Southern California Research Institute

Validation of Standardized Field Sobriety Tests for Use in the Marine Environment

Dary Fiorentino
Bergetta Dietel
Dulcemonica Jimenez

September 1, 2010

Documentation Page

Report No. MAR-165	Report Date April 2010
Title and Subtitle Validation of standardized field sobriety tests for the marine environment	
Author(s) Dary Fiorentino Bergetta Dietel Dulcemonica Jimenez	
Performing Organization Name and Address Southern California Research Institute 7065 Hayvenhurst Avenue Suite 9 Van Nuys, CA 91406	
Managing Organization Name and Address National Association of State Boating Law Administrators 1500 Leestown Road Suite 330 Lexington, KY 40511	
Sponsoring Organization Name United States Coast Guard	

Executive Summary

The objective of this project was to develop sobriety tests that can be administered in the seated position to assist water patrol officers in detecting impairment caused by blood alcohol concentrations (BACs) of .08 and above. As in the roadside tests, the seated tests must be easy to administer, so as to not overburden law enforcement officers, who must continually monitor the environment for their own safety and the safety of the boaters suspected of impairment. The tests must discriminate impaired performance without the knowledge of the individual suspect's baseline performance. Most importantly, the tests must be useful for an arrest/release decision. Unlike the roadside tests, however, the seated tests cannot make use of any measure of equilibrium.

The current project was divided into two studies. The first study was conducted in the laboratory and was aimed at developing a battery of tests suitable for the seated position. The second study was conducted in the field and was aimed at validating the tests developed in the laboratory for the marine environment.

Six seated tests were evaluated in the laboratory. Data were obtained under double-blind conditions at relatively low BACs.

Four of the six tests, horizontal gaze nystagmus (HGN), finger to nose (FTN), palm pat (PP), and hand coordination (HC) were found to reliably discriminate BAC Status (BAC $< .08$ v. BAC $\geq .08$). HGN was the most useful test, followed by FTN, PP, and HC. HGN and any one of FTN, PP, and HC correctly predicted BAC Status in 70.8% of the cases.

In the laboratory study, the overall correct percentages, sensitivity, and specificity of the tests were below what is typically reported in literature on the roadside standardized field sobriety tests (SFSTs). Three possible reasons were identified. First, the wider distribution of BACs in the previous studies may have made the detection of impairment less difficult than in the current study. Second, the impairment decisions were made exclusively on the basis of the tests, without external clues such as smell of alcohol, appearance, speech, and demeanor. Third, some officers were not properly trained to administer and score the test. Nevertheless, the seated tests showed enough promise to warrant a field study.

In the field study, conducted on the Lake of the Ozarks in central Missouri, four sobriety tests were administered on the water to determine their usefulness in detecting impairment caused by BACs of .08 and above. Data were obtained by a team of four marine officers and two civilian observers.

Because officers' training was identified in the laboratory study as a potential factor for the results of that study, considerable attention and time were devoted to ensuring that officers were proficient and comfortable administering and scoring the test prior to data collection.

The four tests predicted BAC Status as follows:

- HGN alone correctly predicted BAC Status in 84.8% of the cases. Sensitivity was .86 and specificity was .84. The positive predictive value was .80 and the negative predictive

value was .89. Positive likelihood ratio and negative likelihood ratio were 5.27 and .16, respectively.

- FTN alone correctly predicted BAC Status in 67.3% of the cases. Sensitivity was .49 and specificity was .81. The positive predictive value was .65 and the negative predictive value was .68. Positive likelihood ratio and negative likelihood ratio were 2.56 and .63, respectively.

- PP alone correctly predicted BAC Status in 65.2% of the cases. Sensitivity was .76 and specificity was .57. The positive predictive value was .57 and the negative predictive value was .77. Positive likelihood ratio and negative likelihood ratio were 1.77 and .41, respectively.

- HC alone correctly predicted BAC Status in 59.4% of the cases. Sensitivity was .62 and specificity was .57. The positive predictive value was .52 and the negative predictive value was .67. Positive likelihood ratio and negative likelihood ratio were 1.46 and .66, respectively.

- HGN and FTN were the best combination of two tests. Combined, they correctly predicted BAC Status in 74.5% of the cases. Sensitivity was .46 and specificity was .96. The positive predictive value was .89 and the negative predictive value was .70. Positive likelihood ratio and negative likelihood ratio were 10.80 and .57, respectively.

- HGN, FTN, and PP was the best combination of three tests. That combination correctly predicted BAC Status in 72.0% of the cases. Sensitivity was .39 and specificity was .97. The positive predictive value was .90 and the negative predictive value was .68. Positive likelihood ratio and negative likelihood ratio were 12.15 and .63, respectively.

- The four tests combined correctly predict BAC Status in 68.1% of the cases. Sensitivity was .28 and specificity was .98. The positive predictive value was .91 and the negative predictive value was .65. Positive likelihood ratio and negative likelihood ratio were 13.16 and .74, respectively.

- HGN and any one of FTN, PP, and HC correctly predicted BAC Status in 84.5% of the cases. Sensitivity was .81 and specificity was .87. The positive predictive value was .82 and the negative predictive value was .86. Positive likelihood ratio and negative likelihood ratio were 6.36 and .22, respectively.

- Without HGN, the best predictor of BAC Status was the combination of FTN, PP and HC, which correctly predicted 66.1% of the cases. Sensitivity was .29 and specificity was .93. The positive predictive value was .76 and the negative predictive value was .64. Positive likelihood ratio and negative likelihood ratio were 4.28 and .76, respectively.

The overall correct percentages, sensitivity, and specificity of the tests were consistent with what is typically reported in literature on the roadside SFSTs. It should be noted that the prevalence of BACs at or above .08 was lower in the current field study (.43) than in the previous field studies on SFSTs (.79, .80, .73).

The tests' reliability was also consistent with what is typically reported in literature on the roadside SFSTs. Note, however, that HGN could not be included in the reliability analyses because, due to space limitations on the patrol boat, the eyes of the BUI suspects were clearly visible to the officers but not to the observers.

It is proposed that marine officers administer HGN, FTN, PP, and HC to all BUI suspects, and then, for each suspect, use the pattern of test results to estimate the probability of BAC ≥ .08, as follows:

- Positive HGN, FTN, PP, and HC tests indicate a .91 probability that the BUI suspect has a BAC ≥ .08.
- Positive HGN, FTN, and PP tests indicate a .90 probability that the BUI suspect has a BAC ≥ .08.
- Positive HGN and FTN tests indicate a .89 probability that the BUI suspect has a BAC ≥ .08.
- Positive FTN, PP and HC tests indicate a .76 probability that the BUI suspect has a BAC ≥ .08.
- Positive FTN and PP tests indicate a .73 probability that the BUI suspect has a BAC ≥ .08.

Acknowledgments

The Southern California Research Institute would like to thank the following individuals who helped with the project.

Arizona Game and Fish Department
Officer Tim Baumgarten

California Highway Patrol
Officer Michael Brown
Officer Christopher Clark
Officer John Fransen
Officer Travis Herbert
Sergeant Danny Lamm
Officer J. D. Owen Jr.
Officer Aaron Rohner
Officer John Wilson

California State Parks
State Park Peace Officer-Ranger Kirk Coon
State Park Peace Officer-Ranger Mark Hofer
State Park Ranger Officer-Ranger John Stephens
State Park Peace Officer-Ranger Pamela Wilson

Clackamas County Sheriff's Department (Oregon)
Lieutenant PJ Steigleder II
Deputy Nathan Thompson

Glendale, CA Police Department
Officer Rich Ulrich

Indiana Law Enforcement Academy
Joe Turner

Intoximeters, Inc.
Mark Gilmer

Los Angeles, CA Police Department
Sergeant Thomas Page (Ret.)
Officer III Michael Boylls
Officer Corey Lapinsky
Officer III C. Todd Mangrum
Officer Michael Wells

Missouri Department of Revenue
Nick Humphrey

Missouri State Water Patrol
Major J. W. Hughes
Captain Matt Walz
Corporal Scott Lance
Patrolman Dean Bartlett
Patrolman Dustin Metzner
Patrolman Eric Mueller

Multnomah County Sheriff (Oregon)
Deputy Ken Yohe

Nevada Department of Wildlife
Game Warden Karen Welden

Oregon State Police
Trooper Robin May

San Bernardino County Sheriff's Department
Detective Robbi DeBois
Detective Mike Landavazo

United States Coast Guard
Petty Officer Chris Hoschak
Petty Officer Terry Jones
Petty Officer Chris Lane
Petty Officer Angela Morris
Petty Officer Francisco Rinaudo
Petty Officer Ryan Whitman

University of California Santa Barbara Police Department
Corporal Zak Ayala

Ventura, CA Police Department
Sergeant Tim Turner

Table of Contents

List of Tables

List of Figures

1 Background

Over the past few decades, the number of recreational boaters has increased steadily (Tseng, Kyle, Shafer, Graefe, Bradle, & Schuett, 2009). There is evidence that alcohol consumption is elevated among recreational boaters (Khiabani, Opdal, & Morland, 2008; Logan, Sacks, Branche, Ryan, & Bender, 1999) and that alcohol consumption significantly increases the risk of dying while boating (Driscoll, Harrison, & Steenkamp, 2004; Lunetta, Penttila, & Sarna, 1998; Smith, Keyl, Hadley, Bartley, Foss, Tolbert, & McKnight, 2001). Some studies indicate that up to 70% of drowning victims test positive for alcohol (Browne, Lewis-Michl, & Stark, 2003; Driscoll et al., 2004).

The responsibility of detecting boating under the influence of alcohol (BUI) falls on water patrol officers. Their job, however, is fraught with difficulties. First, on many of our nation's waterways, it is not illegal to drink alcohol while boating. An open container, therefore, is not probable cause for a stop. Second, on many waterways there are no speed limits or traffic control devices, making excessive speed, failure to maintain a single lane, and other observations not necessarily clues of impairment. Third, environmental conditions (wind, water choppiness, and glare) can make it difficult to determine boaters' impairment. Finally, unlike land-based officers, water patrol officers do not have a validated battery of sobriety tests to be used on water.

To examine the type of tests water patrol officers currently use, the Southern California Research Institute (SCRI) made a nationwide request to all agencies with water patrol duties to provide their BUI arrest records for the previous year. A total of 1,146 BUI reports from agencies in Alaska, Arizona, Connecticut, Delaware, Florida, Georgia, Idaho, Illinois, Nevada, Ohio, Tennessee, Texas, Virginia, and Wisconsin were received and analyzed. With the exception of the three tests that comprise the standardized field sobriety tests (SFSTs), it was found that no test was uniformly administered from state to state or, often, from agency to agency within a state. This lack of standardization may result in uneven application of sanctions and penalties for BUI.

The roadside SFST battery is not suitable for use on the water, because walking and balance tests need to be administered on a firm, flat surface. Marine officers who use these tests must bring the suspect boater to shore and may need to wait a pre-established period of time to get the suspect adapted to being on land (usually 15 minutes). This can be inconvenient for both officers and boaters. Tests that can be administered without bringing the suspect ashore will save time, but safety concerns mandate that they be performed with the suspect seated. Previous efforts examined a variety of seated tests on boats and found encouraging results (Sussman, Needalman, & Mengert, 1990).

In general, sobriety tests are diagnostic tests. That is, they are used to predict impairment caused by blood alcohol concentrations (BACs[1]) of .08 and above. Diagnostic tests can be systematically evaluated by arranging the presence/absence of the condition of interest and positive/negative results of the test to detect the condition of interest in a 2 x 2 table. Figure 1 is a schematic representation of this process, with the associated calculations for evaluation.
The overall percent correct is based on the sum of correct positives and correct negatives. Sensitivity is the proportion of people with the condition of interest who have a positive test. Specificity is the proportion of people without the condition of interest who have a negative test. The positive predictive value (PPV) is the proportion of people with a positive test who have the condition of interest. The negative predictive value (NPV) is the proportion of people with a negative test who do not have the condition of interest.

Note that the overall percent correct, sensitivity, specificity, PPV, and NPV can all be affected by the prevalence of the condition of interest (Alberg, Park, Hager, Brock, & Diener-West, 2004). It is useful, therefore, to examine the likelihood ratios. The positive likelihood ratio (LR+) is the ratio of the probability of a positive test in people with the condition of interest to the probability of a positive test in people without the condition of interest. The negative likelihood ratio (LR-) is the ratio of the probability of a negative test in people with the condition of interest to the probability of a negative test in people without the condition of interest (Grimes & Schulz, 2005).

A positive likelihood ratio of 3.16, for example, indicates that people with BACs $\geq .08$ are three times more likely to test positive than people with BACs $< .08$. A negative likelihood ratio of .20, on the other hand, indicates that a negative test is a fifth as likely in people with BACs $\geq .08$ than in people with BACs $< .08$.

In general, tests with positive likelihood ratios between two and five and negative likelihood ratios between .5 and .2 are considered suggestive but not conclusive; whereas tests with positive likelihood ratios greater than five and negative likelihood ratios smaller than .2 tests are considered conclusive (LeBlond, Brown, DeGowin, 2009).

[1] Alcohol in the body is measured in terms of *weight per volume*. In the US, alcohol in the body is reported as blood alcohol concentration (BAC) in grams of absolute alcohol per deciliter of whole blood (e.g., $.08\ g_{absolute\ alcohol}/dL_{whole\ blood}$) or breath alcohol concentrations (BrACs), in grams of absolute alcohol per 210 liters of breath (e.g., $.08\ g_{absolute\ alcohol}/210\ L_{breath}$). *Per se* laws are typically enacted with both BAC and BrAC as the basis for a charge of driving under the influence.

In this report, statutes and experimental conditions are expressed in BACs. Measurement of alcohol in the body is expressed in either BAC if blood was analyzed or BrAC if breath was analyzed.

		MEASURED	
		Present	Absent
PREDICTED	Positive	*a* **True positive**	*b* **False positive**
	Negative	*c* **False negative**	*d* **True negative**

Overall percent correct = $(a+d)/(a+b+c+d)*100$
Sensitivity = $a/(a+c)$
Specificity = $d/(b+d)$
Positive predictive value (PPV) = $a/(a+b)$
Negative predictive value (NPV) = $d/(c+d)$
Positive likelihood ratio (LR+) = $[a/(a+c)]/[b/(b+d)]$
Negative likelihood ratio (LR-) = $[c/(a+c)]/[d/(b+d)]$

Figure 1. 2x2 table for calculations of overall accuracy, sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio.

1.1 Prior Research on Sobriety Tests

The roadside sobriety tests were developed in the late 70's and early 80's by SCRI. Two studies developed and refined the roadside sobriety tests. The first examined the usefulness of six candidate tests in detecting BACs of .10 and above (Burns & Moskowitz, 1977). In that study, 238 subjects were semi-randomly assigned to one of four BAC Groups: .00, .05, .10, and .15. Note that the positive BAC Groups represented half the legal limit, the legal limit, and 1.5 times the legal limit of the time (.10). Law enforcement officers administered six tests to the subjects. The six tests were One-Leg Stand (OLS), Finger to Nose (FTN), Walk and Turn (WAT), Finger Count (FC), Tracing, and Horizontal Gaze Nystagmus (HGN). Based on the results, the authors recommended a reduced battery of tests which included the OLS, WAT, and HGN.

Nystagmus is an involuntary jerking of the eyeball that occurs when there is a disturbance of the vestibular system (National Highway Traffic Safety Administration, 1999). Nystagmus is a complex phenomenon that can occur for a variety of reasons (Dell'Osso, 1989). Within the impaired driving context, however, HGN specifically refers to a lateral jerking of the eyeball affected by alcohol, certain nervous system depressants, inhalants and phencyclidine. The HGN test is comprised of six clues, three for each eye: lack of smooth pursuit, maximum deviation, and angle of onset (National Highway Traffic Safety Administration, 1999). Four out of six possible clues indicate impairment. The WAT test requires a person to assume a heel-to-toe position on a designated line, arms at the sides, and to listen while instructions are given. The person is then required to make nine heel-to-toe steps along the line, turn around keeping one foot on the line, and return with another nine heel-to-toe steps. Two out of eight possible clues indicate impairment. The OLS test requires a person to stand, heels together, feet at a slight angle, and arms at the sides. The person is then required to raise one leg forward about six inches off the ground. Two out of four possible clues indicate impairment.

In the second study (Tharp, Burns, & Moskowitz, 1981), 297 subjects were administered enough alcohol to reach peak BACs of .00, .05, .11, and .15. Again, .05 was half the legal limit, .11 was slightly above the legal limit, and .15 was 1.5 times the legal limit at the time (.10). As shown in Table 1, a combination of HGN, WAT, and OLS correctly identified 81.2% of the subjects.

Since the development of the roadside sobriety tests, they have been routinely used by law enforcement officers throughout the US to identify BACs at or above the legal limit. Three validation studies have confirmed their usefulness. The first (Burns & Anderson, 1995), is unique because it was conducted in Colorado, which had a two-tiered system, one for drivers with BACs between .05 but less than .10 (now .08), who were charged with driving-while-ability-impaired; and one tier for drivers with BACs of .10 and above (now .08), who were charged with driving-under-the-influence. Thirty-one officers from six law enforcement agencies collected the data, accompanied on approximately half the stops by observers who verified that data were collected according to study procedures. In general, the officers stopped drivers suspected of being BAC .05 and above and administered the three sobriety tests (HGN, WAT, and OLS). The accuracy of the arrest/release decision was verified with a portable breath alcohol screener, which was always administered following the sobriety tests. Complete data were collected from 234 drivers, with BrACs ranging from .00 to .34, with an average BrAC of .15. As shown in Table 1, with the .05 criterion, officers using the sobriety tests had an overall correct percentage of 85.9%, .89 sensitivity, and .76 specificity.

The second validation study was conducted in Florida (Burns & Dioquino, 1997), which already had a .08 statute. Eight officers from Pinellas County Sheriff's Office collected the data, at times accompanied by observers. In general, the procedures were similar to the Colorado study (Burns & Anderson, 1995). Complete data were collected from 256 drivers, with BrACs ranging from .00 to .28. As shown in Table 1, Florida officers using the sobriety tests had an overall correct percentage of 93.0%, .96 sensitivity, and .82 specificity.

The third validation study was conducted in California (Stuster, 2006), which also had a .08 statute. Seven officers from the San Diego Police Department collected the data. In general, the procedures were similar to the Colorado and Florida studies (Burns & Anderson, 1995; Burns & Dioquino, 1997). Complete data were collected from 297 drivers, with an average BrAC of .12. As shown in Table 1, California officers using the sobriety tests had an overall percent correct of 91.2%, .98 sensitivity, and .73 specificity.

Note that the prevalence of BrAC $\geq$ .08 in the field studies is dramatically higher than the prevalence in the earlier laboratory studies. In the laboratory, the distribution of BACs is dictated by practical and ethical considerations. In the field, the range of BACs is much greater.

1.2 Current Project

The objective of this project was to develop sobriety tests that can be administered in the seated position to assist water patrol officers in detecting impairment caused by BACs of .08 and above.

As in the roadside tests, the seated tests must be easy to administer, so as to not overburden law enforcement officers, who must continually monitor the environment for their own safety and the safety of the boaters suspected of impairment. The tests must discriminate impaired performance without the knowledge of the individual suspect's baseline performance. Most importantly, the tests must be useful for an arrest/release decision. Unlike the roadside tests, however, the seated tests cannot make use of any measure of equilibrium.

The current project was divided into two studies. The first study was conducted in the SCRI laboratory and was aimed at developing a battery of tests suitable for the seated position. The second study was conducted in the field and was aimed at validating the battery of tests developed in the laboratory for the marine environment.

Table 1

Selected Prior Studies on Sobriety Tests

Study	Test	Prevalence	% Correct	Sensitivity	Specificity	PPV	NPV	LR+	LR-
Burns & Moskowitz (1977)	Combination[a]	.27	83.2	.69	.89	.69	.89	5.98	.35
	OLS	.27	75.5	.65	.80	.54	.86	3.17	.45
	FTN	.27	70.4	.56	.76	.46	.82	2.31	.58
	FC	.27	67.1	.57	.71	.42	.82	1.96	.61
	WAT	.27	75.1	.60	.81	.54	.84	3.11	.50
	Tracing	.27	76.5	.56	.84	.57	.83	3.57	.53
	HGN	.27	81.8	.68	.87	.66	.88	5.21	.36
Tharp, Burns, and Moskowitz (1981)	Combination[b]	.28	81.2	.64	.88	.68	.86	5.32	.41
Sussman, Needalman, and Mengert (1990)	Combination[c]	.29	82.5	.75	.86	.68	.89	5.18	.29
Burns and Anderson (1995, with .05 statute)	Combination[b]	.79	85.9	.89	.76	.93	.64	3.69	.15
Burns and Dioquino (1997)	Combination[b]	.80	93.0	.96	.82	.96	.82	5.31	.05
Stuster (2006)	Combination[b]	.73	91.2	.98	.73	.91	.94	3.61	.03

Note. PPV = Positive predictive value. NPV = Negative predictive value. LR+ = Positive likelihood ratio. LR- = Negative likelihood ratio. OLS = One leg stand. FTN = Finger to nose. FC = Finger count. WAT = Walk and turn. HGN = Horizontal gaze nystagmus. [a] = OLS, FTN, FC, WAT, Tracing, HGN. [b] = WAT, OLS, HGN. [c] = Interview and observation, alphabet recital, FTN, nystagmus on boat, nystagmus on land, cumulative WAT & OLS.

2 Laboratory Study: Development of Sobriety Tests for the Seated Position

2.1 Introduction

As was done in the development of the roadside tests, the usefulness of six candidate tests in detecting alcohol impairment was first evaluated in a controlled environment. As in the past, subjects were tested at .00 BAC, at half the legal limit, at the legal limit, and at 1.5 times the legal limit.

Subjects were recruited and randomly assigned to one of four BAC conditions: .00 (placebo), .04, .08, and .12. Six tests were administered in counterbalanced order to each study participant by qualified testers. The six tests were selected on the basis of the results of a pilot study which examined a total of 15 tests. The initial 15 tests were selected based on the scientific literature and current usage by law enforcement agencies.

2.2 Method

2.2.1 Subjects

One hundred and fifty-seven subjects participated as paid volunteers, 17 people were recruited but failed to show on testing day, eight people failed to confirm their availability the day before testing, eight people tested positive for illegal drugs and were sent home without testing, and 13 people were sent home without testing because SCRI staff had concerns about their health prior to participating in the study. Subjects who completed the study were paid $100.

2.2.2 Testers

Thirty-three law enforcement officers participated in the study. Officers had an average of 9.7 years of experience administering the roadside SFSTs. Officers' participation spanned four days. Day 1 consisted of a training session on the tests' administration and scoring. Days 2, 3, and 4 were data collection days. Each data collection day lasted about five hours. Officers were paid $100 per day.

2.2.3 Apparatus

The Intox EC/IR and the Alco Sensor FST (Intoximeters, Inc., St. Louis, MO) breath alcohol testing instruments were used to measure the subjects' BAC.

2.2.4 Drug Screeners

All subjects provided a urine specimen and were tested for drug use. Ten types of drugs were screened: methamphetamine, opiates, cocaine, marijuana, PCP, benzodiazepines, barbiturates, methadone, tryciclic antidepressants, and amphetamine.

2.2.5 Pregnancy Tests

Female subjects provided a urine specimen and the specimens were screened for hCG, the pregnancy hormone.

2.2.6 Tests

Six tests were evaluated:

- *Finger to Nose*. The FTN test required the subjects to bring the tip of the index finger to touch the tip of the nose. It was performed with eyes closed and head tilted slightly back.

- *Time Estimation*. The TE test required the subjects to estimate the passage of 30 seconds. It was performed with the eyes closed and the head tilted back. The test was scored as the absolute time deviation from 30 seconds.
- *Finger Count*. The FC test required the subjects to extend one hand forward palm up, and to count to four while touching the tips of each finger with the tip of the thumb. The process was then reversed and the subjects counted backwards. Three complete sets were performed.
- *Hand Coordination*. The HC test required the subjects to perform a series of tasks with their hands. It was loosely adapted from the SFST WAT test.
- *Palm Pat*. The PP test required the subjects to extend one hand, palm up, and to place the other hand on it palm down. The subject was instructed to use the top hand to pat the bottom hand. The top hand rotated 180 degrees thereby alternating the pat between the back and the palm of the hand. The bottom hand remained stationary. The subject counted each pat aloud.
- *Horizontal Gaze Nystagmus*. Each eye was examined for lack of smooth pursuit, angle of onset, and jerking at maximum deviation.

2.2.7 Procedures

Subjects were recruited with newspaper ads, internet postings, flyers, and referrals. An initial telephone interview determined eligibility for the study. Applicants were screened in terms of health history, current health status, and use of alcohol and other drugs. The Quantity-Frequency-Variability scale (QFV; Cahalan, Cisin, & Crossley, 1969) was used to classify applicants into five groups: abstainers, infrequent drinkers, light drinkers, moderate drinkers, and heavy drinkers. Only moderate and heavy drinkers were eligible to participate in the study. Pregnancy, chronic disease, or evidence of substance abuse resulted in exclusion from the study.

Subjects were transported from their residence to the SCRI laboratory and from the SCRI laboratory to their residence via taxi or shuttle. Subjects arrived at the facility in pairs at 9:00 am, 11:00 am, and 12:00 pm. Thus, a maximum of six subjects were tested per day.
Upon arrival at the laboratory, each subject gave informed consent to participate in the study, and each received a copy of the signed Informed Consent and of the Subjects' Bill of Rights. A breath alcohol test, a second administration of the QFV, a pregnancy test for females, and a drug screen confirmed eligibility for the study. Measurements of blood pressure, heart rate, height, and weight were taken next. Cardiovascular measures within acceptable ranges (systolic blood pressure = 120 ± 30 mmHg, diastolic blood pressure = 80 ± 20, heart rate = 70 ± 20) further confirmed eligibility for the study.

Subjects were randomly assigned to one of four BAC groups (.00, .04, .08, and .12), by lottery. No efforts were made to counterbalance drinking practice.

Age, gender, weight, and height were used to calculate the alcohol dose with the Sahlgrenska formula provided by Tzamaloukas, Murata, Vanderjagt, Servilla, and Glew (2003). A drinking period of 30 minutes followed. Subjects were served three equal-sized drinks at 10-minute intervals, and were instructed to pace each drink evenly over the entire 10 minutes. SCRI staff monitored the subjects continually throughout the drinking period.

For subjects in the .04, .08, and .12 BAC groups, the alcohol drink consisted of one part 80 proof vodka and 1.5 part orange juice. For subjects in the .00 BAC group, the placebo drink consisted of one part water and 1.5 part orange juice. The placebo glasses had their rim swabbed with vodka and 10 ml of vodka floated in each of them to produce an initial taste and odor of alcohol.

Twenty minutes after the end of the third drink, BrAC measurements were obtained at 5-minute intervals until the peak BrAC was detected. Peak BrAC was expected 30 minutes after the end of the third drink. For the .00 BAC group, testing occurred at the first available testing window 30 minutes post drink. Subjects were not privy to their target BAC.

When the subjects reached the target BrAC on the descending limb of their BrAC curve, they were brought in the testing room and they were asked to sit down. The battery of tests was administered twice to each subject, each time by a different tester, with only one tester in the testing room at a time.

Testers remained in a separate room and had no interaction with the subjects prior to testing them. A SCRI staff member was present during testing to ensure that the interaction between subject and tester was limited to the administration and scoring of the tests. A BrAC reading was obtained immediately after the testing.

BAC group and test order were counterbalanced. When the subjects' BrACs dropped below .03, they were debriefed, paid $100, and transported home by taxi.

2.3 Results

2.3.1 Sample Characteristics

In general, the recruitment procedures yielded the desired variety in subjects' demographics. Subjects tended to be relatively young, with some college education. Table 2 reports means, standard deviations, minimum and maximum values for age, school years completed, people in household, and personal income.

Table 2
Subjects' Characteristics 1

Variable	*N*	Mean	*SD*	Min	Max
Age	157	32.96	10.70	21	62
School years completed	153	13.58	3.63	1	24
People ≥ 18 years in household	156	2.08	1.27	0	8
People < 18 years in household	152	.51	.96	0	4
Personal income ($)	146	33,197.26	23,720.69	0	149,000

Table 3 reports the percentage of subjects by gender, ethnic background, race, marital status, and employment status. Table 4 shows the number of subjects by gender, age, and drinking practice. In spite of the screening procedures designed to recruit only moderate and heavy drinkers, eight light drinkers (5 females and 3 males) participated in the study.

Table 3 *Subjects' Characteristics 2*	
Percent males	50.3
Percent Latino/a	28.0
Percent African American	20.4
Percent American Indian	.6
Percent Asian	3.2
Percent Caucasian	47.8
Percent Pacific Islander	1.3
Percent other race	10.8
Percent declined to answer	15.9
Percent single	66.2
Percent married	15.9
Percent living together	10.2
Percent divorced	6.4
Percent separated	.6
Percent widowed	.6
Percent employed full time	33.1
Percent employed part time	31.2
Percent unemployed	19.7
Percent retired	1.3
Percent going to school	7.6
Percent other employment status	7.0

Table 4
Number of Subjects by Gender, Age, and Drinking Practice

		Drinking Practice			
Gender	**Age Group**	**Light**	**Moderate**	**Heavy**	**Total**
Female	21-29	3	15	19	37
	30-39	1	9	14	24
	40-49	1	9	2	12
	50-59	5	2	5	7
	60-69	0	0	0	0
	Total	5	35	40	80
Male	21-29	0	9	35	44
	30-39	2	5	6	13
	40-49	0	4	7	11
	50-59	0	4	3	7
	60-69	1	0	1	2
	Total	3	22	52	77

2.3.2 BACs

Of the 157 subjects, 39 were assigned to the zero BAC condition, 40 to the .04 BAC condition, 39 to the .08 BAC condition, and 39 to the .12 BAC condition (see Table 5). Unequal group sizes were the results of some subjects' failure to meet study criteria. Moderate and heavy drinkers were equally divided among the four BAC groups.

In general, the testing BrACs were slightly lower than the target BACs, for two reasons. First, the dosing procedure was aimed at avoiding overdosing the subjects, for obvious health and safety reasons. Second, a bottleneck occasionally resulted when two subjects reached the target BrAC at the same time, which delayed some of the testing. The following analyses were conducted with the average of the pre-test BrACs and the post-test BrACs.

Table 5
Average BrAC by BAC Group

BAC Group	*N*	Mean	*SD*	Min	Max
.00	39	.000	.000	.000	.000
.04	40	.038	.009	.035	.044
.08	39	.079	.004	.063	.085
.12	39	.109	.011	.087	.128
All	157	.057	.042	.000	.128

2.3.3 Differences Across Four BAC Groups

As expected, the differences in BrACs across the four BAC groups were statistically significant, $F(3, 153) = 1527.37$, $p<.001$. The mean scores for the six tests across the four BAC groups are shown in Table 6. In general, mean scores increased with higher BrACs. The differences in mean scores across BAC groups were statistically significant for Finger to Nose, Hand Coordination, Palm Pat, and Horizontal Gaze Nystagmus.

Table 6
Average Total Test Scores by BAC Group

		BAC Group				
Test		**.00**	**.04**	**.08**	**.12**	***F***
Finger to Nose	Mean	7.16	8.65	9.44	11.38	5.13**
	SD	4.06	4.23	4.80	5.99	
	N	38	40	39	39	
Time Estimation	Mean	6.49	7.48	7.49	7.95	.35
	SD	5.30	8.61	5.14	6.29	
	N	39	40	39	39	
Finger Count	Mean	4.74	6.90	7.46	7.90	2.19
	SD	4.64	6.56	5.77	6.45	
	N	39	40	39	39	
Hand Coordination	Mean	2.64	2.40	2.74	3.59	2.91*
	SD	1.86	1.96	1.83	1.96	
	N	39	40	39	39	
Palm Pat	Mean	1.51	1.65	1.67	2.62	5.12**
	SD	1.14	1.59	1.06	1.70	
	N	39	40	39	39	
Horizontal Gaze Nystagmus	Mean	2.03	3.60	4.85	5.28	23.53**
	SD	2.02	2.23	1.74	1.43	
	N	39	40	39	39	

Note. * $p \leq .05$. ** $p \leq .01$.

2.3.4 BAC Status

In the field, marine officers need to assess whether boaters are within the legal BAC limit (below .08) or outside the legal BAC limit (at or above .08). Thus, it is important to examine whether the mean scores for the six tests differ significantly between the two conditions. To that end, a new variable was created by dichotomizing the BrACs of the test battery into BrACs below .08 and BrACs at or above .08. Characteristics of the resulting variable (BAC Status) are shown in Table 7. Note that BAC Status analyses were based on 138 of the 157 cases, because 19 subjects had pre-test and post-test BrACs that were on both sides of the .08 cutoff. As expected, the differences in BrACs across BAC Status were statistically significant, $F(1, 136) = 463.16$, $p<.001$.

Table 7
Average BrAC by BAC Status

BAC Status	*N*	Mean	*SD*	Min	Max
BrAC < .080	85	.023	.024	.000	.077
BrAC ≥ .080	53	.102	.016	.081	.128
All	138	.053	.044	.000	.128

2.3.5 Differences Across BAC Status

The mean scores for the six tests across BAC Status are shown in Table 8. The mean scores across BAC Status were consistent with the results across BAC groups.

Table 8
Average Total Test Scores by BAC Status

Test		BAC Status (BrAC < .080)	BAC Status (BrAC ≥ .080)	*F*
Finger to Nose	Mean	8.11	10.77	9.39**
	SD	4.29	5.87	
	N	84	53	
Time Estimation	Mean	6.91	7.91	.74
	SD	6.90	6.14	
	N	85	53	
Finger Count	Mean	5.98	7.98	3.62
	SD	5.73	6.45	
	N	85	53	
Hand Coordination	Mean	2.54	3.28	4.89*
	SD	1.89	1.96	
	N	85	53	
Palm Pat	Mean	1.64	2.41	9.49**
	SD	1.36	1.57	
	N	85	53	
Horizontal Gaze Nystagmus	Mean	2.98	5.06	32.96**
	SD	2.28	1.67	
	N	85	53	

Note. * $p \leq .05$. ** $p \leq .01$.

2.3.6 Correlations

The correlation between the six tests and the average BrAC, BAC group, and BAC Status are shown in Table 9. The test with the highest correlation to BAC was Horizontal Gaze Nystagmus, followed by Finger to Nose, and Palm Pat.

Table 9
Correlations Between BAC and Six Tests

Test	BrAC	BAC Group	BAC Status
Finger to Nose (FTN)	.286**	.299**	.255**
Time Estimation (TE)	.077	.076	.074
Finger Count (FC)	.201*	.187*	.161
Hand Coordination (HC)	.187*	.185*	.186*
Palm Pat (PP)	.236**	.255**	.255**
Horizontal Gaze Nystagmus (HGN)	.557**	.547**	.442**

Note. * $p \leq .05$. ** $p \leq .01$.

2.3.7 BAC Status Classifications

The question of how the tests would predict BAC Status was addressed next. Several sets of analyses were conducted. In the first, logistic regression was used to predict BAC Status on the basis of the combination of the four tests: HGN Positive/Negative, FTN total score, PP total score, and HC total score as the predictors. Of the four tests, only HGN had a pre-established Positive/Negative criterion. Logistic regression allows prediction of group membership from a set of variables. The outcome is the probability of having one outcome or another based on a nonlinear function of the best linear combination of the predictors (Tabachnick & Fidell, 2007). In statistical software packages, the classification criterion is typically set at .5, but it is possible to change it, up or down, if there are reasons to believe that the classification algorithm can be improved. The next set of analyses was for the individual tests. Finally, a prediction of BAC Status was attempted on the combination of HGN Positive/Negative, FTN Positive/Negative, PP Positive/Negative, and HC Positive/Negative.

Note that because the following analyses build on previous analyses, they may capitalize on chance. The following results, therefore, must be interpreted with caution.

2.3.7.1 Individual Tests

2.3.7.1.1 Horizontal Gaze Nystagmus

Of the four tests, only HGN had a pre-established positive/negative criterion (negative = three or less clues, positive = four or more clues). A test of the full model with HGN Positive/Negative scores against a constant-only model was statistically significant, χ^2 (1, N = 138) = 26.48, $p <$ .001. HGN alone correctly predicted BAC Status in 67.4% of the cases. Sensitivity was .87 and specificity was .55. The positive predictive value was .55 and the negative predictive value was .87. Positive likelihood ratio and negative likelihood ratio were 1.94 and .24, respectively.

2.3.7.1.2 Finger to Nose

The Positive/Negative criterion for FTN was set to nine on the basis of the results shown in Table 10. With that criterion, a test of the full model with FTN against a constant-only model was statistically significant, χ^2 (1, N = 137) = 4.83, $p < .05$. FTN alone correctly predicted BAC

Status in 59.9% of the cases. Sensitivity was .58 and specificity was .61. The positive predictive value was .48 and the negative predictive value was .70. Positive likelihood ratio and negative likelihood ratio were 1.49 and .68, respectively.

Table 10
FTN Test's Characteristics at Various Positive/Negative Criteria

FTN Criterion	Prevalence	% Correct	Sensitivity	Specificity	PPV	NPV	LR+	LR-
6	.39	50.4	.85	.29	.43	.75	1.19	.53
7	.39	50.4	.81	.31	.43	.72	1.18	.61
8	.39	56.2	.70	.48	.46	.71	1.33	.63
9	.39	59.9	.58	.61	.48	.70	1.49	.68
10	.39	65.7	.57	.71	.56	.72	1.98	.61
11	.39	65.7	.43	.80	.58	.69	2.14	.71

2.3.7.1.3 Palm Pat

The Positive/Negative criterion for PP was set to two on the basis of the results shown in Table 11. With that criterion, a test of the full model with PP against a constant-only model was statistically significant, χ^2 (1, N = 138) = 4.83, $p < .05$. PP alone correctly predicted BAC Status in 57.2% of the cases. Sensitivity was .66 and specificity was .52. The positive predictive value was .46 and the negative predictive value was .71. Positive likelihood ratio and negative likelihood ratio were 1.37 and .66, respectively.

Table 11
PP Test's Characteristics at Various Positive/Negative Criteria

PP Criterion	Prevalence	% Correct	Sensitivity	Specificity	PPV	NPV	LR+	LR-
1	.38	50.7	.94	.24	.43	.87	1.23	.24
2	.38	57.2	.66	.52	.46	.71	1.37	.66
3	.38	60.9	.42	.73	.49	.67	1.53	.80
4	.38	65.9	.23	.93	.67	.66	3.21	.83

2.3.7.1.4 Hand Coordination

The Positive/Negative criterion for HC was set to three on the basis of the results shown in Table 12. With that criterion, a test of the full model with HC against a constant-only model was statistically significant, χ^2 (1, N = 138) = 3.87, $p < .05$. HC alone correctly predicted BAC Status in 57.2% of the cases. Sensitivity was .64 and specificity was .53. The positive predictive value was .46 and the negative predictive value was .70. Positive likelihood ratio and negative likelihood ratio were 1.36 and .68, respectively.

Table 12 *HC Test's Characteristics at Various Positive/Negative Criteria*								
HC Criterion	**Prevalence**	**% Correct**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**LR+**	**LR-**
1	.38	47.8	.94	.19	.42	.84	1.16	.30
2	.38	50.0	.79	.32	.42	.71	1.16	.65
3	.38	57.2	.64	.53	.46	.70	1.36	.68
4	.38	58.0	.40	.69	.45	.65	1.30	.87

2.3.7.2 Combined Tests

The previous analyses clearly indicate the HGN was the most useful individual test, followed by FTN, PP, and HC. Diagnostics characteristics for the four individual tests and the various combinations of the four tests are shown in Table 13.

HGN and FTN were the best combination of two tests. That combination correctly predicted BAC Status in 73.0% of the cases, χ^2 (1, N = 137) = 23.49, $p < .001$. Sensitivity was .55 and specificity was .85. The positive predictive value was .69 and the negative predictive value was .75. Positive likelihood ratio and negative likelihood ratio were 3.54 and .54, respectively.

HGN, PP, and HC was the best combination of three tests. That combination correctly predicted BAC Status in 71.7% of the cases, χ^2 (1, N = 138) = 12.44, $p < .001$. Sensitivity was .45 and specificity was .88. The positive predictive value was .71 and the negative predictive value was .72. Positive likelihood ratio and negative likelihood ratio were 3.85 and .62, respectively.

The four tests combined correctly predict BAC Status in 67.4% of the cases, χ^2 (1, N = 138) = 9.57, $p < .01$. Sensitivity was .30 and specificity was .91. The positive predictive value was .67 and the negative predictive value was .68. Positive likelihood ratio and negative likelihood ratio were 3.21 and .77, respectively.

HGN and any one of FTN, PP, and HC correctly predicted BAC Status in 70.8% of the cases, χ^2 (1, N = 137) = 26.28, $p < .001$. Sensitivity was .77 and specificity was .67. The positive predictive value was .59 and the negative predictive value was .82. Positive likelihood ratio and negative likelihood ratio were 2.32 and .34, respectively.

Without HGN, the best predictor of BAC Status was the combination of FTN, PP and HC, which correctly predicted 63.5% of the cases, χ^2 (1, N = 138) = 4.52, $p < .05$. Sensitivity was .34 and specificity was .82. The positive predictive value was .55 and the negative predictive value was .66. Positive likelihood ratio and negative likelihood ratio were 1.90 and .80, respectively.

Table 13
Prediction of BAC Status by Four Tests Alone and in Combination

Test	Prevalence	% Correct	Sensitivity	Specificity	PPV	NPV	LR+	LR-	*p*
1-HGN Positive/Negative	.38	67.4	.87	.55	.55	.87	1.94	.24	.001
2-FTN Positive/Negative	.39	59.9	.58	.61	.48	.70	1.49	.68	.05
3-PP Positive/Negative	.38	57.2	.66	.52	.46	.71	1.37	.66	.05
4-HC Positive/Negative	.38	57.2	.64	.53	.46	.70	1.36	.68	.05
1, 2	.39	73.0	.55	.85	.69	.75	3.54	.54	.001
1, 3	.38	69.6	.57	.78	.61	.74	2.53	.56	.001
1, 4	.38	70.3	.55	.80	.63	.74	2.74	.57	.001
2, 3	.39	62.8	.40	.77	.53	.67	1.75	.78	.05
2, 4	.39	62.0	.25	.86	.52	.64	1.72	.88	n/s
3, 4	.38	63.8	.53	.71	.53	.71	1.80	.67	.05
1, 2, 3	.39	68.6	.36	.89	.68	.69	3.35	.72	.01
1, 2, 4	.39	68.6	.38	.88	.67	.69	3.17	.71	.001
1, 3, 4	.38	71.7	.45	.88	.71	.72	3.85	.62	.001
2, 3, 4	.39	63.5	.34	.82	.55	.66	1.90	.80	.05
1, 2, 3, 4	.39	67.4	.30	.91	.67	.68	3.21	.77	.01
1 and any other one test	.39	70.8	.77	.67	.59	.82	2.32	.34	.001
1 and any other two tests	.39	74.5	.58	.85	.70	.77	3.78	.49	.001
2 and any other one test	.39	62.8	.58	.65	.52	.71	1.69	.63	.01
2 and any other two tests	.39	66.4	.47	.79	.58	.70	2.20	.67	.01
3 and any other one test	.39	59.9	.64	.57	.49	.72	1.50	.63	.05
3 and any other two tests	.39	69.3	.55	.79	.62	.73	2.55	.58	.001
4 and any other one test	.39	58.4	.62	.56	.47	.70	1.41	.67	.05
4 and any other two tests	.39	69.3	.57	.77	.61	.74	2.50	.56	.001

Note. PPV = Positive predictive value. NPV = Negative predictive value. LR+ = Positive likelihood ratio. LR- = Negative likelihood ratio. p = Significance of the test (χ^2) of the full model against a constant-only model. HGN = Horizontal gaze nystagmus, FTN = Finger to nose. PP = Palm pat. HC = Hand coordination.

2.4 Discussion

The objective of this laboratory project was to develop sobriety tests for the marine environment. Six seated tests were evaluated in the laboratory to determine their feasibility for use on the water. Data were obtained under double-blind conditions, at relatively low BACs. The six tests were administered by law enforcement officers with an average 9.7 years of experience administering the roadside SFSTs.

HGN was found to be the most useful of tests in predicting BACs of .08 and above, followed by

FTN, PP, and HC. HGN and any one of FTN, PP, and HC correctly predicted BAC Status in 71% of the cases. Given that HGN is the most useful individual test, it should always be included in the arrest/release decision. HGN and any one of FTN, PP, and HC correctly predicted BAC Status in 71% of the cases. The positive likelihood ratio of 2.32 and the negative likelihood ratio of .34 indicate that this combination is useful in detecting alcohol-related impairment, but not conclusively. Without HGN, the best predictor of BAC Status was the combination of FTN, PP and HC, which correctly predicted 64% of the cases.

The overall correct percentages, sensitivity, and specificity of the tests were below what is typically reported in literature on the roadside SFSTs. Comparisons with prior studies, however, should be made with caution. First, in this study, the average BrACs were considerably lower than in previous studies. In the Burns and Moskowitz (1977) study, for example, 48 subjects were tested at a mean BrAC of .120 and 16 subjects were tested at a mean BrAC of .156. In comparison, in the current study, the highest BAC group was tested at a mean BrAC of .110. The wider distribution of BrACs in the previous studies may have made the detection of impairment less difficult than in the current study.

Second, the impairment decisions were made exclusively on the basis of the tests, without external clues such as smell of alcohol beverages, appearance, speech, and demeanor. Third, although the officers in the current study were required to have prior experience administering the roadside SFSTs, and were, therefore, assumed to have nearly equal proficiency in administering HGN, great differences in proficiency were in fact observed between the officers. Also, given that officers collected data for three days, with six subjects scheduled per day, the maximum number of subjects that could be examined by a single officer was 18, which may not have been enough to master the tests.

Although the tests, as administered and scored by officers in the laboratory, had lower correct percentages, sensitivity, and specificity than is typically reported in the literature, they showed enough promise to warrant a field study.

3 Field Study: Validation of Sobriety Tests for the Marine Environment

3.1 Introduction

In the laboratory portion of the study, it was hypothesized that a seated battery of tests might be of potential usefulness to marine officers in detecting impairment in boaters due to BACs of .08 and above, provided that officers are adequately trained. The purpose of this portion of the project was to test that hypothesis in the field, in real-life conditions.

In prior validation studies of the SFSTs, the general approach has been to have officers stop drivers suspected of driving under the influence of alcohol, administer three standardized sobriety tests, and make an arrest/release decision on the basis of the three tests. The accuracy of the arrest/release decision was verified with a portable breath alcohol screener, which was administered following the sobriety tests by trained civilian observers.

The field portion of this project followed the same approach. Marine officers stopped boaters suspected of BUI, asked them to come aboard the patrol vessel, and administered four sobriety tests. The four sobriety tests were horizontal gaze nystagmus (HGN), finger to nose, (FTN), palm pat (PP), and hand coordination (HC). Lastly, an alcohol breath test was obtained to verify the accuracy of the tests in detecting BACs and/or BrACs of .08 and above.

Unlike previous SFSTs validation studies, the alcohol breath tests were administered by the marine officers, not the civilian observers. This was required for practical and safety reasons due to the small size of the deck on the police vessel. The space limitation made it cumbersome for the marine officer and the observer to switch places in order for the observer to interact with the BUI suspect and administer the alcohol breath test. The switch would have created a potentially unsafe situation in which the officer could not guarantee the safety of the boater and the observer. The role of the observers, therefore, was limited to ensuring that the alcohol breath test consistently followed the four sobriety tests in the examination.

In the laboratory study, we found that officers could have benefited from additional training in test administration and scoring. Prior to initiation of data collection for the field study, therefore, the study officers attended a four-day training period that involved both classroom instruction and field practice. Unlike officers in the laboratory study, field study officers were given clear Positive/Negative criteria for FTN, PP, and HC.

3.2 Method

3.2.1 Study Site

The study was conducted on the Lake of the Ozarks in central Missouri. The Missouri State Water Patrol (MSWP) was the agency that collaborated with SCRI for the purposes of the study. MSWP is based in Jefferson City, but the study site was in Osage Beach.

The Lake of the Ozarks was selected as the study site for two reasons. The first was the cooperation of MSWP, which provided study officers. The second was that the lake is a popular boating destination, with enough cases of BUI to support data collection for the study.

3.2.2 Study Officers

Four marine officers were selected by the MSWP for participation in the study. All four officers had prior experience administering the HGN test.

3.2.3 Officers' Training

Officer training spanned four days, beginning Thursday, June 18, 2009. Day 1 consisted of an 8-hour in-class explanation and demonstration of the four sobriety tests (HGN, FTN, PP, and HC). During that class, conducted by SCRI staff, the officers became familiar with the administration and scoring of the tests. Two volunteers drank until their BrACs were over .08. The four officers then practiced on the volunteers while the SCRI staff provided feedback. Days 2, 3, and 4 consisted of 10-hour shifts, in patrol boats on the water, with the sole purpose of allowing the marine officers to become proficient with the tests.

3.2.4 Civilian Observers

Two SCRI employees were the designated observers for all study activities. They were based in Osage Beach for the duration of the study. For the observers' safety, the officer was always positioned between them and the suspect. The observers were close enough to observe the suspects' performance but far enough as to not interfere (about 5 feet away).

3.2.5 Sobriety Tests

The results of the laboratory portion of this project indicated that four seated tests, HGN, FTN, PP, and HC might be useful to marine officers in detecting impairment due to BACs of .08 and above. The four tests were:

- *Horizontal Gaze Nystagmus*. The HGN test is comprised of three separate checks, administered independently to each eye. Four or more clues indicate impairment due to BAC $\geq$ .08.
- *Finger to Nose*. The FTN test requires the subject to bring the tip of the index finger to touch the tip of the nose. It is performed with eyes closed and head tilted slightly back. Nine or more clues indicate impairment due to BAC $\geq$ .08.
- *Palm Pat*. The PP test requires the subjects to place one hand extended, palm up, out in front of them. The other hand is placed on top of the first with the palm facing down. The top hand then begins to pat the bottom hand. The top hand rotates 180 degrees alternating between the back of the hand and the palm of the hand. The bottom hand remains stationary. The subjects count out loud in relation with each pat. Two or more clues indicate impairment due to BAC $\geq$ .08.
- *Hand Coordination*. The HC test requires the subjects to perform a series of tasks with their hands. It is very loosely adapted from the Walk-And-Turn test performed on land. Three or more clues indicate impairment due to BAC $\geq$ .08.

3.2.6 Equipment

Officers used a pen, pencil, or small flashlight as the stimulus for the HGN test. SCRI provided MSWP with four Alco Sensor FST (Intoximeter, Inc., St. Louis, MO) breath alcohol testing instruments. The observers were required to meet the MSWP's water safety requirements while on the patrol boat.

3.2.7 Study Dates and Shifts

Data were collected from Friday, June 26, 2009 to Monday, September 7, 2009, inclusive. Data were collected during the expected busiest boating days: Fridays, Saturdays, Sundays, and holidays. Shifts started at 12 p.m. and lasted from 10 to 12 hours, depending on the workload.

3.2.8 Procedures

The general procedures for the study were as follows. The officers stopped boaters suspected of BUI and asked them to come aboard the patrol boat. The suspects sat on a bench seat on the stern of the boat. After a few agency-specific questions, the officer administered the sobriety tests in the following order: HGN, FTN, PP, and HC. The tests were scored during administration. Following the tests, two successive alcohol breath tests were administered. At this point, based on evidence from the sobriety tests and the breath alcohol tests, the officer either released or arrested the boater. The observers ensured that the sobriety tests' data were collected prior to the alcohol breath tests.

In case the BUI suspect was released, the officer and the observer resumed patrolling the assigned area. In case of an arrest, the suspect was brought ashore and processed for arrest by the officer. Because that often took some time, the observer often teamed up with another available study officer.

Of the 331 study cases, 251 (76%) were obtained with observers present and 80 (24%) were obtained without observers. When possible, given the limitations of operating in a small space on the patrol boat, the observers also scored some of the sobriety tests while the officers were administering them. Only a portion of the FTN, PP, and HC tests could be scored by the observers. No HGN test could be scored by the observers because it was impossible to clearly see the suspects' eyes from their position on the boat. The observers and the officers never shared their results prior to the administration of the alcohol breath tests.

3.3 Results

3.3.1 Stop Characteristics

3.3.1.1 Stop Times

With observers, data collection hours ranged from (24-hour clock) 13:59 to 06:04. Without observers, data collection hours ranged from 10:20 to 07:04.

Table 14 *Stop Times by Observer Status*		
Time	**Number of Stops with Observers**	**Number of Stops without Observers**
12:00 to 12:59	0	4
13:00 to 13:59	1	2
14:00 to 14:59	8	4
15:00 to 15:59	17	2
16:00 to 16:59	20	6
17:00 to 17:59	53	16
18:00 to 18:59	34	11
19:00 to 19:59	22	7
20:00 to 20:59	34	4
21:00 to 21:59	14	5
22:00 to 22:59	17	6
23:00 to 23:59	17	2
00:00 to 00:59	11	3
01:00 to 01:59	1	1
02:00 to 02:59	1	1
03:00 to 11:59	1	3
Missing	0	3
Total	251	80

3.3.1.2 Type of Stop

There were two types of stops in the study. A probable cause stop involved a boater suspected of BUI by the officer. A checkpoint stop involved a boater selected at random from the flow of boats. With observers, 221 (88%) of the stops were probable cause and 30 (12%) of the stops were checkpoint. Without observers, 41 (51.3%) of the stops were probable cause, 14 (17.5%) were checkpoint, and 25 (31.3%) were unknown.

3.3.1.3 Weather Conditions

When possible, weather conditions, wind conditions, air temperature, water temperature, surface condition, and lighting condition were recorded for each stop.

Table 15
Weather Conditions by Observer Status

Condition		**Number of Stops with Observers**	**Number of Stops without Observers**
Weather	Clear	195	23
	Cloudy	56	12
	Rain	0	1
	Missing	0	44
Wind	0 – 5 mph	203	22
	6 – 15 mph	41	11
	Over 16 mph	1	0
	Missing	0	44
Air Temperature	60s	7	1
	70s	80	12
	80s	129	12
	90s	27	3
	100s	2	0
	Missing	6	52
Water Temperature	70s	48	5
	80s	188	22
	90s	1	0
	Missing	14	54
Surface	Calm	198	33
	Choppy	42	3
	Rough	11	0
	Missing	0	44
Lighting	Daylight	148	25
	Dusk – Down	30	3
	Dark with lights	69	7
	Dark no lights	4	2
	Missing	0	43

3.3.1.4 Sample Characteristics

Occasionally, it was necessary to release control of a boat to a suitable passenger. Some passengers, therefore, were administered the tests to determine their level of impairment. Although the passengers were not tested for the purpose of an arrest/release decision, their data were included in the analyses. Sample characteristics by observer status are shown in Table 16.

Table 16
Sample Characteristics by Observer Status

Characteristic		Number of Stops with Observers	Number of Stops without Observers
Boater Type	Driver	218	50
	Passenger	27	6
	Other	6	0
	Missing	0	24
Watercraft	Jet Ski	17	7
	Pontoon	59	11
	Cruiser	39	9
	Other	136	25
	Missing	0	28
Age	19	2	1
	20s	73	25
	30s	61	12
	40s	70	20
	50s	31	5
	60s	11	2
	70s	1	1
	80s	1	0
	Missing	0	14
Sex	Male	238	62
	Female	13	4
	Missing	0	14
Ethnicity	Latino/a	3	0
	Non Latino/a	248	27
	Missing	0	53
Race	African American	2	0
	American Indian	1	0
	Asian	0	1
	Caucasian	248	56
	Other	0	0
	Missing	0	23
Open Container	Yes	143	28
	No	107	9
	Missing	1	43
Alcohol Odor	No	82	7
	Faint	78	27
	Moderate	72	15
	Strong	19	7
	Missing	0	24
Eyes	Normal	172	26
	Watery	25	3

Table 16
Sample Characteristics by Observer Status

Characteristic		**Number of Stops with Observers**	**Number of Stops without Observers**
	Bloodshot	47	5
	Glassy	7	1
	Staring	0	0
	Other	0	19
	Missing	0	26
Pupils	Normal	250	52
	Dilated	1	2
	Contracted	0	0
	Poor reaction to light	0	0
	Missing	0	26
Balance	Sure	155	21
	Fair	88	21
	Swaying	3	6
	Wobbling	2	0
	Sagging knees	0	0
	Falling	2	0
	Other	1	1
	Missing	0	31
Speech	Coherent	228	44
	Slurred	21	7
	Confused	1	0
	Incoherent	0	0
	Stuttering	0	0
	Mumbling	1	0
	Other	0	1
	Missing	0	28
Clothing	Neat	177	45
	Mussed	8	0
	Other	66	1
	Missing	0	34
Soiled	No	245	19
	Dirt	6	0
	Urine	0	0
	Vomit	0	0
	Saliva	0	0
	Other	0	0
	Missing	0	61
Attitude	Cooperative	197	50
	Indifferent	32	1
	Antagonistic	9	0

Table 16 *Sample Characteristics by Observer Status*			
Characteristic		**Number of Stops with Observers**	**Number of Stops without Observers**
	Cocky	3	0
	Combative	4	1
	Insulting	0	0
	Uncooperative	0	0
	Other	6	1
	Missing	0	53

3.3.1.5 BrACs

In general, BrACs ranged from .000% to .320%. Table 17 shows the frequency of BrACs by observer status.

Table 17 *BrACs by Observer Status*		
BrAC	**Number of Stops with Observers**	**Number of Stops without Observers**
.000	37	11
.001 to .009	8	6
.010 to .019	24	5
.020 to .029	9	8
.030 to .039	18	1
.040 to .049	12	8
.050 to .059	14	3
.060 to .069	8	2
.070 to .079	11	5
.080 to .089	10	6
.090 to .099	13	3
.100 to .109	11	7
.110 to .119	10	3
.120 to .129	13	1
.130 to .139	11	0
.140 to .149	9	3
.150 to .159	9	3
.160 to .169	5	1
.170 to .179	6	2
.180 to .189	3	0
.190 to .199	2	0
.200 to .209	3	0
.210 to .219	1	0
.220 to .229	1	1
.230 to .239	0	0

Table 17
BrACs by Observer Status

BrAC	Number of Stops with Observers	Number of Stops without Observers
.240 to .249	0	1
.250 to .259	0	0
.260 to .269	0	0
.270 to .279	1	0
.280 to .289	0	0
.290 to .299	0	0
.300 to .309	0	0
.310 to .319	0	0
.320 to .329	1	0
Missing	1	0
Total	251	80

3.3.2 Observational Clues

Officers rely on the totality of the evidence in the decision to arrest or release a BUI suspect. Although observational clues were not part of the four tests, they may have provided converging evidence of impairment/no impairment.

To examine this issue, correlation analyses were conducted to determine how observational clues were related to BAC. Of all the observational clues, the most strongly related to BrAC was odor of alcoholic beverage, $r = .68$, $p < .001$; followed by speech characteristic, $r = .23$, $p < .001$; eye condition, $r = .22$, $p < .001$; clothing condition, $r = .17$, $p < .01$; attitude, $r = .16$, $p < .01$; and balance, $r = .16$, $p < .01$.

3.3.3 BrACs and Tests' Differences by BAC Status

As in the laboratory study, the tests were examined by BAC Status (BrAC < .08 v. BrAC ≥ .08). Note that this is a very conservative approach as it classifies cases on the basis of the criterion rather than the behavioral characteristics of the subject. One of the 251 cases from the observer data was missing a BrAC, as the boater refused to provide a breath or blood specimen. That case was dropped from the analyses. One case from the without observer data was missing the HGN test. That case was included in the analyses.

Analysis of variance (ANOVA) was conducted on each of the variables to determine whether BrAC, HGN, FTN, PP, and HC varied as a function of BAC Status. Table 18 reports the means, standard deviation (*SD)* and number of cases for each of those variables by BAC Status and observer status. Because there were only minor differences between the data set obtained with observers and the data set obtained without observers, only the ANOVAs for the combined data set are reported here.

As expected, the differences between lower BrACs ($M = .028$) and higher BrACs ($M = .133$) were statistically significant, $F(1, 328) = 837.36$, $p < .001$. There were statistically significant differences in total scores as a function of BAC Status for all four tests: HGN, $F(1, 327) = 377.10$, $p < .001$; FTN, $F(1, 328) = 34.76$, $p < .001$; PP, $F(1, 328) = 73.15$, $p < .001$; and HC,

$F(1, 328) = 12.03$, $p < .01$.

Table 18
BrAC and HGN, FTN, PP, HC Total Clues by BAC Status and Observer Status

Variable		With Observers		Without Observers		Combined	
		BrAC < .08	**BrAC ≥ .08**	**BrAC < .08**	**BrAC ≥ .08**	**BrAC < .08**	**BrAC ≥ .08**
BrAC	Mean	.028	.134	.028	.125	.028	.133
	SD	.025	.040	.025	.041	.028	.041
	N	141	109	49	31	190	140
HGN	Mean	1.45	4.98	1.94	4.97	1.58	4.98
	SD	1.57	1.53	1.68	1.54	1.61	1.52
	N	141	109	48	31	189	140
FTN	Mean	5.72	8.17	6.29	8.59	5.86	8.26
	SD	3.34	3.72	3.42	5.00	3.36	4.02
	N	141	109	49	31	190	140
PP	Mean	1.30	2.34	1.41	2.35	1.33	2.34
	SD	.93	1.18	1.02	1.28	.95	1.20
	N	141	109	49	31	190	140
HC	Mean	2.34	3.00	2.29	2.52	2.33	2.89
	SD	1.49	1.42	1.62	1.21	1.52	1.39
	N	141	109	49	31	190	140

3.3.4 Correlations Between BrAC, BAC Status and HGN, FTN, PP, and HC

The correlations between the four tests, BrAC, and BAC Status are shown in Table 19. The test with the highest correlation to BAC was HGN, followed by PP, FTN, and HC. There was no statistically significant difference in the HGN-BrAC correlation between data collected with observers and data collected without observers, $z = -.65$, $p = .51$. There was no statistically significant difference in the FTN-BrAC correlation between data collected with observers and data collected without observers, $z = .56$, $p = .58$. There was no statistically significant difference in the PP-BrAC correlation between data collected with observers and data collected without observers, $z = .47$, $p = .64$. There was no statistically significant difference in the HC-BrAC correlation between data collected with observers and data collected without observers, $z = 1.27$, $p = .20$.

Table 19
Correlations between BrAC, BAC Status and HGN, FTN, PP and HC by Observer Status

Variable	With Observers		Without Observers		Combined	
	BrAC	**BAC Status**	**BrAC**	**BAC Status**	**BrAC**	**BAC Status**
HGN	.757**	.750**	.791**	.581**	.761**	.715**
FTN	.396**	.324**	.333**	.228	.375**	.298**
PP	.471**	.428**	.422**	.299*	.458**	.403**
HC	.297**	.207**	.139	.030	.265**	.182**

Note. * $p \leq .05$. ** $p \leq .01$.

3.3.5 Positive/Negative Classifications

Because there were no statistically significant differences between the data collected with observers and the data collected without observers in the correlations between BAC and each of the four tests, it was possible to conduct the classification analyses on the combined data set. Table 20 summarizes the classification analyses.

3.3.5.1 Individual Tests

3.3.5.1.1 Horizontal Gaze Nystagmus

A test of the full model with HGN Positive/Negative scores against a constant-only model was statistically significant, χ^2 (1, N = 329) = 174.31, $p < .001$. HGN alone correctly predicted BAC Status in 84.8% of the cases. Sensitivity was .86 and specificity was .84. The positive predictive value was .80 and the negative predictive value was .89. Positive likelihood ratio and negative likelihood ratio were 5.27 and .16, respectively.

3.3.5.1.2 Finger to Nose

A test of the full model with FTN Positive/Negative scores against a constant-only model was statistically significant, χ^2 (1, N = 330) = 32.85, $p < .001$. FTN alone correctly predicted BAC Status in 67.3% of the cases. Sensitivity was .49 and specificity was .81. The positive predictive value was .65 and the negative predictive value was .68. Positive likelihood ratio and negative likelihood ratio were 2.56 and .63, respectively.

3.3.5.1.3 Palm Pat

A test of the full model with PP Positive/Negative scores against a constant-only model was statistically significant, χ^2 (1, N = 330) = 37.74, $p < .001$. PP alone correctly predicted BAC Status in 65.2% of the cases. Sensitivity was .76 and specificity was .57. The positive predictive value was .57 and the negative predictive value was .77. Positive likelihood ratio and negative likelihood ratio were 1.77 and .41, respectively.

3.3.5.1.4 Hand Coordination

A test of the full model with HC Positive/Negative scores against a constant-only model was statistically significant, χ^2 (1, N = 330) = 12.37, $p < .001$. HC alone correctly predicted BAC Status in 59.4% of the cases. Sensitivity was .62 and specificity was .57. The positive predictive value was .52 and the negative predictive value was .67. Positive likelihood ratio and negative likelihood ratio were 1.46 and .66, respectively.

3.3.5.2 Combined Tests

HGN and FTN were the best combination of two tests. Combined, they correctly predicted BAC Status in 74.5% of the cases, χ^2 (1, N = 329) = 86.44, $p < .001$. Sensitivity was .46 and specificity was .96. The positive predictive value was .89 and the negative predictive value was .70. Positive likelihood ratio and negative likelihood ratio were 10.80 and .57, respectively.

HGN, FTN, and PP was the best combination of three tests. That combination correctly predicted BAC Status in 72.0% of the cases, χ^2 (1, N = 329) = 72.62, $p < .001$. Sensitivity was .39 and specificity was .97. The positive predictive value was .90 and the negative predictive value was .68. Positive likelihood ratio and negative likelihood ratio were 12.15 and .63, respectively.

The four tests combined correctly predicted BAC Status in 68.1% of the cases, χ^2 (1, N = 329) = 50.71, $p < .001$. Sensitivity was .28 and specificity was .98. The positive predictive value was .91 and the negative predictive value was .65. Positive likelihood ratio and negative likelihood ratio were 13.16 and .74, respectively.

HGN and any one of FTN, PP, and HC correctly predicted BAC Status in 84.5% of the cases, χ^2 (1, N = 329) = 165.67, $p < .001$. Sensitivity was .81 and specificity was .87. The positive predictive value was .82 and the negative predictive value was .86. Positive likelihood ratio and negative likelihood ratio were 6.36 and .22, respectively.

Without HGN, the best predictor of BAC Status was the combination of FTN, PP and HC, which correctly predicted 66.1% of the cases, χ^2 (1, N = 330) = 29.99, $p < .001$. Sensitivity was .29 and specificity was .93. The positive predictive value was .76 and the negative predictive value was .64. Positive likelihood ratio and negative likelihood ratio were 4.28 and .76, respectively.

Table 20

Prediction of BAC Status by Four Tests Alone and in Combination

Test	Prevalence	% Correct	Sensitivity	Specificity	PPV	NPV	LR+	LR-	*p*
1-HGN Positive/Negative	.43	84.8	.86	.84	.80	.89	5.27	.16	.000
2-FTN Positive/Negative	.42	67.3	.49	.81	.65	.68	2.56	.63	.000
3-PP Positive/Negative	.42	65.2	.76	.57	.57	.77	1.77	.41	.000
4-HC Positive/Negative	.42	59.4	.62	.57	.52	.67	1.46	.66	.000
1, 2	.43	74.5	.46	.96	.89	.70	10.80	.57	.000
1 ,3	.43	81.5	.70	.90	.84	.80	6.96	.33	.000
1, 4	.43	76.0	.56	.90	.81	.74	5.93	.48	.000
2, 3	.42	68.5	.41	.89	.73	.67	3.68	.67	.000
2, 4	.42	65.8	.34	.89	.70	.65	3.19	.74	.000
3, 4	.42	66.1	.51	.77	.62	.68	2.22	.63	.000
1, 2, 3	.43	72.0	.39	.97	.90	.68	12.15	.63	.000
1, 2, 4	.43	69.0	.31	.97	.88	.66	9.90	.71	.000
1, 3, 4	.43	74.5	.49	.93	.84	.71	7.17	.54	.000
2, 3, 4	.42	66.1	.29	.93	.76	.64	4.28	.76	.000
1, 2, 3, 4	.43	68.1	.28	.98	.91	.65	13.16	.74	.000
1 and any other one test	.43	84.5	.81	.87	.82	.86	6.36	.22	.000
1 and any other two tests	.43	79.3	.64	.91	.84	.77	7.07	.40	.000
2 and any other one test	.43	70.8	.44	.91	.78	.69	4.84	.62	.000
2 and any other two tests	.43	70.8	.44	.91	.78	.69	4.84	.62	.000
3 and any other one test	.43	71.3	.73	.70	.65	.78	2.46	.39	.000
3 and any other two tests	.43	76.3	.61	.87	.78	.75	4.84	.44	.000
4 and any other one test	.43	66.3	.59	.71	.61	.70	2.08	.57	.000
4 and any other two tests	.43	73.3	.54	.87	.76	.72	4.28	.52	.000

Note. PPV = Positive predictive value. NPV = Negative predictive value. LR+ = Positive likelihood ratio. LR- = Negative likelihood ratio. *p* = Significance of the test (χ^2) of the full model against a constant-only model. HGN = Horizontal gaze nystagmus, FTN = Finger to nose. PP = Palm pat. HC = Hand coordination.

3.3.6 Reliability

When possible, the observers scored the FTN, PP, and HC tests while the officer administered the tests to the BUI suspects. HGN could not be scored because the eyes of the BUI suspects were not clearly visible from where the observers were standing on the patrol boat.

Table 21 reports the reliability findings for FTN, PP, and HC. In general, the officers and the observers showed substantial to almost perfect agreement.

Table 21
Tests' Reliability

Test	Correlation (Total Scores)	Kappa (Positive/Negative Scores)	Interpretation
FTN	.838	.727	Substantial agreement
PP	.837	.866	Almost perfect agreement
HC	.822	.840	Almost perfect agreement

3.4 Discussion

The objective of this field project was to validate sobriety tests for the marine environment. Four sobriety tests were administered on the water to determine their usefulness in detecting impairment caused by BACs of .08 and above. Data were obtained by a team of four marine officers and two civilian observers.

Officers' training was identified in the laboratory study as a potential factor in the results of that study. In the field study, extensive training was followed by a period of on-water practice. Only when the officers were proficient and comfortable administering and scoring the test did data collection begin.

The tests were administered at almost all hours of the day; with probable cause or at sobriety checkpoints; under clear or cloudy weather; with and without wind; at various water and air temperatures; on calm, choppy, or rough water surface; and under various lighting conditions.

The sample of boaters was relatively homogeneous, as it consisted predominantly of Caucasian males. Very few women, Latinos, African Americans, and Asians were stopped for the study. Ages ranged from 19 to 80 years.

Study BrACs ranged from .000 to .320. Suspects' odor of alcohol strongly correlated with BrAC ($r = .68$). Other observational clues that correlated with BrAC were speech, eye condition, clothing condition, attitude, and balance.

As in the laboratory study, HGN was found to be the most useful test in predicting BACs of .08 and above, followed by FTN, PP, and HC. A positive HGN test indicates a .80 probability (80%) that the BUI suspect has a BAC ≥ .08. Thus, HGN is a very good predictor of BAC Status (BAC < .08 v. BAC ≥ .08) not only at roadside, but also on the water. Alone, it can correctly identify 84.8% of BUI suspects as either BAC < .08 or BAC ≥ .08. Officers who can properly administer it and score it may confidently rely on it to form their arrest/release decision.

FTN is a moderate predictor of BAC Status (BAC < .08 v. BAC ≥ .08). Alone, it can correctly identify 67.3% of BUI suspects as either BAC < .08 or BAC ≥ .08. A positive FTN test indicates a .65 probability that the BUI suspect has a BAC ≥ .08.

The PP and HC tests are only fair predictors of BAC Status. Alone, they can correctly identify 65.2% and 59.4%, respectively, of BUI suspects as either BAC < .08 or BAC ≥ .08. A positive PP test indicates a .57 probability that the BUI suspect has a BAC ≥ .08. A positive HC test indicates a .52 probability that the BUI suspect has a BAC ≥ .08.

HGN and any one of FTN, PP, and HC correctly predicted BAC Status in 84.5% of the cases. The positive likelihood ratio of 6.36 and the negative likelihood ratio of .22 indicate that this combination is useful in detecting alcohol-related impairment.

Without HGN, the best predictor of BAC Status was the combination of FTN, PP, and HC, which correctly predicted 66.1% of the cases. The positive likelihood ratio of 4.28 and the negative likelihood ratio of .76 indicate that this combination is likely to be moderately useful in detecting alcohol-related impairment.

The overall correct percentages, sensitivity, and specificity of the tests were consistent with what is typically reported in literature on the roadside SFSTs. It should be noted that the prevalence of BACs at or above .08 was lower in the current field study (.43) than in the previous field studies on SFSTs (.79, .80, .73).

The tests' reliability was also consistent with what is typically reported in literature on the roadside SFSTs. Note, however, that HGN could not be included in the reliability analyses because it was impossible for the observers to clearly see the suspects' eyes from their position on the boat.

4 Summary and Conclusions

The current project is the first to systematically examine the usefulness of seated sobriety tests for use in the marine environment. Results from the laboratory study and the field study support the use of four tests for detection of impairment due to BACs of .08 or above in the marine environment, provided that officers are extensively trained to administer and score the tests. The four tests are HGN, FTN, PP, and HC.

It is proposed that marine officers administer HGN, FTN, PP, and HC to all BUI suspects, and then, for each suspect, use the pattern of test results to estimate the probability of BAC ≥ .08, as follows:

- Positive HGN, FTN, PP, and HC tests indicate a .91 probability that the BUI suspect has a BAC ≥ .08.
- Positive HGN, FTN, and PP tests indicate a .90 probability that the BUI suspect has a BAC ≥ .08.
- Positive HGN and FTN tests indicate a .89 probability that the BUI suspect has a BAC ≥ .08.
- Positive FTN, PP and HC tests indicate a .76 probability that the BUI suspect has a BAC ≥ .08.
- Positive FTN and PP tests indicate a .73 probability that the BUI suspect has a BAC ≥ .08.

The usefulness of this approach should be assessed periodically, including a systematic review of the performance of each test, alone and in combination. If necessary, changes in administration and scoring may be required from time to time to maximize the predictive power of the battery.

References

Alberg, A. J., Park, J. W., Hager, B. W., Brock, M. V., & Diener-West, M. (2004). The use of "overall accuracy" to evaluate the validity of screening or diagnostic tests. *Journal of General Internal Medicine, 19,* 460-465.

Browne, M. L., Lewis-Michl, E. L., & Stark, A. D. (2003). Watercraft-related drownings among New York State residents, 1988-1994. *Public Health Reports, 118,* 459-463.

Burns, M., & Anderson, E. W. (1995). *A Colorado Validation Study of the Standardized Field Sobriety Test (SFST) Battery*. Washington, DC: National Highway Traffic Safety Administration.

Burns, M., & Dioquino, T. (1997). *A Florida Validation Study of the Standardized Field Sobriety Test (SFST) Battery*. Washington, DC: National Highway Traffic Safety Administration.

Burns, M., & Moskowitz, H. (1977). *Psychophysical Tests for DWI Arrest* (US Department of Transportation Rep. No. DOT HS-802 424). Washington, DC: National Highway Traffic Safety Administration.

Cahalan, D., Cisin, I. H., & Crossley, H. M. (1969). *American drinking practices; a national study of drinking behavior and attitudes*. New Brunswick, NJ: Rutgers Center of Alcohol Studies.

Dell'Osso, L. F. (1989). Nystagmus and saccadic intrusions and oscillations. In S. Lessell & J. T. W. van Dalen (Eds.), *Current Neuro-Ophthalmology: Vol. 2* (pp. 147-182). Chicago, IL: Year Book Medical Publishers, Inc.

Driscoll, T. R., Harrison, J. A., & Steenkamp, M. (2004). Review of the role of alcohol in drowning associated with recreational aquatic activity. *Injury Prevention, 10*, 107-113.

Grimes, D. A., & Schulz, K. F. (2005). Refining clinical diagnosis with likelihood ratios. *Lancet, 365*, 1500-1505.

Khiabani, H. Z., Opdal, M. S., & Morland, J. (2008). Blood alcohol concentrations in apprehended drivers of cars and boats suspected to be impaired by the police. *Traffic Injury Prevention, 9,* 31-36.

LeBlond, R. F., Brown, D. D., & DeGowin, R. L. (2009). *DeGowin's diagnostic examination* (9th ed., pp. 798-812). New York: McGraw-Hill.

Logan, P., Sacks, J. J., Branche, C. M., Ryan, G. W., & Bender, *P*. (1999). Alcohol-influenced recreational boat operation in the United States, 1994. *American Journal of Preventive medicine, 16,* 278-282.

Lunetta, P., Penttila, A., & Sarna, S. (1998). Water traffic accidents, drowning and alcohol in Finland, 1969-1995. *International Journal of Epidemiology, 27*, 1038-1043.

National Highway Traffic Safety Administration. (1999). *Horizontal Gaze Nystagmus: The Science and the Law*. Washington, DC: Author.

Smith, G. S., Keyl, P. M., Hadley, J. A., Bartley, C. L., Foss, R. D., Tolbert, W. G., & McKnight, J. (2001). Drinking and recreational boating fatalities: a population-based case-control study. *Journal of American Medical Association, 286,* 2974-2980.

Stuster, J. (2006). Validation of the standardized field sobriety test battery at 0.08% blood alcohol concentration. *Human Factors, 48,* 608-614.

Sussman, E. D., Needalman, A., & Mengert, P. H. (1990). *An experimental evaluation of a field sobriety test battery in the marine environment*. Washington, DC: United States Coast Guard.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, MA: Pearson.

Tharp, V., Burns, M., & Moskowitz, H. (1981). *Development and Field Test of Psychophysical Tests for DWI Arrest* (US Department of Transportation Rep. No. DOT-H-805-864). Washington, DC: National Highway Traffic Safety Administration.

Tseng, Y. P., Kyle, G. T., Shafer, C. S., Graefe, A. R., Bradle, T. A., & Schuett, M. A. (2009). Exploring the crowding-satisfaction relationship in recreational boating. *Environmental Management, 43,* 496-507.

Tzamaloukas, A. H., Murata, G. H., Vanderjagt, D. J., Servilla, K. S., & Glew, R. H. (2003). Body composition evaluation in peritoneal dialysis patients using anthropometric formulas estimating body water. *Advances in Peritoneal Dialysis, 19,* 212-216.

Appendix A: Tests' Instructions and Scoring

General Instructions

To ensure that the subjects are stable, give the following instructions to all subjects before starting any of the tests.

- **Please sit straight at the front edge of your seat**
- **Put your arms down at your side**
- **Place your feet shoulder-width so that you are comfortable and stable**
- **Are you comfortable and stable?**
- Wait for response
- **Do not move your feet until the test is over. Stay in this position. Don't do anything else until I tell you to do so. Do you understand?**
- Get acknowledgement of understanding.

Horizontal Gaze Nystagmus

This test is comprised of three separate checks, administered independently to each eye.

Administrative Procedures

- Ask if the subject is wearing contact lenses and note the response. If the subject is wearing eyeglasses, have them removed.
    - "I am going to check your eyes."
    - "Keep your head still and follow this stimulus with your eyes only."
    - "Keep following the stimulus with your eyes until I tell you to stop."
    - "Do you understand?"
- Position the stimulus approximately 12-15 inches from the nose and slightly above eye level.
- Check to see that both pupils are equal in size and for the presence of resting nystagmus.
- Check the subject's eyes for the ability to track together.
    - Move the stimulus smoothly across the subject's entire field of vision. Check to see if the eyes track the stimulus together or one lags behind the other. If the eyes don't track together it could indicate possible medical disorder, injury, or blindness.
- Check both eyes for lack of smooth pursuit.
    - Check the subject's left eye by moving the stimulus to the right. Move the stimulus smoothly, at a speed that requires approximately two seconds to bring the subject's eye as far to the side as it can go. Look at the subject's eye and determine whether it is able to pursue smoothly.
    - Move the stimulus all the way to the left, back across subject's face checking if the right eye pursues smoothly. Movement of the stimulus should take approximately two seconds out and two seconds back for each eye. Repeat the procedure.
- Check the eyes for distinct and sustained nystagmus at maximum deviation.
    - Move the stimulus to the subject's left side until the eye has moved as far to the side as possible. No white will be showing in the corner of the eye at maximum deviation.
    - Hold the eye at that position for a minimum of four seconds, then move the stimulus all the way across the subject's face to check the right eye, holding that position for a minimum of four seconds. Repeat the procedure.
- Check for onset of nystagmus prior to 45 degrees.
    - Start moving the stimulus towards the right at a speed that would take approximately four seconds for the stimulus to reach the edge of the subject's shoulder. Watch the eye carefully for any sign of jerking. When observed, stop and verify that the jerking continues.
    - Move the stimulus to the left at a speed that would take approximately four seconds for the stimulus to reach the edge of the subject's shoulder. Again, when you see jerking, stop and verify that the jerking continues.
    - Repeat the procedure. If the subject's eyes start jerking before 45 degrees, check to see that some white of the eye is still showing in the corner of the eye closest to the ear.
- Check for vertical nystagmus
    - Raise the stimulus upward until the subject's eyes are elevated as far as possible.
    - Hold for approximately four seconds and watch for evidence of jerking.

Documenting the Test

The test requires monitoring of three sets of clues: Lack of smooth pursuit for left and right eye, maximum deviation for left and right eye, and angle of onset for left and right eye.

Lack of Smooth Pursuit

Score whether the subject shows lack of smooth pursuit first in the left eye, then in the right eye.

Maximum Deviation
Score whether there is a distinct and sustained jerking of the eyeballs at maximum deviation for the left and right eye.

Angle of Onset
Estimate whether the angle of onset occurs prior to 45 degrees for both the left and right eye.

Criterion
Four or more clues suggest that the individual being tested is impaired with BAC ≥ .08.

Figure 2. Horizontal Gaze Nystagmus test.

- (Have the subjects remove their eyeglasses)
- (Position the stimulus about 12-15 inches from the subjects' nose)
- Hold your head still and follow the stimulus with your eyes only
- (Check for signs of medical impairment)
- (Check for lack of smooth pursuit)
- (Check for distinct nystagmus at maximum deviation)
- (Check for angle of onset)
- (Check for vertical nystagmus)

Horizontal Gaze Nystagmus			
	Yes	No	
Contact lenses? (If yes, hard or soft?)	□	□	
Equal tracking?	□	□	
Equal pupil size?	□	□	
	Left	Right	Total
Lack of smooth pursuit	□	□	
Distinct nystagmus at maximum deviation	□	□	
Onset prior to 45 degrees	□	□	
Total			

Finger to Nose

This test requires the subjects to bring the tip of the index finger to touch the tip of the nose. It is performed with eyes closed and head tilted slightly back.

This test should be administered in an environment where the subject is stable and is able to tilt their head back with eyes closed without risking personal injury.

Administrative Procedures

- Tell the subjects to make a fist with both hands, extend the index fingers, and turn the palms forward.
- Tell the subjects that when you say BEGIN, they should tilt their head back slightly and close their eyes.
- Demonstrate how head should be tilted back, but do not close your eyes.
- Inform the subjects that you will instruct them to bring the tip of the index finger to touch the tip of the nose.
- The arm is brought directly from the subjects' side in front of the body touching the tip of the nose with the tip of the index finger.
- Demonstrate how the subject is supposed to move the arm and how they are supposed to touch the tip of the nose with the tip of the finger.
- Tell the subjects that as soon as they touch their nose, they must return the arm to their side.
- Tell the subjects that when you say RIGHT they must move the right hand index finger to their nose; and when you say LEFT they must move the left hand finger to their nose.
- Get acknowledgement of understanding.
- Tell the subjects to BEGIN.
- Ensure they tilt their head back and close their eyes. Do not start to give the commands until the subjects are in compliance. If necessary, emphasize to the subjects that they must keep their eyes closed until you say to open them.
- Give the commands in exactly this order: LEFT, RIGHT, LEFT, RIGHT, RIGHT, LEFT.
- Make sure the subjects return the arms to their side immediately after each attempt. Pause about two or three seconds between commands.
- After the sixth attempt, tell the subjects to open their eyes and straighten their head.

Documenting the Test

The test requires monitoring two sets of clues: compliance with instructions and finger-to-nose accuracy.

Compliance with Instructions

1. Inability to follow instructions
2. Started at wrong time
3. Did not tilt head
4. Did not close eyes
5. Moved head during test
6. Opened eyes during test.

Finger-To-Nose Accuracy

1. Wrong hand
2. Wrong finger
3. Hesitated
4. Searched
5. Missed fingertip
6. Missed tip of nose
7. Did not bring hand down.

Criterion

Nine or more clues suggest that the individual being tested is impaired with BAC $\geq$ .08.

Figure 3. Finger-To-Nose test.

- Make a fist with both hands, extend the index fingers, and turn the palms forward
- When I say begin, tilt your head back slightly and close your eyes (demonstrate)
- On my command, touch the tip of your nose with the tip of your index finger and return it to your side (demonstrate arm and hand movement and show tip of finger and tip of nose)
- When I say right, you must move the right hand index finger to your nose; when I say left, you must move the left hand index finger to your nose
- Do you understand?
- Begin
- Left, Right, Left, Right, Right, Left
- Open your eyes and straighten your head

Finger to Nose Compliance	
Unable to follow instructions	☐
Started at wrong time	☐
Did not close eyes	☐
Did not tilt head	☐
Moved head during test	☐
Opened eyes during test	☐
Total	

Finger to Nose Accuracy							
	Left	*Right*	*Left*	*Right*	*Right*	*Left*	Total
Wrong hand	☐	☐	☐	☐	☐	☐	
Wrong finger	☐	☐	☐	☐	☐	☐	
Hesitated	☐	☐	☐	☐	☐	☐	
Searched	☐	☐	☐	☐	☐	☐	
Missed fingertip	☐	☐	☐	☐	☐	☐	
Missed tip of nose	☐	☐	☐	☐	☐	☐	
Did not bring hand down	☐	☐	☐	☐	☐	☐	
Total							

Palm Pat

The Palm Pat FST requires the subjects to place one hand extended, palm up, out in front of them. The other hand is placed on top of the first with the palm facing down. The top hand then begins to pat the bottom hand. The top hand rotates 180 degrees alternating between the back of the hand and the palm of the hand. The bottom hand remains stationary. The subject counts out loud, ONE-TWO, ONE-TWO, etc., in relation with each pat.

Administrative Procedures

- Start by instructing the subjects to put one hand out in front of them with the open palm facing upward. The opposite hand is then placed on top of the first hand with the open palm facing downward
- The hand with the palm facing upward is held in a stationary position. The hand on top with the palm facing downward will be the only hand moving
- When told to begin, the subjects will rotate the top hand 180 degrees and pat the back of the top hand to the palm of the bottom hand simultaneously counting out loud, “one”. The top hand then rotates 180 degrees so the palm of the top hand pats the palm of the bottom hand simultaneously counting out loud, “Two”
- Demonstrate
- The process then repeats. The subjects should start at a slow speed then, gradually increase the speed until a relatively rapid pace is reached.
- If necessary, prompt the subject to increase the speed
- The subject should perform this test for a minimum of 10 seconds but no more than 15 seconds.

Documenting the Test

The test requires monitoring of compliance with instructions.

Compliance with Instructions

1. Starting position
    a. Inability to follow instructions
    b. Started at wrong time
2. Palm pats
    a. Did not count as instructed
    b. Did not pat as instructed
        i. Rolled hands
        ii. Double pat
        iii. Karate chopped
        iv. Other (document)
        v. Did not increase speed
    c. Rotated hands
    d. Stopped before being told.

Criterion

Two or more clues suggest that the individual being tested is impaired with BAC ≥ .08.

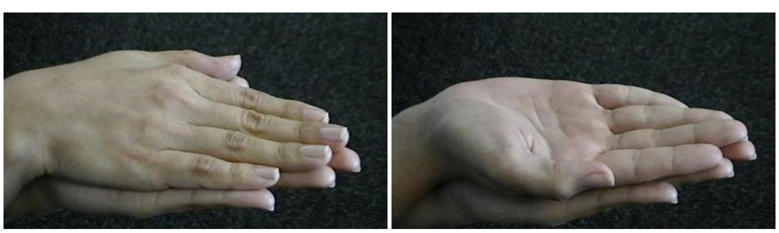

Figure 4. Palm Pat test.

- Place your hands palm to palm like this (demonstrate)
- Remain in this position while I explain the test
- When I tell you to begin, turn the top hand over and count 1, 2, 1, 2, like this (demonstrate)
- Speed up as you go, and do not stop until told
- Do you understand?
- Begin
- (If necessary) speed up…speed up…

Palm Pat	
Starting position	
Unable to follow instructions	□
Started at wrong time	□
Palm Pats	
Did not count as instructed	□
Rolled hands	□
Double pat	□
Chopped pat	□
Other improper pat (document)	□
Did not increase speed	□
Rotated hands	□
Stopped before being told	□
Total	

Hand Coordination

This test requires the subjects to perform a series of tasks with their hands. It is adapted from the Walk-And-Turn test performed on land.

Administrative Procedures

- Tell the subjects to make fists with both hands, place the left fist thumb against the sternum, and the thumb side of the right fist against the fleshy side of the left fist
- Demonstrate
- Tell the subjects to stay in that position
- Tell the subjects that when you say BEGIN they must perform four tasks. The first is to count aloud from one to four, placing one fist in front of the other, in step-like fashion, making sure the thumb side of one fist is touching the fleshy side of the other fist at each step
- Demonstrate
- The second task is to memorize the position of the fists after having counted to four, clap the hands three times (no aloud count required), and return the fists in the memorized position
- Demonstrate
- The third task is to move the fists in step-like fashion in reverse order counting aloud from five to eight, and return the left fist to the chest
- Demonstrate
- Finally, tell the subjects to return their hands, opened and palms down, to their laps
- Get acknowledgement of understanding
- Say BEGIN.

Documenting the Test

The test requires monitoring for compliance with instructions.

Compliance with Instructions

1. Starting position
    a. Inability to follow instructions
    b. Started at wrong time
2. Forward steps
    a. Improper count
    b. Improper touch
    c. Did not perform
3. Hand clapping
    a. Improper count
    b. Improper touch
    c. Improper return
    d. Did not perform
4. Return steps
    a. Improper count
    b. Improper touch
    c. Did not perform
5. End position
    a. Improper position
    b. Did not perform.

Criterion

Three or more clues suggest that the individual being tested is impaired with BAC $\geq .08$.

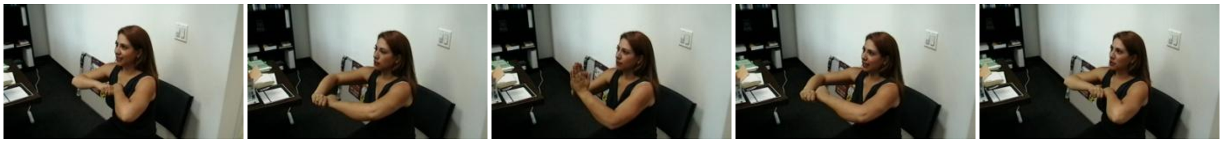

Figure 5. Hand coordination test. From left to right, top to bottom, fists at chest, fists out after 1-4 count, hand clapping, fists out after hands clapping, and fists back to chest (end position not shown).

- Make fists with both hands and place them like this (demonstrate)
- Remain in this position while I explain the test
- When I say begin, you must perform four tasks: the first is to count aloud from one to four placing one fist in front of the other, in step-like fashion making contact between the fists from one step to the other (demonstrate)
- The second task is to memorize the position of your fists after you have counted to four, clap your hands three times, and return the fists to the position you have memorized (demonstrate)
- The third task is to move the fists in step-like fashion in reverse order—counting aloud from five to eight—and return the left fist to the chest (demonstrate)
- Finally, I want you to return your hands, opened and palm-down to your lap
- Do you understand?
- BEGIN

Hand Coordination	
Starting position	
Unable to follow instructions	□
Started at wrong time	□
Forward steps	
Improper count	□
Improper touch	□
Did not perform	□
Hand clapping	
Improper count	□
Improper touch	□
Improper return	□
Did not perform	□
Return steps	
Improper count	□
Improper touch	□
Did not return fist to chest	□
Did not perform	□
End position	
Improper position	□
Did not perform	□
Total	