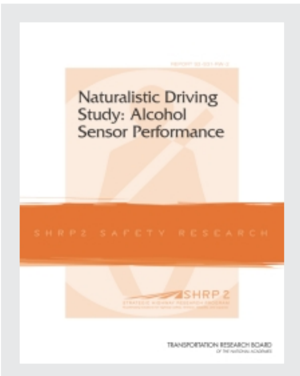


Naturalistic Driving Study: Alcohol Sensor Performance

DETAILS

35 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-43284-9 | DOI 10.17226/22230

AUTHORS

Ryan C. Smith, Zachary Doerzaph, and Jon Hankey; Strategic Highway Research Program Safety Focus Area; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine

The Second

STRATEGIC HIGHWAY RESEARCH PROGRAM

SHRP 2 REPORT S2-S31-RW-2

Naturalistic Driving Study: Alcohol Sensor Performance

RYAN C. SMITH, ZACHARY DOERZAPH, AND JON HANKEY
Virginia Tech Transportation Institute
Blacksburg, Virginia

TRANSPORTATION RESEARCH BOARD
WASHINGTON, D.C.
2015
www.TRB.org

Subject Areas

Data and Information Technology
Highways
Safety and Human Factors
Vehicles and Equipment

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-S31-RW-2

ISBN: 978-0-309-31492-3

Copyright Information

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

Notice

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

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

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

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

SHRP 2 Reports

Available by subscription and through the TRB online bookstore:

www.mytrb.org/store

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 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. C. D. (Dan) Mote, Jr., 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. Victor J. Dzau is president of the Institute of Medicine.

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

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

www.national-academies.org

SHRP 2 STAFF

ACKNOWLEDGMENTS

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

This research was made possible through the various and vast contributions of many individuals at the Virginia Tech Transportation Institute (VTTI). In particular, Jon Hankey and Zachary Doerzaph provided expert review throughout the research process. They helped with research design, execution of the project, and evaluation of the final results. The six-member VTTI data reduction team performed a superb job of carefully evaluating video evidence from SHRP 2 trips to make determinations of in-vehicle alcohol intoxication and the visual identification of substances that could trigger the alcohol sensor. This effort was led by Julie McClafferty and Timothy Walker.

Research Associates from the Center for Advanced Automotive Research at VTTI also made significant contributions. Tom Gorman helped with the research design and played a critical role in Phase 1 of the overall research effort. Miao Song helped code the alcohol-detection algorithm that allowed for validation against SHRP 2 trip files. The Center for Technology Development at VTTI provided the engineering background for the in-vehicle experimentation, including the design and development of the mechanical breather, the Boozooka.

FOREWORD

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

This report analyzes the performance of a passive alcohol sensor included in the head unit of the data acquisition system used in the SHRP 2 Naturalistic Driving Study (NDS). Driver impairment is a critical issue in traffic safety, and the ability to identify alcohol-impaired drivers would be valuable for users of the NDS data. The sensor responds to the presence of alcohol in the cabin air. A positive sensor reading can come from many sources: alcohol from the breath of a driver or other occupant, an open container of an alcoholic beverage, aftershave lotion or perfume, windshield wiper fluid, and even some fast food. On the other hand, open windows may dissipate alcohol from an impaired driver's breath before it reaches the sensor. Thus, the sensor can produce a positive reading when the driver is sober and can produce a negative reading for an alcohol-positive driver. The objective of this report is to evaluate the sensor performance under several scenarios with known driver alcohol levels and to investigate the feasibility of developing an algorithm to identify potentially alcohol-impaired drivers based on the sensor output.

The SHRP 2 NDS is the first large-scale study focused on collision prevention (as opposed to injury prevention once a collision occurs) since the Indiana Tri-Level Study (*Tri-Level Study of the Causes of Traffic Accidents: Final Report,* DOT HS-805 085, U.S. Department of Transportation, May 1979). Vehicle use was recorded continuously during the SHRP 2 NDS. Information on vehicle travel, or exposure, can be extracted at the same level of detail as for safety-related events, such as crashes and near crashes. Hence, the SHRP 2 NDS is the first large-scale study to support detailed estimates of collision risk. Moreover, crashes are a leading cause of nonrecurring congestion, so collision prevention has added benefits in terms of reduced delay, fuel consumption, and emissions. The NDS provides objective information on the role of driver behavior and performance in traffic collisions and on the interrelationship of the driver with vehicle, roadway, and environmental factors.

The SHRP 2 Safety research program was carried out under the guidance of the Safety Technical Coordinating Committee (TCC), which was composed of volunteer experts. The Safety TCC developed and approved all project descriptions and budgets and met semiannually to review progress and approve any program modifications. The Oversight Committee approved all budget allocations and contract awards. Assistance was provided by expert task groups, which developed requests for proposals, evaluated proposals and recommended contractors, and provided expert guidance on many issues, such as data access policies and procedures. The decisions and recommendations of the governing committees were implemented by the SHRP 2 staff as they carried out day-to-day management of the research projects.

CONTENTS

Executive Summary

The second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS) offers a unique glimpse into alcohol-impaired driving through the inclusion of an alcohol sensor within the standard instrumentation package. This research effort developed and evaluated an alcohol-detection algorithm using the sensor through two approaches: an experimental in-vehicle testing regimen and an examination of a subset of SHRP 2 NDS trips.

For the experimental in-vehicle testing, a sedan was instrumented with two SHRP 2 alcohol sensors. During 50 15-minute trials, controlled levels of alcohol were introduced into the cabin and alcohol sensor readings were recorded. The sensitivity of the algorithm at detecting alcohol presence during these trips was 96.6% and the specificity was 100%.

A data set was created using SHRP 2 trips that were visually reduced via manual video coding for the presence of imbibed and unimbibed alcohol ($n = 659$). It provided insight into many unimbibed alcohol sources that can lead to misclassification as imbibed alcohol. Confusion matrices were conducted on the data set and a subset comprising trips with moderately impaired drivers and "normal" driving trips. The results indicated a sensitivity ranging from 92.2% to 93.7% and a specificity ranging from 36.9% to 100%. The large variance in specificity is due to one of the data sets intentionally oversampling "positive" cases.

The results indicate that an alcohol-detection algorithm can be a useful tool for identifying potential alcohol-impaired trips in the SHRP 2 database. However, trained data reductionists should also be used to make final impairment determinations due to the large number of unimbibed alcohol sources that can affect the sensor.

CHAPTER 1

Background

Overview

Alcohol-impaired driving continues to be a significant public health concern. In 2011, 9,878 people were killed as a result of drunk driving in the United States, representing 31% of all traffic fatalities (*1*). Increased traffic crash risk results from drivers who experience pronounced and systematic physiological impairment while under the influence of alcohol (*2, 3*). Additionally, the intoxication level of passengers is linked to heightened crash risk through increased distraction of drivers (*3, 4*). Therefore, the presence of alcohol within a vehicle is a measure of significant interest when investigating crash causation.

The second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS) has collected 50 million vehicle miles of naturalistic driving data on more than 3,000 drivers. The instrumentation of the vehicles used in this study included prototype alcohol-sensing technology that continuously detected the amount of ethanol in the vehicle's interior. However, this sensor outputs a raw value that is difficult for those unfamiliar with the technology to interpret. In addition, the presence of uncharacterized noise affects the sensor's ability to detect the presence of alcohol.

Several previous efforts explored the efficacy of the SHRP 2 alcohol sensor to detect the presence of imbibed alcohol. These included an experimental pilot study and Phase 1 (proof of concept) of the current project. The current effort, Phase 2, was built on previous attempts to develop an alcohol-detection algorithm that could be used to flag potential alcohol-involved cases in the SHRP 2 data set.

Background

Detailed descriptions of the SHRP 2 database, alcohol sensor, and previous research efforts follow.

SHRP 2 Database

The SHRP 2 database provided the primary data set used for early algorithm efforts and is also the primary database for this project. The SHRP 2 NDS is the largest naturalistic driving study of its kind. As noted, it comprises 50 million miles of data, 5 million trip files, and more than 3,000 primary drivers. The data set contains video and kinematic information from each of the specially instrumented SHRP 2 vehicles. The ability to detect the presence of alcohol using SHRP 2 sensors will help answer a multitude of research questions and will aid in evaluating the impact of alcohol on driver errors and crash likelihood.

Hardware Configuration

The alcohol sensor installed in the SHRP 2 vehicles was a model HS130D from Sencera Co., Ltd. (Figure 1.1). This sensor is a tin dioxide semiconductor gas sensor designed to quickly detect alcohol vapors at high relative humidity.

The alcohol sensor was manufactured to detect the presence of ethanol in the air. This could include alcohol vapors released into the cabin of a vehicle from the natural breathing of an individual who had consumed alcohol or a variety of unimbibed sources (e.g., hand sanitizer, perfume, mouthwash). The sensor returned a raw value in millivolts (mV). For the SHRP 2 project, the alcohol sensor was installed on the underside of the head unit, which was mounted near the vehicle's rearview mirror mount. This central location meant that it was equally able to detect alcohol vapors from the breathing of both the driver and the passenger.

Pilot Investigation

A pilot study of the alcohol sensor was conducted to examine the accuracy of the sensor under a variety of conditions. This

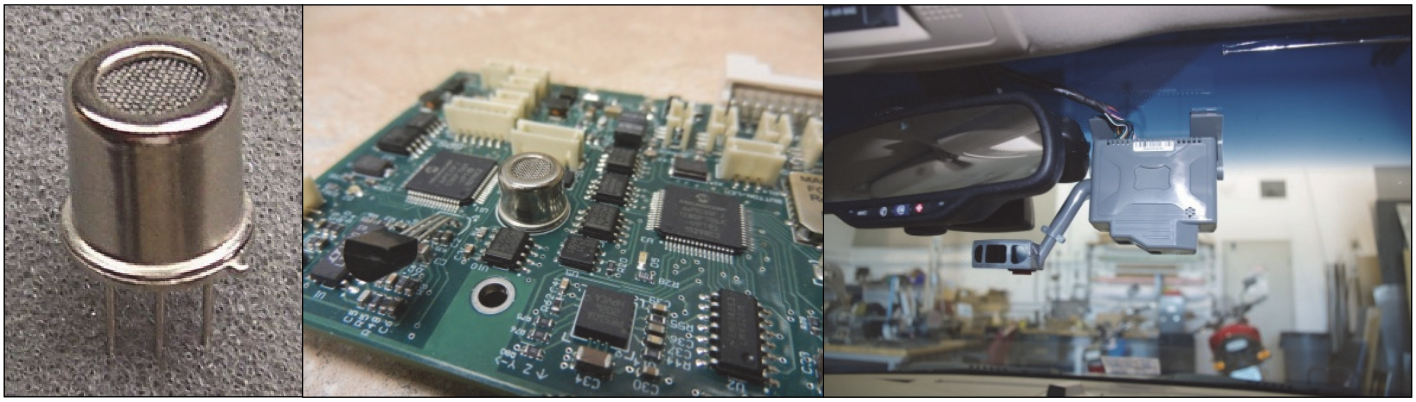


Figure 1.1. Alcohol sensor (*left*), sensor on circuit board (*middle*), and head unit as installed in SHRP 2 fleet (*right*).

research involved administering controlled doses of alcohol to researchers and having them sit in a vehicle equipped with an alcohol sensor. Researchers were administered alcohol to achieve target breath alcohol content (BrAC) readings of 0.05 grams per deciliter (g/dL) and 0.08 g/dL. These researchers then sat in either a front or back seat. Test conditions varied, based on whether the vehicle was moving, the air inside the vehicle was set to recirculate, or the windows were up or down.

Results demonstrated that the prototype alcohol sensor configuration installed in the vehicle at that time could detect and distinguish varying levels of participant intoxication in several of these experimental conditions. In particular, the most accurate sensor readings occurred when the heating, ventilation, and air conditioning (HVAC) was set to recirculate air and the windows were up. The sensor had difficulty detecting all but relatively high levels of alcohol when the HVAC system was on fresh air settings.

Furthermore, the study also found differences in alcohol sensor readings for imbibed versus unimbibed alcohol. The primary difference was a steady, continually decreasing sensor reading for imbibed alcohol versus a sharper spike in sensor readings for unimbibed alcohol. These results suggested that a suitably designed algorithm could potentially differentiate between imbibed versus unimbibed alcohol.

There were several significant changes between the alcohol sensor used in the pilot and the final SHRP 2 instrumentation. These include sensor position, the addition of a fan to circulate air across the sensor in the SHRP 2 instrumentation, and significant software changes. Thus, while the results from the pilot research are useful, they cannot be generalized to alcohol sensor readings in the actual SHRP 2 data set.

Phase 1: Alcohol-Detection Algorithm Proof of Concept

The first phase of research examined the likelihood of developing an alcohol-detection algorithm from the SHRP 2 data set. This was a limited exploratory effort that sought to develop an alcohol-detection algorithm from the SHRP 2 alcohol sensor and apply that algorithm to a selection of SHRP 2 trips. It also included limited sensor testing using a human participant with the alcohol sensor in the same location as the final SHRP 2 sensor instrumentation.

The human participant testing indicated that sensor readings in the SHRP 2 location decreased as a function of BrAC. Figure 1.2 shows the filtered alcohol sensor readings over time for occupants with differing BrAC levels. The vertical red lines represent the times when the windows were rolled up or down; the first red line marks windows being rolled down, the second red line shows windows being rolled up, and the third red line marks windows being rolled down again. This shows that alcohol sensor readings were sensitive to windows being rolled up or down and that sensors may, in fact, be able to detect the presence of alcohol because the change in sensor readings corresponds directly to the introduction of fresh air into the vehicle cabin.

Despite the strong results, a successful algorithm was not developed in this initial phase that could reliably detect driver and passenger impairment in the SHRP 2 data set. This was partly due to the limited time, scope, and unanticipated challenges that arose during this effort. For example, there was no true "gold standard" data set of known alcohol-impaired trips to use as a model for alcohol sensor algorithm development. Also, a number of unanticipated substances were identified that seemed to strongly influence alcohol sensor readings (e.g., windshield wiper fluid and fast food). Given these outcomes, a second phase of research was initiated to continue development of the alcohol-detection method.

Phase 2: Development and Evaluation of an Alcohol-Detection Algorithm

The current effort, Phase 2, built on the results and overcame many of the challenges of the previous efforts. This

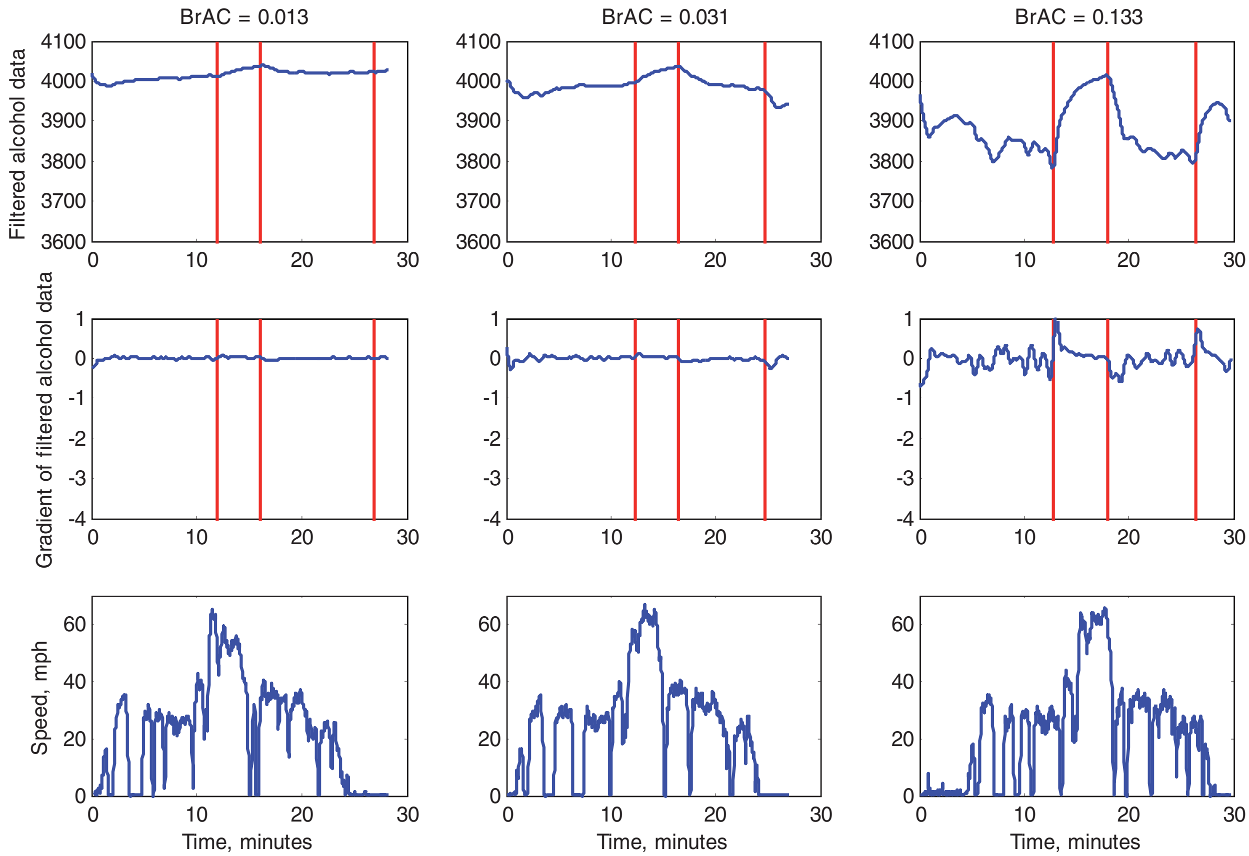


Figure 1.2. Alcohol sensor readings decrease as BrAC increases.

was primarily accomplished through the development of a gold standard data set through experimental in-vehicle testing and extensive data reduction of SHRP 2 trips for signs of impairment. The objectives of the effort were to develop an alcohol-detection algorithm to be applied to the SHRP 2 database and evaluate the accuracy of the algorithm at detecting and differentiating imbibed and unimbibed alcohol in the SHRP 2 database.

CHAPTER 2

Research Approach

The present research study built on the pilot investigation and Phase 1. Those earlier exploratory efforts established a framework for understanding the nature of the alcohol sensor, provided significant insight into the necessary direction for Phase 2, and demonstrated several research needs. The primary need was a gold standard data set to better understand alcohol sensor readings and more precisely evaluate the accuracy of any developed alcohol-detection algorithm using this sensor. Additionally, an accurate assessment of intoxication based on video evidence was necessary to evaluate driver and passenger impairment in existing SHRP 2 trips.

To meet these diverse objectives, two research approaches were undertaken. Each of these approaches met different objectives as follows:

1. Mechanical Breather Approach: A gold standard data set was critical to understanding sensor readings and accurately evaluating an alcohol-detection algorithm using this signal. A carefully controlled experimental data set was created by instrumenting a sedan with the standard SHRP 2 instrumentation—including the alcohol sensor—and controlling the amount of alcohol present within the vehicle cabin. This was accomplished through a variety of methods that are more fully described in the Mechanical Breather Approach section.
2. Naturalistic Approach: The ultimate objective of the research was to validate the efficacy of the alcohol sensor at identifying imbibed alcohol trips within the SHRP 2 database. Specifically, the naturalistic data approach was designed to determine whether the sensor could
    a. Differentiate between imbibed and unimbibed alcohol; and
    b. Differentiate between cases of suspected moderate alcohol use and cases with no suspected alcohol use.

    The naturalistic approach met this objective by coding trips for alcohol involvement from the SHRP 2 database by manually performing visual data inspection to identify both alcohol impairment of the driver or passenger and potential unimbibed alcohol sources.

A data set was created from the SHRP 2 database for the naturalistic approach. This naturalistic test data set was designed to investigate false positives and the ability of the alcohol sensor to differentiate between imbibed and unimbibed alcohol. Given the purpose of this data set, the vast majority of included trips were selected on the basis of positive alcohol sensor values. This overrepresentation of positive alcohol sensor readings made this data set a poor source for evaluating the sensor's performance at discriminating trips involving alcohol from those trips not involving alcohol. This sampling approach inflated the number of false positives compared with the SHRP 2 database. Thus, a subset of this data set, called the "impaired data set," was created; the subset was limited to control trips and trips in which a passenger or driver was deemed at least "moderately" impaired by the data reductionists. Specifically, the impaired data set included only trips in which data reductionists suspected at least a moderate impairment of a driver or passenger and control trips that were not initially selected because of alcohol sensor flags or time of day. The sampling, trip selection, and manual coding are described in more detail in the Naturalistic Approach section of this report.

Finally, known impaired trips from the SHRP 2 data set were examined to develop the alcohol-detection algorithm. These trips were chosen independently from the trip files used in the naturalistic test data set. These were primarily identified by drivers admitting they were driving while impaired or from trips in which alcohol impairment was identified when the trip was being evaluated for another purpose.

Mechanical Breather Approach

One important objective of this research was to identify whether the alcohol sensor could reasonably detect alcohol and how it responded to varying levels. This was necessary to determine if the sensor was accurately capturing data and to aid in developing an algorithm. Since the purpose of the sensor was to detect imbibed alcohol, ideally a study could have been conducted in which drivers with different levels of blood alcohol content (BAC) would drive a vehicle (or at least be in the driver seat of a vehicle) instrumented with the alcohol sensor. With data from such a study, the sensor's response to varying BACs could have been assessed. Unfortunately, because of both cost reasons and Institutional Review Board (IRB) challenges, this approach was not practical within the allotted time frame. Given these constraints, the best solution was to develop a mechanical alcohol breather based on a human's natural breath at varying BACs. This mechanical breather was then used to produce different levels of alcohol in the vehicle while it was being driven.

Development of the "Boozooka" Mechanical Breather

Collecting data from intoxicated individuals is time consuming and complex given the extra steps needed to protect human participants and control BrAC. To circumvent these problems, a mechanical breather, fondly called the Boozooka, was developed. The Boozooka—consisting of an air compressor, regulator, alcohol chamber, mixing valve, and breath alcohol tester that released controlled amounts of alcohol vapors into the cabin—diffused alcohol into the air in a manner similar to human breathing (Figure 2.1).

The compressor pumped air through a regulator that was tuned to represent a typical human breathing rate of 10 liters of air per minute. The air was then split into two streams, one of which passed air through an alcohol chamber filled with cotton balls saturated with 80 proof vodka (cotton balls helped control splashing in the chamber while the vehicle was in motion). A needle valve allowed the alcohol concentration to be tuned by controlling the amount of air passing through the chamber relative to the bypass. During trips, the compressor was powered using a cigarette lighter plug-in.

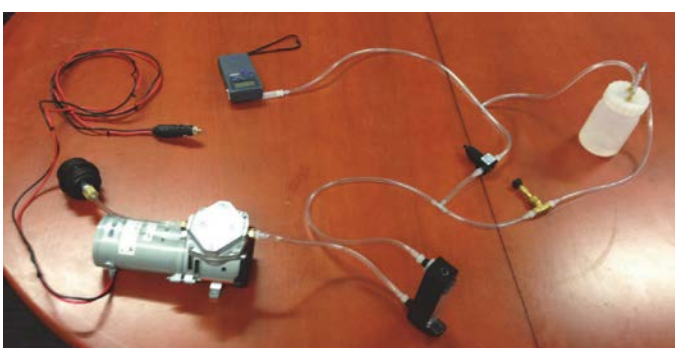


Figure 2.1. Boozooka mechanical breather used for in-vehicle alcohol testing.

The exact BrAC was measured using a Lifeloc FC20 breath alcohol tester, which is accurate to ±0.005 g/dL. Three Lifeloc FC20 breath alcohol testers were rotated between trips to ensure none of the units became saturated. All units were factory recertified by Lifeloc Technologies immediately before data collection. During each experiment, the Boozooka released alcohol vapors of varying BrAC concentrations according to the test plan.

Even though the Boozooka did not emulate certain characteristics of human breath (e.g., humidity, temperature, carbon monoxide and other gases) it was still functionally equivalent for the alcohol sensor. The Boozooka "breathed" the alcohol at a controlled BrAC and volume. Furthermore, the Boozooka was designed to emulate how a human breathes in the vehicle. Since the output from the Boozooka did not blow directly onto the alcohol sensor, it was expected that the various gases of humans in the vehicle and the Boozooka were mixed by the time they reached the alcohol sensor. Finally, the Boozooka was also tested against an alcohol-impaired individual to ensure that the system responded qualitatively in a similar manner.

Experimental Vehicle

For the in-vehicle testing of the mechanical breather, an M35 Infinity sedan was specially instrumented for the project. It was imperative that this instrumentation mirror the SHRP 2 instrumentation setup as closely as possible. The M35 had two SHRP 2 head units, encompassing the alcohol sensor, mounted just under the rearview mirror. This positioning not only matched that of the single head unit of the SHRP 2 instrumentation but also allowed for the collection of readings from two alcohol sensors simultaneously, doubling the amount of data collected and providing a reliability estimate of alcohol sensor values. Specifically, this allowed for the direct comparison of two alcohol sensor readings under the exact same conditions. The vehicle was also equipped with a small data acquisition system (MiniDAS) to collect the alcohol sensor data. Images of the experimental equipment are shown in Figure 2.2.

Data-Recording Interface

The vehicle data acquisition system (DAS) used a data-recording program that ran on a custom embedded Linux operating system. In addition to choreographing the collection of synchronized data, this program also provided an interface for researchers to enter experimental conditions and variables into the in-vehicle data stream (a screenshot of the user interface is provided in Appendix A). In particular, study variables could be directly associated with corresponding alcohol sensor readings from the two head units. For example, the researchers kept a record of the in-vehicle BrAC, experimental conditions, air currents, and trip number. Data

Figure 2.2. M35 (upper left), outside view of the two head units and MiniDAS (upper right), MiniDAS (lower left), and inside view of two head units and MiniDAS (lower right).

were collected and stored at a rate of 10 Hz, the same rate used by the alcohol sensor in the SHRP 2 study.

Experimental Procedure

Experimental tests were performed on a predetermined study route that included several types of roadways and speeds ranging from 25 mph to 65 mph and that lasted approximately 15 minutes. Fifty trials were conducted under three experimental conditions. Each trial included only one condition. The conditions and number of trials are as follows:

1. Control testing (no BrAC), 6 trials;
2. Human testing, 11 trials; and
3. Boozooka testing, 33 trials.

The heaviest emphasis was on Boozooka testing; the human testing was done primarily for validation.

Across all experimental conditions within the vehicle, one researcher drove and another researcher entered data. During Boozooka testing, a third researcher operated the Boozooka and confirmed its BrAC output using the breath alcohol testers. In the human testing condition, an additional researcher who had consumed alcohol sat in the passenger seat. The control trials were performed primarily with the driver and the researcher who entered data, occasionally with a third researcher observing from the back seat.

Ten trials were conducted using recirculating air, and 40 trials used fresh air settings. The pilot study research indicated fresh air would provide a more challenging environment for detecting alcohol presence, so more fresh air trials were chosen. Windows were rolled up for all trips since previous research indicated that the alcohol sensor could not detect alcohol presence with the windows rolled down. Accordingly, trips with windows rolled down were not conducted.

Between each experimental trial for all test conditions, the car was parked and the windows or doors were opened for approximately 10 to 15 minutes to allow the alcohol sensor to return to baseline.

Before each Boozooka trial, the level of alcohol within the alcohol chamber was examined. If the level was too low, then alcohol was added. The Boozooka operated for at least 10 minutes before beginning a trip to ensure stable BrAC

Table 2.1. Mechanical Breather (Gold Standard) Data Set

Conditions	Experimental Trials	Total Samples
Control (no BrAC)	6	12
Human Impaired	11	22
Boozooka	33	66
Total	50	100

output, and a trip did not start until multiple consecutive readings were within approximately 0.04 g/dL. Readings on Boozooka BrAC were taken at least every 5 minutes of the trip using the Lifeloc FC20 breath alcohol testers. For the human testing, a BrAC reading was taken at the beginning and end of each trip since slight changes in BrAC were possible during the course of a trip.

Mechanical Breather Data Set

The mechanical breather data set included a total of 50 15-minute experimental trials. Two alcohol sensors in the instrumentation provided the independent readings represented in Table 2.1, doubling the number of samples for analysis. These trips constituted the gold standard data set because of the experimentally controlled presence of alcohol within the vehicle.

Naturalistic Approach

The ultimate goal of a successful alcohol-impairment algorithm was to detect impaired trips within the SHRP 2 data set while minimizing false detections due to the presence of unimbibed alcohol. It was important to make sure the sample had an adequate number of true positives in which imbibed alcohol was likely (most important), false positives in which other sources of alcohol were likely, and a control group in which alcohol presence, though possible, was unlikely.

To enable testing with SHRP 2 data, a reference data set was developed in which impairment was independently determined. Significant effort went into the development of a behavioral checklist for impairment using visual cues that could be identified using the SHRP 2 cameras. Similarly, the SHRP 2 trips included in the sample and the data reduction team in charge of reducing the trip files were carefully selected.

Note that trip files, or trips, for the purpose of the naturalistic driving data are defined as the time from the vehicle being started to it being turned off. Up to a 60-second delay occurred between vehicle ignition and data collection due to the startup of the DAS. All analyses made use of the entire trip to identify the possible presence of unimbibed and imbibed alcohol.

Selection of SHRP 2 Trip Files

At the beginning of the project, it was impossible to determine the number of total impaired trips in the SHRP 2 database; the number of substances or events that would trigger the alcohol sensor; and the success of the algorithm in detecting and differentiating types of alcohol. This necessitated a trip selection approach that would maximize the probability of finding impaired trips and populate the data set with trips of interest for evaluation. Thus, the first and largest ($n = 562$) batch of trips was selected to maximize the likelihood of positive (i.e., sensor-flagged) trips. All trips in this batch were selected from trips occurring between 12:00 a.m. and 4:00 a.m. in the driver's local time, a time range expected to have a high likelihood of trips containing imbibed alcohol. Additionally, trips were selected on the basis of alcohol sensor values to maximize the probability of finding impaired trips and trips with unimbibed alcohol. Trips that crossed a certain threshold for spikes, standard deviations, or averages for sensor readings were flagged and set aside for the reduction team. This process created an ideal data set for evaluating false positives and testing the algorithm's ability to discriminate between imbibed and unimbibed alcohol within a vehicle.

Once a large number of impaired and unimbibed alcohol trips were identified through manual video review, a second batch of trips was chosen. These trip files were chosen at random, irrespective of time of day or sensor readings, to select a group of trips during which imbibed alcohol presence was unlikely. While the first batch of trips achieved the objective of exploring the sensor's ability to differentiate imbibed from unimbibed alcohol presence, the second batch of trips ($n = 97$) successfully evaluated the alcohol sensor's ability to differentiate alcohol-involved trips from trips without alcohol presence (i.e., false alarms). This was primarily accomplished through the inclusion of the control trips.

The final data set included a total of 659 trips. Originally, 692 trips were sent to the data reduction team across the two batches. However, 33 trips were excluded because (1) they did not go through the initial data ingestion quality assurance process required for any trip used in the SHRP 2 data set, (2) the consented driver was not present for the trip, or (3) the videos or alcohol sensor data did not properly load. The final data set of 659 trips served as the naturalistic test data set.

Alcohol-Impairment Behavioral Checklist and Data Reduction Checklist

Other than the alcohol sensor and the kinematic sensors, the camera views of the standard SHRP 2 instrumentation package provided the most reliable method for determining intoxication. Camera views captured the forward roadway, rear view, driver's face, and driver's hands. Reductionists looked at all camera views to determine impairment.

Data reductionists checked for and coded signs of impairment across several broad categories that included visually seeing alcohol sources within a vehicle (both imbibed and unimbibed), behavioral cues from the face-view camera, and driving performance. The data reductionists observed and coded these categories of data and were then asked to make subjective evaluations of driver impairment and degree of impairment.

The behavioral checklist was made after carefully reviewing and synthesizing information from literature related to visual signs of impairment (*5–10*). From this literature more than 90 behavioral and visual cues were identified for assessing alcohol intoxication. Out of this broader list of identified cues, 64 cues were deemed observable using SHRP 2 video. These cues and operational definitions are provided in Appendix B.

Reductionists also coded signs or presence of unimbibed alcohol. They were given a protocol with a list of known alcohol sensor triggers: fast food, hand sanitizer, perfume, cologne, cigarettes, marijuana, other drugs, chewing gum, and windshield wiper fluid. Reductionists had the option of noting in the data file any other substances that they deemed suspicious or that might contain any form of alcohol.

Data Reduction Procedure

Six specially trained data reductionists from the Virginia Tech Transportation Institute (VTTI), all of whom had experience working with intoxicated individuals in a field setting for other research efforts, reduced the trip files. In previous endeavors they collected BrAC data from pedestrians in a bar-centric downtown district; on average, each individual had spent over 100 hours conducting research in this field setting with intoxicated individuals (*11–13*). Furthermore, each of the data reductionists was trained specifically for this project by a data reduction supervisor.

Four of the data reductionists served as the primary data analysts for trip files. These reductionists filled out a check sheet that coded behavioral cues for impairment (see Appendix B), noted any observed types of unimbibed alcohol, noted indicators of poor driving performance, and made determinations about the driver's and passenger's level of intoxication. These data reductionists were allowed to examine all video views but were instructed not to look at alcohol sensor readings to remove potential bias. Reduction of each video took approximately 10 to 15 minutes with the data reductionists watching the first 2 minutes, last 2 minutes, and 3 minutes at random points in between. Additionally, the data reductionists watched most of the video at 10× speed to find possible sources of unimbibed alcohol and identify critical events that could be particularly informative in determining impairment.

One data reductionist served as the alcohol sensor validator and, to avoid bias, was the only individual allowed to look at the alcohol sensor readings. This individual was instructed to consider the alcohol sensor and all video views to further examine points in the trip when the alcohol sensor drastically changed. This reductionist would then code events happening around that time period in an Excel file separate from the one used by the four primary reductionists. This procedure ensured that sources of unimbibed alcohol were not missed since it was difficult for the primary reductionists to catch every substance used, particularly during longer trips. Unimbibed alcohol was caught most often because it usually produced sharp spikes in sensor readings. However, this procedure was also able to identify intoxicated individuals entering the vehicle in designated driver (DD) scenarios, alcoholic beverages being poured within a vehicle, and other similar types of impaired trips. The alcohol sensor validator was not allowed to edit the Excel log of the primary data reductionists and did not determine impairment. At the conclusion of the data reduction effort, a senior member of the research team updated the primary data file to include substances identified by the alcohol sensor validator. Again, only the timestamps and identification of substances were updated, not decisions regarding intoxication.

The final data reductionist served as the quality assurance reviewer who went through and validated the work of the four primary reductionists. This reductionist added comments to the data file when discrepancies of judgment were identified, which caused the original reductionists to review their comments. When the data reductionists agreed with the changes, they were entered into the data file. In the event of disagreement, the data reduction supervisor or senior member of the research team would view the trip and make the final determination.

Naturalistic Test Data Set

As shown in Figure 2.3, the naturalistic test data set includes 659 trips. Of these trips, 562 were initially chosen because of a preliminary positive alcohol sensor reading and their late-night occurrence when alcohol use was more likely (Batch 1). The remaining 97 trips served as a control group and were selected randomly, irrespective of the time of day (Batch 2). All 659 trips were reduced via manual video coding to identify possible impairment and possible sources of alcohol, including potential false positives.

Although the test data set did contain control trips, it contained over five times more trips with likely alcohol presence. This ratio is much different from what would be expected in the SHRP 2 database. The SHRP 2 database would be expected to have far fewer trips with alcohol present than trips without

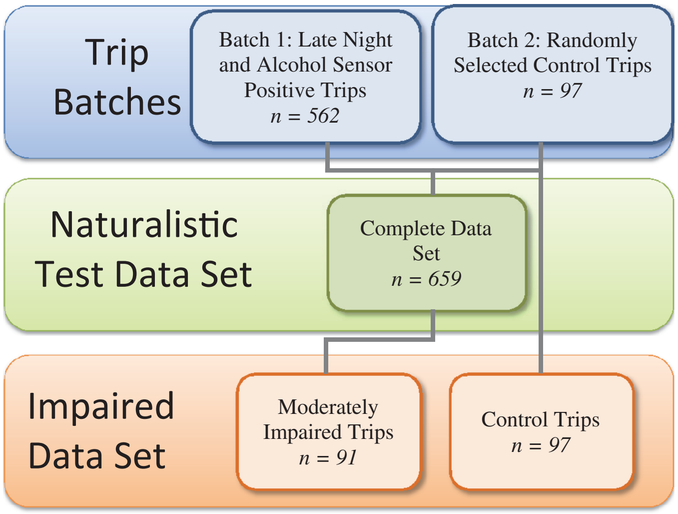


Figure 2.3. Distribution of trip files into relevant analysis databases.

alcohol present. Therefore, this data set was not designed to determine the sensitivity and specificity that an algorithm would have when compared with the SHRP 2 database—it was too biased toward positive alcohol readings. Instead, this data set was designed to investigate the accuracy of the alcohol-detection algorithm in differentiating imbibed from unimbibed alcohol because reliably separating false positives from true positives was a primary objective of the study.

A second objective of the naturalistic test data set was to determine how well impaired driving trips could be differentiated from unimpaired driving trips. As shown in Figure 2.3, a subset of the naturalistic test data set, the impaired data set, was used to test this objective. The impaired data set included all the randomly selected control trips (Batch 2) and trips from the naturalistic test data set identified by reductionists as having at least one vehicle occupant who was moderately impaired. Thus the impaired data set comprised 188 trip files, including 91 trips in which video reductionists rated at least one vehicle occupant as being "moderately" impaired and 97 randomly selected trips serving as a controlled baseline.

The impaired data set was designed specifically to show how well the algorithm could differentiate moderately impaired driving from normal driving. Analyses were conducted separately on both the complete naturalistic test data set and the impaired data set derived from this sample.

Neither the naturalistic test data set nor its subset, the impaired data set, provided a representative estimate of the algorithm's effectiveness across the entire SHRP 2 database, as developing a data set for this purpose would have been cost prohibitive. For example, out of the 97 control trips, only one indicated potential driver impairment. Thus, a large number of randomly selected trips would need to be evaluated to find a sufficient number of potentially impaired trips to assess this algorithm, an effort that was outside the budgetary constraints of the project.

CHAPTER 3

Findings and Applications

Algorithm Considerations and Development

Previous research efforts for the alcohol sensor and an examination of trips known to have imbibed and unimbibed alcohol presence created the framework for the development of an alcohol-detection algorithm using alcohol sensor data. These considerations and examples are detailed below along with a description of their translation into the final nonproprietary algorithm used in this research.

Critical Considerations for Algorithm Development

As shown in Figure 3.1, alcohol sensor readings are almost always defined by certain characteristics. In particular, the figure shows an unimpaired trip free from the presence of unimbibed alcohol; it also shows the warm-up period and "shadow" that characterize the alcohol sensor output. Almost all readings, even in the presence of alcohol, have a warm-up period and shadow. The shadow is seen in the figure as a set of sensor readings that mirror the primary readings and is typically 15 mV below the primary sensor output.

Figure 3.1 also demonstrates a warm-up period before the alcohol sensor readings stabilize, which can range from a couple of seconds to a couple of minutes. In this and other graphs showing alcohol sensor readings, the y-axis represents alcohol sensor readings in mV. The x-axis represents time in milliseconds (ms). In the unimpaired trip represented by this graph, the baseline reading can also be observed when no alcohol is present. These baseline readings are most often above 4,000 mV. The presence of alcohol causes the sensor reading to drop.

The sensor was also sensitive to a variety of unimbibed alcohol and other substances. These included windshield wiper fluid, hand sanitizer, chewing gum, fast food, cologne, perfume, cigarettes, aerosol spray, glass cleaner, mouthwash, and other substances containing alcohol or an alcohol base. Many of these substances have very similar effects on the sensor. An example of sensor readings under various unimbibed substances can be seen in Figure 3.2. Again, the y-axis represents mV, and the x-axis represents time (ms). This figure shows that sensor readings typically drop sharply at the introduction of unimbibed substances and have a slow, often gradual, return to baseline. Fast food differs slightly, with a slow and less pronounced drop in sensor values. Additional investigation was not warranted: it is unknown what components of the fast food excited the sensor or the breadth of food types that may have had an influence.

Additionally, several positive cases of impairment were identified. These included clear visual evidence based on driver performance and video, drivers verbally reporting alcohol involvement using the critical incident button, video footage of police arresting a driver for driving under the influence (DUI), or visual confirmation of alcoholic beverages in the vehicle. These trips were useful in developing algorithms that could detect alcohol-impaired drivers.

Figure 3.3 shows an example of a trip with impaired passengers. During this trip, several intoxicated individuals got into the back seat of a vehicle while at a stop sign. The individuals had plastic cups with what was later reported verbally by the driver to be alcohol. In the associated video, the driver, who had not been drinking, pressed the critical incident button once the passengers exited to report that they smelled strongly of alcohol. The figure shows a dip in sensor readings when the intoxicated passengers entered the vehicle followed by a rise to baseline when they exited the vehicle, showing how sensor readings typically drop under the influence of intoxicated individuals.

From these confirmed alcohol-imbibed trips, it was evident the alcohol sensor had the potential to detect the presence of imbibed alcohol regardless of the number of individuals in the vehicle who consumed alcohol or their position within the

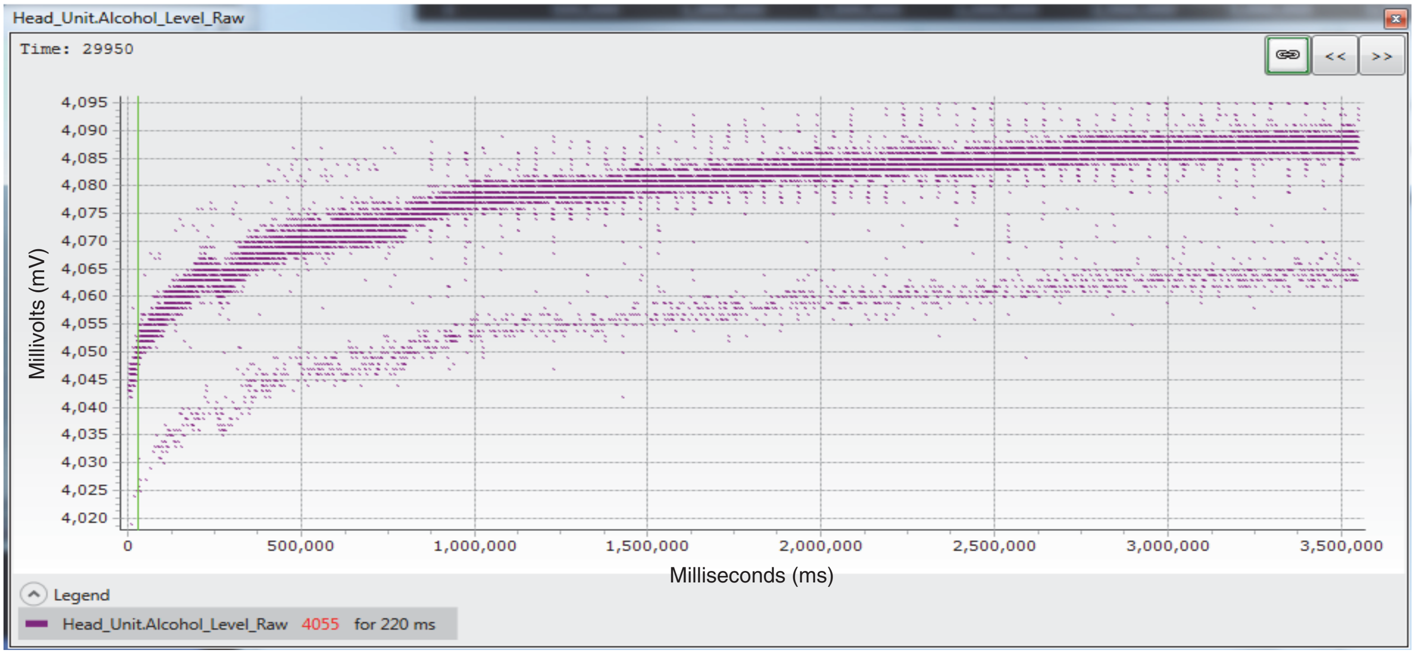


Figure 3.1. Alcohol sensor shadow effect and example of warm-up period.

vehicle. The readings appeared to indicate a stronger sensor response with a decrease in readings when the individual who consumed alcohol was in the driver or passenger seat rather than in the back seat. However, a similar pattern of sensor readings was observed outside the magnitude of the drop in sensor readings based on a front-seat versus back-seat individual having a positive BrAC. From these cases there did not appear to be a difference in sensor response whether the individual who consumed alcohol was in the driver seat or the passenger seat. A scenario in which the driver with a positive BrAC was the sole occupant of a vehicle would produce sensor readings similar to a scenario in which the front-seat passenger was intoxicated and the driver was sober. These observations were consistent with the alcohol sensor setup. As a result of the sensor's location near the rearview mirror, it was more exposed to the breath of the front-seat passenger and driver rather than back-seat passengers.

Algorithm Development

The information from the graphs in Figure 3.3 was used to create an alcohol-detection algorithm using the alcohol sensor to detect impaired individuals within a vehicle and to differentiate imbibed from unimbibed sources of alcohol. To remove bias from the performance metrics of the final algorithm, none of the trips used for development of the algorithm were included in any of the evaluation data sets. Across the files, it was observed that both the slope and absolute value of alcohol sensor readings were critical to understanding alcohol presence within a vehicle. Researchers considered the following while developing the alcohol algorithm, which should be used as guidance in future development:

- A moving average allowed the sensor shadow to be eliminated by removing all points that deviated strongly from the average.
- The absolute value of sensor readings was shown to relate to the presence of imbibed alcohol. While the strength of this effect can potentially vary from sensor to sensor, a threshold of 3,965 mV was established for the alcohol-detection algorithm. Any trip with an average below this level after controlling for unimbibed alcohol was classified as containing imbibed alcohol.
- A quick change, or steep slope, in alcohol sensor readings was most often the result of unimbibed alcohol and should be considered in reducing false positives.
- There appeared to be a large amount of variance in whether or not a particular trip had a warm-up period and the length of that period. For the purposes of this alcohol-detection algorithm, a warm-up period of 50 seconds was set. This helped ensure that short trips were not thrown out while removing some of the anomalous readings that could occur at the beginning of trip files.

Mechanical Breather Boozooka Validation

Before assessing the accuracy of the alcohol-detection algorithm on the experimental gold standard data set, the Boozooka needed to be validated for representativeness and accuracy. This was accomplished by comparing alcohol sensor readings

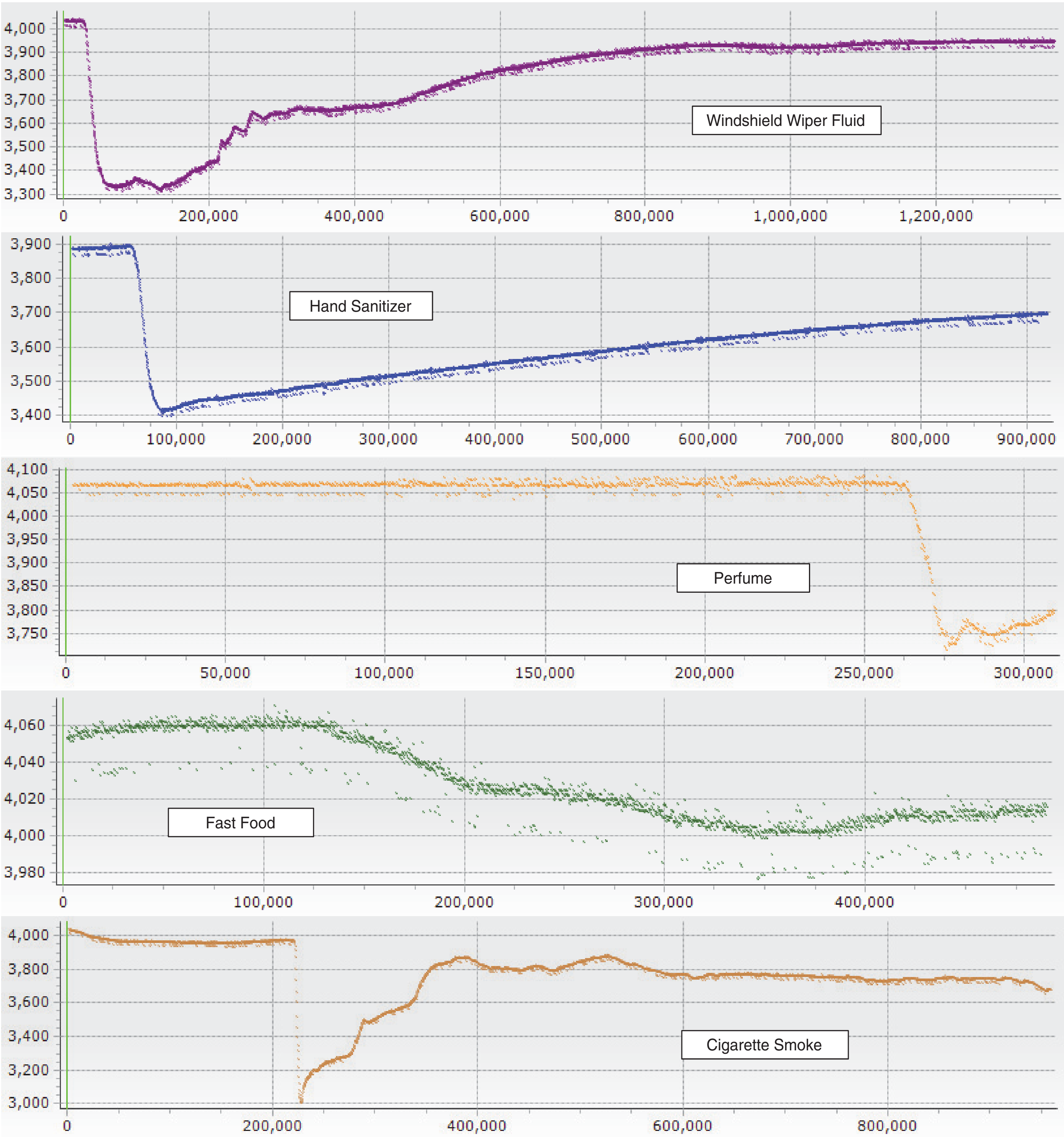


Figure 3.2. Effect of various types of unimbibed alcohol on alcohol sensor.

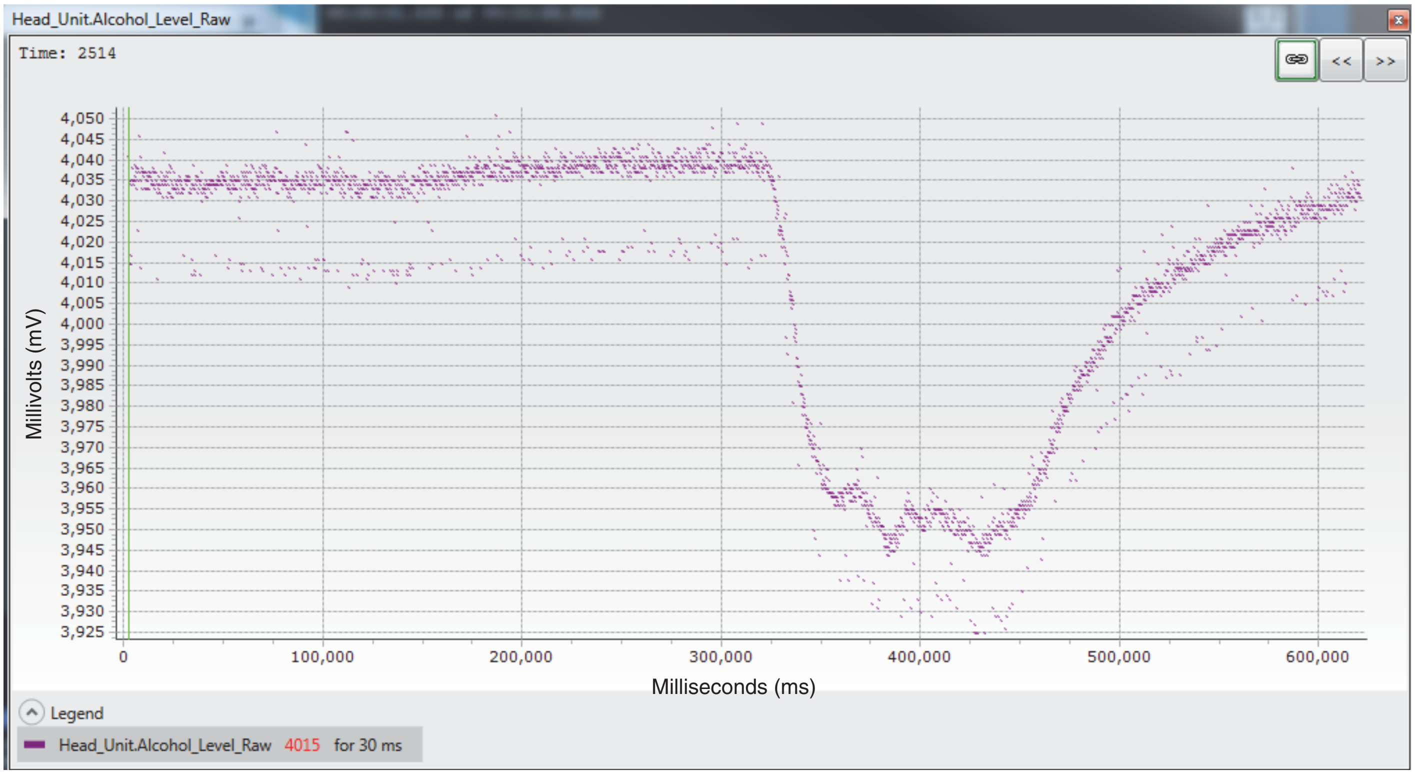


Figure 3.3. Alcohol-involved trip.

in the presence of the Boozooka with the controlled trips with intoxicated individuals.

As shown in Figure 3.4, values for the first alcohol sensor (top) and second alcohol sensor (bottom) changed consistently as a function of the Boozooka's activity. This operational control indicated the Boozooka worked at a fundamental level. Indeed, the sensor readings changed in a way that was consistent with other known impaired trips and were highly correlated.

Direct comparisons between trips with an impaired passenger and the Boozooka were also conducted. These tests matched the BrAC levels of the individual to the Boozooka for direct comparisons. Figure 3.5 shows average alcohol sensor readings at various BrAC categories broken down by Boozooka and human testing. The pattern of sensor readings was consistent for both conditions—voltage decreased as BrAC increased. However, the average value of sensor readings was substantially lower for the Boozooka condition than for human testing. This was likely the result of setting the Boozooka to "breathe" 10 liters per minute, the average rate for an individual of medium-to-large build. The researcher used for human testing was substantially smaller in weight and likely also breathed at a much lower rate, resulting in lower volumes of alcohol entering the cabin at a similar BrAC. The consistent trend of sensor readings created confidence in the Boozooka as a reasonable proxy for an intoxicated individual. Albeit, the discrepancy served as a reminder that many factors can influence the volume of alcohol that enters a cabin. Future research could expand the Boozooka testing by looking at the influence of breathing rate across a larger spectrum of representative values.

Mechanical Breather Gold Standard Data Results

The dual sensor setup (i.e., two head units each with an alcohol sensor installed in the same vehicle) provided a unique opportunity to investigate how two sensors operated under virtually identical environmental conditions. Across all trips, the Pearson correlation between Alcohol Sensor 1 (AS1) and Alcohol Sensor 2 (AS2) was .989, $p < 0.01$. While this did not necessarily demonstrate that both sensors worked perfectly at assessing alcohol presence, it did indicate that the sensors behaved similarly across various levels of alcohol presence and trips. It should be noted that correlations are a measure of similarity of rank order. Thus, this strong correlation between sensors did not necessarily mean they provided identical readings but rather that their readings varied to similar degrees across time periods.

The gold standard data set also allowed for the examination of whether or not BrAC differences could be detected using the SHRP 2 alcohol sensor. Pearson correlations between AS1 and BrAC were −.526, $p < 0.01$, and between AS2 and BrAC were −.534, $p < 0.01$. This indicated that differences in BrAC could potentially be detected within a given sensor.

Figure 3.6 shows the average readings for each alcohol sensor across all trips categorized by ranges of BrAC; it also

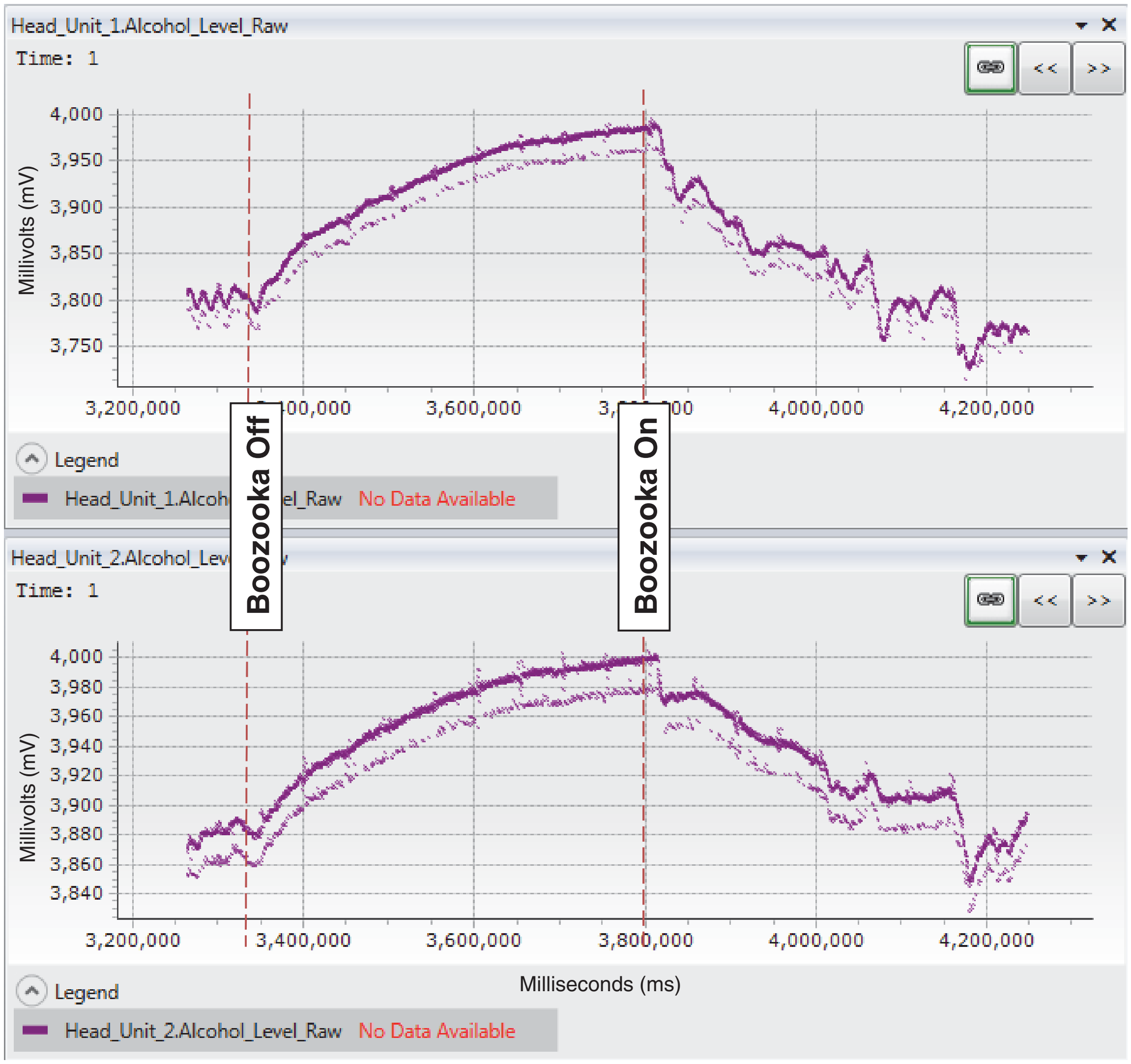


Figure 3.4. Change in sensor based on Boozooka on/off status (functional control).

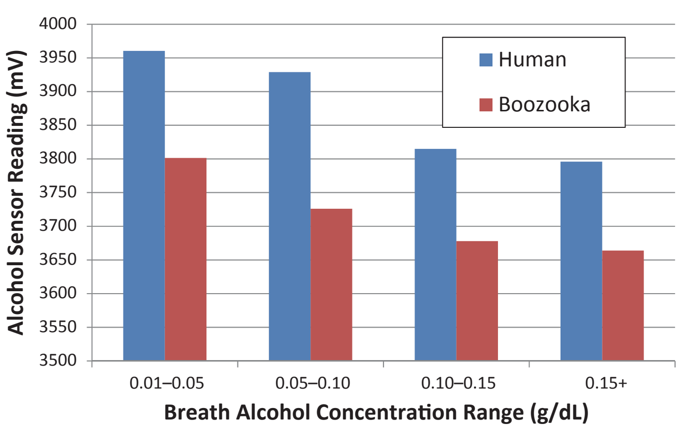


Figure 3.5. Alcohol sensor readings as a function of BrAC for Boozooka and human testing.

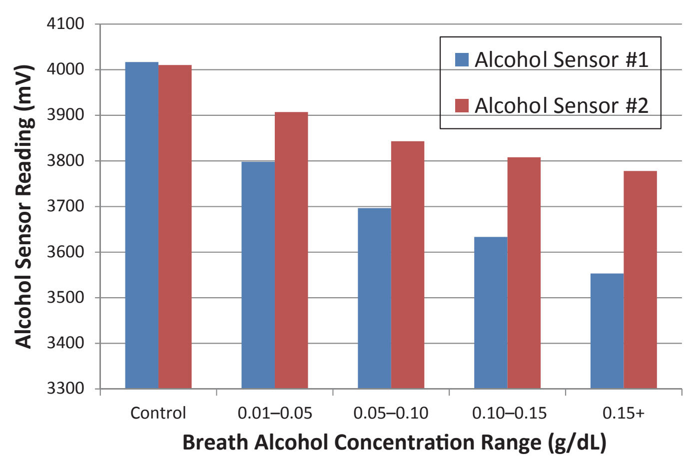


Figure 3.6. Alcohol sensor readings as a function of BrAC by sensor.

		True State	
		Alcohol present	No alcohol present
Algorithm Response	Alcohol present	True positive	False positive
	No alcohol present	False negative	True negative

Figure 3.7. Confusion matrix.

shows that alcohol sensor readings dropped for both sensors as BrAC increased. However, although the absolute value of the alcohol sensor readings for AS1 and AS2 were similar under no alcohol presence, AS1 had a steeper decrease in sensor readings with increased BrAC. Thus, while sensor readings were sensitive to BrAC, the sensitivity appeared to differ across sensors. Without standardizing sensor readings or knowing the calibration for a given sensor, BrAC estimates from sensor values are likely not possible. Additional research with several sensors would be necessary to confirm this finding.

In addition to providing a better basic understanding of the alcohol sensors, the gold standard data set also helped evaluate an algorithm to detect in-vehicle alcohol presence by using a confusion matrix based on the principles of Signal Detection Theory. In this confusion matrix, each trip was classified according to whether alcohol was actually present within a vehicle and the algorithm's estimate of alcohol presence. For example, if alcohol was present in the vehicle but the algorithm estimated that no alcohol was present, then that instance would be classified as a false negative. A sample confusion matrix indicating four possible outcomes is displayed in Figure 3.7.

Results for the alcohol-detection algorithm against the gold standard data set are presented in Figure 3.8. As shown in the figure, the sensitivity of the algorithm was 96.6%. The specificity of the algorithm was 100%. Overall, only three of the 100 trips were incorrectly categorized by the alcohol-detection algorithm. These trips were at low BrACs, and sensor readings were close to the threshold for detection. The results for the gold standard data set indicate that the algorithm performed well at differentiating trips based on alcohol presence.

		True State	
		Alcohol present	No alcohol present
Algorithm Response	Alcohol present	85	0
	No alcohol present	3	12

Figure 3.8. Confusion matrix for gold standard data set.

Naturalistic Test Data Results

The naturalistic test data set contained 659 trips and was designed to be overrepresented with cases in which the alcohol sensor indicated a positive reading. This allowed researchers to examine the ability of the alcohol sensor to differentiate imbibed from unimbibed alcohol.

Out of these 659 trips, unimbibed alcohol was found in 290 trips. While the percentage of trips with unimbibed alcohol was inflated because of how trips were select, the relative frequency of various unimbibed substances in relation to each other was likely to be representative of the broader SHRP 2 database. Of the 290 trips, 58% ($n = 167$) contained windshield wiper fluid. This was followed by hand sanitizer ($n = 26$; 9%), cologne/perfume ($n = 24$; 8%), multiple substances ($n = 18$; 6%), cigarettes/other drugs ($n = 36$; 13%), fast food ($n = 12$; 4%), glass cleaner ($n = 4$; 1%), and chewing gum ($n = 3$; 1%).

Since this data set oversampled positive sensor readings, it was most useful for evaluating the alcohol-detection algorithm's accuracy at differentiating between these positive readings. Table 3.1 is a contingency table depicting the accuracy of the algorithm at differentiating all possible combinations of alcohol presence. The true state is shown in the columns of the table, and the algorithm response is shown in the rows. A chi-square test on the contingency table was statistically significant, $\chi^2(9) = 365.4$, $p < 0.01$. This result indicated that the alcohol-detection algorithm performed better than chance at estimating and differentiating alcohol presence within a vehicle on this data set. For the aforementioned reasons, the number of false positives was inflated in this data set. However, the chi-square test shows that the alcohol-detection algorithm was moderately accurate at differentiating alcohol presence. While the alcohol-detection algorithm was 91.7% accurate when there was imbibed alcohol, it was only 22.6% accurate at predicting unimbibed alcohol alone and 35.5% accurate at predicting the presence of both imbibed and unimbibed alcohol.

Even though the number of false positives was inflated in this data set, it was possible to explore the alcohol algorithm's accuracy at detecting alcohol presence in a confusion matrix. Though this bias toward alcohol-positive cases should be considered when interpreting the results, a confusion matrix was created by classifying trips as imbibed alcohol present or no imbibed alcohol present. This was

Table 3.1. Contingency Table Showing Algorithm Accuracy on the Full Naturalistic Test Data Set

Algorithm Response	True State				Total	Percent Accurate
	No Alcohol Present	Imbibed Alcohol	Unimbibed Alcohol	Both Types of Alcohol		
No Alcohol Present	101	1	1	0	103	98.1%
Imbibed Alcohol	99	133	107	33	372	35.7%
Unimbibed Alcohol	12	5	53	7	77	68.8%
Both Types of Alcohol	5	6	74	22	107	20.5%
Total	217	145	235	62	659	
Percent Accurate	46.5%	91.7%	22.6%	35.5%		

		True State	
		Imbibed alcohol present	No imbibed alcohol present
Algorithm Response	Imbibed alcohol present	194	285
	No alcohol present	13	167

Figure 3.9. Confusion matrix for alcohol-detection algorithm against naturalistic test data set.

done by categorizing trips coded as "imbibed alcohol only" and "both imbibed and unimbibed alcohol present" in an "alcohol present" classification. Similarly, trips that were coded as "unimbibed alcohol only" or "neither imbibed nor unimbibed alcohol present" were categorized as "no alcohol present." The results of the confusion matrix are depicted in Figure 3.9. The sensitivity was 93.7% and the specificity was 36.9%. Again, this range was likely the result of the inflated number of false positives, and the algorithm likely performed worse than it would on the entire SHRP 2 database.

As mentioned, a subset of the naturalistic test data set was extracted to form the impaired data set. The ability of the alcohol sensor algorithm to differentiate the type of alcohol presence within the vehicle was evaluated using a chi-square test on the impaired data set. The chi-square contingency table is shown in Table 3.2, with rows representing the algorithm response and columns representing the estimated true state. The overall chi-square test was significant, $\chi^2(9) = 259.95$, $p < 0.01$, indicating the alcohol sensor algorithm performed better than chance at estimating alcohol presence. It should be noted that the accuracy of the alcohol-detection algorithm led to many cells in the contingency table having fewer than five observations, which can bias chi-square values. The table shows that the alcohol-detection algorithm was particularly accurate at estimating no alcohol presence and impaired driver or passenger trips.

Table 3.2. Contingency Table Showing Algorithm Accuracy on the Impaired Data Set

Algorithm Response	True State				Total	Percent Accurate
	No Alcohol Present	Imbibed Alcohol	Unimbibed Alcohol	Both Types of Alcohol		
No Alcohol Present	96	1	0	0	97	99.0%
Imbibed Alcohol	0	54	0	16	70	77.1%
Unimbibed Alcohol	0	3	2	3	8	25.0%
Both Types of Alcohol	0	3	0	10	13	76.9%
Total	96	61	2	29	188	
Percent Accurate	100.0%	88.5%	100.0%	34.5%		

		True State	
		Alcohol present	No alcohol present
Algorithm Response	Alcohol present	83	0
	No alcohol present	7	98

Figure 3.10. Confusion matrix for impaired data set.

The confusion matrix for the impaired data set is shown in Figure 3.10. The sensitivity of the algorithm against this data set was 92% and the specificity was 100%. This indicated the alcohol-detection algorithm performed extremely well when trips were classified with greater certainty of imbibed alcohol involvement. In particular, the algorithm was not prone to false negatives, suggesting that alcohol-detection performance increased as the observable signs of intoxication increased.

CHAPTER 4

Conclusions and Suggested Research

Researchers used a multifaceted approach to further investigate the standard alcohol sensor in SHRP 2 vehicles and develop an initial alcohol-detection algorithm based on the sensor data. The second phase of research was designed with this goal in mind. Specifically, the objectives were to (1) determine the necessary considerations in an alcohol-detection algorithm to be applied to the SHRP 2 database and (2) evaluate the accuracy of an algorithm using these considerations at detecting and differentiating between imbibed and unimbibed alcohol in the SHRP 2 database. Through an examination of SHRP 2 data and several experimental manipulations, an alcohol-detection algorithm was developed and investigated. This allowed for a basic evaluation of an algorithm and an assessment of the usefulness of the sensor for identifying imbibed alcohol. A summary of conclusions and recommendations from this research follows.

Alcohol-Detection Algorithm Accuracy

To maximize its utility, this type of algorithm must detect cases of imbibed alcohol and also filter out the effects of other substances that can affect the alcohol sensor and produce false detections. After all, the sensor was only designed to detect general alcohol presence within the cabin. At the conception of this project, little was known about the alcohol sensor's reaction in the presence of various types of alcohol. Thus, characterization of the sensor's response to both imbibed and unimbibed alcohol was necessary.

Sensor's Ability to Detect Alcohol Presence

To answer the fundamental question of the sensor's alcohol-detection accuracy, an experimental gold standard data set was created. While it was difficult to confirm alcohol impairment in the SHRP 2 data set, the gold standard data set allowed for the careful distribution of alcohol into the cabin of a vehicle using a mechanical breather. This provided an ideal metric for evaluating the alcohol-detection algorithm. When tested against this data set, the algorithm was over 95% accurate at differentiating alcohol presence from no alcohol presence. The only inaccuracies were a few trips with low doses of alcohol that barely missed the detection threshold.

Differentiating Unimbibed and Imbibed Alcohol Presence

A variety of substances that are naturally introduced into motor vehicles contain alcohol. To reduce false alerts, an alcohol-detection algorithm must be able to differentiate these unimbibed forms of alcohol from the presence of humans in the vehicle who have imbibed alcohol. This ability was assessed by creating a naturalistic test data set from the SHRP 2 database that was heavily weighted toward positive alcohol sensor readings.

A variety of substances were identified that had an impact on the alcohol sensor readings. The most common substance was windshield wiper fluid—although other substances were shown to affect the sensor when introduced into the vehicle. These substances generally had a similar effect: a steep drop in alcohol sensor readings followed by a gradual return to baseline for the given trip. However, this was not always the case. For example, fast food often resulted in a slow, less severe drop in alcohol sensor readings. As another example, when windshield wiper fluid was used to melt ice on the windshield, it often had a lingering, constant presence with an unusually slow or even absent return of alcohol sensor readings to baseline. It is hypothesized that this occurred because ice saturated by windshield wiper fluid remained on the windshield. This introduced a constant stream of alcohol vapors into the cabin, thus making it difficult to differentiate imbibed from unimbibed alcohol.

While the alcohol-detection algorithm was highly accurate in determining alcohol presence within a vehicle, it was only weakly to moderately able to differentiate imbibed versus unimbibed alcohol. In particular, the alcohol-detection algorithm tended to classify unimbibed alcohol trips as imbibed trips, which meant a tendency for false positives. For most analyses, false positives are better than false negatives. Visual data validation can remove trips that were actually the result of unimbibed substances, but false negatives cannot be feasibly reduced through a manual validation process (all video would need to be reduced for signs of alcohol).

The difficulty in differentiating imbibed from unimbibed alcohol stemmed from a variety of sources. The effects of unimbibed alcohol on sensor readings can be long lasting and, thus, similar to the signature left by imbibed alcohol. Additionally, many sources of unimbibed alcohol are introduced into the vehicle at the beginning of a trip, placing the steep spike in sensor readings (a strong differentiating characteristic between imbibed and unimbibed alcohol) in the sensor warm-up period. As a result, this spike is not detected or assessed by the alcohol-detection algorithm. The warm-up period is occasionally marked by a spike in sensor readings even without the presence of unimbibed or imbibed alcohol. Thus, the characteristic spike of unimbibed alcohol is masked by this warm-up period. Finally, from the data on the DAS, there is no way to determine many of the changes in the cabin that can influence the sensor and further mask unimbibed alcohol (e.g., HVAC settings and window position).

Detecting BrAC Differences

The primary purpose of the alcohol-detection algorithm was to provide a yes or no estimate of whether or not an impaired individual was within a vehicle. Substantial benefits emerge from this binary rating. However, it was also worth examining whether the alcohol sensor could be used to estimate level of intoxication. Results from the gold standard data set revealed a moderate negative correlation between BrAC and alcohol sensor values. Unfortunately, while readings within a sensor could reflect level of intoxication (albeit with a large margin of error), these readings were not consistent across sensors and could not be interpreted to measure BrAC.

The potential remains for a rough calibration of each sensor individually to attempt concentration assessment. This would require fairly accurate estimates of intoxication within a vehicle and using that as a basis for interpreting other readings on the same sensor. This could potentially be achieved by calculating the average alcohol sensor reading per trip and standardizing those readings across trips for a given alcohol sensor. From this information, a trip could be assessed on the basis of the standard deviation of the alcohol sensor reading for that trip compared with other trips in the same vehicle (i.e., the same sensor). While this may potentially provide a very rough estimate of intoxication, it remains uncertain if such an approach would be feasible or functional.

Ultimately, other challenges with calculating BrAC made this calibration unrealistic, and future research should not consider trying to calculate BrAC from the SHRP 2 database. For example, the absolute value of sensor readings could change depending on the size and position of the intoxicated individual, the number of intoxicated individuals, the presence of other sources of alcohol within the vehicle (i.e., unimbibed alcohol), air circulation, windows being up or down, humidity, vehicle cabin volume, and a number of other factors. It is not feasible to assess BrAC with reasonable accuracy from the SHRP 2 sensor. Imbibed alcohol detection should be thought of as a binary classification rather than as a measure of concentration.

The Algorithm's Ability to Detect Moderately Impaired Drivers from Baseline Driving

Ultimately, the alcohol-detection algorithm must be accurate at determining the presence of alcohol in SHRP 2 trips. Since this is naturalistic data, it is unknown how much alcohol, if any, someone has consumed. The ground truth measure used for assessing alcohol consumption came from video review of the drivers. It was difficult to confirm alcohol impairment in SHRP 2 trips via video review when there was potentially minor impairment with few or no behavioral cues. Therefore, some trips may have been misclassified in the test data set as "no imbibed alcohol" when in fact imbibed alcohol was present but the video reduction team could not see it.

In 91 trips, the driver was judged to be moderately impaired. These trips were combined with the 97 control trips to assess how well the algorithm would do at differentiating moderately impaired trips from "normal" driving. In this case, the algorithm performed quite well, indicating that the sensor and an associated impairment algorithm could be used to identify trips in which moderate impairment was likely. In general, the algorithm appeared to perform better as observable impairment increased. This result supports running an algorithm across the SHRP 2 data to isolate potential imbibed alcohol use. However, false alarms due to other alcohol sources appear to be quite common in the data set and will continue to be a problem. Therefore, running this algorithm and accepting a positive result for imbibed alcohol without further verification is not recommended.

Recommendations

These results suggest many future directions for research and provide insight into future use of the alcohol sensor with SHRP 2 data. Several key recommendations follow.

Recommendation 1: Consider the Following Criteria in an Alcohol-Detection Algorithm for Use with the SHRP 2 Data

Any algorithm that is developed for use on the alcohol sensor should consider the following criteria:

- There is large variance in the warm-up period for this sensor before it begins to produce stable readings. This can last as long as 2 minutes.
- An artifact in the data effectively creates a sensor shadow (e.g., 15 mV offset) that should be filtered out or ignored.
- The absolute value of sensor readings can potentially vary from sensor to sensor. Nominally, a threshold of 3,965 mV should be established for the alcohol-detection algorithm (assuming no unimbibed alcohol). A more sensitive approach would be to obtain a stable baseline from each sensor.
- Unimbibed alcohol should be considered in reducing false positives. A quick change in alcohol sensor readings (i.e., steep slopes) is often associated with unimbibed alcohol.
- Many factors can influence this sensor, including temperature change, air flow and circulation, and number of passengers.

Recommendation 2: Broadly Use the Alcohol-Detection Algorithm to Find Impaired Trips

The applied alcohol-detection algorithm was not without error, yet it regularly performed better than chance across multiple research efforts and data sets. Indeed, the sensitivity was over 90% across all of the various data sets and approaches. This suggests that the alcohol sensor data are useful at identifying alcohol-impaired trips and providing an initial indicator of the potential for alcohol involvement within a trip. The SHRP 2 data set provides a unique glimpse into alcohol-impaired driving. Considering the high hit rate of the alcohol-detection algorithm, it could be of the utmost importance in gleaning information on impaired driving from the SHRP 2 database. Identified trips can then be further explored using trained data reductionists to differentiate driver impairment from other potential unimbibed alcohol sources.

Recommendation 3: Always Accompany Use of the Algorithm with Trained Data Reductionists

While the alcohol-detection algorithm performed well at identifying impaired trips, many known barriers restricted its accuracy. In particular, sources of unimbibed alcohol can produce significant errors in interpreting the alcohol sensor readings. These sources appear to be quite common in the SHRP 2 database. Many of these sources of unimbibed alcohol can be visually identified by a trained data reductionist. In addition to finding sources of unimbibed alcohol, highly trained data reductionists can and should be used to validate the results of any alcohol-detection algorithm using the SHRP 2 alcohol sensor. The accuracy of the results of an algorithm could be enhanced with confirmation by trained data reductionists. The misclassification of other alcohol sources as intoxication by the algorithm might provide erroneous conclusions regarding alcohol-impaired driving in the SHRP 2 database if trips are not visually validated by specially trained data reductionists.

Recommendation 4: Do Not Disregard the Impact of Unimbibed Alcohol Presence

Much of this report discusses unimbibed alcohol as a primary source of error in the alcohol-detection algorithm. However, some of these substances (e.g., cigarettes, fast food, hand sanitizer) may involve distracted driving that has an impact on driver behavior and performance. Thus, a secondary benefit of the alcohol-detection algorithm may be to identify other substances that may also affect driving performance.

Overall Conclusion

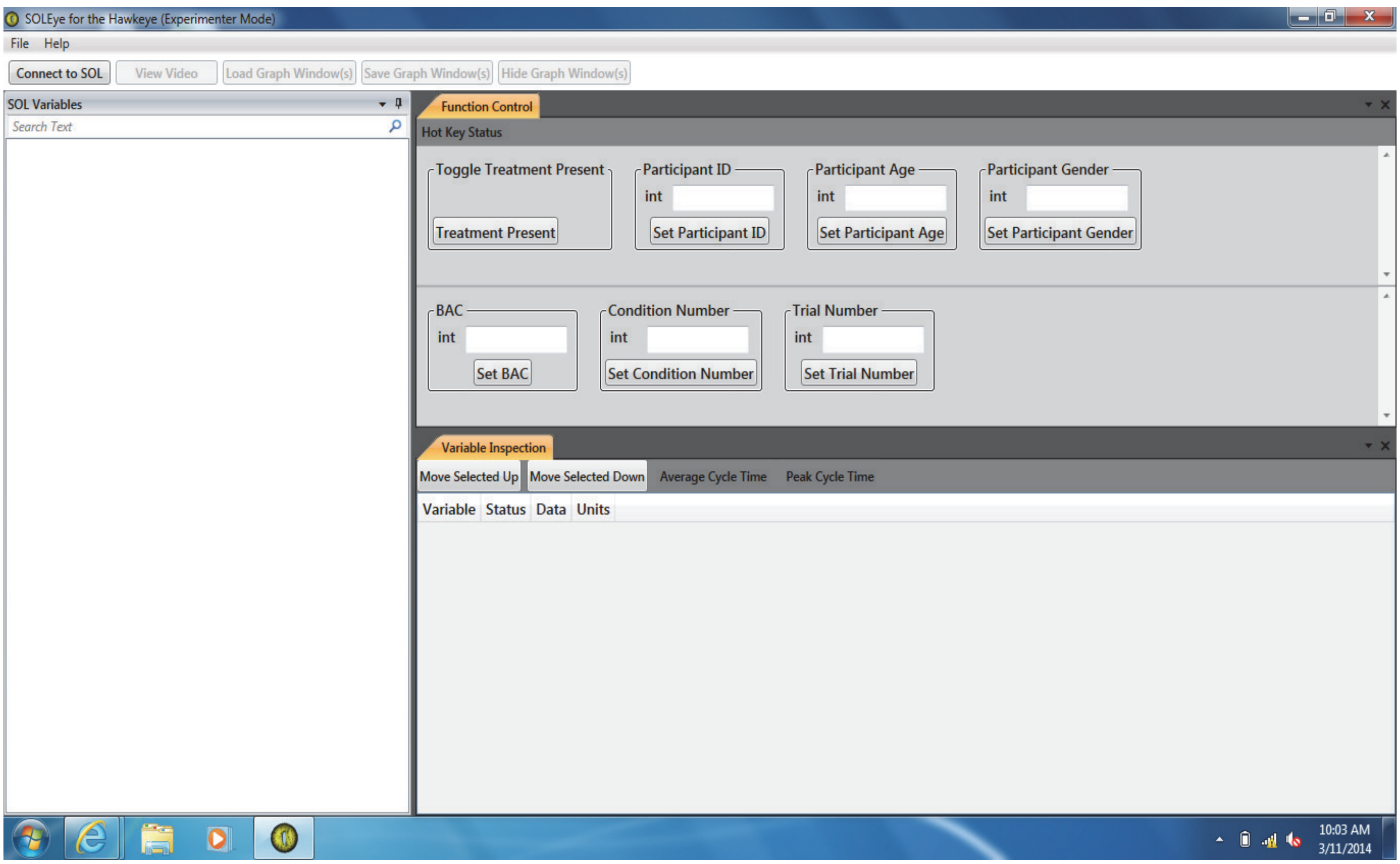
The alcohol sensor of the standard SHRP 2 instrumentation was designed to detect alcohol vapors within a cabin. At the beginning of this effort, it was uncertain if the alcohol sensor accurately performed this function. Assuming the function was fulfilled, it was even more uncertain if alcohol sensor readings could reliably differentiate imbibed versus unimbibed alcohol within a vehicle. To answer these questions, the mechanical breather and SHRP 2 data sets were used to explore the accuracy of an alcohol-detection algorithm. Considering the scope and detail of the SHRP 2 data set, an alcohol-detection algorithm could shed valuable light on alcohol-impaired driving. However, many challenges remain for the broad implementation of an algorithm using this sensor. Despite its relatively high success rate, care should be taken when using an algorithm and this sensor. Other substances can result in false positives, and visual inspection of SHRP 2 data should almost always accompany the algorithm's application in scientific endeavors.

References

1. National Highway Traffic Safety Administration (NHTSA). 2012. *Traffic Safety Facts: Alcohol-Impaired Driving.* DOT HS 811 700. U.S. Department of Transportation.
2. Carpenter, J. A. 1962. Effects of Alcohol on Some Psychological Processes: A Critical Review with Special Reference to Automobile Driving Skill. *Quarterly Journal of Studies on Alcohol,* Vol. 23, pp. 274–314.
3. Rothe, P. J., and L. J. Carroll. 2009. Hazards Faced by Young Designated Drivers: In-Car Risks of Driving Drunken Passengers. *International Journal of Environmental Research and Public Health,* Vol. 6, pp. 1760–1777.
4. Sakar, S. A., M. Anderas, and F. de Faria. 2005. Who Uses Safe Ride Programs: An Examination of the Dynamics of Individuals Who Use a Safe Ride Program Instead of Driving Home While Drunk. *American Journal of Drug and Alcohol Abuse,* Vol. 31, pp. 305–325.
5. Harris, D. H., R. A. Dick, S. M. Casey, and C. J. Jarosz. 1980. *The Visual Detection of Driving While Intoxicated: Field Test of Visual Cues and Detection Methods.* DOT HS-7-1538. U.S. Department of Transportation.
6. McKnight, J. A. 1991. Factors Influencing the Effectiveness of Server-Intervention Education. *Journal of Studies on Alcohol,* Vol. 52, No. 5, pp. 389–397.
7. McKnight, J. A., E. A. Langston, P. R. Marques, and A. S. Tippetts. 1997. Estimating Blood Alcohol from Observable Signs. *Accident Analysis and Prevention,* Vol. 29, No. 2, pp. 247–255.
8. McKnight, J. A., E. A. Langston, S. A. McKnight, and J. E. Lange. 2002. Sobriety Tests for Low Blood Alcohol Concentrations. *Accident Analysis and Prevention,* Vol. 34, pp. 305–311.
9. McKnight, J. A., and P. R. Marques. 1990. *Host and Server Determination of Alcohol Intoxication Level* (DOT HS 807 639). U.S. Department of Transportation.
10. Rubenzer, S. 2011. Judging Intoxication. *Behavioral Sciences & the Law,* Vol. 29, No. 1, pp. 116–137.
11. Smith, R. C., Z. Robinson, A. Bazdar, and E. S. Geller. In press. Intervening to Decrease the Probability of Alcohol-Impaired Driving: Impact of Novel Field Sobriety Tests. *Journal of Prevention and Intervention in the Community.*
12. Smith, R. C., and E. S. Geller. 2014. A Field Investigation of College-Student Alcohol Intoxication and Return Transportation from At-Risk Drinking Locations. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2425, pp. 67–73.
13. Smith, R. C., P. T. Coyle, C. Baldner, B. Bray, and E. S. Geller. 2013. A Field Investigation of Perceived Behavioral Control and Blood Alcohol Content: A Pattern-Oriented Approach. *Addictive Behaviors,* Vol. 38, pp. 2080–2083. doi:10.1016/j.addbeh.2013.01.003.

APPENDIX A

SOL-Eye User Interface Screenshot



APPENDIX B

Operational Definition of Behavioral Cues

Category	Behavior	Operational Definition
Appearance		
1. Eyes	Blinking (flitting)	Rapid blinking. Excessively frequent blinking.
	Lids, heavy	Prolonged, slow blinks. Drooping eyelids. Eyes half shut.
	Dozing	Similar to "nodding off." Eyes completely shut for prolonged periods of time. Head drops and jerks back up.
2. Hands	Shaking	Hands twitching/shaking. Hands unsteady.
3. Face	Flushed	Face is red. Face is blotchy or has dark spots on cheeks.
	Sweating	Beads of sweat on face. Hair appears wet.
	Drooling	Saliva appears around mouth. Drool may run down face. Driver frequently wipes mouth.
4. Hair	Disheveled	Hair disorganized. Hair sticking up or ruffled. Hair not neat.
5. Clothes	Shirt unkempt	Shirt partially untucked. Shirt wrinkled excessively. Shirt buttoned incorrectly. Collar turned up or partially turned up. Shirt generally fitting or being worn improperly.
	Loosening/taking off	Shirt unbuttoned or partially buttoned. Sleeves rolled up. Untucking or loosening shirt. Removing articles of clothing. Missing articles of clothing.
6. Body tremors/ shaking	Body tremors/shaking	Body is shaking uncontrollably. Slight shaking of the body. Body is experiencing tremors.
7. Sensation	Rubbing head (like for a headache)	Driver is using hands to rub the face or temples as if they have a headache. Massages temples, face, or back of neck.
	Rubbing face	Rubs face for a second or more. Does not include brushing objects off of face.
	Dizzy/swaying/leaning against window	Body appears to be uncontrolled. Head spinning. Body leaning to one side or not upright. Body leaning against window.
	Nauseous	Driver appears to be on the verge of getting sick. May make motion as if to throw up or gag. Covers mouth with hand like they are about to vomit. Face appears sick.
8. Memory	Repetitive action	Engages in any repetitive action. Examples: Playing with hair, adjusting clothing, rubbing eyes, etc.
	Gets lost	Looks around as if not knowing where s/he is.
Affective		
9. Nervousness	Nervous	Driver looks uncomfortable or nervous.
	Restless	Driver seems fidgety. Frequent movements. May seem uncomfortable.
	Agitated	Appears annoyed. May seem short with passengers. Driver may display frustrated facial emotions or give rude hand gestures.
	Relaxed	Driver appears overly calm.

Category	Behavior	Operational Definition
10. Mood	Crying	Face wet from tears. Wipes eyes.
	Exhilarated	Driver seems full of energy. Bouncing in seat. Drumming on steering wheel. Singing to radio. Dancing in seat.
	Rapid changes in mood	Driver experiences a multitude of emotions in a very short period. Emotions quickly fluctuating.
	Hostile	Angry look on driver's face. Driver makes rude or aggressive gestures. Driver appears to be yelling. Driver is forceful with gestures.
	Distracted	Not paying attention to the road. Excessively looking around in and/or outside the vehicle.
	Extremely friendly	Driver may appear overly talkative, use excessive or exaggerated hand gestures, or have an unusually open body posture.
	Talkative	Driver is constantly talking.
	Sexually aggressive	Excessive touching, forced contact, or strong sexual gestures.
	Confused	Driver does not appear to know what is going on or where s/he is. Has confused look on face. Looking around as if to gain clues.
Motor		
11. Coordination	Posture/can't sit up straight	Slumped in seat. Leaning to one side of seat. Driver attempts to sit up straight but fails.
	Drops/spills/knocks things over	Attempts to reach for something and instead knocks it over. While holding something, drops it without intent.
12. Nervous system	Hiccups	Determined by throat constricting, shoulders abruptly moving up and then down.
	Belching	Determined by mouth open, head moving forward.
	Vomiting	Throwing up.
	Seizures, convulsions	Violent jerking motions of entire body.
	Asleep	Eyes completely closed with body relaxed.
	Breathing fast	Frequent and quick rising and dropping of chest and/or shoulders.
	Yawning	Prolonged wide open mouth.
	Stupor	Driver has blank glance. Appears "zoned out." Not paying attention to surroundings. Steady forward gaze without focusing eyes or scanning driving scene.

Related SHRP 2 Research

- Naturalistic Driving Study: Development of the Roadway Information Database (S04A)
- Design of the In-Vehicle Driving Behavior and Crash Risk Study (S05)
- Naturalistic Driving Study: Technical Coordination and Quality Control (S06)
- Naturalistic Driving Study: Collecting Data on Cell Phone Use (S06)
- Naturalistic Driving Study: Field Data Collection (S07)
- Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk (S08A)
- Analysis of Naturalistic Driving Study Data: Offset Left-Turn Lanes (S08B)
- Analysis of Naturalistic Driving Study Data: Roadway Departures on Rural Two-Lane Curves (S08D)
- Naturalistic Driving Study: Descriptive Comparison of the Study Sample with National Data (S31)
- Naturalistic Driving Study: Linking the Study Data to the Roadway Information Database (S31)