

November 2021

Walking the Line: Understanding the History and Development of the NHTSA Standardized Field Sobriety Tests

Jon D. Gurney
University of South Florida

Follow this and additional works at: <https://digitalcommons.usf.edu/etd>

 Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Scholar Commons Citation

Gurney, Jon D., "Walking the Line: Understanding the History and Development of the NHTSA Standardized Field Sobriety Tests" (2021). *USF Tampa Graduate Theses and Dissertations*.
<https://digitalcommons.usf.edu/etd/9677>

This Dissertation is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact scholarcommons@usf.edu.

Walking the Line: Understanding the History and Development of the NHTSA Standardized
Field Sobriety Tests

by

Jon D. Gurney

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Education in Educational Program Development
with a concentration in Educational Innovation
Department of Language, Literacy, Ed.D., Exceptional Education, and Physical Education
College of Education
University of South Florida

Major Professor: Elizabeth Shaunessy-Dedrick, Ph.D.
J. Howard Johnston, Ph.D.
Mary Beth Corace, Ph.D.
John Mann, Ed.D.

Date of Approval
November 4, 2021

Keywords: drunk-driving, alcohol impairment

Copyright © 2021, Jon D. Gurney

Dedication

I dedicate this work to my mother, Frances Gurney, and to my sister, Claire Gurney.

Acknowledgments

I would like to thank Dr. Elizabeth Shaunessy-Dedrick for all of her guidance and support throughout my doctoral program. Elizabeth, as my major professor, you made the doctoral journey an absolutely wonderful experience. Thank you for being a mentor and a friend! Additionally, I would like to thank the other members of my doctoral committee: Dr. Howard Johnston, Dr. Mary Beth Corace, and Dr. John Mann. I am so appreciative of your valuable thoughts and perspectives and remain ever so grateful to each of you for serving on my doctoral committee.

Table of Contents

List of Tables	iii
Abstract.....	iv
Chapter 1: Introduction	1
The Crime of Driving Under the Influence of Alcohol (DUI) in the United States	1
Conceptualization of the Moral Entrepreneur.....	3
Statement of the Research Problem	6
Chapter 2: Literature Review.....	7
The Formulation of the Standardized Field Sobriety Tests	7
The Horizontal Gaze Nystagmus Test	7
The Walk and Turn Test	8
The One Leg Stand Test	8
Field Sobriety Test Research Funded by NHTSA.....	8
Chapter 3: Methodology	15
Study Questions	15
Defining Exploratory Research.....	15
Document Analysis as a Qualitative Method.....	15
Document Analysis as a Stand-Alone Qualitative Research Method.....	16
Advantages and Limitations of Document Analysis	17
The Document Analysis Process	17
Reflexivity Statement.....	18
Data Collection	20
Chapter 4: Discussion	21
Proving the Crime Driving Under the Influence of Alcohol	21
Use of Video Recording in the DUI Investigation Process	22
Research Findings Supportive of the NHTSA Standardized Field Sobriety Tests.....	23
Burns and Moskowitz (1977) Laboratory Study	23
Tharp et al. (1981) Laboratory Study	24
McKnight et al. (2002) Laboratory Study.....	25
Tharp et al. (1981) Field Study	25
Anderson et al. (1983) Field Study	26
Burns and Anderson (1995) Field Study	27
Stuster and Burns (1998) Field Study.....	28
Burns and Dioquino (1998) Field Study.....	29
Resources Demonstrating Criticism of NHTSA Standardized Field Sobriety Tests.....	32
Judicial Interpretations of the NHTSA Field Sobriety Tests	40

The Competing Interests in a DUI Investigation	44
Findings.....	46
Chapter 5: Summary and Recommendations.....	50
Putting It All Together	50
Recommendations for Future Research & Policy Development	51
References.....	56
Appendix A: Reflection	60
Appendix B: DUI Statutes	62

List of Tables

Table 1: Summary of Early Studies Demonstrating Support for NHTSA Standardized Field Sobriety Tests	31
---	----

Abstract

In the early 1980's within the United States, state legislatures criminalized the practice of driving under the influence of alcohol due to the efforts of moral entrepreneurs such as Mothers Against Drunk Driving that served to vilify the drunk-driver. During the political culture of the early 1980's, the National Highway Traffic Safety Administration sought to develop field sobriety tests that would assist law enforcement in the detection and arrest of drunk drivers.

This study provided a brief historical account of the development of the NHTSA Standardized Field Sobriety Test battery, which consisted of the Horizontal Gaze Nystagmus test, the Walk and Turn test, and the One Leg Stand test, as a tool used in the investigation and prosecution of suspects charged with the crime of driving under the influence of alcohol.

The study utilized document analysis in order to demonstrate (1) how the NHTSA Standardized Field Sobriety Tests were pilot tested for use in driving under the influence of alcohol cases and (2) to explore various criticisms levied against the development and use of the NHTSA Standardized Field Sobriety Tests for the investigation of the crime of driving under the influence of alcohol.

This study concluded that although many of the early government funded pilot studies suggested that the NHTSA Standardized Field Sobriety Tests were valid at predicting a breath alcohol concentration at or above the legal limit, there are multiple criticisms surrounding these tests that call into question the validity of these tests when used in the investigation of driving under the influence of alcohol offenses. Brief recommendations for future research and policy development in the investigation of driving under the influence of alcohol cases were provided.

Chapter 1: Introduction

The Crime of Driving Under the Influence of Alcohol (DUI) in the United States

Concern over drinking and driving in the United States has come to the forefront of political circles in recent times. Like many other social problems, the extent to which the government reacts by punishing a particular behavior is largely a product of the pressure placed on government by its citizens. The case of drinking and driving is no exception. In the United States during the early 1980's, an anti-drunk-driving movement came about as a product of citizen discontent with regard to the manner in which government was handling cases of drunk-driving (Reinarman, 1988).

Prior to the early 1980's, drunk-driving was regarded as a "folk crime." A folk crime refers to an offense that is routinely committed by many a "good citizen" (Gusfield, 1981; Reinarman, 1988). However, in the early 1980's this conceptualization began to change as a result of a sentiment of public outrage towards drunk-drivers. This sentiment was channeled to the government via advocacy groups at multiple levels in an effort to form a collective voice and effectuate change in policies and laws governing drunk-driving cases in the United States.

While there are certainly many advocacy groups that can take credit in the strength behind the anti-drunk-driving movement, Reinarman (1988) posits that one anti-drunk-driving group stands out above all the rest. Mothers Against Drunk-Driving (referred to as MADD) formed in 1980 as a collective voice for those who have lost loved ones as a result of an automobile accident caused by a drunk-driver. He suggests that the political clout possessed by

MADD can be attributed to its successful role in the anti-drunk-driving movement in the early 1980's in the United States.

There are two factors which led to the success of MADD as an effective anti-drunk-driving advocacy organization. First, Reinerman (1988) stated that it is the credibility of the claims made by its members. The fact that they are speaking publicly following the loss of a loved one in an accident caused by a drunk driver attests to the sincerity of their message that is being communicated. Secondly, he suggested that the historical context of the early 1980's played a role in terms of the acceptability of MADD's message. The early 1980's was a period of time in the United States when conservative views resonated, and the anti-drunk-driving message espoused by MADD fell on willing ears.

The political culture of the early 1980's provided fertile ground for an anti-drunk-driving movement. As noted earlier, the political claims of MADD existed in accord with the prevailing political views of the time. At a time when the criminal justice system took a "get tough" stance on crime and the rights of the victim were of increased importance politically, MADD's vilification of the drunk driver came at a perfect time. As part of the message, the very term "MADD" appealed to a sense of moral anger. MADD members used the media to rally against the "killer drunk" and publicly criticized government as being neglectful of "drunk-driving -the nation's most frequently committed violent crime and only socially acceptable form of homicide." (Reinerman, 1988, p. 105).

Given the political clout of the anti-drunk-driving movement in the early 1980's in the United States, organizations such as MADD took on a unique role. With the power of their anti-drunk-driving message MADD took on a cause that few could argue against. In an attempt to vilify the drunk-driver responsible for the death of a loved one, MADD's ideology was

understood in many political circles and gained stronger support. Such an ideology appealed to politicians because the vilification of the drunk driver while championing the rights of the victim was a cause that could be exploited politically at the time.

Conceptualization of the Moral Entrepreneur

Howard Becker (1963) coined the term “moral entrepreneur” to describe those who interact within a larger context to take the initiative and generate a moral enterprise. During this process, Becker explains, the moral entrepreneur will act by either serving as a moral crusader or as a rule enforcer. The identification of MADD as a moral entrepreneur seems appropriate in that it became an organization that crusaded against the drunk-driver and portrayed such person as one who threatens to take human life. This moral entrepreneurial appeal movement combined with the historical context of the early 1980’s meshed in “a peculiar conjecture of trends favorable to claims in both the alcohol arena and the larger political culture” (Reinarman, 1988, p. 101).

The image portrayed by MADD had much to do with its success in the anti-drunk-driving movement. As “the voice of the victim” Reinarman (1988, p. 105) suggests that the message of MADD lent itself well to a narrow and conservative focus. This appealing message, combined with the law-and-order political climate of the day, called for an end to plea bargaining and judicial discretion. Newspaper accounts of MADD often cite the national punishment mentality in the very punitive era of the 1980’s as a factor in the organization’s rise (Cohn, 1984).

The MADD agenda has resulted in multiple measures that are designed to increase the levels of punishment for one convicted of drunk-driving.¹ These measures, while beginning in

¹ Some jurisdictions employ the term DUI (Driving Under the Influence), while other jurisdictions use the term DWI (Driving While Impaired). Both of these are statutory constructions of the crime of drunk-driving and carry similar penalties.

the 1980's, have carried on over time to influence law makers across the United States.

According to a national newsletter published by MADD, these measures include the elimination of plea bargaining in DUI cases, reduction of the blood alcohol content (BAC) standard from .10 to .08, mandatory implied consent laws for all drivers to provide a breath sample while being investigated for driving under the influence, "per se" license suspensions on the scene for drivers failing to consent to a blood alcohol content test, mandatory jail sentences, higher minimum fines, pre-sentence probation investigations of driving records, and roadblocks/sobriety checkpoints in order to stop vehicles randomly (Reinarman, 1988).

Since its inception MADD grew to over 400 chapters by 1990 and, by the year 2000, consisted of over 600 chapters in the United States. MADD chapters are now located in all 50 states plus Guam, Puerto Rico and Canada. However, the success that MADD has attained with its anti-drunk-driving message does not fully explain why drunk-driving laws became harsher in the 1980's and have continued to become increasingly punitive. Reinarman (1988) emphasized the role of MADD as a moral entrepreneur in his account of changing laws about drunk-driving in the United States during the 1980's. However, factors other than moral entrepreneurs may have also played a role during the 1980s, the time when the use of field sobriety tests to prove drunk driving became institutionalized. Dickson (1968) reminds us that agencies pursue their own organizational interests in ways that affect law. The NHTSA, as an agency of the United States Department of Transportation, actually started developing information on drunk-driving detection in the mid 1970's before MADD was formed.

The organizational position of NHTSA arguably was enhanced when it began addressing drunk driving by formulating field sobriety tests because such tests had a direct effect on drunk-driving enforcement practices at the state and local levels. As part of the anti-drunk-driving

movement, the NHTSA launched a fortified attempt in 1981 to devise a standardized system of tests and procedures used by police officers to determine when a driver is impaired by alcohol. Such an attempt to devise a uniform battery of field sobriety tests represented the role of the federal government in working closely with state and local jurisdictions across the United States in order to arrest and prosecute drunk drivers. During the 1980s President Reagan was looking at federal agencies with an eye toward limiting federal government and helping the states. The field tests represented tangible ways in which NHTSA, a federal agency, could help local governments make a difference in protecting the public from drunk drivers. Consistent with the argument advanced by Dickson (1983) about marijuana and federal drug agencies, the attempt to standardize field sobriety tests is one way in which the NHTSA perpetuated its own cause and need for existence.

MADD's influence and NHTSA's organizational interests meshed within a larger conservative political climate represented by the Reagan administration. Included as part of that conservative climate were two other themes that help establish the context. One was the victim rights movement. Wisconsin enacted the first Victim Bill of Rights in 1979 and in the next 10 years 44 states passed similar legislation (Eastman, 2003). Recall that Reinerman (1988) noted that MADD became the voice of the victims; it was organized by mothers of victims of drunk driving. The second was the overall "get tough" approach to crime that resulted in changes in sentencing schemes and tougher sentences for many kinds of offenders in addition to drunk drivers. For example, starting in the 1980s almost all legislatures amended their juvenile codes to "get tough" on juvenile offenders (Bishop et al., 1998). MADD, the NHTSA, and other forces converged in a symbiotic relationship wherein each made contributions to campaign against drunk driving.

Statement of the Research Problem

This study examined the social climate surrounding drunk driving as reflected in the anti-drunk-driving movement in the United States during the 1980's. Next, the study examined the role of the United States Department of Transportation - National Highway Traffic Safety Administration (NHTSA) in establishing research findings about drunk driving and developing practices designed to enhance public safety, specifically field sobriety tests used to establish drunk driving. The NHTSA performed much of the groundwork for introducing standardized field sobriety tests in the 1981. It was instrumental in developing the field sobriety tests, disseminating information about them, and encouraging their adoption by state and local law enforcement agencies. However, over the years, the NHTSA standardized field sobriety test battery have all received a fair amount of criticism. Both support for and criticism of the NHTSA standardized field sobriety test will be explored.

This study sought to provide an analysis of the NHTSA DUI field sobriety tests through asking the following four research questions:

- (1) How did the NHTSA DUI Field Sobriety Tests originate?
- (2) How were the NHTSA DUI Field Sobriety Tests developed and pilot tested?
- (3) How have the NHTSA DUI Field Sobriety Tests been evaluated since they were initially developed and piloted?
- (4) To what extent are the NHTSA Field Sobriety Tests supported by empirical evidence?

Chapter 2: Literature Review

The Formulation of the Standardized Field Sobriety Tests

The NHTSA, in 1981, developed a battery of standardized field sobriety tests which consist of the following three tests: (1) Horizontal Gaze Nystagmus, (2) The Walk and Turn, and (3) The One Leg Stand. The results of these tests serve as a basis to establish probable cause for an arrest for the crime of driving under the influence of alcohol. Each of the field sobriety tests will be explained below briefly utilizing their description found in the NHTSA standardized field sobriety test training program management system (NHTSA, 2001).

The Horizontal Gaze Nystagmus Test

The Horizontal Gaze Nystagmus (HGN) is an involuntary jerking of the eye that occurs naturally as the eyes gaze to the side. Under normal circumstances, nystagmus occurs when the eyes are rotated at high peripheral angles. However, when a person is impaired by alcohol, nystagmus is exaggerated and may occur at lesser angles. An alcohol impaired person will also often have difficulty smoothly tracking a moving object. In the HGN test, the officer observes the eyes of a suspect as the suspect follows a slowly moving object such as a pen or a small flashlight, horizontally with his or her eyes. The examiner looks for three indicators of impairment in each eye: (1) if the eye can not follow a moving object smoothly, (2) if jerking is distinct when the eye is at a maximum deviation, and (3) if the angle of onset of jerking is within 45 degrees of center (NHTSA, 2001).

The Walk and Turn Test

The subject is directed to take nine steps, heel to toe, along a straight line. After taking the steps, the suspect must turn on one foot and return in the same manner in the opposite direction. The examiner looks for eight indicators of impairment: (1) if the suspect cannot keep balance while listening to the instructions, (2) begins performing before the tests are finished, (3) stops while walking to regain balance, (4) does not touch heel to toe, (5) steps off the line, (6) uses arms to balance, (7) makes an improper turn, or (8) takes an incorrect number of steps (NHTSA, 2001).

The One Leg Stand Test

The suspect is instructed to stand with one foot approximately six inches off the ground and count aloud by thousands (one thousand-one, one thousand-two, etc.) until told to put the foot down. The officer times the subject for 30 seconds. The officer looks for four indicators of impairment, including (1) swaying and balancing, (2) using arms to balance, (3) hopping to maintain balance, and (4) putting the foot down (NHTSA, 2001).

Field Sobriety Test Research Funded by NHTSA

The NHTSA, in 1975, as part of its mission to decrease the number of alcohol related fatalities, contracted with researchers who rode along with police officers in the field in order to develop a Driving While Impaired detection guide that listed 20 driving cues and the probability that a driver exhibiting a cue would have a BAC of at least .10 percent (Harris et al., 1980). This endeavor was an attempt to develop a list of possible tests that accurately detect alcohol impairment.

In 1977, the NHTSA began to sponsor laboratory evaluations of several of the tests that were being used inconsistently by law enforcement officers as part of their drunk-driving

enforcement efforts based on the findings of the 1975 ride along evaluations (Burns & Moskowitz, 1977). The selected tests were evaluated in laboratory experiments using 238 test subjects and 10 police officers who evaluated the subjects using the tests and had to decide whether subjects should be arrested or released had the tests been performed at roadside, assuming a legal threshold of .10 BAC as the basis for arrest (Burns & Moskowitz, 1977). Such laboratory evaluations at that time included the finger to nose, walk and turn, finger count, finger tracing, and the one leg stand (NHTSA, 2001). In addition, the researchers began preliminary evaluations of the Horizontal Gaze Nystagmus test based on the reliability it was found to have when predicting BAC in a study conducted in Finland (Pentilla et al., 1971).

The subjects were all alcohol consumers and were instructed not to eat for four hours prior to the experiments. They were given measured doses of alcohol such that they would have BACs ranging from 0 to .15%, but the tests subjects did not know the amount of the dose each received. Officers had to administer the test package and determine if the person should likely be arrested for having a BAC at or above .10%. (Burns & Moskowitz, 1977).

According to the NHTSA, all of the field sobriety tests evaluated in the 1977 Burns and Moskowitz study were found to be sensitive to BAC in varying degrees in laboratory conditions, but the walk and turn, one leg stand, and the horizontal gaze nystagmus tests were the most effective at correlating blood alcohol levels of .10 or more (NHTSA, 2001). The goal of the 1977 NHTSA study was to evaluate currently used physical coordination tests to determine their relationship to alcohol impairment and to develop more sensitive tests that would provide more reliable evidence of alcohol impairment (Burns & Moskowitz, 1977). As a result, the one leg stand, walk and turn, and the horizontal gaze nystagmus were selected by the NHTSA for further review.

Following the Burns and Moskowitz (1977), the NHTSA conducted another study in 1981 in an effort to standardize the test administration and scoring procedures and conduct further laboratory and field evaluations of the new battery of the three tests. The results of this study were compiled and documented in a technical report entitled, *Development and Field Test of Psychophysical Tests for DWI Arrests* (Tharp et al., 1981). The data in this study showed that in 77% of the cases the Horizontal Gaze Nystagmus could predict a BAC of .10 or higher. The Walk and Turn test could predict a BAC of .10 or higher in 68% of the cases and the One Leg Stand could predict a BAC of .10 or higher in 65% of the cases. The three tests, when combined, could predict a BAC of .10 or higher in 81% of the cases (NHTSA, 2001). Shortly after the release of this study, the NHTSA states that the study was cited across the United States to establish “scientific validity” of the field sobriety test battery and to support officers’ testimony in court, despite the fact that, at best, the tests were accurate at predicting a BAC of .10 in only 81% of the cases.

During the 1980’s law enforcement officers began using the “standardized” field sobriety tests at roadside to help determine whether motorists suspected of drunk driving have blood alcohol concentrations greater than .10 percent (NHTSA, 2001). This was an attempt to provide uniformity in the way in which drivers suspected of DUI would be tested.² Prior to this time, various tests were employed by law enforcement agencies when investigating suspects of driving under the influence. Given that some of these tests did not become part of the standardized field sobriety tests employed by the NHTSA, many law enforcement agencies converted to the NHTSA field sobriety tests given the fact that they were deemed to be “standardized.” The

² The adoption of NHTSA standardized field sobriety tests was done in a piece-meal fashion by law enforcement agencies across the United States. This has been an evolutionary process across jurisdictions and such tests may not be used by every law enforcement agency in the United States.

assumption is that such “standardization” provided validity to the field sobriety test administration process.

Since 1981, many states have implemented laws that define as DUI a BAC below a .10. (It should be noted, however, that the legal limit today is .08% in all states except the State of Utah where it is now .05%). NHTSA sponsored additional research after the 1981 development of the Standardized Field Sobriety Tests in an attempt to further evaluate the accuracy of the standardized field sobriety test battery to discriminate a BAC above or below a .08 and above or below a .04 percent (NHTSA, 2001).

The NHTSA sponsored additional validation studies in Colorado, San Diego, and Florida in 1995, 1997, and 1998 respectively (Burns, 2003). The goal in each of these studies was to ascertain the accuracy of an officer’s determination of driver impairment based upon the NHTSA Standardized Field Sobriety Tests. Unlike the 1981 study, officers in the subsequent validation studies were specifically selected based on criteria set forth by the NHTSA.³ The officers in the subsequent validation studies belonged to DUI/DWI enforcement units within their respective law enforcement agencies. Additionally, each officer possessed years of experience and training in the administration of the NHTSA Standardized Field Sobriety Tests prior to their participation in the subsequent validation studies (Stuster & Burns, 1998).

The following account will discuss the differences between the 1981 NHTSA sponsored study and the subsequent standardized field sobriety test (SFST) studies conducted in 1995, 1997, 1998 NHTSA sponsored study in Colorado, San Diego, and Florida, respectively. In each of these three studies, actual traffic stops using the SFSTs were investigated. As noted previously,

³ The fact that the 1998 validation study used officers’ skilled in DUI detection, while the 1981 validation study did not, certainly requires one to question whether a comparison of these studies has anything to do with the actual validity of the field sobriety tests, or rather more to do with the skill level of the officer to detect DUI.

officers had previous training and experience in administering standardized field sobriety tests (in the Florida study, all 16 were SFST instructors) and received "refresher" training before beginning data collection (Rubenzer, 2008).

In the Colorado and Florida studies, observers from the study (either researchers or participating police officers) monitored about half of the stops to ensure they observed the study protocols (no use of portable breath tests or PBTs until after the field sobriety tests were given and scored) and that the standardized field sobriety tests were administered correctly.

In the Colorado and Florida studies, researchers obtained PBTs on the majority of drivers who were tested but released. This allowed an estimate of false negatives-failures to make an arrest when warranted. The different studies investigated the standardized field sobriety test performance at determining BAC levels of .05% and .08%. All three studies reported that correct arrest decisions based on the SFSTs exceeded 90% (Rubenzer, 2008). In an effort to draw a comparison between the 1981 NHTSA study, the 1998 San Diego study will be discussed below.

Using many of the same researchers that were employed in the previous NHTSA funded study in 1981, the results in the San Diego study found the Standardized Field Sobriety Tests to be extremely accurate (NHTSA, 2001). In the San Diego validation study, the officers' estimates were correct 91 percent of the time when estimating a motorist's BAC as being above or below a .08 and estimates of whether a motorist's BAC was above a .04 but under a .08 were accurate in 94 percent of the decisions to arrest (Stuster & Burns, 1998).

It is important to point out the differences in officer determination of impairment across the 1981 and the 1998 validation studies. Tharp et al. (1981) stated that the percentage of correct decisions by officers for determining driver impairment at or above the .10 BAC level was 77% for the Horizontal Gaze Nystagmus test, while in the 1998 San Diego site study by Stuster and

Burns the percentage of correct impairment determinations at the .08 BAC level was increased to 88%.

For the Horizontal Gaze Nystagmus test in the 1981 study, the percentage of correct decisions for determining driver impairment at or above the .10 BAC level was 77%, while in the 1998 San Diego sample there was an increase to 79% correct in determining driver impairment at or above the .08 BAC level. For the Walk and Turn test in the 1981 study, the percentage of correct decisions for determining driver impairment at or above the .10 BAC level was 68%, while in the 1998 San Diego sample there was an increase to 88% correct in determining driver impairment at or above the .08 BAC level. For the One Leg Stand test in the 1981 study, the percentage of correct decisions by officers for determining driver impairment at or above the .10 BAC level was 65%, while in the 1998 San Diego sample there was an increase to 83% correct in determining driver impairment at or above the .08 BAC level (NHTSA, 2001). Thus, when the BAC level was lowered under law from .10 to .08, the predictive ability of the officers in the follow up study increased.

Clearly, there was an alleged increase in the accuracy of officers' determination of impairment when the 1981 sample is compared to the 1998 sample. The NHTSA cites the most salient difference between the results of the 1981 study and the 1998 validation study as being the substantial increase in accuracy of the officers' decisions, despite the lower BAC criterion of .08 in the 1998 study. The greater accuracies of the standardized field sobriety battery and component tests during the 1998 study are attributable to the differential experience of the officers who participated in the two studies.

The officers who participated in the original research study in 1981 learned the procedures as part of the 1981 laboratory study. In contrast, the officers who participated in the

1998 study each used the NHTSA Standardized Field Sobriety Tests for several years to help make arrest decisions under operational conditions (NHTSA, 2001). Given this fact, the NHTSA states that the levels of accuracy observed during the 1998 study reflect current conditions and should be considered the validated measures of standardized field sobriety test accuracy.

Chapter 3: Methodology

This study answered the four questions set forth below:

Study Questions

- (1) How did the NHTSA DUI Standardized Field Sobriety Tests originate?
- (2) How were the NHTSA DUI Standardized Field Sobriety Tests developed and pilot tested?
- (3) How have the NHTSA DUI Standardized Field Sobriety Tests been evaluated since they were initially developed and piloted?
- (4) To what extent are the NHTSA Standardized Field Sobriety Tests supported by empirical evidence?

Defining Exploratory Research

From the onset, it is critical to note that this research endeavor was best classified as exploratory at its most basic core. As Babbie (2008) indicated, an exploratory study is one in which the researcher attempts to understand a particular subject or topic about which very little is known in order to satisfy the researcher's curiosity and desire for better understanding. Additionally, Babbie noted that exploratory research studies are quite valuable in social scientific research studies in that they yield new insights into a topic for future research.

Document Analysis as a Qualitative Method

In order to shed light on the history, development, and use of the NHTSA field sobriety tests in the United States, the qualitative method of document analysis that was employed in this research as the vehicle in which to facilitate this exploratory study. Document analysis is a

qualitative research method that is a systematic procedure for reviewing or evaluating documents. Like other analytical methods in qualitative research, document analysis requires that the data be examined and interpreted in order to elicit meaning, gain understanding, and develop empirical knowledge (Corbin & Strauss, 2008).

Document Analysis as a Stand-Alone Qualitative Research Method

As Bowen (2009) noted, document analysis has served mostly as a complement to other research methods, however, it also exists as a stand-alone method, whereby some forms of qualitative research rely solely on the analysis of documents. Accordingly, it is acceptable to engage in document analysis as a sole qualitative research method for a study. In assessing the value of document analysis as a qualitative research method, there can be several very specific uses for this methodology. Bowen (2009) set forth several uses of documents that may be beneficial to a social researcher. First, documents provide background information, as well as historical insight. This helps a researcher understand the historical roots of specific issues and can indicate the conditions that impinge upon the phenomena under investigation. Secondly, information found in documents may suggest some questions that need to be asked and situations that need to be observed. Third, documents provide supplementary research data in that information derived from documents may be valuable additions to a knowledge base. Fourth, documents provide a means of tracking change and development. In this regard, analyzing periodic or final reports of an organization may paint a picture of how an organization or program fared over time. Fifth, documents can be analyzed as a way to verify findings or corroborate evidence from other sources.

Advantages and Limitations of Document Analysis

Document analysis has several advantages in relation to other qualitative research methods. As stated by Yin (1994) document analysis is exact and provides broad coverage along different periods of time, places, and settings. Additionally, documents are stable in that the investigator's presence does not alter what is being studied, are not reactive to the research process, are efficient, and are a cost-effective method of data collection (Bowen, 2009; Merriam, 1988).

With respect to the limitations of document analysis as a qualitative research method, Yin (1994) points out that low retrievability and biased selectivity may be causes of concern to the social researcher engaging in document analysis. Biased selectivity may be due to an incomplete collection of documents within an organization or documents that only align with the agenda and principles of the organization. Additionally, low retrievability may become an issue when specific documents are either blocked or retrievability is difficult (Bowen, 2009).

The Document Analysis Process

Once a researcher selects document to analyze, the analytical process may take several forms. This involves skimming the document (superficial examination), reading the entire document (thorough examination), and finally, interpretation of the document (Bowen, 2009). Many researchers engage in thematic analysis of the documents to determine if there is a pattern recognition within the data, wherein emerging themes develop as categories for analysis (Fereday & Muir-Cochrane, 2006).

It is incumbent upon the researcher engaging in document analysis to look at documents with a critical eye. Documents should not be treated as being necessarily precise, accurate, or complete recordings of an event or phenomenon. The original purpose of the document and the

target audience of the document should be considered (Bowen, 2009). Thus, when engaging in document analysis, the researcher must keep all of these factors in mind and consider how these factors may affect the quality of the data.

Reflexivity Statement

This study was deeply rooted and entrenched in my past professional experiences as a practicing attorney and university instructor, teaching criminal justice and legal studies courses for over 15 years. As a former prosecuting attorney in two different states, I realize the significance of the topics addressed in this applied dissertation. This study shines light on two competing interests: the necessity of the government to enforce its drunk-driving laws for the safety of society versus the rights of those accused of the crime of drunk driving to be entitled to a fair and transparent investigation process, while all the while, ensuring that basic due process rights afforded under the constitution are preserved.

As a former prosecuting attorney, one might think that I inherently take a ‘get tough’ on crime position. While I do have a deep-rooted respect for the role of the criminal law in society, the law, as written, is only as good as the extent to which it is applied fairly and with transparency. I want to make clear that nothing within this study should lead one to the facile assumption that I support drunk-driving or seek to minimize the seriousness of the crime of driving under the influence of alcohol. Any characterization of my positions set forth in this study as supporting drunk-driving or minimizing the seriousness of this criminal offense are simply misguided and incorrect. What I do support, however, is a comprehensive DUI investigation process that is rooted in fairness and complete transparency. As a former prosecuting attorney, I can think of no ethical criminal prosecutor that would want anything less than a completely just investigatory and trial process.

This study does not look like a dissertation that one would see in a Doctor of Philosophy degree program. In short, that is because it is not and does not attempt to mirror a Ph.D. dissertation. A few years ago, when seeking to apply to doctoral programs that are designed for practicing professionals, I discovered the Doctor of Education degree. The Ed.D. is a degree that enables practitioners to take on a scholarly role in the development of a “problem of practice” within their respective professional settings. The most beautiful thing about the structure of the Ed.D. degree program is that it allows one to select a problem of practice and look for tangible, real-world solutions to the problem of practice. In understanding the importance of this structural framework, I sought to shed light on my selected topic so that all can read about this problem of practice with ease and, hopefully, hold a conversation with others about the history of the crime of DUI in the United States and the governmental response to it. If this occurs, then I have succeeded in my endeavor to educate others on this topic of vital importance.

In writing this applied dissertation, one of my dissertation committee members previously asked me a very important question about this topic. The question was – who is your target population for this study? In other words, who will benefit from reading it? My immediate response was- every person that drives a motor vehicle in the United States. I say this because, although an individual can guarantee that they will not operate a motor vehicle while under the influence of alcohol, they simply cannot guarantee that they will not be investigated and arrested for the crime of driving under the influence of alcohol. If the potential loss of liberty is not a critical problem of practice, I shudder to think of what is. It is my sincere hope that this study sheds light on a subject that otherwise might just exist in the shadows.

Data Collection

This study collected data via a systematic review of documentary sources, including the United States Department of Transportation NHTSA Standardized Field Sobriety Test agency reports from the 1970's to the present time, law review articles and judicial case law pertaining to the admissibility and use of field sobriety tests, along with books and scholarly articles relating to the crime of DUI. In this research endeavor, various data bases were utilized in order to attain documentary evidence. These data bases included JSTOR, Google Scholar, and LexisNexis. In conducting document searches on these data bases, various search terms were utilized in order to find relevant documents for this study. These terms included, moral entrepreneur, Mothers Against Drunk Driving, drunk driving, driving under the influence, driving while intoxicated, NHTSA Standardized Field Sobriety Tests, roadside sobriety tests, field sobriety test administration, field sobriety test validation, field sobriety test studies, field sobriety test case law, and field sobriety test judicial decisions.

This study provides an analysis of the NHTSA Standardized Field Sobriety Tests used by law enforcement in the investigation of driving under the influence of alcohol cases, including how and why they came into existence, how they are tested and evaluated since their inception, the extent to which the NHTSA Standardized Field Sobriety Tests are supported/ unsupported by empirical evidence, how courts across the United States have dealt with the admissibility of NHTSA Field Sobriety Test evidence, and finally, how all of this affects those that are investigated for the crime of driving under the influence of alcohol. Through an analysis of the documents indicated above, the goal of this study was to look for themes that emerge from the documents that were reviewed. The themes identified via this document analysis provided grounded guidance in answering the questions set forth within this study.

Chapter 4: Discussion

Introduction

The purpose of this chapter is to engage in a review existing documentary evidence relating to the development and use of the NHTSA Standardized Field Sobriety Tests. This included a review of resources that demonstrate support for use of the NHTSA Standardized Field Sobriety Tests, resources that contain criticisms of the NHTSA Standardized Field Sobriety Tests, and a brief review of judicial case law pertaining to the use of field sobriety tests in a court of law gathered from various United States jurisdictions. Through engaging in a document analysis of these source materials, general themes developed in this process. In short, this chapter bifurcates these themes into those that are supportive of the NHTSA DUI Standardized Field Sobriety Tests and those that demonstrate criticism of these tests.

Proving the Crime Driving Under the Influence of Alcohol

At the inception of this chapter, it is important to clearly point out that the prosecution of a Driving Under the Influence charge (DUI) typically occurs via two potential methods under existing state criminal law. A DUI charge may be brought forth by the state under two different statutory provisions. The first is typically referred to as the physical impairment method. This is where a DUI suspect is prosecuted by the state using evidence such as observations of the suspect vehicle in motion before the DUI traffic stop, evidence of the suspect's appearance and demeanor at initial contact with the officer, and the suspect's performance on the NHTSA Standardized Field Sobriety Tests. Using this method of prosecution, a state typically proves that a driver is less safe to drive or is impaired to the extent that his normal faculties are

compromised. The second method of DUI prosecution under most state criminal law is generally referred to as a *per-se violation*. This method is generally prosecuted via the use of a blood or breath alcohol concentration test which provides a specific numeric value reading, which if over the legal BAC limit, would be ground for criminal prosecution even in the absence of a demonstration of physical impairment (See Appendix A – The State of Florida DUI statute as an example).

This applied dissertation was not concerned with the mechanism for criminal prosecution under the per-se violation approach. To be sure, there are many legitimate concerns with the science of blood and breath testing in DUI cases. However, such a method of prosecution is beyond the scope of the discussion in this applied dissertation and will not be addressed. This dissertation examined literature pertaining to the development and use of the NHTSA Standardized Field Sobriety Tests in the investigation and prosecution of the crime of DUI under statutory provisions pertaining to physical and mental impairment of normal faculties due to alcohol, as opposed to a per-se method of DUI prosecution through the admission of blood or breath test results used to demonstrate alcohol concentration in blood or breath samples.

Use of Video Recording in the DUI Investigation Process

It is important to point out that, at the present time, only one state legally requires that the DUI investigation process be video-taped. Currently, the State of South Carolina is the only state that requires law enforcement to videotape the DUI investigation process. This requirement, pursuant to the applicable statute, requires video evidence that commences with the officer activating the blue overhead lights to initiate the traffic stop all the way through the arrest and administration of the breath test to the suspect (See Appendix B containing the South Carolina DUI video tape requirement statute). With limited exceptions indicated in the South Carolina

statute, failure of the law enforcement agency to provide such video will result in a dismissal of the DUI charge against the accused.

Research Findings Supportive of the NHTSA Standardized Field Sobriety Tests

Below is a brief recitation of the laboratory studies and field studies supporting the alleged validity of the NHTSA Standardized Field Sobriety Tests as set forth by Burns (2003), wherein a description of each study is provided.

Burns and Moskowitz (1977) Laboratory Study

This study was the precursor to the 1981 study that would lead to the development of the NHTSA Standardized Field Sobriety Tests in 1981. This study was initiated in 1975 to evaluate physical coordination tests that were being used by police officers and to develop tests that would provide more reliable evidence of impairment, with the eventual goal of standardizing the administration procedures for those tests.

An initial set of potential tests were piloted in the laboratory and six were found to most fully meet the test criteria. These six tests included the following: The One Leg Stand, Walk and Turn, Finger to Nose, Finger Count, Gaze Nystagmus, and a Tracing test. An additional set of alternate tests included the Romberg test, Subtraction, Counting Backwards, and Letter Cancellation. The study used ten traffic officers from city, county, and the highway patrol as examiners. Each officer was instructed in standard test administration and scoring during a four-hour training session and subsequently tested twenty to thirty subjects during two sessions.

The study was performed double blind with neither subjects nor examiners knowing the alcohol content of the individual drinks administered, the range of target BAC, or the measured BAC. Examiners did not observe or interact with the subjects prior to testing, as the subjects were sequestered in a private room while consuming alcoholic beverages. The examiner scored

each subject on a 0-10 scale with scores increasing as a function of poor performance. Examiners then made a decision to arrest/not arrest based upon the subjective belief that the subject's BAC was .10% or higher given that this was the statutory limit for arrest at the time the study was conducted.

In the study, officers recorded false arrests for 47 individuals with measured BAC concentrations below .10%. It was determined that 24 of those cases had a BAC of just below .10%. The results demonstrated that eighty-four percent of officers' arrest decisions, based on six tests, and seventy-three percent of their release decisions were correct. Using stepwise discriminant analysis, the research team determined that the best test battery consisted of three tests: Walk and Turn, One Leg Stand, and Nystagmus. The Nystagmus test was identified as the single best test. The researcher team concluded that this three-test battery yielded as much information about alcohol impairment as can be obtained in a reasonable time period under roadside conditions.

Tharp et al. (1981) Laboratory Study

This study followed the initial Burns and Moskowitz (1977) study listed above. The purpose of this study was to standardize the administration of the three-test battery and to examine the reliability and validity of the battery in the laboratory and to assess its feasibility, utility, and validity in the field. This experiment paralleled the design used by Burns and Moskowitz (1977), utilizing the three-test battery: Horizontal Gaze Nystagmus, Walk and Turn, and One Leg Stand. The study utilized 10 examiners (police officers) to administer the field sobriety test battery to 296 subjects, ranging in age from 21 to over 65 years old. In this study, the examiners' correctly classified eighty-one percent of the subjects as above or below a .10%

BAC. The examiners' incorrectly classified nine percent as being above and ten percent as being below the criterion level.

In this study, the examiners were also concerned with the issue of test-retest liability. 145 subjects were tested at a second session after being given alcohol to produce the same BAC's they reached at the first testing session. Testing by the same examiner from the previous session, the test-retest reliability coefficients were near .70, which according to the research team, meets the reliability standard for psychomotor tests. However, it was also noted that reliability coefficients were lower for testing by different officers.

These findings were reported to NHTSA and that agency developed a training curriculum and scoring guidelines for the three tests, designating them as the Standardized Field Sobriety Test battery. This test battery continues to be used to the present day by law enforcement agencies across the United States.

McKnight et al. (2002) Laboratory Study

This was a third study of roadside sobriety tests in a laboratory setting that utilized 16 police officer examiners administering a battery of tests to 30 drinking subjects at various BAC's and to 4 subjects that received no alcohol at all. A total of 231 observations were made in this study, including 217 from alcohol dosed subjects at BAC's in the range of .00% to .10% and 14 from placebo subjects. The research team concluded that the test battery serves as well to discriminate above and below a .08% BAC as above and below a .10% BAC.

Tharp et al. (1981) Field Study

This field study involved an examination of the effect of training with the test battery on the number of arrests, officers' discrimination of .10% BAC, and measured BACs of arrestees as the final stage in this study. Sheriff's deputies were randomly assigned to an experimental or

control group for which the independent variable was trained vs. untrained. Over the course of this study, 15 deputies worked 685.5 shifts during the study and completed 3,128 data forms for all traffic stops during three phases of data collection. The first phase took place prior to training officers to administer the three-test battery. The second phase involved half of the officers remaining untrained, while the other half were trained. In the third phase, all officers were trained to administer the three-test battery. Over the course of this study, research staff rode with each officer one or more times during each study phase to observe the use of the tests and to obtain BAC's from drivers that were released. Use of the test battery was associated with a 20% increase in arrests and more accurate arrest decisions by officers. However, it was noted that the data did not support definitive conclusions about field use of the tests. Additionally, the research team noted that some deputies and supervisory personnel did not cooperate fully with the study requirements.

Anderson et al. (1983) Field Study

In this study, field data were collected in 4 jurisdictions near Washington, D.C. The traffic officers in this study administered the standardized battery to drivers that were stopped on suspicion of driving while intoxicated (DWI) during a 4-month period. The number of stops were indicated as follows: Arlington County Police conducted 345 stops, Maryland State Police conducted 451 stops, North Carolina State Police conducted 434 stops, and Washington, D.C. police conducted 276 stops. The research team concluded that the tests effectively assess whether a drivers' BAC is above or below a .10% and that the Horizontal Gaze Nystagmus (HGN) is the most powerful of the three tests. The research team indicated that caution needs to be taken with these findings because officers in three of the four agencies indicated within the study had preliminary breath test instruments (commonly referred to as a PBT instrument) and knew the

driver's BAC prior to arrest. Thus, they noted that it is not possible to determine how test performance contributed to their decisions.

Burns and Anderson (1995) Field Study

In this study, data were collected using 7 law enforcement agencies (ranging from small municipality police agencies to the state patrol) within the State of Colorado. The research question in this study was phrased as: How accurate are officers' arrest and release decisions when the Standardized Field Sobriety Tests are used by trained and experienced officers? The research team noted that a control group was not necessary to answer such question. The study used 31 officers that were routinely assigned to traffic patrol or specialized DUI units, all of whom were trained under NHTSA guidelines to administer Standardized Field Sobriety Tests to collect the data. Thirty-one other individuals who were trained with the NHTSA test battery rode with the officers on forty-one percent of the data collection work shifts. This was done in an effort to monitor testing and ensure compliance with standardized procedures, and they obtained BAC's from drivers after officers decided to release the suspect.

Arrest records and the results of breath or blood tests were obtained for 305 administrations of the test battery during the period spanning from February through July 1995. BACs were unknown for drivers that either refused to provide a breath sample or were released when an observer was not present. The data in this study was analyzed in two ways. One analysis was limited to 234 cases with a measured BAC. The other analysis included all 305 cases and made the following assumptions: (a) Officers' decisions to arrest three drivers on drug charges were correct; (b) Drivers who refused to provide a breath or blood specimen were at or above the legal limit, and their arrests were correct decisions; (c) The BAC distribution was the same for released drivers with no breath test as for measured BAC's.

The results of the breath or blood test administered after administering the NHTSA test battery supported 93% of the arrest decisions. Specimens from the other 7% of arrested drivers did not contain alcohol, however, it is not known whether the arrestee was impaired by some other substance given that drug evaluations were not available. Based upon these findings, the research team concluded that the NHTSA test battery, as used by trained and experienced police officers, yields valid indices of the presence of alcohol.

Stuster and Burns (1998) Field Study

In this study, data were collected from 7 San Diego Police Department officers in order to evaluate the accuracy of the NHTSA Standardized Field Sobriety Tests for roadside decisions at BAC's below .10%. All 7 police officer participants were members of the alcohol enforcement unit of the San Diego Police Department and all officers were previously trained with the NHTSA curriculum. Additionally, the officers attended a four-hour session at which the standardized procedures were reviewed, and a slightly modified scoring was introduced. The modified scoring specified that two Horizontal Gaze Nystagmus signs would be the criterion for .04% BAC and four for .08% BAC. This was due to the fact that the State of California statutory limit was a .08% BAC and a .04% BAC for commercial drivers.

During the six-month study, the officers completed a study data form for each administration of the test battery. The officers recorded a score for each of the three tests and a BAC estimate. The final step required the officers to obtain a BAC measurement with an evidentiary breath test instrument. The results of this study indicated that officers' estimates of whether a driver's BAC was above or below a .08% were over 90% correct for both .04% and .08% BAC's. Additionally, the officers' mean BAC estimates differed from the mean measured BAC's by only .005%.

Burns and Dioquino (1998) Field Study

In this study, data were collected from 8 deputies of the Pinellas County Sheriff's Office in the State of Florida, where the statutory BAC limit for driving is .08%. The research team hypothesized that DUI arrest decisions by officers trained under NHTSA guidelines and experienced with the Standardized Field Sobriety Tests are greater than 90% correct when they use the tests to examine drivers at the roadside.

The study was conducted from June 1 through September 4, 1997, with Pinellas County Sheriff deputies that were assigned to a Selective Traffic Enforcement Program. The study procedures replicated those of the 1995 study conducted in Colorado. In this study, 8 deputies made a total of 379 vehicle stops, and one of 17 observers were present during the 242 (64%) of those stops. More than 95% of the arrest decisions and 82% of their release decisions were correctly indicated the presence of alcohol. The Horizontal Gaze Nystagmus score was recorded for drivers with a .08% or higher BAC. The data in this study demonstrated that officers may have given the HGN test more weight in the arrest decision, compared to the Walk and Turn Test and the One Leg Stand Test. Overall, however, the research team concluded that the test battery contributed to correct arrest decisions with high accuracy and that the tests aided them in correctly assessing nonimpaired drivers.

As indicated in Table 1 below, Burns and associates were involved in the vast majority of the lab and field studies of the NHTSA Standardized Field Sobriety Tests between 1977 and 1998. All of these studies conducted by Burns and associates were funded by governmental sources, including the United States Department of Transportation, with some of the later field studies funded by the states. The studies by Burns and associates were published in government reports, but never published in peer reviewed publications. Each of the studies conducted by

Burns and associates found that the NHTSA Standardized Field Sobriety Tests were a valid measure for correctly classifying a BAC level at or above the legal limit. Two studies listed within Table 1 were published by researchers other than Burns.

The two additional studies included in Table 1 that were conducted by researchers other than Burns found that the field sobriety test battery was empirically supported. The first is a journal published lab study wherein, Mcknight et al. (2002) found support for the HGN test and the One Leg Stand test; however, they did not recommend the Walk and Turn test to be included in the field sobriety test battery, opting instead for an observation of red eyes to be included.

The second is a field study by Anderson et al. (1983) finding that the Standardized Field Sobriety Tests were effective at predicting a driver's BAC level at or above .10%. However, as noted by Burns (2003), a significant concern with this study is the fact that officers in three out of the four participating law enforcement agencies had preliminary breath test instruments and knew the driver's BAC before making an arrest decision. Given this, it is impossible to know whether the decision to arrest was based on the driver's field sobriety test performance or the portable breath test reading known by the officer.

Table 1.

Summary of Early Studies Demonstrating Support for NHTSA Standardized Field Sobriety Tests

Authors and Year Performed	Setting of the Study	Sample Utilized in Study	Was the Study Published?	Was the Study Gov. Funded?	Was the Study Peer Reviewed?	Findings of the Study
Burns and Moskowitz (1977)	Lab	10 traffic officers tested 20-30 subjects each.	In a report by the U.S. Department of Transportation	Yes	No	83% of the subjects were correctly classified at .10% BAC
Tharp et al. (1981)	Lab	10 police officers tested 296 subjects, with 145 subjects tested a second time.	In a report by the U.S. Department of Transportation	Yes	No	81% of the subjects were correctly classified at .10% BAC
Mcknight et al. (2002)	Lab	16 police officers tested 30 drinking subjects and 4 non-drinking subjects.	Journal: Accident Analysis & Prevention, 34, 305-311	No	Yes	SFT battery serves as well to discriminate above and below a .08% as above and below a .10%. No % correct was given
Tharp et al. (1981)	Field	15 deputies engaging in 3,128 traffic stops.	Unpublished due to lack of cooperation by deputies	Yes	No	Study duration did not allow for sufficient data collection
Anderson et al. (1983)	Field	Number of participating officers across 4 jurisdictions was not given. Total of 1,506 traffic stops.	In a report by the U.S. Department of Transportation	Yes	No	The FST's effectively assess whether a driver's BAC is above or below a .10%; the HGN is the most powerful of the three tests
Burns and Anderson (1995)	Field	31 officers assigned to traffic patrol administered test battery to 305 subjects.	Report given to Colorado Department of Transportation	Yes	No	More than 90% of arrest decisions by officers trained in the SFST battery were supported by the results of blood or breath tests
Stuster and Burns (1998)	Field	7 officers in the alcohol enforcement unit administered test battery to 380 drivers.	In a report by the U.S. Department of Transportation	Yes	No	Officers' estimates of whether a driver's BAC was above or below .08% were more than 90% correct for both .04% and .08% BAC
Burns & Dioquino	Field	8 deputies conducted 379 traffic stops	Report given to Florida Department of Health	Yes	No	More than 95% of deputies were correct at measuring .08% or above BAC

Resources Demonstrating Criticism of NHTSA Standardized Field Sobriety Tests

Psychologist Steven Rubenzer, Ph.D., a leading critic of the NHTSA Standardized Field Sobriety Test studies, suggests many flaws with respect to the methodology utilized by the NHTSA when conducting tests of the NHTSA Standardized Field Sobriety Tests. Rubenzer's criticism was explored in detail in order to demonstrate the possible extent to which the NHTSA Standardized Field Sobriety Tests may be flawed in their ability to accurately detect those who are driving under the influence of alcohol to the extent that their normal faculties are impaired by alcohol.

Rubenzer (2003a) notes that there are several significant defects in the NHTSA field sobriety test validation studies. Each defect, as noted by Rubenzer, will be described below: First, the field studies validated the arrest decisions of the officers, not the field sobriety tests. Because officers had access to driver behavior and demeanor, the field sobriety tests did not function as stand-alone tests. Driver behavior and demeanor possibly functioned as the basis for the decision to arrest, rather than performance on the field sobriety tests (Rubenzer, 2003a).

Second, the police officers and the degree of supervision are not typical of DUI stops. The primary problem was that these officers were "highly motivated volunteers" (this term refers to officers who volunteered for the study and possessed a significant amount of prior DUI investigation experience) who do not reflect the typical officer in the field. The typical officer in the field may have very little experience in the administration of standardized field sobriety tests and often times there may be no other officers present to observe an officer administer the field sobriety tests to a suspect (Rubenzer, 2003a).

Third, the studies are insufficiently documented for scientific papers, and none discusses the weaknesses or limitations of the studies, as is customary among published works. This may

lead the uninformed reader to believe that such studies have no weaknesses or limitations (Rubenzer, 2003a). Fourth, as noted by Rubenzer (2003a), the authors did not report the accuracy of arrest decisions for stops that were observed versus those that were not, or for field sobriety tests performed under adverse climate conditions. Given that climate conditions play such an integral role in the proper administration of field sobriety tests, it is necessary to discuss how the validity of such tests can be compromised in adverse climates. Fifth, none of the field sobriety test studies have been published in peer reviewed scientific journals; they were simply written up and submitted to various state Department of Transportation agencies. Thus, the methods and conclusions of these studies have not been subjected to scrutiny by other professionals prior to being reported to the government and accepted as truth (Rubenzer, 2003a).

The following concerns, as summarized below, are all noted by Rubenzer (2003b) as being particularly problematic with the respect to the standardized administration of the NHTSA Standardized Field Sobriety Tests.

1. Screening questions for possible medical conditions should be standardized and validated. As it now stands, persons who are suspected of DUI may not always have an opportunity to discuss medical conditions which may result in a false positive determination of alcohol impairment with officers prior to field sobriety testing (Rubenzer, 2003b).

2. Changes in field sobriety exercise administration instructions have occurred over time and this may affect how police officers are trained. The concern here is that DUI certification courses at the local level may not update their course manuals in a timely fashion, thus instructing their officers in instructions and procedures that may not be the state of the art in the field (Rubenzer, 2003b).

3. Field sobriety test training does not emphasize rigorous adherence to instructions. This is problematic in that in jurisdictions across the nation, standardized field sobriety tests will be administered in uniquely subjective ways, thus affecting the overall validity of the test battery. The NHTSA cannot control for this and, of course, this is no fault of the organization. However, on this point, the NHTSA should apparently take more care to point out that instructions should be communicated in a uniform manner. One way in which to accomplish this would be to distribute video-taped demonstrations wherein law enforcement agencies across the country may become aware of proper oratory skills when administering instructions for the field sobriety tests (Rubenzer, 2003b).

4. Training manuals do not address how instructions are to be delivered. This point is related to point three above. The tone and delivery of instructions matter. Therefore, the NHTSA should take special care to explain this to each agency conducting DUI training seminars. Rubenzer (2003b) suggested that the tone of the officer administering the tests may affect the validity of tests themselves.

5. For the walk and turn test, a variety of lines are utilized. Rubenzer (2003b) noted that there is currently no research on the effects of using an imaginary line, a crooked line, or lines on uneven surfaces. The problem here is that officers are creating a fiction in the field every time the walk and turn test is conducted. Add to this the fact that there may be poor lighting conditions or weather conditions contributing even further to an already poorly constructed test.

6. Scoring rules are often inadequately specified. This is a point where the inherent subjectivity of the field sobriety test administration process can be observed. What one officer determines to be an error may not be considered an error by another officer. Rubenzer (2003b) pointed out that this is especially common on the walk and turn exercise where officers are

forced to make subjective interpretations regarding inappropriate turns and the like. This point has implications on other field sobriety tests, as well, and is not merely limited to the walk and turn test.

7. The criteria to determine a failing score are ambiguous. Rubenzer (2003b) indicated that the NHTSA manual provides cut off scores for each test, plus a grid for the combination of the horizontal gaze nystagmus test and the walk and turn test. However, it does not say which field sobriety test score is to be considered the primary score. As Rubenzer noted, a suspect can fail in four different ways due to scoring from each of the three tests and from the combination of the walk and turn. The implication here, of course, is that an officer can conceivably manipulate the tests and scores in order to achieve the desired outcome with respect to the defendant ultimately passing or failing.

8. Officers are not directed to record their observations immediately. Thus, as Rubenzer (2003b) stated, their memories may fade if too much time passes between the officer observing the defendant perform the tests and subsequently preparing the arrest report and DUI case investigation report. This is problematic in that it may contribute to officer fabrication on official documents or engaging in collusion with other officers in order to copy their reports.

9. The field sobriety tests have not been subjected to rigorous blind assessments of their validity. As a result, past experiences of the participating officers may skew the true measure of validity of the field sobriety tests. This is especially true in the 1998 study where all officers were volunteers that possessed a significant amount of prior DUI enforcement experience. (Rubenzer, 2003b).

10. The effects of fatigue and drowsiness on field sobriety test performance has not been analyzed. Rubenzer (2003b) noted that this may have the greatest effect on the horizontal gaze

nystagmus test due to the fact that the onset of nystagmus will differ up to five degrees depending on the time of day.

11. Drivers suspected of DUI may be highly anxious. This increased level of anxiety may affect field sobriety test performance. Rubenzer (2003b) noted that no study has shown fear or anxiety to be irrelevant with regard to performance on the field sobriety test, thus the true effect of fear and anxiety on field sobriety test performance is still unknown. To this end, Rubenzer recommends that norms based on large, representative samples, broken down by relevant factors, should be developed.

12. The one leg stand and the walk and turn lack documentation of actual reliability and validity. For example, Rubenzer (2003b) indicated that on the walk and turn it is possible that some clues are more valid than others, but there is no data published that is directly on point. Thus, it is conceivable that a defendant may fail certain clues, but not others, and still be arrested. However, are the clues that the defendant failed actually valid measures or not?

13. Reliability Data are lacking or below accepted standards for psychological tests used to make decisions about individuals. Rubenzer (2003b) suggested that tests such as these should have .90 reliability coefficient as a bare minimum result; the NHTSA field sobriety tests do not even come close to this target number.

14. Standard errors of measurement are not provided. Rubenzer (2003b) noted that NHTSA field sobriety test studies do not include basic descriptive statistics of the data. Thus, mean and standard deviations are not calculated. Also, given that there is no standard error of measurement reported, confidence intervals cannot be calculated either.

15. NHTSA Standardized Field Sobriety Tests have not been normed on sober people. Never have the tests been given to a large representative group of sober people. Rubenzer

(2003b) stated that when most psychological tests are developed, they test large samples in order to determine what is “normal.” With respect to the NHTSA field sobriety tests, there is no normal score available for this purpose.

16. The data on field sobriety exercise performance on people under the age of 21 and over the age of 50 is limited. Rubenzer (2003b) noted that in the NHTSA reported sample, only 3.1% of the sample was older than the age of 55. Implicit in this is that there may be a generalizability problem with respect to the manner in which the field studies were conducted. How middle-aged persons perform on these tests may be vastly different from how younger drivers or older drivers may perform.

17. NHTSA field sobriety tests have questionable validity for those who are elderly, in poor physical condition, or overweight. Rubenzer (2003b) stated that the physical size of the defendant performing the test may affect his or her ability to perform.

18. Even NHTSA claims that the field sobriety tests, when optimally used, are about 80% accurate. Such accuracy may be suitable for probable cause to require a BAC reading, but it is alarming when one considers this as the basis for arrest and to prove guilt of intoxication beyond a reasonable doubt (Rubenzer, 2003b).

The above criticisms suggest that despite the fact that the NHTSA considers field sobriety tests to be “standardized”, much still needs to be done in the way of test and retest before such a label may be utilized. In any situation where the organization reporting the research has funded the research, the results must be critically examined. The NHTSA studies were conducted “in house” and, as stated by Rubenzer (2003b), were not subject to a peer review process. Further, NHTSA does not utilize effective descriptive statistical methods in order to interpret the data of their studies. Merely collecting some data and announcing unsupported

conclusions does not signify sound research practice. Curiously, however, courts around this nation have, by and large, accepted the results of the NHTSA field sobriety exercise studies or simply determined that two of the three field sobriety tests (The One Leg Stand and the Walk and Turn) are deemed lay person observations, thus not subject to the judicial scrutiny for admitting scientific evidence.

Other researchers have criticized the NHTSA field sobriety tests for multiple reasons. In reviewing the Standardized field sobriety tests promulgated by NHTSA in 1981, Cole and Nowaczyk (1994), stated that even when an individual does not consume alcohol, the number of errors made by individuals performing the field sobriety tests was sufficient for officers to judge that the individuals had too much to drink. Therefore, this situation might lead to the arrest of a suspect that is not impaired by alcohol.

Even though the HGN is regarded as the best predictor of a BAC level out of the three-test battery, it is not without its fair share of criticism. Pangman (1987) provided a number of causes of nystagmus due to conditions other than alcohol consumption. Such causes included medical conditions, distractions from traffic, and poor lighting. Although police officers may be trained to administer the HGN test consistent with NHTSA standards, it certainly is not reasonable to expect that they have sufficient expert training in the fields of optometry and neurology, such that they can make nuanced judgments as to the etiology of nystagmus. Moreover, even optometrists have admitted to having very little training or experience in eye testing for determining sobriety (Rubenzer, 2008).

In assessing the scientific foundation of the Horizontal Gaze Nystagmus test, Honts and Amato-Henderson (1995) labeled the foundation as being, at best, weak. The authors indicated that “although NHTSA’s work claims to have found a correlation between BAC and HGN

performance, this research has not been replicated by independent investigators. The lack of evidence makes the continued use of HGN in the courtroom questionable” (p.694). In a strong rebuke of the Horizontal Gaze Nystagmus test, Booker (2004) stated that “the United States Department of Transportation engaged in deliberate fraud in order to mislead the law enforcement and legal communities into believing that the HGN test was scientifically meritorious and overvaluing its worth in the context of criminal evidence.” (Booker, 2004, p. 134).

All of this notwithstanding, Honts and Amato-Anderson (1995) demonstrated that courts across the country generally took one of three positions on the admissibility of the results from the HGN test at trial. The first position taken is that although HGN is scientific in nature, it has gained general acceptance in the scientific community and therefore satisfied the admissibility requirements set forth in *Frye v. United States* (1923) as a matter of law. The second position taken on this issue is that HGN is not scientific in nature, thereby declaring it admissible in the absence of a scientific foundation. This position essentially holds that HGN is no more scientific than other field sobriety tests. The third position taken on this issue is that HGN evidence is scientific in nature and expert testimony is required to demonstrate general acceptance within the relevant scientific community prior to admission into evidence at trial. Since the publication of this article, it is important to note that many state courts and the federal courts have replaced the *Frye* standard for admissibility of scientific evidence with the more recent *Daubert* standard articulated by the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993). The *Daubert* standard requires that the trial court judge serve a gatekeeper function, wherein the trial judge must engage in a line of inquiry pertaining to the use of the scientific theory or technique,

whether it has been subjected to peer review, and whether it has attracted widespread acceptance within a relevant scientific community.

Judicial Interpretations of the NHTSA Field Sobriety Tests

In *Florida v. Meador* (1996), a defendant was arrested for drunk driving and asked to perform a series of field sobriety tests, including the NHTSA standardized field sobriety test battery. The defendant challenged the admissibility of the field sobriety tests on the ground that they lacked scientific validity and probative value. The Florida Fourth District Court of Appeals held that testimony concerning the performance on psychomotor field sobriety tests is sufficiently reliable as a lay person observation of impairment to be relevant in proving impairment; the fact that such tests may be unfairly prejudicial did not outweigh their probative value so as to find it necessary to exclude their admissibility into evidence at trial. In essence, the court argued that the field sobriety tests, except for the horizontal gaze nystagmus, are simply lay observations of impairment, rather than scientific tests.

It should be noted, however, that many courts treat the horizontal gaze nystagmus test differently than the walk and turn and the one leg stand test. This is because the horizontal gaze nystagmus test is deemed to be “scientific” and, as such, must meet certain conditions under state rules of evidence in order to be admissible at trial. Most states in the United States follow the holding of either *Frye v. United States* (1923) which required that the admissibility of novel scientific be generally accepted by the relevant scientific community or *Daubert v. Merrell Dow Pharmaceuticals* (1993) which required that an independent judicial assessment of reliability of novel scientific evidence be made by a judge at trial. Under this standard, the judge serves a gatekeeper function and determines which scientific evidence may be admitted at trial.

In *Ohio v. Homan* (2000), the Ohio Supreme Court held that, while field sobriety tests are admissible into evidence, if field sobriety tests are conducted “in a manner that departs from the methods established by the NHTSA, then they are inherently unreliable” and thus inadmissible as evidence at trial. According to the American Prosecutors Research Institute (APRI), the Homan decision was based on a portion of the NHTSA student manual that states, “if any one of the standardized field sobriety test elements is changed, the validity is compromised.” Clearly, the standard set forth in *Homan* (which, incidentally, as an Ohio case does not serve as binding precedent in other jurisdictions) is a standard which protects the rights of the accused from poor investigatory skills on the part of the police (NDAA, 2005).

Interestingly, however, the APRI advises prosecutors to argue that an officer’s failure to follow the NHTSA guidelines on field sobriety tests should not be a basis for excluding such evidence. The APRI advises prosecutors to argue that deviations in following the instructions should go to the weight of the evidence, and not its admissibility (NDAA, 2005). APRI states that this is consistent with the wording of the NHTSA manual in that the manual describes how field sobriety tests should be administered under “ideal conditions.” However, no literature seems to suggest what exactly constitutes an ideal condition for conducting field sobriety tests.

It is noteworthy that many other states do not agree with the position taken in *Homan*. In fact, according to the APRI, Florida, Georgia, Hawaii, Illinois, Pennsylvania, Texas, Tennessee, Wisconsin, and Wyoming have all reviewed the admissibility of field sobriety tests that assess physical coordination. These jurisdictions have held that deviations in the administration of these tests, including departure from NHTSA guidelines, will not result in suppression of such test results. These courts have consistently held, rather, that the field sobriety tests (other than the

horizontal gaze nystagmus) are simple physical dexterity tests that can be interpreted by the fact finder through common sense and experience (NDAA, 2005).

Another important concern for an individual accused of DUI is whether or not the field sobriety test results are to be considered testimonial in nature. If they are, certain procedural safeguards under *Miranda v. Arizona* (1966), may be invoked. In *Pennsylvania v. Muniz* (1990), a defendant was stopped after being suspected of drunk driving. An officer administered three field sobriety tests to the defendant on the side of the road and at the police station. At the station Muniz was told that his actions and voice were to be recorded, however, he had not been advised of his Miranda rights. He was also asked a series of questions pertaining to his age, address, weight and date of birth as part of the DUI investigation. In particular, Muniz was asked “do you know the date of your sixth birthday?” Muniz had difficulty with articulating answers to some of the questions. He was arrested and given his Miranda rights at that time. Muniz was subsequently convicted and appealed the conviction. The U.S. Supreme Court held that “Muniz’s answers are not rendered inadmissible by Miranda merely because the slurred nature of the speech is incriminating. The physical inability to articulate words in a clear manner due to the lack of muscular coordination of his tongue and mouth is not in itself a testimonial component of the response.”

The court in *Muniz* pointed out that in *Schmerber v. California* (1966), a distinction was drawn between testimonial and physical evidence regarding the privilege against self-incrimination. Further, in *Holt v. United States* (1910), the court stated that both federal and state courts generally offer no protection against compulsion to provide fingerprinting, photographing, or to write or speak for identification. Thus, the court opined, a person suspected of DUI can be required to provide a blood or breath sample because such sample was outside of the privilege

and the sample was procured in a manner by which the petitioner's testimonial capacities were not implicated. In addition, if a subject accused of DUI does not consent to such a test, their license may be suspended under the "Implied Consent" laws of the state.

Given such a rationale, it is important to determine if any additional aspects of the field sobriety test administration process will invoke 5th amendment concerns in the context of Miranda. For example, at what point does the field sobriety exercise administration process become a custodial interrogation? In *Berkemer v. McCarthy* (1984), the United States Supreme Court held that in the context of a DUI traffic stop a suspect is merely engaged in a short investigatory encounter with the police and is not yet deemed to be in police custody. Accordingly, the fact that such an encounter is deemed to be non-custodial relieves the state's responsibility to issue the Miranda warning in this context and all incriminating statements made by the suspect in the pre-arrest encounter are admissible at trial. However, in *Allred v. Florida* (1993), the Florida Supreme Court held that the recitation of the alphabet during a DUI investigation constituted a custodial interrogation. In that case, the court cited Article 1, Section 9 of the Florida Constitution, a clause that requires that suspects be informed of their right against self-incrimination. Further, the court stated that it is the content of the speech that is being introduced into evidence, not the physical act, such as slurred speech. The court stated that "the incriminating inference is drawn from the testimonial act." Thus, there appears to be a limit on what may be asked during a DUI investigation in some states for the purposes of determining if the exchange constitutes incriminating evidence based on a suspect's testimony. Thus, some state supreme courts are willing to afford additional procedural protections that the United States Supreme Court has not granted.

While the horizontal gaze nystagmus has been admitted in most jurisdictions under the *Frye* or *Daubert* standards for novel scientific evidence, it should be noted that not all police officers may proffer testimony as to the results of this test. Only officers who have been trained as drug recognition experts may proffer testimony regarding a defendant's performance on the HGN test in many jurisdictions. Thus, in cases where an officer is not certified in drug recognition, the results of such test may be inadmissible against the accused at trial in many jurisdictions.

The Competing Interests in a DUI Investigation

One must look at the DUI investigation process as being based on a set of competing interests. These interests include, but are not limited to, the constitutional rights of the accused drunk-driver versus the right of the state to arrest and prosecute such persons in an effort to make the roadways safe to travel upon for all members of society. Clearly, there is a social harm that is being addressed by DUI legislation. It is the protection of our society from drunk drivers. It should come as no surprise, then, that DUI legislation is constantly a source of debate in the legislatures across the United States.

A legal fiction is defined as an assumption or supposition of law that something which is or may be false is true (Black's Law Dictionary, 1999). The NHTSA has conceptualized field sobriety tests as an empirical assessment of one's impairment due to alcohol, without adequately demonstrating that, in fact, the field sobriety tests are an accurate measure of such a condition. Such fictions, it has been argued, create pathology within the law (Fuller, 1967). That is, although the legal fiction leads to a conclusion that appeals to us as right, they also may produce mistakes that create grave injustice (Fuller, 1967). Such fictions, as in the case of field sobriety tests, serve an ideological function, by providing both the intellectual superstructure to

rationalize the actions of lawmakers and by appealing through the appeal of metaphor (Fuller, 1967). This is to say that through the advancement and alleged standardization of field sobriety tests, such tests can be rationalized and legitimated by government institutions and thus applied to and used against those accused of drunk driving. Furthermore, the conceptualization of the drunk driver as a social threat serves to facilitate the process by which the legal fiction can be applied, as Fuller suggested, with regard to the appeal of metaphor in advancing the legal fiction itself.

The most basic concern in the case of field sobriety tests as they apply to the crime of drunk driving is one of due process and liberty. Under the 5th and 14th amendments to the United States Constitution, all citizens are entitled to due process under the law. Under the 14th amendment to the constitution of the United States, states must provide each citizen with due process of law. From a substantive and procedural due process standpoint, this stands for the proposition that a court must protect the accused against erroneous convictions. The state has the burden of proof in a criminal trial and must prove every element of a criminal offense beyond a reasonable doubt. Implicit in this fact is the notion that a court must ensure the validity and reliability of evidence in the fact-finding process.

Considering the limited scientific support for the NHTSA Standardized Field Sobriety Tests based on the available literature, classifying the one leg stand and the walk and turn field sobriety tests as non-scientific appears to facilitate the admissibility of these tests without any further inquiry by courts relating to such tests as scientific evidence. Such a technique, Fuller would argue, allows a court to escape the consequences of an existing, specific rule of law (Fuller, 1967). In the case of HGN, courts have taken more pro-active measures to ensure that the admissibility of HGN meets standards for the admissibility of scientific evidence.

Findings

(1) How did the NHTSA Standardized Field Sobriety Tests originate?

Moral entrepreneurs such as MADD were a driving force that served to shape legislative and judicial mandates relating to the crime of drunk driving. It should be recognized, however, that no single force can take credit for such an enterprise. To assert that one moral entrepreneur such as MADD can take credit for the 1980's anti-drunk-driving movement would be both naïve and analytically incorrect. MADD entered the arena at a point in time where the United States conservative political climate was able and willing to receive their message. Further, the NHTSA, which was already in a nascent stage of assessing tests for apprehending drunk-drivers, had a vested interest in establishing and maintaining its interest in the anti-drunk-driving movement of the 1980's and beyond. The NHTSA responded in a concerted effort through the 1980's and 1990's by developing a comprehensive methodology in the form of a field sobriety test battery which would assist the government in proving cases of drunk driving. These factors, in an aggregate sense, all served to facilitate the anti-drunk-driving movement of the 1980's and its subsequent punitive effects to the present day.

(2). How were the NHTSA Standardized Field Sobriety Tests developed and pilot tested?

The NHTSA Standardized Field Sobriety Tests were developed through a research process commencing in the mid 1970's conducted by Marcelline Burns of the Southern California Research Institute and various colleagues in an effort to develop tests that would provide more reliable evidence of impairment and to standardize the administration procedures for those tests. Tharp et al. (1981) evaluated six tests for further evaluation and ultimately determined that the Horizontal Gaze Nystagmus test, the Walk and Turn test, and the One Leg Stand test were the three tests most effective at correctly classifying a BAC level as above or

below a .10%. Upon the conclusion of the study by Tharp et al. (1981), the NHTSA developed a training curriculum and scoring guidelines for the three-test battery and the NHTSA Standardized Field Sobriety Tests were born. Training in the NHTSA 3 test battery commenced in the early 1980's across the United States and continues through the present day. Subsequent pilot test studies were funded by NHTSA. These studies took place in the field utilizing officers from various law enforcement agencies in Colorado (Burns & Anderson, 1995), San Diego (Stuster & Burns, 1998), and Florida (Burns & Dioquino, 1998). The researchers in each of these pilot field studies indicated that the NHTSA Standardized Field Sobriety Tests were empirically supported, although such assertion is the subject of intense debate.

(3). How have the NHTSA Standardized Field Sobriety Tests been evaluated since they were initially developed and piloted?

The NHTSA Standardized Field Sobriety have been evaluated by several entities subsequent to the last pilot study in 1998. Of significance, however, is that NHTSA has not funded additional validation studies since the last Standardized Field Sobriety Test pilot study completed by Burns and Dioquino in 1998. Much of the subsequent literature evaluating the NHTSA Field Sobriety Test battery comes in the form of critical evaluations of the test battery from psychologists and lawyers. While some of these criticisms were published in peer reviewed journals, a variety of others were published in DUI defense-oriented publications. Accordingly, as with any type of publication, it is the responsibility of the reader to look at the source of the material and question the practices and research methods employed by the author of each study.

What is critically important to point out, as well, is that much of the work by Burns and her colleagues were published in the form of reports given to government agencies including NHTSA (a division of the United States Department of Transportation), the Florida Department

of Health, and the Colorado Department of Transportation. While this fact alone certainly does not discount the value of the data necessarily, it may cause some to question the veracity of the data when it is being performed for the government agency that funds the research endeavor. This is especially true when one considers the criticism that the research in question was not subject to peer review, as it would be if submitted to a refereed academic journal (Rubenzer, 2003b). In short, the argument can easily be made that additional, peer reviewed, studies of the NHTSA Field Sobriety Tests should be an immediate priority in order to better understand the value, or lack of value, that these tests possess going forward.

(4). To what extent are the NHTSA Standardized Field Sobriety Tests supported by empirical evidence?

A review of the located literature on the NHTSA Standardized Field Sobriety Tests suggests that this is a difficult question to answer. On the one hand, the various NHTSA funded studies performed by Burns and her associates would lead the average person to believe that the NHTSA Standardized Field Sobriety Tests are empirically supported by scientific evidence. In support of this, Burns (2003) published an article synthesizing the field sobriety test research up until 2003. Interestingly, the vast majority of articles cited by Burns in support of the NHTSA Standardized Field Sobriety Tests consist of her own research; two of the three laboratory studies and four of the five field studies involved her participation as a lead researcher. In assessing the empirical literature in support of the NHTSA Standardized FST's, Burns stated, "Studies indicate that the Standardized Field Sobriety test battery yields valid and reliable indices of the presence of alcohol. Traffic officers' arrests of alcohol-impaired drivers are largely correct when they rely on the tests within the context of other information and evidence available at the roadside." (Burns, 2003, p. 1198). Curiously, however, this statement does not opine on the ability of the

standardized field sobriety tests to predict a BAC level of .08% or higher. Simply being “valid and reliable indices of the presence of alcohol” falls short considering that the presence of alcohol in the system is not necessarily illegal. Burns then went on to acknowledge that officers arrest decisions are largely correct when they rely on the field sobriety tests within the context of other information and evidence that is available at the roadside. Such statements seemingly take on a concessionary tone with respect to the alleged validity of the standardized field sobriety test battery.

As noted previously, critics argue that a review of the literature demonstrates a lack of peer reviewed research on the NHTSA Standardized Field Sobriety Tests as a predictor of a BAC of .08% or higher. Given that peer review is a hallmark of the scholarly enterprise, many might argue that the NHTSA test battery has no empirical support at all, given that the validation process was not subject to peer review by the scientific community. Additionally, NHTSA has not funded additional Standardized Field Sobriety Test studies since the validation studies conducted by Burns and her associates.

Based on the data that I have located, there have been no additional independent studies to date that attempt to validate the NHTSA Standardized Field Sobriety test battery for the presence of alcohol. The majority of recent literature on the NHTSA Standardized Field Sobriety test battery relate to the use of these tests in cases of drug impairment, rather than alcohol impairment. Because drug impairment was not the focus of this study, those studies were not addressed here. Until such time as additional studies can be performed to validate the NHTSA Standardized Field Sobriety Tests in driving under the influence of alcohol cases, the empirical support for these tests remains limited and subject to intense debate.

Chapter 5: Summary and Recommendations

Putting It All Together

This study demonstrated a complex set of competing interests at play between the government's need to strictly enforce DUI laws and the rights of those accused of DUI to a fair investigative process. In this chapter, brief recommendations are offered in an effort to fortify the DUI investigation process so that it is more empirically sound, so that evidence is accurately collected and stored for later use in the courtroom, and so that the DUI investigation process is performed in a competent and professional manner by law enforcement officers.

Groups such as MADD espoused a strong message in the early 1980's in the United States that was warmly received during a conservative, tough on crime era in American history. The National Highway Traffic Safety Administration, a division of the United States Department of Transportation, was compelled to spring into action and devise a plan to assist states with the prosecution of the dangerous drunk driver. Development of the NHTSA Standardized Field Sobriety Tests occurred in 1981 after several years of pilot testing various psychomotor tests for the detection of alcohol impairment. Shortly thereafter, the NHTSA Standardized Field Sobriety Tests became the new standard used by law enforcement officers and prosecutors in the investigation and prosecution of DUI cases. Since that time, these tests have received a fair share of praise and criticism. The praise came from those that developed these tests and the law enforcement and criminal prosecution community. The criticism of these tests came from many constituencies, including psychologists and the criminal defense bar.

Recognizing some of the limitations of the past, future policy related to field sobriety tests as a method to prove drunk-driving cases should seek to ensure that valid and reliable measures to determine driver impairment are employed during the drunk-driving investigation process. Such a process necessitates further development of valid and reliable field sobriety tests by NHTSA in a peer reviewed format, comprehensive police officer training and education courses at the local, state, and federal levels to increase the skill and knowledge base of each police officer administering field sobriety tests, and the education of judges, prosecutors, and criminal defense attorneys with respect to the field sobriety investigation process in order to ensure that the drunk-driving investigation process is fair and due process rights of the accused are protected. Such initiatives will add both legitimacy and procedural fairness to the drunk driving investigatory process.

Recommendations for Future Research & Policy Development

In order to fortify and strengthen the DUI investigation process, the following recommendations are made:

1. If the NHTSA Standardized Field Sobriety Tests continue to be used in the future, additional research as to their validity and reliability must be performed in a peer reviewed forum. As noted by many experts, these tests face multiple forms of criticism from various constituencies. This criticism should be acknowledged and addressed in an effort to produce a wider body of scientific research on each of the three standardized field sobriety tests. It might be wise for NHTSA to also fund additional research on the three-test battery given that decades have passed since the last validation study was performed. Many critics have previously called for this to occur. NHTSA should be receptive to this message to fund additional validations studies. Additionally, social

scientists would be well served to develop a comprehensive, peer reviewed study of the NHTSA Standardized Field Sobriety Test battery.

2. State law should require that all aspects of the DUI investigation process be videotaped by a patrol car dash cam. The video tape should be initiated by the officer at the time the officers first observe the driver operating a vehicle in a manner that is contrary to law. The driving pattern of the vehicle should be recorded by the dash cam in an effort to clearly demonstrate the basis for the traffic stop. Upon making the traffic stop, all aspects of the officer-subject encounter should be in plain view of the camera. A microphone connected to the officer's uniform should enable the conversation with the suspect to be audible in the video. Based on the subjective determination of the officer, the suspect may be asked to step out of their vehicle. Should this occur, the officer should direct the suspect to stand directly in front of the officer's patrol car so that the entirety of the interaction is included within the video. Should the officer ask the suspect to perform field sobriety tests, all aspects of the field sobriety tests should be included. This means that the officer should be certain that the entire field sobriety test administration process is included in the video. This should begin with the officer asking questions relating to the subject's fitness to engage in the test battery, through the detailed instruction process for each field sobriety test, and through the performance of each of the field sobriety tests by the subject (if the subject refuses to take the field sobriety tests, such refusal should be recorded on the video as well). If the officer finds probable cause to arrest, the dash cam should remain on while the Miranda Warning is being read to the suspect and during the process of transporting the suspect to jail until arrival at the jail. Once in the jail, the officer should continue to video tape the proceeding either through a bodycam or video

camera at the jail facility, including during the wait time that is required before administering the breath test and should record the entirety of the breath test process. Doing this will ensure that any future observer (prosecutor, judge, or juror) can see the investigation process in its entirety and make an informed judgment based upon the entirety of the evidence recorded on the video tape. At present, only the State of South Carolina has a videotape requirement for DUI investigations. Going forward, every jurisdiction within the United States should require that the DUI investigation process be recorded on video through the use of dashboard cameras and police officer body cameras. Suppression of this evidence at trial would be a sufficient remedy for failure to videotape the DUI investigation, absent good cause.

3. All police officers that are members of general patrol or of a specialized traffic unit should complete the basic NHTSA DWI Detection and Standardized Field Sobriety course through their respective departments. Additionally, each department should require officers to take the NHTSA DWI Detection and Standardized Field Sobriety refresher course each year. Law enforcement agencies should then require additional in service DWI detection and investigation training for officers that belong to specialized traffic and DUI enforcement units. This additional specialized training should include mandatory coverage of HGN administration in an effort to familiarize the officer with the intricacies of this test, as it is far more technical than the Walk and Turn and the One Leg Stand. Furthermore, each law enforcement agency should mandate that all officers belonging to a traffic or DUI unit be certified as drug recognition experts, as much of the HGN course training pertains to how alcohol and drugs affect eye nystagmus. Such

additional training will produce an officer that is more skilled in understanding various indicators of impairment by alcohol.

4. Law enforcement agencies should recognize the criticism of the NHTSA Standardized Field Sobriety Tests and stress to officers the need to engage in a complete and holistic investigation process. Be certain that a suspect can physically perform the field sobriety tests. Ask relevant questions pertaining to prior medical history, age, and physical condition. If the officer has a reasonable belief that the suspect cannot perform the field sobriety tests based on these factors, then the tests should not be administered. Additionally, officers should make certain that suspects clearly understand each test through giving clear instructions about how to perform the test so that a suspect understands what is expected in each of the tests. This would include both a demonstration of the test and an understanding of how the test is scored. Furthermore, it might be wise to incorporate the Standardized Field Sobriety tests into a high school driver education course curriculum. In this way, students will learn about each test and understand what is expected in each test and how it is scored well before they attain their license to operate a motor vehicle. This will provide a foundational understanding of the field sobriety tests in a classroom environment.

The officer should then focus on all other aspects of the investigation to serve as probable cause to arrest (or release). This includes the driving patterns of the vehicle prior to the traffic stop, the suspect's appearance and demeanor, along with any voluntary statements made during the encounter. It may very well be better to use the solid evidence that is present in the absence of administering field sobriety tests, rather than risking injury to the subject. Should the matter proceed to trial, a judge or juror may very

well question why an officer would make a suspect complete the field sobriety tests when the suspect provides a reasonable basis for not being able to perform the tests.

The goal of this study is to illuminate how DUI emerged as a social problem in the United States in the 1980's and to provide an example of one concrete way in which government responded to this social problem. The creation of the NHTSA Standardized Field Sobriety Tests was a meaningful response by the government to assist in combatting the drunk driving problem. While well intentioned as this response may have been, there are unintended consequences that came about during an innovative development process. Going forward it is imperative that additional testing of the NHTSA Standardized Field Sobriety Tests be performed in a peer reviewed forum so that sound empirical conclusions may be drawn as to the true benefit of these tests as a tool in the DUI investigation process. Additionally, law enforcement officers conducting DUI investigations should make certain that they use all available investigative tools at their disposal and take careful attention to avoid over-reliance on the NHTSA Standardized Field Sobriety Test battery in an effort to conduct a thorough investigation, while at the same time, maintaining a fair and just investigatory process for those suspected of the crime of driving under the influence of alcohol.

References

- Allred v. Florida, 622 So. 2d 984. (1993).
- Anderson, T., Schweitz, R., & Snyder, M. (1983). *Field Evaluation of a behavioral test battery for DWI*. (DOT HS -806-475) Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.
- Babbie, E. (2008). *The Basics of Social Research* (4th ed). Thomson Wadsworth.
- Becker, H. (1963). *The Outsiders: Studies in the Sociology of Deviance*. New York, Free Press.
- Berkemer v. McCarty, 468 U.S. 420 (1984).
- Bishop, D. M., Lanza-Kaduce, L., & Frazier, C. E. (1998). Juvenile Justice Under Attack: an Analysis of the Causes and Impact of Recent Reforms. *University of Florida Journal of Law and Public Policy*, 10(1), 129–156.
- Black's Law Dictionary. (1999). 7th Edition, West Group.
- Booker, J. (2004). The Horizontal Gaze Nystagmus Test: Fraudulent Science in the American Courts. *Science & Justice*, 44(3), 133–139.
- Bowen, G. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*, 9(2), 27-40.
- Burns, M. (2003). An Overview of Field Sobriety Test Research. *Perceptual and Motor Skills*, 97(3_suppl), 1187–1199.
- Burns, M., & Anderson, E. (1995). *Field Evaluation of the Standardized Field Sobriety Test Battery*, Colorado Department of Transportation.

- Burns, M., & Dioquino, T. (1998). *A Florida Validation Study of the Standardized Field Sobriety Test (SFST) Battery*. Florida Department of Health.
- Burns, M., & Moskowitz, H. (1977). U.S. Department of Transportation, National Highway Traffic Safety Administration: Psychophysiological Tests for DWI Arrest, Final Report (DOT-HS-802-424). Washington, D.C.
- Cohn, D. (1984, August 26). Reformers of a New Breed Pour It on Against Alcohol, The Los Angeles Times.
- Cole, S., & Nowaczyk, R. H. (1994). Field Sobriety Tests: Are They Designed for Failure? *Perceptual and Motor Skills*, 79(1), 99–104.
- Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (3rd ed.). Sage.
- Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).
- Dickson, D. (1968). Bureaucracy and Morality; An Organizational Perspective on a Moral Crusade. *Social Problems*, 16, 143-156.
- Eastman, J. (2003). A Constitutional Amendment for Victims: The Unexplored Possibility, *Victimology: A Study of Crime Victims and Their Role* (pp. 333-346). Prentice-Hall.
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods*, 5(1), 80–92. Retrieved 25 June 2020, from http://www.ualberta.ca/~iiqm/backissues/5_1/pdf/fereday.pdf.
- Florida v. Meador, 674 So. 2d 826 (1996).
- Frye v. United States, 293 F. 1013 (1923).
- Fuller, L. (1967). *Legal Fictions*. City of Stanford Press.

- Gusfield, J. (1981). *The Culture of Public Problems: Drinking-Driving and the Symbolic Order*, University of Chicago Press.
- Harris, D., Dick, R., Casey, S., & Jarosz, C. (1980). U.S. Department of Transportation: The Visual Detection of Driving While Intoxicated: National Highway Traffic Safety Administration (DOT-HS-7-1538). Washington, D.C.
- Holt v. United States, 218 U.S. 295 (1910).
- Honts, C. R., & Amato-Henderson, S. L. (1995). Horizontal Gaze Nystagmus Test: The State of The Science in 1995. *North Dakota Law Review*, 71(3), 671–700.
- McKnight, A. J., Langston, E. A., McKnight, A. S., & Lange, J. E. (2002). Sobriety Tests for Low Blood Alcohol Concentrations. *Accident Analysis and Prevention*, 34(3), 305–311.
- Merriam, S. B. (1988). *Case Study Research in Education: A Qualitative Approach*. Jossey-Bass.
- Miranda v. Arizona, 384 U.S. 436 (1966).
- Ohio v. Homan, 732 N.E. 2d 952 (2000).
- Pangman, W. A. (1987). Horizontal Gaze Nystagmus: Voodoo science. *DWI Journal: Law & Science*, 2(3), 1-6.
- Pennsylvania v. Muniz, 496 U.S. 582 (1990).
- Pentilla, A., Tenhu, M., & Kataja, M. (1971). Clinical Examination for Intoxication in Cases of Suspected Drunken Driving, Statistical Research Bureau of TALJA, Helsinki, Finland.
- Reinarman, C. (1988). The Social Construction of an Alcohol Problem: The Case of Mothers Against Drunk-Drivers and Social Control in the 1980's, *Theory and Society*, 17(1), 91-120.
- Rubenzler, S. J. (2003a). The Psychometrics and Science of the Standardized Field Sobriety Tests, Part 1. *The Champion*, 27(4), 48-54.

- Rubenzler, S. J. (2003b). The Psychometrics and Science of the Standardized Field Sobriety Tests, Part 2. *The Champion*, 27(5), 40-44.
- Rubenzler, S. J. (2008). The Standardized Field Sobriety Tests: A Review of Scientific and Legal Issues. *Law & Human Behavior (Springer Science & Business Media B.V.)*, 32(4), 293–313.
- Schmerber v. California, 384 U.S. 757 (1966).
- Stuster, J., & Burns, M. (1998). Validation of the Standardized Field Sobriety Test Battery at BAC's Below .10. United States Department of Transportation, National Highway Traffic Safety Administration (DOT-HS-808-654).
- Tharp, V., Burns, M., & Moskowitz, H. (1981). Development and Field Test of Psychophysiological Tests for DWI Arrest. United States Department of Transportation, National Highway Traffic Safety Administration (DOT-HS-805-864).
- United States Department of Transportation, National Highway Traffic Safety Administration. (2001). Development of a Field Sobriety Test Training Management System (DOT-HS-89-400).
- Yin, R. K. (1994). *Case Study Research: Design and Methods* (2nd ed.). Sage.

Appendix A: Reflection

This dissertation project is rooted in my early professional experience in criminal law as a prosecuting attorney. The crime of DUI has always fascinated me for a variety of reasons because there are so many component parts to a DUI investigation. Additionally, DUI is a crime where the law and science intersect in a significant way. This dissertation did not focus on the “per se” method of DUI prosecution, but rather on the impairment method of DUI prosecution.

Through engaging in the research and writing process for this dissertation, much of what I already knew in a professional context was challenged. For example, as a young prosecutor, I always worked under the assumption that the DUI field sobriety tests used by police were tested and approved through the rigors of the scientific method. My facile understanding of the tests at the time was superficial in nature and failed to look beneath the surface in an effort to understand how the NHTSA Standardized Field Sobriety tests were developed and tested.

The goal in this research was to educate others in an effort to make certain that this component of the DUI investigation is what it purports to be. Future research on the NHTSA Standardized Field Sobriety tests is warranted and overdue. Additionally, researchers are urged to develop their own research studies on these tests in order to understand if these tests are, in fact, valid and reliable measures. If not, additional research should be done to determine if there is a better alternative to these tests or if DUI investigations should focus more intensively on “per se” methods of prosecution utilizing breath and blood testing.

The DUI investigation process is a very personalized experience that potentially has major consequences. I realize that the administration of these tests at the roadside by police

officers can be a very frightening experience for anyone. In light of recent events involving police officers and minorities in the United States, I also realize that there are factors beyond the scope of this study that are of critical importance. In reflecting upon this, I thought intensively about how minority suspects might feel if pulled over and investigated by police officers. Would this lack of trust and fear of the police affect their ability to perform the Standardized Field Sobriety tests? The obvious answer to this question is that surely it must. Furthermore, if minorities are stopped at a higher rate by police than others, then they may very well be more frequently asked to perform field sobriety tests on the roadside. Thus, this issue has consequences for communities that lack trust in the police as a social institution. The point to be made here is that the focus of this study is but a small part of a larger and complex set of circumstances within the criminal justice system in the United States. While the situational factors mentioned above are outside the limited scope of inquiry in this study, it is incumbent to acknowledge their existence and work towards creating a criminal justice system that is more transparent and protects the rights of all citizens.

Appendix B: DUI Statutes

State of South Carolina, Title 56, Motor Vehicles

SECTION 56-5-2953. Incident site and breath test site video recording.

(A) A person who violates Section 56-5-2930, 56-5-2933, or 56-5-2945 must have his conduct at the incident site and the breath test site video recorded.

(1)(a) The video recording at the incident site must:

(i) not begin later than the activation of the officer's blue lights;

(ii) include any field sobriety tests administered; and

(iii) include the arrest of a person for a violation of Section 56-5-2930 or Section 56-5-2933, or a probable cause determination in that the person violated Section 56-5-2945, and show the person being advised of his Miranda rights.

(b) A refusal to take a field sobriety test does not constitute disobeying a police command.

(2) The video recording at the breath test site must:

(a) include the entire breath test procedure, the person being informed that he is being video recorded, and that he has the right to refuse the test;

(b) include the person taking or refusing the breath test and the actions of the breath test operator while conducting the test; and

(c) also include the person's conduct during the required twenty-minute pre-test waiting period, unless the officer submits a sworn affidavit certifying that it was physically impossible to video record this waiting period.

(3) The video recordings of the incident site and of the breath test site are admissible pursuant to the South Carolina Rules of Evidence in a criminal, administrative, or civil proceeding by any party to the action.

(B) Nothing in this section may be construed as prohibiting the introduction of other relevant evidence in the trial of a violation of Section 56-5-2930, 56-5-2933, or 56-5-2945. Failure by the arresting officer to produce the video recording required by this section is not alone a ground for dismissal of any charge made pursuant to Section 56-5-2930, 56-5-2933, or 56-5-2945 if the arresting officer submits a sworn affidavit certifying that the video recording equipment at the time of the arrest or probable cause determination, or video equipment at the breath test facility was in an inoperable condition, stating which reasonable efforts have been made to maintain the

equipment in an operable condition, and certifying that there was no other operable breath test facility available in the county or, in the alternative, submits a sworn affidavit certifying that it was physically impossible to produce the video recording because the person needed emergency medical treatment, or exigent circumstances existed. In circumstances including, but not limited to, road blocks, traffic accident investigations, and citizens' arrests, where an arrest has been made and the video recording equipment has not been activated by blue lights, the failure by the arresting officer to produce the video recordings required by this section is not alone a ground for dismissal. However, as soon as video recording is practicable in these circumstances, video recording must begin and conform with the provisions of this section. Nothing in this section prohibits the court from considering any other valid reason for the failure to produce the video recording based upon the totality of the circumstances; nor do the provisions of this section prohibit the person from offering evidence relating to the arresting law enforcement officer's failure to produce the video recording.

(C) A video recording must not be disposed of in any manner except for its transfer to a master recording for consolidation purposes until the results of any legal proceeding in which it may be involved are finally determined.

(D) SLED is responsible for purchasing, maintaining, and supplying all necessary video recording equipment for use at the breath test sites. SLED also is responsible for monitoring all breath test sites to ensure the proper maintenance of video recording equipment. The Department of Public Safety is responsible for purchasing, maintaining, and supplying all videotaping equipment for use in all law enforcement vehicles used for traffic enforcement. The Department of Public Safety also is responsible for monitoring all law enforcement vehicles used for traffic enforcement to ensure proper maintenance of video recording equipment.

(E) Beginning one month from the effective date of this section, all of the funds received in accordance with Section 14-1-208(C)(9) must be expended by SLED to equip all breath test sites with video recording devices and supplies. Once all breath test sites have been equipped fully with video recording devices and supplies, eighty-seven and one-half percent of the funds received in accordance with Section 14-1-208(C)(9) must be expended by the Department of Public Safety to purchase, maintain, and supply video recording equipment for vehicles used for traffic enforcement. The remaining twelve and one-half percent of the funds received in accordance with Section 14-1-208(C)(9) must be expended by SLED to purchase, maintain, and supply video recording equipment for the breath test sites. Funds must be distributed by the State Treasurer to the Department of Public Safety and SLED on a monthly basis. The Department of Public Safety and SLED are authorized to carry forward any unexpended funds received in accordance with Section 14-1-208(C)(9) as of June thirtieth of each year and to expend these carried forward funds for the purchase, maintenance, and supply of video recording equipment. The Department of Public Safety and SLED must report the revenue received under this section and the expenditures for which the revenue was used as required in the department's and SLED's annual appropriation request to the General Assembly.

(F) The Department of Public Safety and SLED must promulgate regulations necessary to implement the provisions of this section.

(G) The provisions contained in Section 56-5-2953(A), (B), and (C) take effect for each law enforcement vehicle used for traffic enforcement once the law enforcement vehicle is equipped with a video recording device. The provisions contained in Section 56-5-2953(A), (B), and (C) take effect for a breath test site once the breath test site is equipped with a video recording. HISTORY: 1998 Act No. 434, Section 9; 2000 Act No. 390, Section 23; 2003 Act No. 61, Section 8; 2008 Act No. 201, Section 11, eff February 10, 2009.

DUI Statute Example

The State of Florida Driving Under the Influence Statute (2021)

316.193 Driving under the influence; penalties.

(1) A person is guilty of the offense of driving under the influence and is subject to punishment as provided in subsection (2) if the person is driving or in actual physical control of a vehicle within this state and:

(a) The person is under the influence of alcoholic beverages, any chemical substance set forth in s. 877.111, or any substance controlled under chapter 893, when affected to the extent that the person's normal faculties are impaired;

(b) The person has a blood-alcohol level of 0.08 or more grams of alcohol per 100 milliliters of blood; or

(c) The person has a breath-alcohol level of 0.08 or more grams of alcohol per 210 liters of breath.

(2)(a) Except as provided in paragraph (b), subsection (3), or subsection (4), any person who is convicted of a violation of subsection (1) shall be punished:

1. By a fine of:

a. Not less than \$500 or more than \$1,000 for a first conviction.

b. Not less than \$1,000 or more than \$2,000 for a second conviction; and

2. By imprisonment for:

a. Not more than 6 months for a first conviction.

b. Not more than 9 months for a second conviction.