conference calls hosted by S06 and including S07 managers. The frequency of these meetings can be adjusted up or down as the project moves forward. Topics will include, at a minimum, communication from the S06 manager on overall study issues and status as well as the same from each S07 manager, and action item assignment and follow-up.

### *Critical Issues Committee*

A daily meeting time will be established where a critical issues committee can be quickly and reliably pulled together to address critical issues, which may include unexpected or abnormal situations that pose a risk to either an individual or group of participants, an individual or group of data-collection sites, or the study in its entirety. This team will be composed of principal members of the S06 contractor and TRB staff. In cases where it may be relevant, an individual or group of S07 site contractors may also be invited, although it is not expected that the site contractors will routinely be required for this meeting. Critical issues committee meetings will only be called to order when there is a critical issue that requires the committee's attention. However, the time slot should be kept available, as often as possible, in the event it is required. Also, only committee members who need to address the issue at hand will be required to attend. A process will be used to determine whether or not to convene a Critical Issues Committee meeting and whom to invite (Figure 4.2).

### *Issue-Tracking Software*

The issue-tracking software, Request Tracker (RT), will be in place for the NDS. This is necessary to successfully and consistently manage issues that arise during the course of the

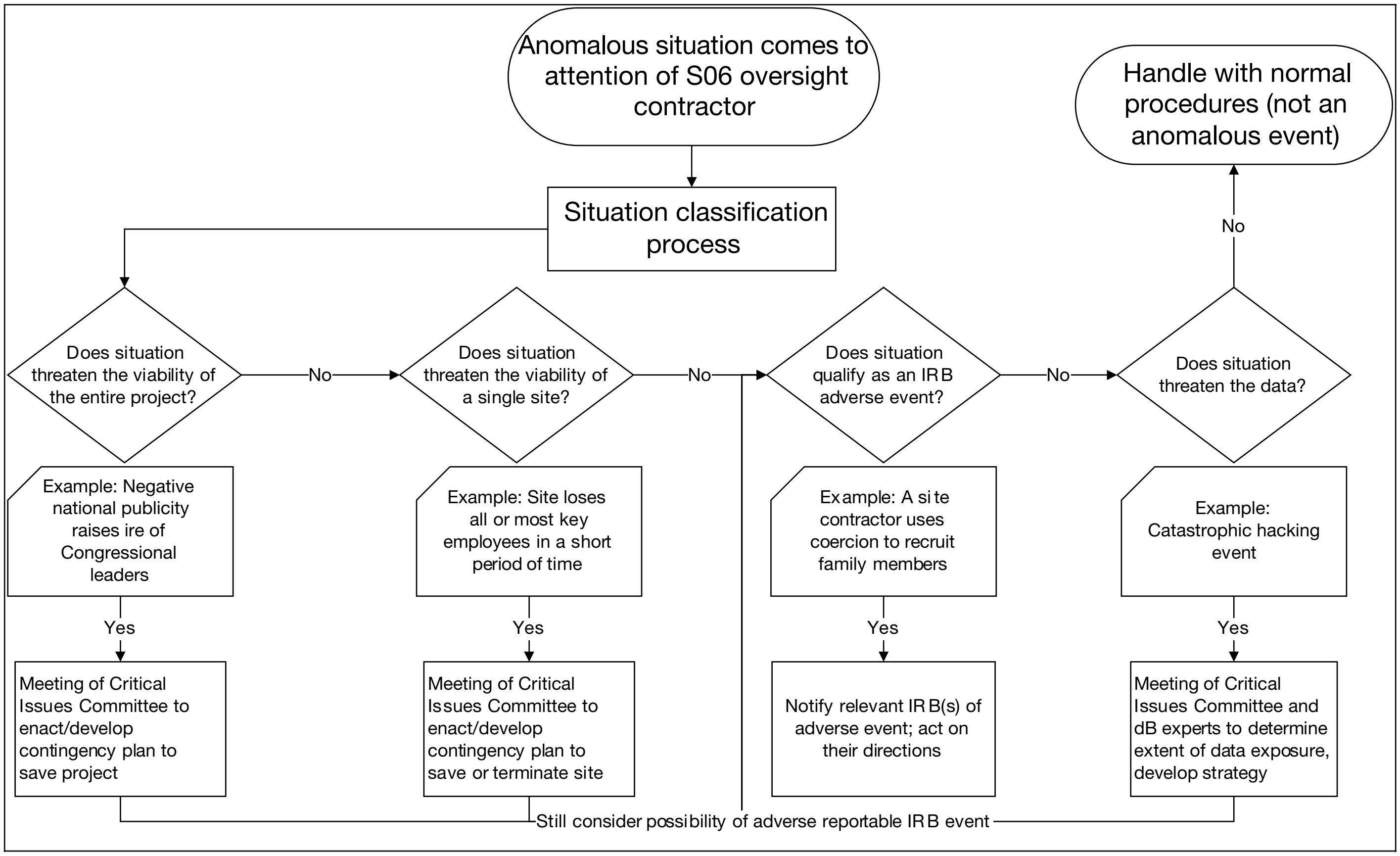


***Figure 4.2. Critical issues management.***

study. The key elements are the ability to track, assign, and report on issues that otherwise may be lost in formats not easily or systematically tracked, such as e-mail, phone calls, and messages, or Post-it notes. Issues can be created by any authorized user (e.g., TRB staff, S06 and S07 personnel), and every issue will be assigned to an appropriate individual to resolve in a timely manner. Issue status will be tracked throughout its life cycle from creation through resolution. Via this approach, reports can be generated on issues by any combination of elements, such as S07 site, issue category (e.g., DAS hardware problem), participant, issue status, and creation date. It is anticipated that the issue tracker will also yield valuable experience that can be used to populate a knowledge base or wiki that will be available to and beneficial to all project personnel.

CHAPTER 5

# Data Management

## Human Subjects Protection

Federal regulations and good research practice call for protection of persons who participate in research studies ("human subjects"). The Office for Human Research Protections (OHRP) in the U.S. Department of Health and Human Services (HHS) provides leadership in the protection of the rights, welfare, and well-being of subjects involved in research. OHRP does this by providing clarification and guidance, developing educational programs and materials, maintaining regulatory oversight, and providing advice on ethical and regulatory issues in biomedical and behavioral research. These protective policies are enforced at the local level by an organization's institutional review board (IRB), an entity required by federal regulations.

Of paramount concern in the design of the SHRP 2 NDS was the need to maintain close coordination with nearly all of the other project tasks, as virtually all tasks have some impact on the safety or privacy of the participants or their data. Key issues include protection of participant confidentiality, protection of unconsented passengers (e.g., no continuous audio recording can be employed since it may capture the conversations of unconsented passengers), informed consent (and assent/parental consent for minor participants), protection of potentially identifying information (e.g., face video and geospatial identifying data), and the continued protection of participant confidentiality once the data are stored in a database for post hoc analyses.

## Institutional Review Boards and Certificate of Confidentiality

Human subjects protection in the SHRP 2 NDS will be ensured by the review and approval of eight separate IRBs: those of the S06 contractor, the six S07 contractors, and the National Academy of Sciences (NAS). To prepare as well as possible for the human subjects protection review expected in the full-scale field study, the protocols for the S05 pilot studies underwent the full board review process at Virginia Tech (VT). This allowed a wider range of reviewers to see the complete protocol and raise human-participant concerns and issues prior to running the NDS. The combination of full board review at VT and full board review at NAS resulted in a very robust protocol that serves as a good starting point for the NDS.

Additionally, a Certificate of Confidentiality (CC) was secured from the National Institutes of Health (NIH) for the S05 pilot study. A CC helps researchers protect the privacy of subjects in biomedical, behavioral, clinical, or other research projects against compulsory legal demands (e.g., court orders and subpoenas) that seek the names or other identifying characteristics of a research subject. The CC covers the collection of sensitive research information for a defined time period (the term of the project); however, personally identifiable information obtained about subjects enrolled while the CC is in effect is protected in perpetuity. A CC will also be requested for the full-scale NDS. On the basis of the approval of the S05 CC, a timely approval for the SHRP 2 NDS CC is anticipated.

Upon NDS inception, one of the first sets of tasks relates to securing the IRB approvals from the S06 IRB, the NAS IRB, and each S07 IRB before proceeding with any and all aspects of the research involving human participants. Similarly, all S06 and S07 project personnel who will interact with participants or their data must certify that they have passed an approved IRB course or a course on protecting human participants.

Each individual site contractor (except any that have chosen to formally rely on the VT IRB) will have to receive approval from its own IRB on the basis of the research protocol and participant-consent documents approved by the VT IRB. It is likely that modifications to the standard set of documents to meet local needs will be reviewed, but these are not expected to fundamentally change. IRB-related submission materials were shared with the various stakeholder IRB personnel early in the process, including during a meeting of these stakeholders in Washington, D.C., in the summer of 2009. IRB approval will be sought for the call center recruitment separately from that for the main study activities.

## Collection Process from Vehicle to Server

The data collected during the NDS will include participant-identifying data and other sensitive personal information that must be protected. Consequently, every effort will be taken to protect all data from unauthorized access. The video data will be encrypted on the DAS and will remain encrypted until the data transfer process to the S06 server has been successfully completed. Once data quality processes have been applied, the video data will be reencrypted for storage. The data collection process is illustrated at a high level in Figure 5.1.

The hard drive on the DAS has a single copy of the data. As those data are transferred to the S07 server, they are replicated and stored on an array of HDs configured in a RAID (Redundant Array of Independent Disks). The RAID configurations on the S07 and S06 servers allow system administrators to completely restore a full copy of the data in the event of an HD failure on the server. Furthermore, once the data have been successfully transferred to the S06 site, an additional copy of the data will be stored for archival purposes.

Data will be encrypted onboard the DAS by way of Advanced Encryption Standard (128- or 256-bit AES) symmetric encryption. The key used for the encryption will be randomly generated for each trip, and that key is encrypted using the Rivest, Shamir, Adleman (RSA) public key of a public/private key pair. The encrypted key file is stored with the same naming convention alongside the encrypted data and video files. This scenario provides the security of having a private key that will not be onboard the system, while allowing the data and video to be encrypted with the much faster symmetric encryption.

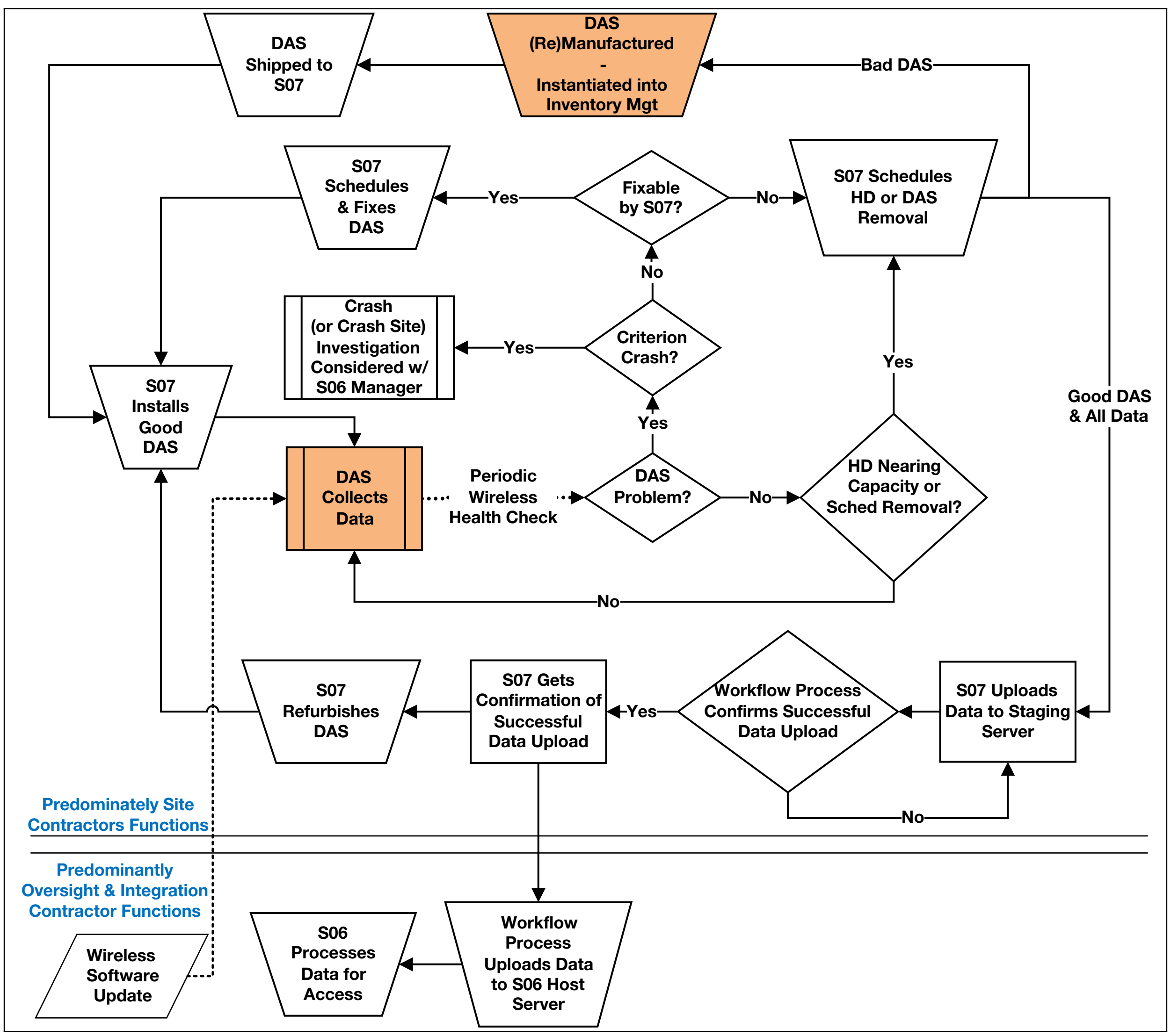


*Figure 5.1. Data collection process.*

# Data Processing

## Data Upload to Data Storage Server

When uploading the data from the S07 staging server to the S06 server, checksum analyses will be performed to ensure the integrity of the uploaded file. After uploading the data to the S06 database and decrypting, multiple quality checks will be done for each trip. These will be similar to but more sophisticated than those done during the routine health checks. Specifically, due to the amount of data and its contiguous nature (i.e., each trip file should begin at or near the same GPS coordinates where the previous trip file ended), more sophisticated comparisons between variables can be made to isolate potential problems within a trip. Analyses will also be conducted to compare trips to ensure that data are not being lost. For example, is the GPS location at the beginning of a trip the same location as the end of the previous trip? When a problem has been identified by the data-quality algorithms, any questionable data will be marked as such. At a minimum, the annotation will include a start sync, end sync, and metadata describing the test the variable failed. As resources are available, fixes may be applied to the data where such is possible (e.g., where it can be determined that a particular sensor was generating data that were off by a known constant value). S06 quality personnel will review the problems to try to determine the root cause (i.e., on the DAS or otherwise). The S06 contractor may need to work with the individual S07 contractor to isolate the problem and determine the best course of corrective action. Quality personnel will also conduct random spot checks by remotely requesting data snippets.

Note that any additional processing required to get the data into a format to answer specific research questions is outside the scope of the current S06 project. However, it is believed that providing access to these data to researchers early on is paramount to the success of this project because it lets stakeholders at all levels begin to see results and the value of the project early on, without waiting for all data to be collected some 28 months later.

## Data Acquisition System Data Processing

The purpose of the data processing is to get the highest-quality data as is feasible in the database and in a form usable by researchers. Several processes will be performed once the data arrive at the S06 server.

### *Backups*

The data will be housed at the VT Data Center. RAID 6 protection is also employed at this facility to guard against loss of data due to server failure. Archival backups of the data will be stored at a different physical location.

### *Trip Summary Data*

Summary data will be extracted for each trip. These data include mileage driven, duration, start time, average speed, maximum speed, number of stationary epochs, maximum deceleration, driver identification (where possible), etc. This summary will help with quality processes, and it will provide a useful first look at the data for researchers.

### *Data Standardization*

Data will be standardized into common formats. Because data are being collected on different vehicle makes, models, and countries of origin, it is possible that the DAS may collect data from a single sensor with different units, scales, axes, sample rates, or coding. It will be important to transform the data into standard units to assist researchers when they attempt to analyze the data across vehicles. This is also important if any algorithms are to be applied across the entire fleet consistently. The raw data will also be stored in the event that any researcher ever wants to review or analyze them. Also, some of the vehicle models (e.g., those equipped with LDW systems) may generate higher-resolution data (i.e., in time or the measured dimension) than others. Using steering information as an example, this higher-frequency data would be of great interest to researchers looking at steering reversals to investigate workload, drowsiness, steering entropy, or the performance of the onboard LDW.

# Expected Data Magnitude

Data that are staged on S07 servers and then transferred to the central S06 data server could often exceed 100Mbps. This requires the use of high-performance research-caliber networks, such as Internet2 or National Lambda Rail. With almost 2,000 DAS units simultaneously collecting video and other sensor data for 2 years each, as well as a projected data life span of up to 30 years, the magnitude of data storage and criticality of adequate infrastructure cannot be overstated. Specifically, the NDS database will house information from several sources, including video and sensor trip data, crash data, health check (i.e., system) data, management information (i.e., inventory data and participant enrollment data), participant demographic and assessment data, vehicle inventory data, and analysis data (i.e., aggregated or reduced), as well as other external sources such as PARs, GIS, weather data, maps, and roadway information. In total, it is anticipated that 2 years of data collection will create a volume of approximately 1 petabyte of data comprising approximately 60–80 million miles and approximately 1.5 to 1.7 million hours of driving data (i.e., a data volume that would require approximately the storage capacity of one million 1-gigabyte USB flash drives). To characterize this volume of data in a different context, it would take approximately 70 million copies of the King James Version of the Bible to fill 1 petabyte of storage capacity.

CHAPTER 6

# Data Provisioning

The primary data-sharing concern will be to ensure that the consent form language clearly states the conditions under which the data may be shared after the study is complete. The language should spell out whether the data to be shared will be identifiable and the conditions for each type of data sharing. Data-sharing agreements may be needed for identifiable data but will likely not be needed for de-identified data that are publicly available. However, some form of IRB approval may still be required for all data access. The terms of this data access are still being determined as of this report date.

## Data Dictionaries

To assist all researchers, data dictionaries are being developed that will identify each data item (or data stream) collected during the study. These data dictionaries include operational definitions as well as database references. These may be especially helpful for those not involved in the original data collection. Additionally, it is anticipated that custom data dictionaries will be necessary to support data located through research-specific queries. Specifically, if new operational definitions are used to develop new variables, the definition and the approach must be documented both for the original researchers looking at these data and for future researchers.

## Role-Based Access

The ability to provide scalable user access control to the combined data sets that will be collected and generated in the course of the study is governed by a role-based security protocol and is dependent on the researcher requesting the data successfully obtaining IRB approval. Role-based security protocols authorize access to data elements based upon explicit security roles that users may be granted. Without any assigned roles, the default authorization granted to users who successfully "authenticate" (successfully log in) is to deny access to any secured data.

Role-based security has successfully demonstrated its ability to cover a range of access requirements. Different roles can be defined to have different security access as appropriate. Security can be defined in terms of row-level access (e.g., where a user has access to all the data within a given region), column-level access (e.g., where a user may have access to all the data except GPS location information), or a combination of the two (e.g., cell-level access where a user may have access to GPS location data within a specified region). It should be noted that this security protocol provides protection to both database and file-based (e.g., video) data, as illustrated in Figure 6.1.

## Required Software

It is generally expected that most external access will be provided through a website or web service. However, provisions will be in place to allow qualified users to obtain their access via direct file server or database server access using the role-based access described previously. For example, in these cases, remote users may be able to obtain access through the use of typical Virtual Private Network (VPN) technologies. Using a role-based approach as defined accommodates and supports the use of a website/service to provide general data access. As such, commercially available data analysis tools (e.g., SAS and Matlab) can be used by researchers interested in conducting analysis on the data. A shareware-based Community Viewer will be provided also, but no specialized proprietary software will be required.

## Coordination and Linking with Roadway Information

SHRP 2 Safety Project S04A, Roadway Information Database Development and Technical Coordination and Quality Assurance of the Mobile Data Collection Project, and Safety Project S04B, Mobile Data Collection, will provide valuable roadway-

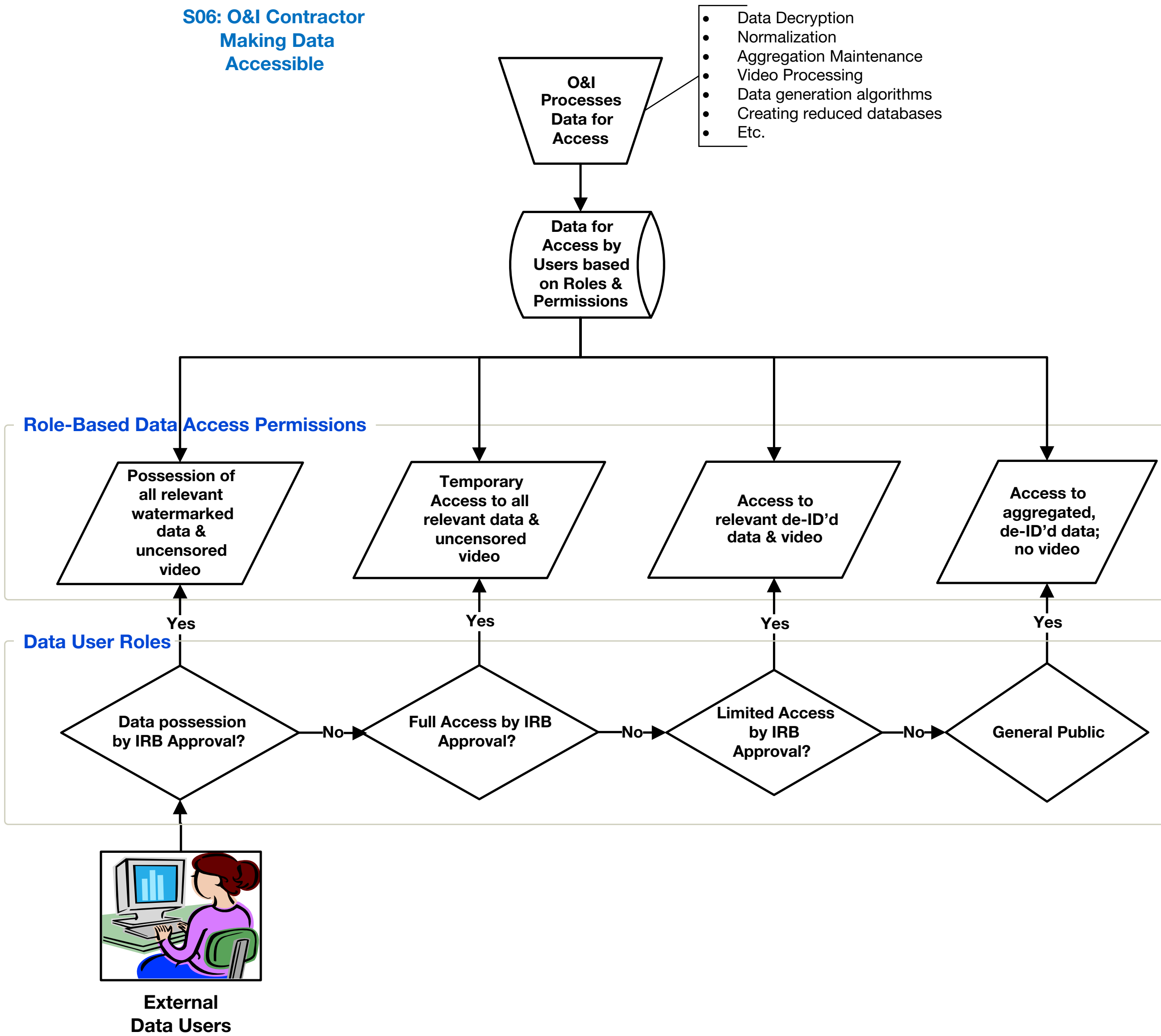


***Figure 6.1. Role-based data access diagram.***

based data describing in detail sections of roadway traveled by the instrumented vehicles included in the NDS. Roadway data will be extracted from existing inventory databases maintained by state and local transportation departments and supplemented with other inventory data collected by highly instrumented on-road data collection vans. Successful integration of these two data sets will support many research questions that involve the interaction between the driver, the vehicle, and the roadway and appurtenances. Recognizing that these two efforts will be executed somewhat independently but simultaneously, actions have been taken in advance to ensure that integration of these data sets will be possible and feasible so as to support the desired research.

Roadway inventory data in currently existing files are generally specified using a zero point; subsequent stations are measured in length along the traveled roadway path from this zero point. Vehicle data are generally located geographically using latitude, longitude, and sometimes elevation, collected throughout a vehicle trip. Roadway inventory data collected by instrumented vans can be located in either fashion—either as a distance from a zero point or geographically. Integration of these data will require transformation of one or both types of data to determine at what position in the roadway data the vehicle is during travel on a measured roadway. Communication between S03 contractors and NDS contractors has been conducted to determine what will be required to make these transformations and what adjustments may be required during data collection to support these transformations. A GPS measure would be required at each zero point but, though valuable, it may not be required at subsequent stations. Vehicle

GPS data can be used to identify when the vehicle passes a zero point, and then vehicle speed and time data could be integrated to estimate when the vehicle passes subsequent stations. The degree of accuracy of candidate approaches will be evaluated as they are identified.

Collection of roadway data will not be feasible on all roadways on which NDS participant vehicles travel. To prioritize the measurement of roadway segments, NDS data will be processed regularly to identify roadway segments that are traveled frequently by more than one participant in the NDS effort. In this way, segments that can provide replications to support research will be identified and provided to the S04 contractors to guide collection. On the basis of investigation of previous naturalistic data collections (Dingus et al. 2006), these high-frequency locations are generally coincident with high-volume corridors within a study area and places of ingress and egress to these corridors and also highways. The accumulating actual naturalistic data will be used to narrow these general rules to determine the specific segments traveled frequently by several participants.

CHAPTER 7

# Conclusions

This final report provides a summary of the key areas of the plan that has been developed in support of the SHRP 2 NDS that focuses on Safety Project S06 (Technical Coordination and Independent Quality Assurance for Field Study) and Safety Project S07 (In-Vehicle Driving Behavior Field Study).

Preventing or reducing the severity of highway crashes by understanding driver behavior is important to the nation's economic system and quality of life. The SHRP 2 NDS promises to yield high payoffs in the safety arena, both during the life of the current project and for many years after it is completed, as it facilitates the answering of many of the key research questions identified in the early phases of the project. By understanding how risk factors influence safety on our roadways, innovative countermeasures can be employed to improve our ability to design and build safer roadways and vehicles, navigate environmental conditions, and teach future driving generations how to use safer driving practices. The SHRP 2 NDS will contribute to these ends by generating a wealth of naturalistic driving data—on the order of approximately 60 to 80 million miles and approximately 1.5 to 1.7 million hours of driving data originating from six different regions around the continental United States. During the course of the study, it is expected that detailed information about approximately 1,000 crashes of all severity levels, many of them serious, will be captured. In addition, it is also expected that an order of magnitude more near crashes will be detected and recorded. The naturalistic method is the only way to consistently record the details of many crashes. It is also the only method that allows identification and detailed examination of near crashes. Finally, it provides the exposure data by which metrics of relative risk can be estimated.

This undertaking is ambitious; however, this S05 project has developed a comprehensive study design addressing the data collection system as well as the project management and quality plans to successfully address the challenges of this important national safety program.

# References


Boyle, L. N., J. D. Lee, D. M. Neyens, D. V. McGhee, S. Hallmark, N. J. Ward. Forthcoming. Appendix A: Global Research Question Priorities. *SHRP 2 Report S2-S02-RW-1: Integration of Analysis Methods and Development of Analysis Plan: Phase 1 Report.* Transportation Research Board of the National Academies, Washington, D.C.

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, and J. Jermeland. 2006. *The 100-Car Naturalistic Driving Study, Phase II—Results of the 100-Car Field Experiment.* Report DTNH22-00-C-07007. National Highway Traffic Safety Administration, U.S. Department of Transportation.

Federal Highway Administration, U.S. Department of Transportation (FHWA). 2010. Online Analysis Tool. National Household Travel Survey, 2001. http://nhts.ornl.gov/tools.shtml. Accessed Nov. 11, 2010.

Najm, W. G., J. D. Smith, and M. Yanagisawa. 2007. *Pre-Crash Scenario Typology for Crash Avoidance Research.* Report DOT-VNTSC-NHTSA-06-02. National Highway Traffic Safety Administration, U.S. Department of Transportation.

National Highway Traffic Safety Administration, U.S. Department of Transportation (NHTSA). 2007. *Traffic Safety Facts 2007: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System.* Report DOT HS 811 002. www-nrd.nhtsa.dot.gov/Pubs/811002.pdf. Accessed Nov. 17, 2010.

———. 2008. *National Motor Vehicle Crash Causation Survey: Report to Congress.* Report DOT HS 811 059. www-nrd.nhtsa.dot.gov/Pubs/811059.pdf. Accessed Nov. 17, 2010.

———. 2010. Fatality Analysis Reporting System Encyclopedia. www.fars.nhtsa.dot.gov/Main/index.aspx. Accessed Nov. 11, 2010.

Royall, D. R., E. C. Lauterbach, J. L. Cummings, A. Reeve, T. A. Rummans, D. I. Kaufer, W. C. LaFrance, Jr., C. E. Coffey. 2002. Executive Control Function: A Review of Its Promise and Challenges for Clinical Research. *Journal of Neuropsychiatry and Clinical Neurosciences,* Vol. 14, No. 4, pp. 377–405.

Treat, J. R., N. S. Tumbas, S. T. McDonald, D. Shinar, R. D. Hume, R. E. Mayer, R. L. Stansifer, N. J. Castellan. 1979. *Tri-Level Study of the Causes of Traffic Accidents Final Report.* Report DOT HS-034-3-535, National Highway Traffic Safety Administration, U.S. Department of Transportation.

Ward's Automotive Group. 2010. U.S. Total Vehicle Sales Market Share by Company, 1961–2009. http://wardsauto.com/keydata/historical/UsaSa28summary. Accessed Nov. 9, 2010.

## TRB OVERSIGHT COMMITTEE FOR THE STRATEGIC HIGHWAY RESEARCH PROGRAM 2*

Chair: **Kirk T. Steudle,** *Director, Michigan Department of Transportation*

### MEMBERS

**H. Norman Abramson,** *Executive Vice President (Retired), Southwest Research Institute*
**Anne P. Canby,** *President, Surface Transportation Policy Partnership*
**Alan C. Clark,** *MPO Director, Houston-Galveston Area Council*
**Frank L. Danchetz,** *Vice President, ARCADIS G&M, Inc.*
**Dan Flowers,** *Director, Arkansas State Highway and Transportation Department*
**Stanley Gee,** *Executive Deputy Commissioner, New York State Department of Transportation*
**Michael P. Lewis,** *Director, Rhode Island Department of Transportation*
**Susan Martinovich,** *Director, Nevada Department of Transportation*
**John R. Njord,** *Executive Director, Utah Department of Transportation*
**Charles F. Potts,** *Chief Executive Officer, Heritage Construction and Materials*
**Gerald Ross,** *Chief Engineer, Georgia Department of Transportation*
**George E. Schoener,** *Executive Director, I-95 Corridor Coalition*
**Kumares C. Sinha,** *Olson Distinguished Professor of Civil Engineering, Purdue University*

### EX OFFICIO MEMBERS

**Victor M. Mendez,** *Administrator, Federal Highway Administration*
**Ron Medford,** *Acting Administrator, National Highway Transportation Safety Administration*
**John C. Horsley,** *Executive Director, American Association of State Highway and Transportation Officials*

### LIAISONS

**Tony Kane,** *Director, Engineering and Technical Services, American Association of State Highway and Transportation Officials*
**Jeffrey F. Paniati,** *Executive Director, Federal Highway Administration*
**John Pearson,** *Program Director, Council of Deputy Ministers Responsible for Transportation and Highway Safety, Canada*
**Margie Sheriff,** *Director, SHRP 2 Implementation Team, Office of Corporate Research, Technology, and Innovation Management, Federal Highway Administration*
**Michael F. Trentacoste,** *Associate Administrator, Research, Development, and Technology, Federal Highway Administration*

## SAFETY TECHNICAL COORDINATING COMMITTEE*

Chair: **Forrest M. Council,** *Senior Research Scientist, Highway Safety Research Center, University of North Carolina*

### MEMBERS

**David L. Banks,** *Professor, Practice of Statistics, Department of Statistical Science, Duke University*
**James A. Bonneson,** *Senior Research Engineer, Texas Transportation Institute, Texas A&M University*
**Richard K. Deering,** *President, RK Deering & Associates, Inc.*
**Leanna Depue,** *Director, Highway Safety Division, Missouri Department of Transportation*
**Joanne L. Harbluk,** *Human Factors Specialist, Transport Canada*
**James H. Hedlund,** *Principal, Highway Safety North*
**Bruce A. Ibarguen,** *Engineer of Traffic, Maine Department of Transportation*
**Mavis Johnson,** *President, Canadian Traffic Safety Institute*
**Lawrence H. Orcutt,** *Chief, Division of Research and Innovation, California Department of Transportation*
**J. Scott Osberg,** *Principal, Social Science Ink*
**Robert W. Schomber,** *Regional Manager, Florida Power & Light Company*
**David Shinar,** *Professor, Department of Industrial Engineering and Management, Ben Gurion University of the Negev*
**Alison Smiley,** *President, Human Factors North, Inc.*
**Thomas M. Welch,** *State Transportation Safety Engineer, Office of Traffic and Safety, Iowa Department of Transportation*
**Terecia W. Wilson,** *Strategic Highway Safety Plan Program Manager, South Carolina Department of Transportation*

### AASHTO LIAISONS

**Kelly Hardy,** *Safety Program Manager, American Association of State Highway and Transportation Officials*
**Jim McDonnell,** *Program Director for Engineering, American Association of State Highway and Transportation Officials*

### FHWA LIAISONS

**Monique Evans,** *Director, Office of Safety Research and Development, Federal Highway Administration*
**Michael Griffith,** *Director, Office of Safety Integration, Federal Highway Administration*
**Margie Sheriff,** *Director, SHRP 2 Implementation Team, Office of Corporate Research, Federal Highway Administration*

### AUTO INDUSTRY LIAISONS

**Michael Cammisa,** *Director, Safety, Association of International Automobile Manufactures, Inc.*
**Scott Schmidt,** *Director, Safety and Regulatory Affairs, Alliance of Automobile Manufacturers*

### CANADA LIAISON

**Kent Speiran,** *Manager, Road Safety, Nova Scotia Department of Transportation and Infrastructure Renewal*

### EUROPEAN SAFETY LIAISON

**Fred Wegman,** *Managing Director, SWOV Institute for Road Safety Research, Netherlands*

### FMCSA LIAISON

**Martin Walker,** *Chief, Research Division, Federal Motor Carrier Safety Administration*

### NHTSA LIAISONS

**Richard Compton,** *Director, Office of Behavioral Safety Research, National Highway Traffic Safety Administration*
**Tim Johnson,** *Director, Office of Human-Vehicle Performance Research, National Highway Traffic Safety Administration*

*Membership as of November 2010.

## Related SHRP 2 Research

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

Mobile Data Collection (S04B)

Technical Coordination and Quality Control (S06)

In-Vehicle Driving Behavior Field Study (S07)

Analysis of the SHRP 2 Naturalistic Driving Study Data (S08)

WWW.TRB.ORG/SHRP2

THE NATIONAL ACADEMIES™

*Advisers to the Nation on Science, Engineering, and Medicine*

The nation turns to the National Academies—National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council—for independent, objective advice on issues that affect people's lives worldwide.

www.national-academies.org

ISBN 978-0-309-12895-7
90000
9 780309 128957