

# Do “Standardized” Field Sobriety Tests



# Reliably Predict Intoxication? *Knowledge for drunk driving litigation*

With the advent of the Breathalyzer and its offspring, the DataMaster, field sobriety tests (FSTs) took a back seat in drunk driving prosecutions. However, now that the new drunk driving laws no longer contain the presumptions against impairment at blood alcohol levels above or below .08,<sup>1</sup> the “observation testimony” of the arresting officer, including the suspect’s performance on FSTs, has suddenly become a much more important part of the prosecution’s case. Now more than ever, competent drunk driving defense requires counsel to be knowledgeable regarding both the proper procedures for administering the FSTs and the “science” behind them.

## The Science Behind Standardized FSTs

### The NHTSA Study

In the late 1970s the National Highway Transportation Safety Administration (NHTSA) commissioned a study by the Southern California Research Institute (SCRI) to determine which FSTs were best out of the dozens being used around the country.<sup>2</sup> SCRI narrowed the list to six tests it thought were the most feasible. It then recruited ten police officers to observe several hundred people who were given varying amounts of alcohol in a double-blind study (neither the officers nor the subjects knew how much alcohol they had been given). The officers’ only task was to determine whether each subject had a blood-alcohol level greater than .10 percent. That study resulted in NHTSA

recommending use of the three tests considered standard today: the horizontal gaze nystagmus (HGN), the walk and turn, and the one-leg stand.

Although that first NHTSA study recommended police departments begin to use only the three standard tests, NHTSA acknowledged the error rate of the officers in the study was 47 percent. In other words, the officers’ ability to detect which subjects had blood-alcohol contents greater than .10 percent was almost no better than flipping a coin.

Because of this rather large error rate, NHTSA commissioned SCRI to do another study in 1981.<sup>3</sup> In that study, NHTSA’s goal was to standardize the procedures for each test to see if the error rate could be lessened. According to NHTSA, standardizing the procedures improved the “success” rate in the laboratory. NHTSA claims that the laboratory test data found that the HGN by itself was accurate at detecting those whose blood-alcohol level was greater than .10 percent 77 percent of the time, the walk-and-turn was accurate 68 percent of the time, and the one-leg stand was accurate 65 percent of the time. When using all three of the tests, NHTSA claimed that the officers were correct 82 percent of the time.

SCRI then took the tests out of the laboratory and put them into the field. NHTSA concluded the field test data would not support the laboratory analysis, but decided the tests were good enough based on the following “favorable trends”: (1) after training on the test battery, officers tended to make more

DWI arrests; and (2) trained officers were more accurate in identifying suspects whose blood-alcohol contents are above .10.

In 1983<sup>4</sup> NHTSA hired SCRI to conduct a more in depth field study. This resulted in another report that concluded the HGN was 77 percent accurate, the walk-and-turn was 68 percent accurate, and the one-leg stand was 65 percent accurate. It is interesting to note these were the exact same percentages found in the laboratory study in 1981.

The SCRI findings have been subject to considerable criticism in the scientific community. For example, in an article discussing the history of the NHTSA studies, Alabama attorney Phillip Price quotes extensively from the sworn trial testimony of the lead author of the study from SCRI, Marcelline Burns. According to Mr. Price, Dr. Burns had to admit to a number of short-comings in all of the NHTSA studies. Among them was the fact that the error rates of 47 percent in the 1977 study and 32 percent in the 1981 study were unacceptably high as a scientific principle.<sup>5</sup>

In attempting to mitigate the high error rates of the 1977 and 1981 study, Dr. Burns pointed to the inexperience of the officers, but she used inexperienced officers again in her 1981 study. Worse yet, Dr. Burns also attributed the 1981 error rate (especially the fact that 18 percent of subjects with no alcohol were identified as above .10 percent) to the fact that the study was done “next to the drug capital of the world,” perhaps suggesting the test subjects might have been high even though they had no alcohol in their system.

Still, Dr. Burns did not see fit to either invalidate the entire study, or start over with drug screened subjects.

Further problems with both the NHTSA protocol and findings were discussed by Dr. Spurgeon Cole, a researcher at Clemson University. Dr. Cole makes much of one significant problem, which is the “dosing differential” of the subjects in the 1981 NHTSA study, i.e., the differences in the amount of alcohol given to each subject.<sup>6</sup> In the 1981 NHTSA study, two-thirds of the subjects were given either a very high amount of alcohol or a very low amount of alcohol (.15 percent versus .05 percent). Subjects with those amounts of alcohol should have been relatively easy to pick out as either really drunk or really sober. The officers were only asked to determine whether an individual was above or below .10 percent, so the error rates should have been much better because two-thirds of the subjects should have been obviously above or obviously below the threshold. Since the number of people in the easy-to-detect ranges (above .15 percent or below .05 percent) went up in the 1981 study compared to the 1977 study, the accuracy rate should have automatically gone up. Thus, even though the accuracy appeared to improve between the two studies, the increase in accuracy may be entirely explained by the dosing differential.

### The Clemson Study

Because of these concerns, Dr. Cole performed his own study and published his findings.<sup>7</sup> Dr. Cole’s findings are alarming. In Dr. Cole’s study, his methodology was to videotape 21 completely sober people performing six FSTs: walk-and-turn; the alphabet; one-leg stand; one-leg stand with head tilted backward, eyes closed, and finger touching nose; one-leg stand while counting; and one-leg stand with leg extended outward. When 14 police officers with a median experience level of 11.7 years were asked to view the tapes, they identified the subjects as too drunk to drive 46 percent of the time. All of the officers had completed a state-mandated DUI detection training program and all had field experience in DUI detection.

The subjects were also given four “normal ability” tests to perform: counting from 1 to 10; reciting one’s social security number, driv-

**Fast Facts:**  
**The three standard tests are the horizontal gaze nystagmus, the walk and turn, and the one-leg stand.**  
**At least one study has shown that officers in the field perform the HGN test incorrectly 95 percent of the time.**  
**Individuals over 65 years of age or with back, leg, or middle ear problems may have trouble performing the walk and turn and the one-leg stand.**

er’s license number, or date of birth; reciting one’s home address and phone number; and walking in a normal manner, turning around, and walking back to the starting point. When the officers viewed the videotapes of the normal ability tests, they incorrectly identified subjects as too drunk to drive 15 percent of the time.

Dr. Cole concludes from this study that FSTs that require subjects to perform unfamiliar and unpracticed motor sequences put the subjects at an unfair disadvantage. He advocates that the “science” justifying the use of FSTs in court is very misleading and disingenuous, and ought to be thoroughly reexamined.

### Challenging the Officer

#### The Horizontal Gaze Nystagmus

The HGN poses the most problems for defendants because a defendant cannot see his or her own eyes, and therefore has no idea whether he or she “passed” or not. A videotape will also not show the defendant’s eyes, so the jury cannot make an independent evaluation. Moreover, **the testifying officer must demonstrate his or her expertise and training in the area and explain the**

**basis of the tests, their administration, and their results in the particular case, in order for the results to be legally meaningful.** However, most officers are poorly trained in how to administer the HGN and at least one study has shown that officers in the field perform the test incorrectly 95 percent of the time.<sup>8</sup> It is impossible for a judge or jury to know this without expert testimony.

In Michigan, the court of appeals has ruled that the HGN is admissible without any expert testimony to determine the presence of alcohol (but not the amount) so long as the test was properly performed and the officer was properly qualified to administer it.<sup>9</sup> However, courts in other states are starting to look again at HGN and wonder whether it really does meet generally accepted scientific principles. Many states have found it does not, and have ruled it inadmissible before the fact finder.<sup>10</sup>

In administering the HGN, the officer should position a stimulus 12 to 15 inches away from the subject’s nose and slightly above eye level.<sup>11</sup> He or she should then make a total of seven passes with the stimulus. These are divided into four groups, with three of the four groups getting two passes for each eye. A complete pass is defined by NHTSA as moving the stimulus from the center all the way to the subject’s right, then all the way to the subject’s left, then back to center.<sup>12</sup>

During these sequences the officer first looks for equal tracking (making one complete pass). This is to confirm equal tracking of the eyes and equal pupil size. The officer moves the stimulus from the center to the person’s far left, to the person’s far right, and back to the center, taking at least two seconds. If equal tracking is not observed, then the test should be terminated because there is a possible medical disorder, injury, or blindness.<sup>13</sup> The officer next looks for lack of smooth pursuit (each eye is checked twice), or stated differently, whether the subject’s eyes track smoothly from side to side. The stimulus is moved from the center position to the far left and back to the center position for each eye, taking approximately two seconds from the center to the side, and two seconds from the side back to the center for each eye (a total of eight seconds for each complete pass or 16 seconds total).



Additionally, the officer is looking for distinct nystagmus at maximum deviation (two complete passes for each eye). This is designed to determine whether the person has distinct and sustained nystagmus at maximum deviation, which is as far as the eye can go to the side with little or no white showing. The stimulus is moved from the center position to the subject's far left, and held at maximum deviation for at least four seconds, and then moved back to the center, taking at least two seconds. Finally, the officer looks for the onset of nystagmus prior to 45 degrees (two complete passes). This is to determine whether the angle of onset of nystagmus occurs prior to the eye moving 45 degrees to the side. (It should be noted that it is difficult to determine this 45 degree angle, and there is little or no training in the practitioner course for this determination.) The stimulus must be moved slowly to detect this, at least four seconds from the center to the predetermined 45 degree angle. Once the onset of nystagmus is detected prior to 45 degrees, the officer must stop the movement of the stimulus to confirm that the

there are at least 35 other causes of nystagmus, including drowsiness.<sup>15</sup>

Finally, if there is a videotape of the stop, a challenge to the HGN is easier because counsel can then observe how the officer performed it, and determine if it was administered correctly. In making this determination, counsel should understand there are only three clues associated with the HGN test: lack of smooth pursuit, distinct nystagmus at maximum deviation, and onset of nystagmus prior to 45 degrees. These are the only clues for the HGN test and they have to be looked for in that order. The minimum amount of time for administration of the HGN test according to the NHTSA standardized protocol (checking each eye two times) excluding vertical gaze nystagmus (VGN) is 50 seconds (58 seconds including VGN). If onset of nystagmus prior to 45 degrees is detected, the minimum administration time may be higher. Each eye is checked two times in the HGN test. This is a total of two complete passes for each eye, for a total of six total passes plus one pass to check for equal tracking, for a total of seven passes, ex-

vided into two phases, the instruction phase and the walking phase. During the instruction phase, the officer should require the subject to place his or her left foot on the line and then place his or her right foot on the line in front of the left foot, while the officer reads the instructions and asks if the subject understood the instructions. The officer must then demonstrate the test, but only a few steps and the turn. The officer should then tell the person to take nine heel-to-toe steps, keeping eyes on feet and counting each step out loud, then turn using a series of small steps, and take nine heel-to-toe steps back. Finally, the officer should instruct "once you start walking do not stop until you have completed the task." There are eight "clues" associated with this test, and if the officer observes two of the eight, the subject will be considered to have failed the test.<sup>17</sup>

The most common mistake officers make with the walk and turn is not providing a designated straight line to walk. The 2002<sup>18</sup> manual specifically says, "walk-and-turn test requires a designated straight line." Moreover, the manual notes, "some people have difficulty with balance even when sober." The 2002 manual also states that "the original research indicated that individuals over 65 years of age, back, leg or middle ear problems had difficulty performing this test."<sup>19</sup> The 1995 manual also said that wind/weather conditions, suspect's age, weight, and suspect's footwear "may interfere with the suspect's performance."<sup>20</sup>

### The One-Leg Stand

The last of the SFSTs is the one-leg stand. In this exercise, the subject is instructed to raise one leg approximately six inches off the ground, keeping it straight with toes pointed, and count out loud for thirty seconds using "one thousand and one, one thousand and two" until told to stop. There are four clues associated with this test, and like the walk-and-turn, if two clues are observed, the subject will be considered to have failed. The four clues are: uses arms for balance, sways, puts foot down, and hops.<sup>21</sup>

The most common mistakes with the one-leg stand are not scoring the test properly (indicating other than the standardized clues)

*...with drunk driving laws becoming ever more punitive, and law enforcement more aggressive, both prosecutors and defense attorneys must have a thorough understanding of the science and law behind field testing.*

cluding the two passes for VGN. Including VGN, the HGN test will have a total of nine passes. Analyzing videotapes and/or testimony with this level of scrutiny will allow counsel to conclude the test was not administered according to the NHTSA protocol. This may lay a foundation for an evidentiary challenge and perhaps even lead to suppression at trial.

### The Walk and Turn

The instructions for the walk and turn are contained in the NHTSA manuals.<sup>16</sup> The 2002 manual indicates the test is di-

cluding the two passes for VGN. Including VGN, the HGN test will have a total of nine passes. Analyzing videotapes and/or testimony with this level of scrutiny will allow counsel to conclude the test was not administered according to the NHTSA protocol. This may lay a foundation for an evidentiary challenge and perhaps even lead to suppression at trial.

Officers can make a number of mistakes when performing the HGN. The most common mistakes are an incorrect number of passes, a failure to follow the timing protocol relative to each pass or set of passes, and a failure to properly estimate a true 45 degree angle. HGN is not generally accepted among psychologists, primarily because of the inherent difficulty in properly estimating the angle of onset.<sup>14</sup> Counsel should also be aware

and not providing a reasonably dry, hard, level, and non-slippery surface.<sup>22</sup> The 2002 manual also states: “the original research indicates that certain individuals over 65 years of age, back, leg or middle ear problems, or people who are overweight by 50 or more pounds had difficulty performing this test. Individuals wearing heels more than two inches high should be given the opportunity to remove their shoes.”<sup>23</sup>

## Conclusions

A review of the validation studies indicates NHTSA has failed to proffer a definition of either the term “standardized” or “normal.” NHTSA has also changed the protocols over time, and when the protocols have been changed, they were not re-validated, or otherwise subjected to rigorous studies (since 1977). This alone might be sufficient to call the underlying science into question, and significantly lessens their forensic value.

Nevertheless, with drunk driving laws becoming ever more punitive, and law enforcement more aggressive, both prosecutors and defense attorneys must have a thorough understanding of the science and law behind field testing. Effective advocacy by both sides demands such understanding, particularly as this area of law becomes increasingly specialized in nature. ♦



*Patrick T. Barone is a sole practitioner devoting his law practice exclusively to drunk driving defense. He is a general member of the National College for DUI defense, a member of the National Association for Criminal Defense Attorneys and the Criminal*

*Law Section of the State Bar of Michigan. His office is located in Birmingham, and he appears in courts throughout the state. He can be reached at (248) 594-4554, or at <http://www.mid3.net>.*



*Jeffery S. Crampton is a litigator with Koernke & Crampton, PC in Grand Rapids. He concentrates his practice on white-collar criminal defense and drunk driving defense. He is a member of the National Association of Criminal Defense*

*Lawyers, and the Criminal Defense Attorneys of Michigan organization. He is the past president of the James Madison College Alumni Association. He can be reached at (616) 458-7900, or at <http://www.mid3.net>.*

## Footnotes

1. MCLA 257.625 et seq., as amended effective September 30, 2003.
2. Burns & Maskowitz, *Psychophysical Tests for DWI Arrests* (1977) Final Report, DOT-HS-802-424, NHTSA.

3. Tharp, Burns & Maskowitz, *Development and Field Test of Psychophysical Tests for DWI Arrest* (1981) Final Report, DOT-HS-805-864, NHTSA.
4. Anderson, T. E., Schweitz, M. B., *Field Evaluation of a Behavioral Battery for DWI* (1983) Final Report, DOT-HS-806476.
5. *Field Sobriety Testing*, Phillip B. Price, Sr.
6. Nowacyk, R. H., Dr. and Cole, S., Dr., *Separating Myth from Fact: A Review of Research on the Field Sobriety Tests* (*The Champion*, August 1995).
7. Nowacyk, R. H., Dr. and Cole, S., Dr., *Field Sobriety Tests: Are They Designed for Failure? (Perceptual and Motor Skills)*, 1994, 79, 99-104).
8. J. L. Booker, “The Horizontal Gaze Nystagmus Test: Fraudulent Science in the American Courts,” *Science & Justice*, Vol 44, p 133-139 (2004).
9. *People v Berger*, 217 Mich App 213, 551 NW2d 421 (1996).
10. See *United States v Horn*, 185 F Supp 2d 530 (D Md 2002), and *State v Chastain*, 960 P2d 756 (Kan 1998) (HGN does not satisfy *Frye*); *State v Witte*, 836 P2d 1110 (Kan 1992) (HGN must satisfy *Frye* to be admitted); *Young v City of Brookhaven*, 693 So 2d 1355 (Miss 1997) (HGN does not satisfy *Frye* but may be used for probable cause); *State v Doriguzzi*, 760 A2d 336 (NJ Super Ct App Div 2000) (HGN must meet *Frye* to be admissible); *State v Torres*, P2d 20 (NM 1999) (state must first satisfy *Daubert* and show the validity and reliability of the test before HGN may be admitted); *Commonwealth v Apollo*, 603 A2d 1023 (Pa Super Ct 1992) (affirmed trial court’s exclusion of HGN for failure to meet *Frye*); *State v Murphy*, 953 SW2d 200 (Tenn 1997) (must satisfy *Daubert* to be admissible).
11. NHTSA *DWI Detection and Standardized Field Sobriety Testing Participant Manual* (2002) ch VIII, p 6.
12. NHTSA *DWI Detection and Standardized Field Sobriety Testing Instructor Manual* (2002) ch VIII, p 18.
13. NHTSA *DWI Detection and Standardized Field Sobriety Testing Instructor Manual* (2002) ch VIII, p 6.
14. *Horn*, supra.
15. *Schultz v State*, 106 Md App 145, 664 A2d 60 (1995).
16. NHTSA, *DWI Detection and Standardized Field Sobriety Testing Participant Manual* (2002) ch VIII, p 9.
17. NHTSA, *DWI Detection and Standardized Field Sobriety Testing Participant Manual* (2002) ch VIII, pp 10, 11.
18. NHTSA, *DWI Detection and Standardized Field Sobriety Testing Participant Manual* (2002) ch VIII, p 11.
19. *Id.*
20. NHTSA, *DWI Detection and Standardized Field Sobriety Testing Student Manual* (1995) ch VIII, p 28.
21. NHTSA, *DWI Detection and Standardized Field Sobriety Testing Participant Manual* (2002) ch VIII, p 13.
22. *Id.*
23. NHTSA, *DWI Detection and Standardized Field Sobriety Testing Participant Manual* (2002) ch VIII, p 14.