

EXHIBIT B



ATTORNEYS

Michael J. Nichols
Wendy M. Schiller-Nichols
Karen L. Phillips

Joshua M. Covert

3452 E. Lake Lansing Road
E. Lansing, MI 48823

Ithaca office:
P.O. Box 188
Ithaca, MI 48847-0188

(517) 432-9000
(800) 550-5892
(517) 203-4448 • facsimile
www.nicholslawyers.com

185 F.Supp.2d 530
United States District Court,
D. Maryland.

UNITED STATES,
v.
Eric D. HORN.

No. CRIM A. 00-946--PWG. | Jan. 31, 2002.

Defendant charged, under Assimilative Crimes Act, with driving while intoxicated (DWI) moved to exclude evidence of his performance on field sobriety tests. The District Court, Grimm, United States Magistrate Judge, held that: (1) test results were admissible on issue of whether there was probable cause for arrest, but not for purpose of proving specific blood alcohol content (BAC); (2) arresting officer could testify with respect to his observations of defendant's performance of tests, but could not suggest that tests were objective indicators of intoxication; (3) if government introduced evidence that defendant exhibited nystagmus, defendant could bring out fact that there were many causes of nystagmus other than alcohol ingestion; and (4) arresting officer could give lay opinion testimony that defendant was driving while intoxicated.

Ordered accordingly.

West Headnotes (10)

^[1] **Automobiles**
⊕Evidence of Sobriety Tests

Results of properly conducted standard field sobriety tests may be considered to determine whether probable cause exists to charge driver with driving while intoxicated (DWI) or driving under influence of alcohol (DUI).

8 Cases that cite this headnote

^[2] **Automobiles**
⊕Evidence of Sobriety Tests

Results of standard field sobriety tests, either individually or collectively, are not admissible for purpose of proving specific blood alcohol content (BAC) of a driver charged with driving while intoxicated (DWI) or driving under influence of alcohol (DUI). Fed.Rules Evid.Rule 702, 28 U.S.C.A.

11 Cases that cite this headnote

^[3] **Arrest**
⊕What Constitutes Such Cause in General

All that is required to establish probable cause to arrest suspect is reasonably trustworthy information that would support reasonable belief that suspect committed offense. U.S.C.A. Const.Amend. 4.

^[4] **Criminal Law**
⊕Physiological Facts
Criminal Law
⊕Cause and Effect

There is well-recognized, but by no means exclusive, causal connection between ingestion of alcohol and detectable presence of exaggerated horizontal gaze nystagmus in person's eyes, which may be judicially noticed by court, proved by expert testimony or otherwise established. Fed.Rules Evid.Rule 201, 28 U.S.C.A.

5 Cases that cite this headnote

^[5] **Automobiles**
⊕Evidence of Sobriety Tests

Criminal Law

⊕In General; Subjects of Opinion Evidence

Criminal Law

⊕Intoxication

Police officer trained and qualified to perform standard field sobriety tests may testify with respect to his or her observations of subject's performance of these tests, if properly administered, to include observation of nystagmus, and these observations are admissible as circumstantial evidence that defendant was driving while intoxicated (DWI) or driving under influence of alcohol (DUI); but officer may not use value-added descriptive language to characterize subject's performance of tests, such as saying that subject "failed test" or "exhibited" certain number of "standardized clues" during test. Fed.Rules Evid.Rule 702, 28 U.S.C.A.

10 Cases that cite this headnote

alcohol; officer may not, however, bolster such testimony by reference to any scientific, technical or specialized information learned from law enforcement or traffic safety instruction, and must confine his or her testimony to helpful firsthand observations of defendant. Fed.Rules Evid.Rule 701, 28 U.S.C.A.

9 Cases that cite this headnote

^[8] **Criminal Law**

⊕Motions in Limine

Rules of evidence, except those dealing with privileges, are inapplicable during pretrial hearings on admissibility of evidence. Fed.Rules Evid.Rules 104(a), 1101(d)(1), 28 U.S.C.A.

^[6] **Criminal Law**

⊕Judicial Notice

Criminal Law

⊕Cross-Examination

If government introduces evidence that driving while intoxicated (DWI) or driving under influence of alcohol (DUI) defendant exhibited nystagmus when officer performed horizontal gaze nystagmus test, defendant may bring out either during cross examination of prosecution witnesses or by asking court to take judicial notice of fact that there are many causes of nystagmus other than alcohol ingestion.

5 Cases that cite this headnote

^[9] **Criminal Law**

⊕Necessity and Sufficiency

Proffered expert testimony must be excluded if it is not product of reliable methods or principles that reliably have been applied to facts of particular case. Fed.Rules Evid.Rule 702, 28 U.S.C.A.

2 Cases that cite this headnote

^[10] **Criminal Law**

⊕Judicial Notice

Doctrine of judicial notice is predicated upon assumption that source materials from which court takes judicial notice are reliable.

^[7] **Criminal Law**

⊕Intoxication

If otherwise admissible, police officer may give lay opinion testimony that defendant was driving while intoxicated or under influence of

Attorneys and Law Firms

*532 Sasha Natapoff, Asst.Fed.Public, Defender,
Baltimore, MD, for Eric D. Horn.

Opinion

MEMORANDUM AND ORDER

GRIMM, United States Magistrate Judge.

At approximately 10:35 p.m. on June 28, 2000, Sergeant Eric D. Horn attempted to enter the Harford Road gate of the Army facility located at Aberdeen Proving Ground, Maryland. Officer Daniel L. Jarrell stopped Horn's vehicle for an identification check. As a result of his observations of Horn, Jarrell suspected that Horn was driving under the influence of alcohol, and he was detained and questioned. Three standard field sobriety tests ("SFSTs") were administered: the "walk and turn" test, the "one leg stand" test and the "horizontal gaze nystagmus" test.¹ As a result of his performance on these tests, Horn was charged with driving while intoxicated under Md.Code Ann., Transp. II § 21-902 (1999 Repl. Vol.),² as assimilated by 18 U.S.C. §§ 7, 13, the Assimilative Crimes Act, a Class A misdemeanor.

[1] [2] [3] [4] [5] [6] [7] Horn has filed a motion *in limine* to exclude the evidence of his performance on the field sobriety tests, asserting that it is inadmissible under newly revised Fed.R.Evid. 702 and the *Daubert/Kumho Tire* decisions.³ The Government has filed an opposition, and Horn has filed a reply. In addition, a two day evidentiary hearing was held, pursuant to Fed.R.Evid. 104(a), on November 19 and 20, 2001, and additional testimonial and documentary evidence was received, which is discussed in detail below. At the conclusion of this hearing, the following ruling was made from the bench, the Court also announcing its intention subsequently to issue a written opinion on this case of first impression:⁴

(1) The results of properly conducted SFSTs may be considered to determine *533 whether probable cause exists to charge a driver with driving while intoxicated ("DWI") or under the influence of alcohol ("DUI");⁵

(2) The results of the SFSTs, either individually or collectively, are not admissible for the purpose of proving the specific blood alcohol content ("BAC") of a driver

charged with DWI/DUI;⁶

(3) There is a well-recognized, but by no means exclusive, causal connection between the ingestion of alcohol and the detectable presence of exaggerated horizontal gaze nystagmus in a person's eyes,⁷ which may be judicially noticed by the Court pursuant to Fed.R.Evid. 201, proved by expert testimony or otherwise;

(4) A police officer trained and qualified to perform SFSTs may testify with respect to his or her observations of a subject's performance of these tests, if properly administered, to include the observation of nystagmus, and these observations are admissible as circumstantial evidence that the defendant was driving while intoxicated or under the influence. In so doing, however, the officer may not use value-added descriptive language to characterize the subject's performance of the SFSTs, such as saying that the subject "failed the test" or "exhibited" a certain number of "standardized clues" during the test;

(5) If the Government introduces evidence that a defendant exhibited nystagmus when the officer performed the horizontal gaze nystagmus test, the defendant may bring out either during cross examination of the prosecution witnesses or by asking the Court to take judicial notice of the fact that there are many causes of nystagmus other than alcohol ingestion; and

(6) If otherwise admissible under Fed.R.Evid. 701, a police officer may give lay opinion testimony that a defendant was driving while intoxicated or under the influence of alcohol. In doing so, however, the officer may not bolster the lay opinion testimony by reference to any scientific, *534 technical or specialized information learned from law enforcement or traffic safety instruction, but must confine his or her testimony to helpful firsthand observations of the defendant.

The issues addressed in this case likely will recur, given the large number of Class A and B misdemeanors prosecuted in this district under the Assimilative Crimes Act. Moreover, the admissibility of SFSTs implicates recent changes to the federal rules of evidence, as well as a large body of state cases on this topic, primarily decided under a different evidentiary standard than that governing the admissibility of the results of SFSTs in federal court.⁸ Accordingly, this opinion will discuss the basis for the above rulings in more detail below.

1. Applicable Rules of Evidence

^[8] Fed. R. of Evid. 104(a) requires the Court to make preliminary determinations regarding the admissibility of evidence, the qualifications of witnesses and the existence of privileges, and Rule 104(a) now permits the Court to make definitive pretrial evidentiary rulings *in limine*. During Rule 104(a) hearings the rules of evidence, except those dealing with privileges, are inapplicable, permitting the Court greater latitude to consider affidavits such as those filed by Horn and the Government. Fed. Rules of Evid. 104(a), 1101(d)(1).

Whether the results of SFSTs are admissible depends first on the purpose for which they are offered. Fed. Rule of Evid. 105. Second, the SFSTs must be relevant and not excessively prejudicial for the purposes offered. Fed. Rules of Evid. 401, 403. Third, if the SFSTs are introduced by the testimony of a sponsoring witness who is testifying as to scientific, technical or specialized matters, the admissibility of the SFSTs is dependent on whether the witness's testimony meets the requirements of newly revised Fed. Rule of Evid. 702 and the *Daubert/Kumho Tire* standards. Finally, Fed. Rule of Evid. 102 emphasizes that interpretations of the rules of evidence should be made with an eye towards promptly, fairly, efficiently and inexpensively adjudicating cases.

In this case, the results of SFSTs potentially could be offered for the following purposes: (1) to establish probable cause to arrest and charge a defendant with DWI/DUI, (2) as direct evidence of the specific BAC of a defendant who performed the SFSTs or (3) as circumstantial proof that a defendant was driving while intoxicated or under the influence of alcohol. Horn has acknowledged that the tests may be used to determine probable cause, as the overwhelming majority of cases have held,⁹ and the Government acknowledges that they are not admissible to prove the defendant's specific BAC, a conclusion almost universally reached by state courts, including Maryland.¹⁰ Accordingly, the task at hand is to determine to what extent the results of SFSTs are admissible as circumstantial proof that a driver has consumed alcohol and was driving while intoxicated or under its influence. Because the results of the SFSTs invariably are introduced by the testimony of an arresting *535 police officer, and, as will be seen, may involve application of scientific, technical or other specialized information, the requirements of Rule 702, as recently revised, are of paramount importance.

Rule 702 permits testimony in the form of an opinion or

otherwise regarding scientific, technical or specialized matters from a qualified expert, provided the testimony is based on (a) sufficient facts or data, (b) is the result of methods or principles that are reliable and (c) is the result of reliable application of the methods or principles to the facts of the particular case. These three requirements, added in December 2000, are complimentary to, but not identical with, the four non-exclusive evaluative factors identified by the Supreme Court in the *Daubert/Kumho Tire* cases: (a) whether the opinions offered are testable; (b) whether the methods or principles used to reach the opinions have been subject to peer review evaluation; (c) whether a known error rate can be identified with respect to the methods or principles underlying the opinion, and, finally, (d) whether the opinion rests on methodology that is generally accepted within the relevant scientific or technical community.¹¹

As further will be seen, almost the entire universe of published case law regarding the admissibility of SFST evidence comes from the state courts, as would be expected, given the fact that there is no uniform federal traffic code, and DWI/DUI cases in federal court usually come about as a result of assimilating state drunk driving laws under 18 U.S.C. §§ 7 and 13. This is significant because the vast majority of the state cases that have analyzed this issue have done so under the *Frye*¹² standard for admitting scientific or technical evidence: whether the methods or principles have gained general acceptance within the relevant scientific or technical community.¹³ While this test has continued vitality as one of the four *Daubert/Kumho Tire* factors, a federal court must do more in determining the admissibility of scientific, technical or specialized evidence than focus on general acceptance.

The starting point for this analysis is the SFSTs themselves, followed by a discussion of the evidence produced by the parties in this case regarding their reliability and then a consideration of the state cases that have focused on this issue.

2. The SFSTs

The three SFSTs that are the subject of this case were developed on behalf of the National Highway Traffic Safety Administration ("NHTSA") beginning in the 1970's. They are discussed in detail by a series of NHTSA publications, including:

* a student manual for DWI detection and standardized field sobriety testing;

* a June 1977 final report prepared for NHTSA by Marcelline Burns, Ph.D.¹⁴ *536 and Herbert Moskowitz, Ph.D. of the Southern California Research Institute ("SCRI") titled "Psychophysical Tests for DWI Arrests" (the "1977 Report");

* a March 1981 final report prepared for NHTSA by Dr. Burns and the SCRI titled "Development and Field Test of Psychophysical Tests for DWI Arrest" (the "1981 Final Report");

* a September 1983 NHTSA Technical Report, authored by Theodore E. Anderson, Robert M. Schweitz and Monroe B. Snyder, titled "Field Evaluation Of A Behavioral Test Battery For DWI" (the "1983 Field Evaluation");

* a November 1995 study of the SFSTs funded by NHTSA and conducted by Dr. Burns and the Pitkin County Sheriff's Office, Colorado, titled "A Colorado Validation Study of the Standardized Field Sobriety Test (SFST) Battery" (the "1995 Colorado Validation Study"); and

* an undated study, authored by Dr. Burns and a sergeant of the Pinellas County Sheriff's Office, Florida, titled "A Florida Validation Study of the Standardized Field Sobriety Test (S.F.S.T.) Battery (the "Florida Validation Study")".

(Gov't. Opposition Memo. Exhs. 2-7).

These studies are very significant, as they have been cited repeatedly by the state courts in their opinions regarding the admissibility of SFSTs in connection with assessment of the reliability of the SFSTs and their general acceptance within the law enforcement and traffic safety communities. They also are important in this case because they have been the subject of critical analysis by Horn's experts, who provided detailed testimony regarding the limitations of these studies and the extent to which the SFSTs are reliable and valid tests for driver intoxication or alcohol impairment.¹⁵

The three SFSTs developed by the research sponsored by NHTSA are summarized in the NHTSA student manual. (Gov't. Opposition Memo., Ex.2). The manual describes the tests and evaluations conducted to develop the SFSTs, then provides detailed instruction on how to administer and score each of the three tests.

*537 The most "scientific" or "technical" of the three is the Horizontal Gaze Nystagmus Test ("HGN Test"). Nystagmus is "the involuntary jerking of the eyes, occurring as the eyes gaze toward the side. Also, nystagmus is a natural, normal phenomenon. Alcohol and certain other drugs do not cause this phenomenon, they merely exaggerate it or magnify it." *Id.* at VIII-12. Horizontal gaze nystagmus "occurs as the eyes move to the side." *Id.* at VIII-13. The HGN SFST requires the investigating officer to look for three "clues": (1) the inability of the suspect to follow a slowly moving stimulus smoothly with his or her eyes, (2) the presence of "distinct" nystagmus when the suspect has moved his or her eyes as far to the left or right as possible (referred to as holding the eyes at "maximum deviation") and held them in this position for approximately four seconds and (3) the presence of nystagmus before the eyes have moved 45 degrees to the left or right (which, the manual states, usually means that the subject has a BAC above 0.10). *Id.* at VIII-14-15. The officer is trained to look for each of the above three "clues" for each of the suspect's eyes, meaning there are six possible "clues." If the officer observes four or more clues the manual asserts that "it is likely that the suspect's BAC is above 0.10[and][u]sing this criterion [one] will be able to classify correctly about 77% of [one's] suspects with respect to whether they are above 0.10." *Id.* at VIII-17. If the results of the HGN test are offered to establish that the suspect's BAC is above 0.10,¹⁶ it is readily apparent that much depends on the investigating officer properly performing the HGN test procedures and on his or her subjective evaluation of the presence of the "standardized clues." Indeed, the manual itself cautions with respect to each of the SFSTs:

[the tests are valid] *only* when ... administered in the prescribed, standardized manner; and *only* when the standardized clues are used to assess the suspect's performance; and, *only* when the standardized criteria are employed to interpret that performance. If any one of the standardized field sobriety test elements is changed, the validity is compromised.

Id. at VIII-12 (emphasis in original).

The Walk and Turn ("WAT") test requires the suspect to place his feet in the heel-to-toe stance on a straight line. The subject then is instructed to place his right foot on the line ahead of the left foot, with the heel of the right foot against the toe of the left. The suspect also is told to keep his arms down at his side and to maintain this position until the officer instructs him to begin the test. *Id.* at VIII-18. Once told to start, the suspect is to take nine

heel-to-toe steps down the line, then to turn around in a prescribed manner, and take nine heel-to-toe steps back up the line. *Id.* While walking, the suspect is to keep his hands at his side, watch his feet, and count his steps out loud. *Id.* at VIII-19. Also, the suspect is told not to stop the test until completed, once told to start. *Id.*

As with the HGN test, the Manual asserts that there are standardized clues, *538 eight in all,¹⁷ that “[r]esearch ... has demonstrated ... are the most likely to be observed in someone with a BAC above 0.10.” *Id.* at VIII-19. Further, it states “[i]f the suspect exhibits two or more distinct clues on this test or fails to complete it, classify the suspect’s BAC as above 0.10. Using this criterion, you will be able to correctly classify about 68% of your suspects.” *Id.* at VIII-21. Once again, it is the officer’s subjective evaluation of the suspect that results in the determination of whether a “clue” is present or not, and, if only two of the eight “standardized clues” are detected, NHTSA asserts that the suspect’s BAC is 0.10 or more.

The third SFST is the One Leg Stand (“OLS”) test. In this test the suspect is told to stand with her feet together, arms at her sides. She then is told not to start the test until told to do so. To perform the OLS test, the suspect must raise whichever leg she chooses, approximately six inches from the ground, toes pointed out. *Id.* at VIII-23. While holding this position, the suspect then must count out loud for thirty seconds, by saying “one-one thousand, two-one thousand,” etc. *Id.* The NHTSA manual identifies four “standardized clues” for the OLS test¹⁸ and instructs law enforcement officers that “[i]f an individual shows two or more clues or fails to complete the [test] ... there is a good chance the BAC is above 0.10. Using that criterion, [one] will correctly classify about 65% of the people [one] test[s] as to whether their BACs are above or below 0.10.” *Id.* at VIII-24.

The NHTSA Manual advises that when the WAT and HGN tests are combined, using a decision matrix developed for NHTSA, an officer can “achieve 80% accuracy” in differentiating suspects with BACs in excess of 0.10. *Id.* at VIII-5. These conclusions are supported, it is claimed, by the results of research and testing done by Dr. Burns and her company that was reported in the 1981 Final Report, the 1983 Field Evaluation, the 1995 Colorado Validation Study and the Florida Validation Study.¹⁹ *Id.* at Exs. 4-8.

As next will be seen, Horn’s experts have challenged the reliability, validity and relevance of the SFSTs to prove driver intoxication and are sharply critical of the claims of

accuracy advanced in the NHTSA publications and the so-called validation studies. They have framed these objections in terms of the factors discussed in the *Daubert/Kumho Tire* decisions, as amplified by this Court in *Samuel v. Ford Motor Co.*, 96 F.Supp.2d 491 (D.Md.2000).

3. Horn’s Challenges to the Reliability/Validity of SFST Evidence

¹⁹ Rule 702 prohibits expert testimony if it is not the product of *reliable* methods or principles that *reliably* have been applied to the facts of the particular case. In the context of scientific or technical *539 testing, such as may be the case with SFSTs, reliability means the ability of a test to be duplicated, producing the same or substantially same results when successively performed under the same conditions. *Daubert*, 509 U.S. at 595, 113 S.Ct. 2786; *Samuel*, 96 F.Supp.2d at 494. Thus, for the SFSTs, if reliable, it would be expected that different officers, viewing the same suspect performing the SFSTs, would reach the same conclusion regarding the level of the suspect’s impairment or intoxication. Alternatively, the same officer re-testing the same suspect with the same BAC as when first tested would reach the same conclusion.

A related, though distinct concept, deals with the *validity* of a test. A test is valid if it has a logical nexus with the issue to be determined in a case. *Daubert*, 509 U.S. at 591, 113 S.Ct. 2786; *Samuel*, 96 F.Supp.2d at 494. In the context of SFSTs, they are valid if there is a logical nexus between what the tests measure and the true ability of a driver safely to operate a motor vehicle. Thus, for example, does the fact that a suspect missed two “cues” in the WAT test mean that the driver cannot safely drive a car, or does it simply mean that the driver has some inability to perform the test that is unrelated to his or her ability to drive? Horn has challenged both the reliability and validity of the SFSTs.

During the Rule 104(a) proceedings, Horn produced four experts, three of whom submitted affidavits, and two of whom also testified: Yale Caplan, Ph.D. (former chief toxicologist for the State of Maryland and former scientific director of the Maryland Alcohol Testing Program); Spurgeon Cole, Ph.D. (Professor of Psychology, Clemson University and author of a series of articles critical of the SFSTs); Harold P. Brull (a licensed

psychologist and consultant specializing in industrial/organizational psychology, particularly the definition and measurement of human attributes in employment and related settings); and Joel Wiesen, Ph.D. (an industrial psychologist with special expertise in experimental psychology, psychometrics and statistics. Dr. Wiesen worked for more than ten years for the Massachusetts Division of Personnel Administration, developing and validating civil service examinations and is an independent consultant in the field of development and validation of human performance tests).

In his testimony and published writings, Dr. Cole was highly critical of the reliability of the SFSTs if used to prove the precise level of a suspect's alcohol intoxication or impairment. His 1994 article "Field Sobriety Tests: Are They Designed for Failure?," published in the journal *Perceptual and Motor Skills*, analyzed the 1977 Report, the 1981 Final Report, and the 1983 Field Evaluation report published by NHTSA regarding the SFSTs. (Def's.Memo, Ex. C.).

Dr. Cole observed the following:

- (1) 47% of the subjects tested in the 1977 NHTSA laboratory study who would have been arrested by the testing officers for driving while intoxicated (BAC of 0.10 or greater) actually had BACs below 0.10;
- (2) in the 1981 Final Report, 32% of the participants in the lab study were incorrectly judged by the testing officers as having BACs of 0.10 or greater; and
- (3) the accepted reliability coefficient for standardized clinical tests is .85 or higher, yet the reliability coefficients for the SFSTs, as reported in the NHTSA studies, ranged from .61 to .72 for the individual tests and .77 for individuals that were tested on two different occasions while dosed to the exact same BAC. More alarmingly, inter-rater reliability *540 rates (where different officers score each subject) ranged from .34 to .60, with an over-all rate of .57.

Id. at 100.

Dr. Cole theorized that the SFSTs, particularly the WAT and OLS tests, required subjects to perform unfamiliar, unpracticed motions and noted that a very few miscues result in a conclusion that the subject failed and had a BAC in excess of 0.10. *Id.* His hypothesis was that individuals could be classified as intoxicated/impaired as

a result of unfamiliarity with the test, rather than actual BAC. *Id.* He tested this hypothesis by videotaping twenty-one completely sober individuals performing either "normal-abilities tests" (such as reciting their addresses or phone numbers or walking in a normal manner) or the WAT and OLS tests. *Id.* at 99-102. The results of the study were that 46% of the officers that viewed the videotape of the sober individuals performing the SFSTs rated the subjects as having had too much to drink, as compared to only 15% reaching this decision after seeing the videotape of the subjects performing the normal-abilities tests. *Id.* at 102. Dr. Cole concluded:

[The SFSTs] must be held to the same standards the scientific community would expect of any reliable and valid test of behavior. This study brings the validity of field sobriety tests into question. If law enforcement officials and the courts wish to continue to use field sobriety tests as evidence of driving impairment, then further study needs to be conducted addressing the direct relationship of performance on these and other tests with driving. To date, research has concentrated on the relationship between test performance and BAC and officers' perception of impairment. This study indicates that these perceptions may be faulty.

Id. at 103.

During his testimony at the Rule 104(a) hearing, Dr. Cole repeated his criticism of the reliability of the 1977, 1981 and 1983 studies but also testified about the Colorado, Florida and San Diego studies performed by Dr. Burns, styled as "field validation studies." This testimony echoed Dr. Cole's written criticisms about the SFSTs' reliability as precise predictors of the level of alcohol intoxication and the SFST's validity as a measure of driver impairment in his 1994 article, co-authored with Ronald H. Nowaczyk, titled "Separating Myth from Fact: A Review of Research on the Field Sobriety Tests" and published in the *Champion* journal of the South Carolina Bar Association. Def's. Reply Memo, Exh. 1.

Dr. Cole's primary criticisms, as discussed in his 1994 article, include, first, that the 1981 Final Report published

by NHTSA claims an 80% accuracy rate for users of the SFSTs. This is misleading because when the actual data is examined with respect to the success rate of using the SFSTs to differentiate between drivers with BACs above 0.10 and those without, the critical population, the officers had "a 50/50 chance of being correct just on the basis of guessing." *Id.* at 539.

Second, the SFSTs have a combined test-retest reliability rates of .77, while the scientific community "expects reliability coefficients to be in the upper .80s or .90 for a test to be scientifically reliable." *Id.* at 540. When different officers tested the same subjects at the same BAC dose level on different days the reliability was only .59—a 41% error rate. Dr. Cole contrasted these substandard reliability coefficients with that of the BAC machine, which is .96 or 96% reliable. *Id.* at 540–41.

Third, Dr. Cole argued that in order for the SFSTs to be valid predictors of BAC *541 they must "not only identify individuals above a BAC level of 0.10 as 'failing', but also identify individuals below .10 as 'passing'." *Id.* at 541. The data from the NHTSA 1977 Report, however, shows that the validity of the HGN, OLS and WAT SFSTs was ".67, .48, and .55, respectively, with a combined validity coefficient of .67." *Id.* This means that use of the SFSTs results in an unacceptably high erroneous arrest rate, if the tests are used by the officer to make arrest decisions based on BAC levels being in excess of .10.

Fourth, Dr. Cole was particularly critical of claims that the NHTSA SFSTs have been "validated" in a "field setting." In this regard, he stated that the 1977 and 1981 NHTSA studies were done in a laboratory setting, and the difference in conditions in a controlled lab are dramatically dissimilar from field conditions that can be expected when officers employ SFSTs at all times of day and night in widely disparate weather and traffic conditions and where issues of officer safety may influence how the test is performed.²⁰ *Id.* at 542. Dr. Cole stated that the NHTSA 1983 Field Evaluation purported to be a field validation study, but it failed to meet the recommendations of the authors of the NHTSA 1981 Final Report that the SFSTs be validated in the field for eighteen months in locations across the country. *Id.* Dr. Cole also stated that Dr. Burns herself has testified that the SFSTs have not been adequately field tested.²¹ *Id.*

Finally, Dr. Cole disputed the claims of proponents of the SFSTs that the studies regarding them have been published in peer review journals. The 1977 and 1981

field studies were published in technical reports by NHTSA, but those reports excluded the "methods and results" sections because they were thought to be too lengthy. *Id.* at 543. Cole concluded "[i]t is difficult to see how the NHTSA could claim that the FST is accepted in the scientific community, when results of studies on the validation of the FST have never appeared in a scientific peer reviewed journal, which is a basic requirement for acceptance by the scientific community." *Id.* Cole concluded:

Because of its widespread use, the FST battery has been assumed to be a reliable and valid predictor of driving impairment. NHTSA has done little to dispel that assumption. Law enforcement cannot be blamed for its use of the FST battery. Training documents refer to NHTSA reports and provide what appears to be supporting evidence for the validity of the FST battery. In addition, there is little doubt that individuals who have high BAC levels will *542 have difficulty in performing the FST battery. However, what the law enforcement community and the courts fail to realize is that the FST battery may mislead the officer on the road to incorrectly judge individuals who are not impaired. The FST battery to be valid must discriminate accurately between the impaired and non-impaired driver. NHTSA's own research on that issue ... has not been subjected to peer review by the scientific community. In addition, a careful reading of the reports themselves provides support for the inadequacy of the FST battery. The reports include low reliability estimates for the tests, false arrest rates between 32 and 46.5 percent, and a field test of the FST that was flawed because the officers in many cases had breathalyzer results at the time of the arrest. NHTSA clearly ignored the printed recommendations of its

own researchers in conducting that field study.

Id. at 546. (Emphasis in original).

Horn also introduced the affidavit of Joel P. Wiesen, Ph.D. Dr. Wiesen is an industrial psychologist with special expertise in experimental psychology, psychometrics and statistics. His experience includes more than ten years working with the Commonwealth of Massachusetts developing civil service examinations and an equal number of years as an independent consultant in the area of test development and validation. In addition, he is a published author of a mechanical aptitude test used nationwide. Although he is most familiar with written tests, he does have experience in the development of human performance tests. Def's. Reply Memo, Exh.6 at 1.

Dr. Wiesen reviewed the NHTSA 1977 Report, the 1981 Final Report, the 1983 Field Evaluation, the 1995 Colorado Validation Study, the undated Florida Validation Study, and the NHTSA student manual for the SFSTs. He was highly critical of these studies, as the following summary illustrates:²²

²²*543 Dr. Wiesen concluded his evaluation of the SFST reports with the following observation:

the studies give only a general indication of the level of potential validity of the tests as described in the NHTSA manual.... Rather than the five studies supporting each other, they evaluate somewhat different combinations of test content and test scoring. The differences are large enough to change the validity and accuracy of the tests. The older studies are probably less germane, due to the changes in test content and scoring over time. The reports for the newer studies are grossly inadequate. Given this, and in light of the specific critiques above (which are not exhaustive), I can only conclude that the field sobriety tests do not meet reasonable professional and

scientific standards.

Id. at 12-13.

Harold P. Brull testified on behalf of Horn and supplied an affidavit as well. Mr. Brull is a licensed psychologist with many years experience consulting in connection with the design and implementation of procedures to measure human attributes, especially in employment settings. He has designed and evaluated tests and procedures measuring human *544 characteristics for over twenty years. Def's. Reply Memo, Exh. 5 at 2.

Mr. Brull reviewed the NHTSA 1977 Report, the 1981 Final Report, the 1983 Field Evaluation, the 1995 Colorado Validation Study, the Florida Validation Study, and the NHTSA officer training manual. Among his general observations of these materials was the opinion that there was a complete absence of evidence "which would allow one to predict a known error rate in the field," where there is no ability to control the performance of the SFSTs like there is in a laboratory setting. Def's. Reply Memo, Exh. 4 at 6. He was especially critical of the assertions in the Florida and Colorado studies regarding the reliability of the SFSTs, primarily because of their use of lower BAC thresholds (0.05 and above instead of 0.10), the fact that the population of drivers evaluated were those stopped because of unsafe driving and the complete absence of any data in the reports to enable meaningful evaluation. *Id.* at 6-7. He further expressed the opinion that none of the reports was published in peer review literature. While Brull was not critical of the methodology used in the 1977 and 1981 laboratory studies, he stated that the results from these studies were inconclusive, and the subsequent field tests "simply do not contain sufficient detail or rigor to support any hypothesis that field sobriety studies, as conducted by police officers in the field, are valid and reliable." *Id.* at 7.

Brull's evaluation of the data contained in the 1977 and 1981 reports was consistent with that of Dr. Cole and Dr. Wiesen. Regarding the 1981 Final Report, he observed that "the degree of predictive error in the field appeared to be substantially larger than in the laboratory," and that "[w]hile training clearly brought about improvement, it does not compare favorably to the laboratory condition and is [sic] a margin of error substantially higher than one would find acceptable for predicting with any degree of certainty." *Id.* at 11.

Brull was most critical of the Colorado and Florida

“validation” studies. He noted that they “are merely summary reports, without foundation, of findings,” and suffered from a “serious methodological flaw,” in that the tests were done on actual motorists stopped by officers because their driving was unsafe, leading the officers automatically to suspect that they were intoxicated. *Id.* Use of this population likely will produce results that Brull characterized as “highly inflated.” *Id.* He further noted that these field studies predicted 90% accuracy in identifying drivers with BAC’s above 0.05, a level only one half that used in the earlier tests and below the level of legal intoxication. While the validation studies provided no data to assess the accuracy of the SFSTs in identifying drivers with BACs of 0.10 or higher, Brull suspected that the accuracy rate would be far lower than 90%. *Id.* at 12.

Brull’s final conclusions were summarized as follows:

- (1) the laboratory studies that form the foundation of the SFSTs (the 1977 and 1981 studies) were well designed;
- (2) the accuracy of the SFSTs, even under laboratory conditions, is less than desired and below the level expected for tests of human performance;
- (3) the field studies were not well documented, produced unknown error rates, but which, if known, likely would have been unacceptable in real world situations;²³

*545 (4) the error rate of SFSTs as actually performed by officers in the field is unknown;

- (5) the only peer review article analyzing the SFST’s was written by Dr. Cole and is highly critical of the accuracy of the SFSTs.

Id. at 14.

Finally, Horn offered the affidavit of Yale H. Caplan, Ph.D., Defs.’ Motion, Ex. E. Dr. Caplan has more than thirty years experience in the field of forensic toxicology and alcohol and drug testing. He served for many years as the chief toxicologist for the Maryland Medical Examiner’s office and now is a consultant in the field of toxicology. *Id.* Dr. Caplan stated that a determination that a person is impaired by alcohol consumption may be made in one of two fashions: by direct evidence of impairment derived from the chemical analysis of a breath or blood specimen; or indirectly by assessing performance indicators of the subject through field sobriety tests. *Id.*

With respect to the latter, Dr. Caplan stated:

Although physiological assessments (e.g. standardized field sobriety tests) when coupled with the odor of alcohol on breath and alcohol’s relatively high epidemiological prevalence in drivers may suggest alcohol as the causative agent, the use of drugs or the concomitant use of alcohol and drugs or other medical conditions must be considered as causes for the impairment. In fact, field sobriety tests alone were never designed for or demonstrated to be unequivocally capable of indicating alcohol impairment.

Id. He expressed the following opinions: (1) that field sobriety tests can be used to define impairment but that a specific blood/breath alcohol test is needed to confirm that the cause of the impairment is alcohol ingestion; (2) that an alcohol test of a suspect’s breath or blood can alone be used to establish impairment, but field sobriety tests alone cannot establish alcohol impairment “with absolute certainty.” *Id.*

4. The Government’s Evidence

In response to the evidence submitted by Horn, the Government introduced the affidavit of Officer Jarrell, the arresting officer, describing the stop, detention and arrest of Horn and the SFSTs administered to him. The Government also introduced the 1977, 1981, and 1983 NHTSA reports, the California and Florida “validation studies,” the NHTSA student manual regarding the SFSTs, and an article titled “Horizontal Gaze Nystagmus: The Science & the Law,” published by the American Prosecutors Research Institute’s National Traffic Law Center (“NITLC”).²⁴ Govt’s. Opposition Memo, Exhs. 1–7.

*546 Additionally, the Government introduced the affidavit of Lieutenant Colonel Jeff C. Rabin, O.D., Ph.D., a licensed optometrist on active duty in the Army, assigned as the Director of Refractive Research at the Walter Reed Army Institute for Research, Walter Reed Army Medical Center.²⁵ *Id.* Exh. 8. Colonel Rabin, who

also testified at the Rule 104(a) hearing, has testified as an expert witness on the effects of alcohol and drugs on eye movements, given presentations to Army doctors and optometrists on this subject and reviewed the NHTSA publications regarding the HGN and other SFSTs. *Id.* Exhs. 8, 9. His affidavit and trial testimony confirmed the fact that alcohol ingestion can enhance the presence of nystagmus in the human eye at BAC levels as low as .04. He expressed the opinion that "there is a very good correlation between the results of the ... [HGN] test and breath analysis for intoxication." *Id.* He also stated that the three "clues" that officers are taught to look for in connection with the HGN SFST "are indicative of alcohol consumption with possible intoxication." *Id.* Colonel Rabin expressed his belief that police officers could be trained adequately to administer the HGN test and interpret its results.

Colonel Rabin's testimony was consistent with his affidavit. He did acknowledge, however, that he acquired his knowledge of, and formed his opinions about, the SFSTs in connection with performing duties as an expert witness for Army prosecutors in two courts martial, not as a result of any independent research that he had done as an optometrist. It further was acknowledged that Colonel Rabin was not asked to analyze in any detail the reliability and validity of the NHTSA SFST studies, and he had no opinion on this subject. Further, the references to the HGN SFST that he read in peer review literature published by the American Journal of Optometry was based primarily on the NHTSA studies, rather than any independent research by that organization. He also acknowledged, in response to questions from the Court, that there are many causes of exaggerated nystagmus in the human eye that are unrelated to the ingestion of alcohol.

DISCUSSION

A. The State Case Law

State courts have wrestled with the admissibility of SFST results in drunk driving cases since 1986, when the Supreme Court of Arizona decided *State v. Superior Court*, 149 Ariz. 269, 718 P.2d 171 (1986). In that decision, based on the testimony before the trial court by Dr. Burns and three police officers, and using the *Frye*²⁶

test, the court held that the results of a HGN test were sufficiently reliable to be used to establish probable cause to arrest a motorist for DWI/DUI, and that it had achieved general acceptance among behavioral psychologists, highway safety experts, neurologists and law enforcement personnel. *Id.*, 718 P.2d at 180. The court therefore held that HGN evidence was admissible to prove driver intoxication/impairment.²⁷ *Id.* at 181.

*547 Since the 1986 Arizona decision, a majority of the states have ruled on the admissibility of HGN and SFST evidence. A reading of these cases reveals that there are a core of decisions that have attempted to undertake a thorough review of the facts relating to admissibility of SFST evidence. Other state courts have relied more on the rulings of courts that previously had addressed the issue than on their own independent evaluation. It would unnecessarily lengthen this opinion to discuss all the state cases in detail. Thus, the Appendix attached to this opinion includes a chart that identifies the majority of state cases and briefly summarizes their holdings.²⁸ I will, however, discuss certain of the state cases in this opinion, as they are essential to understanding the rulings reached herein.

Maryland's appellate cases discussing the admissibility of HGN and other SFST evidence fall into the category of state court cases that have undertaken a comprehensive evaluation of the admissibility of this evidence. The principal case, *Schultz v. State*, 106 Md.App. 145, 664 A.2d 60 (1995), has been cited repeatedly by other state courts in support of their own rulings on the admissibility of SFST evidence.

The defendant in *Schultz* was convicted of DUI. At the trial in the circuit court, the state's only evidence that the driver was driving under the influence of alcohol came from the arresting officer. Accordingly, the Court of Special Appeals was deprived of any evidence of record regarding the reliability of the HGN test. Its decision in *Schultz* was based on the court's own evaluation of other cases and the published literature regarding the HGN test from which the court took judicial notice of its reliability and general acceptance. *Id.*, 664 A.2d at 69-74. In doing so, the court observed that under *548 Rule 5-70229 of the Maryland Rules of Evidence, it was required to apply the *Frye* test, adopted in Maryland in *Reed v. State*, 283 Md. 374, 391 A.2d 364 (1978).³⁰ In doing so, the court used a three prong test to determine whether HGN evidence satisfied the *Frye/Reed* test: (1) whether the scientific theory underlying the HGN test was *reliable*; (2) whether the *methods* used in connection with the HGN

test had been accepted by scientists familiar with the test and its use; and (3) whether the police officer in the case at bar properly had been trained to administer the test and administered it properly.³¹ *Id.*, 664 A.2d at 64. The *Schultz* court based its findings regarding the HGN test on the Arizona Court's decision in *State v. Superior Court*, the decisions of other state courts, as well as its reading of various studies and articles. *Id.* at 72–73. Its consideration regarding the reliability of the HGN test, however, is most significant with respect to the ruling made in this decision. Because it lacked the robust evidentiary record available to this court regarding the reliability of the HGN, OLS, WAT tests, the Court of Special Appeals was required to look at case law and published materials to determine whether the HGN test was reliable and generally accepted. The primary bases for its conclusion that it was, and that it therefore could take judicial notice of this fact, were a decision by the Texas Supreme Court in *Emerson v. State*, 880 S.W.2d 759 (Tex.Crim.App.1994), a 1986 article authored by Edward B. Tenney and published in the *New Hampshire Bar Journal*,³² and the NHTSA 1983 Field Evaluation. *Id.* at 73 and n. 12.

In *Emerson*, the Texas court based its conclusions regarding the reliability of the HGN test on the NHTSA studies. *Emerson*, 880 S.W.2d at 766–67. The Tenney article cited only the NHTSA studies regarding the scientific basis for the HGN test and reached the conclusion that “[i]f the State of New Hampshire is still a true *Frye* jurisdiction, then the likelihood that results from horizontal gaze nystagmus testing will be admitted into evidence in this state is extremely thin,”³³ making it a questionable source to cite for the reliability *549 HGN testing. Finally, the conclusions of the NHTSA 1983 Field Evaluation have been aggressively challenged by Horn's experts in this case. In short, the foundation of the Court of Special Appeals' decision that the HGN test was sufficiently reliable and generally accepted rests on taking judicial notice of studies and articles that, at the time of their publication, had not been subject to the type of critical evaluation presented in this case.

^[10] The doctrine of judicial notice is predicated upon the assumption that the source materials from which the court takes judicial notice are reliable.³⁴ Where, as here, that reliability has been challenged, the court cannot disregard the challenge, simply because a legion of earlier court decisions reached conclusions based on reference to the same then-unchallenged authority. For the reasons that will be explained below, on the record before me, I cannot agree that the HGN, WAT and OLS tests, singly or in

combination, have been shown to be as reliable as asserted by Dr. Burns, the NHTSA publications, and the publications of the communities of law enforcement officers and state prosecutors. While I ultimately agree, in large part, with the conclusions reached by the vast majority of state courts that the results of the HGN tests are admissible as circumstantial evidence of alcohol consumption, I must do so by recognizing their limited reliability and with substantial doubts about the degree of their general acceptance within an unbiased scientific or technical community.

This is not to say that I am critical of the decisions in *Schultz* or the other state courts. To the contrary, they are, for the most part, well-reasoned and written, based on the information then available to the deciding courts and the inherent limitations of the process by which courts receive proof—either from evidence introduced by the parties themselves or by the taking of judicial notice from decisions of other courts or published materials. The *550 Court of Special Appeals itself noted the danger inherent in such a process:

We note with some caution the dissent in *Emerson, supra*, which initially noted that, by taking judicial notice of the reliability of HGN testing and technique, the appellate court had relieved the State of its burden of establishing the reliability of the test at trial. We acknowledge that we, in taking judicial notice of the reliability of the test ... are likewise relieving the State of that burden. We shall, nevertheless, take judicial notice that HGN testing, a scientific test, is sufficiently reliable and generally accepted in the relevant scientific community.... To do otherwise at this stage in the development of the science would leave to individual courts within the twenty-three jurisdictions of this State (and the various courts and judges within each jurisdiction) to determine, on a case-by-case basis, the scientific reliability of the test. In each of the various jurisdictions, the determination of the reliability and acceptability of such evidence

would depend upon the competence, energy, and schedules (and even budgets) of the various prosecutors throughout the State in obtaining, and producing the attendance of experts at the thousands of trials involving alcohol related offenses in which HGN testing is sought to be admitted. Disparate results and decisions might result in many instances, not from the actual scientific reliability of the tests themselves, but from the differing abilities and resources of prosecutors and the availability of witnesses from the scientific community.

Schultz, 664 A.2d at 74.

The practical truth of the above reasoning cannot be denied. None today can doubt the serious public safety concerns related to driving by intoxicated or impaired motorists or the magnitude of this problem.³⁵ Neither can it be disputed that, given the volume of DWI/DUI cases, the press of other criminal cases, and the limited resources and time of prosecutors to prepare them for trial, it is highly desirable to have available a simple, inexpensive, and reliable test that can be administered by police officers on the road, which would facilitate a prompt and inexpensive trial. Indeed, Rule 102 would militate in favor of interpreting the rules of evidence in such a fashion as to accomplish this end, if fairly possible. What cannot be lost in the process, however, is the requirement that the trial be a fair one and that the sum of the evidence introduced against the defendant must be sufficiently probative to prove guilt beyond a reasonable doubt.³⁶ Expedient as it may be for courts to take judicial notice of scientific or technical matters to resolve the crush of DWI/DUI cases, this cannot be done in the face of legitimate challenges to the reliability and accuracy of the tests sought to be judicially *551 noticed. As will be seen, there is a place in the prosecutor's arsenal for SFST evidence, but it must not be cloaked in an aura of false reliability, lest the fact finder, like the protagonist in the Thomas Dolby song, be "blinded by science" or "hit by technology."³⁷

From a review of the state court decisions regarding the admissibility of HGN evidence in particular, and SFST

evidence in general, a number of observations may be made. First, most of the states that have ruled that HGN evidence is admissible have not allowed it to be used to prove specific BAC but instead only as circumstantial proof of intoxication or impairment. *See, e.g., Ballard v. State*, 955 P.2d 931 (Alaska Ct.App.1998); *State v. City Court of the City of Mesa*, 165 Ariz. 514, 799 P.2d 855 (1990); *State v. Ruthardt*, 680 A.2d 349 (Del.Super.Ct.1996); *State v. Garrett*, 119 Idaho 878, 811 P.2d 488 (1991); *State v. Buening*, 229 Ill.App.3d 538, 170 Ill.Dec. 542, 592 N.E.2d 1222 (1992); *State v. Taylor*, 694 A.2d 907 (Me.1997); *Wilson v. State*, 124 Md.App. 543, 723 A.2d 494 (1999); *State v. Baue*, 258 Neb. 968, 607 N.W.2d 191 (2000); *City of Fargo v. McLaughlin*, 512 N.W.2d 700 (N.D.1994); *State v. Bresson*, 51 Ohio St.3d 123, 554 N.E.2d 1330 (1990); *State v. O'Key*, 321 Or. 285, 899 P.2d 663 (1995); *State v. Sullivan*, 310 S.C. 311, 426 S.E.2d 766 (1993); *Emerson v. State*, 880 S.W.2d 759 (Tex.Crim.App.1994).

Second, most of the states that have ruled that HGN evidence is admissible have employed the *Frye* standard requiring general acceptance of the test within the relevant scientific or technical community. *See, e.g., Malone v. City of Silverhill*, 575 So.2d 101 (Ala.Crim.App.1989); *State v. Superior Court*, 149 Ariz. 269, 718 P.2d 171 (1986); *People v. Leahy*, 8 Cal.4th 587, 34 Cal.Rptr.2d 663, 882 P.2d 321 (1994); *Williams v. State*, 710 So.2d 24 (Fla.Dist.Ct.App.1998); *Hawkins v. State*, 223 Ga.App. 34, 476 S.E.2d 803 (1996); *Garrett*, 119 Idaho 878, 811 P.2d 488 (1991); *State v. Buening*, 229 Ill.App.3d 538, 170 Ill.Dec. 542, 592 N.E.2d 1222 (1992); *State v. Witte*, 251 Kan. 313, 836 P.2d 1110 (1992); *State v. Armstrong*, 561 So.2d 883 (La.Ct.App.1990); *Schultz*, 106 Md.App. 145, 664 A.2d 60 (1995); *People v. Berger*, 217 Mich.App. 213, 551 N.W.2d 421 (1996); *State v. Klawitter*, 518 N.W.2d 577 (Minn.1994); *State v. Baue*, 258 Neb. 968, 607 N.W.2d 191 (2000); *State v. Cissne*, 72 Wash.App. 677, 865 P.2d 564 (1994). Some courts, however, have used other evidentiary standards. *See, e.g., Connecticut v. Russo*, 62 Conn.App. 129, 773 A.2d 965 (2001) (remanding case to trial court to evaluate admissibility of HGN evidence under *Daubert* standard adopted by the Connecticut Supreme Court in 1997); *State v. Ito*, 90 Hawai'i 225, 978 P.2d 191 (App.1999); *Hulse v. State*, 289 Mont. 1, 961 P.2d 75 (1998);³⁸ *New Hampshire v. *552 Duffy*, 778 A.2d 415 (N.H.2001) (using state evidence Rule 702 that requires showing of reliability before HGN evidence can be admitted; remanding to trial court to hold a hearing on the test's reliability); *State v. Torres*,³⁹ 127 N.M. 20, 976 P.2d 20 (1999) (reversing trial court's ruling that HGN

evidence was admissible, remanding for hearing using *Daubert* test).⁴⁰

Third, of the state cases where the courts undertook the task of evaluating the admissibility of HGN evidence, the NHTSA studies and, in many instances, the testimony of Dr. Burns, figured prominently in their conclusions that the HGN tests were admissible as evidence of intoxication or impairment. See, e.g., *Ballard v. State*, 955 P.2d 931 (Alaska Ct.App.1998)(court relied on trial testimony of Dr. Burns, NHTSA training video and testimony of state trooper. Defendant called a psychology professor and neuro-ophthalmologist); *State v. Superior Court*, 149 Ariz. 269, 718 P.2d 171 (1986) (court considered trial court testimony of Dr. Burns, two police officers, NHTSA studies, and published articles on HGN test); *People v. Joehnk*, 35 Cal.App.4th 1488, 42 Cal.Rptr.2d 6 (1995)(court considered trial testimony of Dr. Burns, NHTSA studies, testimony of a "criminalist" and a toxicologist. Defendant called an emergency room doctor to testify); *State v. Ruthardt*, 680 A.2d 349 (Del.Super.Ct.1996) (court considered trial testimony of Dr. Burns, NHTSA studies, testimony of police officer, behavioral optometrist and neuro-ophthalmologist, defense introduced testimony of Dr. Cole, one of the defense witnesses in the pending case); *Williams v. State*, 710 So.2d 24 (Fla.Ct.App.1998) (Dr. Burns, a neurologist and three state doctors called as witnesses by the state); *Hawkins v. State*, 223 Ga.App. 34, 476 S.E.2d 803 (1996) (court relied on NHTSA studies, other state court rulings and articles); *State v. Hill*, 865 S.W.2d 702 (Mo.Ct.App.1993) *553 (Dr. Burns only witness called at trial on HGN test); *State v. O'Key*, 321 Or. 285, 899 P.2d 663 (1995)(court considered testimony of Dr. Burns, an optometrist, police officer and NHTSA studies).

Finally, those courts that did not undertake an independent evaluation of the admissibility of HGN evidence tended simply to cite to the decisions of other state courts. See, e.g., *Malone v. City of Silverhill*, 575 So.2d 101 (Ala.Crim.App.1989); *Hawkins v. State*, 223 Ga.App. 34, 476 S.E.2d 803 (1996); *State v. Garrett*, 119 Idaho 878, 811 P.2d 488 (1991); *State v. Buening*, 229 Ill.App.3d 538, 170 Ill.Dec. 542, 592 N.E.2d 1222 (1992); *State v. Murphy*, 451 N.W.2d 154 (Iowa 1990); *State v. Breitung*, 623 So.2d 23 (La.Ct.App.1993); *State v. Bresson*, 51 Ohio St.3d 123, 554 N.E.2d 1330 (1990); *State v. Cissne*, 72 Wash.App. 677, 865 P.2d 564 (1994); *State v. Zivcic*, 229 Wis.2d 119, 598 N.W.2d 565 (1999).

B. Difference between *Daubert/Kumho Tire/New Rule 702* and *Frye*.

The difference in approach between the *Daubert/Kumho Tire* /New Rule 702 and the *Frye* tests reveals an unmistakable irony. The *Frye* approach to admissibility of scientific evidence was criticized widely as being too "rigid" because it would deny admissibility to evidence that was the result of new scientific discovery that, while factually sound and methodologically reliable, had not yet gained general acceptance. Christopher Mueller & Laird Kirkpatrick, Evidence § 7.8 (4th ed.1995); 29 Charles Alan Wright & Victor James Gold, Federal Practice and Procedure § 6266 (1997). Under the *Daubert* test, however, general acceptance was but one of the evaluative factors and, provided the evidence at issue was subject to being tested, did not suffer from an unacceptably high error rate and favorably had been peer reviewed, the evidence would be admitted because it was reliable. Under *Daubert*, therefore, it was expected that it would be easier to admit evidence that was the product of new science or technology.

In practice, however, it often seems as though the opposite has occurred—application of *Daubert/Kumho Tire* analysis results in the exclusion of evidence that might otherwise have been admitted under *Frye*. Although this may have been an unexpected outcome, it can be explained by the difference in methodology undertaken by the trial courts when measuring proffered evidence under *Daubert/Kumho Tire*, as opposed to *Frye*. Under *Daubert*, the parties and the trial court are forced to reckon with the factors that really do determine whether the evidence is reliable, relevant and "fits" the case at issue. Focusing on the tests used to develop the evidence, the error rates involved, what the learned publications in the field have said when evaluating it critically, and then, finally, whether it has come to be generally accepted, is a difficult task. But, if undertaken as intended, it does expose evidentiary weaknesses that otherwise would be overlooked if, following the dictates of *Frye*, all that is needed to admit the evidence is the testimony of one or more experts in the field that the evidence at issue derives from methods or procedures that have become generally accepted. Wright & Gold, 29 Federal Practice and Procedure § 6266 ("Daubert's focus upon multiple criteria for scientific validity compels the lower courts to abandon long existing per se rules of admissibility or inadmissibility grounded upon the Frye standard.").

Daubert's challenge is unmistakable. While courts may be skilled at research and analysis, the task of deciding

the admissibility of new or difficult scientific or technical evidence involves subject matters that are highly specialized, and there is a *554 risk that the court, forced to resolve an issue without the luxury of unlimited time to reflect on it, will get it wrong. This is especially true because judges do not determine the reliability of scientific or technical issues in the abstract but rather in the context of deciding a specific dispute.⁴¹

The principle shortcoming of *Frye* was that it excused the court from even having to try to understand the evidence at issue. 4 Jack B. Weinstein & Margaret A. Berger, Weinstein's Federal Evidence, § 702.05[1] (2d ed. 1997) (Under *Frye* "[t]he court itself did not have to comprehend the science involved ... [it] only had to assure itself that among the people involved in the field, the technique was acceptable as reliable."). Further, given the impact of the *stare decisis* doctrine, once a court, relying on *Frye*, had ruled that a doctrine or principle had attained general acceptance, it was all too easy for subsequent courts simply to follow suit. Before long, a body of case law could develop stating that a methodology had achieved general acceptance without there ever having been a contested, detailed examination of the underpinnings of that methodology. The admissibility of SFST evidence illustrates this hazard, as a review of the state cases reveals that, despite more than sixteen years of case law relating to this evidence, the number of instances where there have been factually well-developed and detailed challenges to the reliability and validity of the tests is extremely small.

Following the *Kumho Tire* decision and the December 2000 changes to Rule 702, a detailed analysis of the factual sufficiency and reliability of the methodology underlying expert testimony is required for all scientific, technical or specialized evidence, not just "novel scientific" evidence. This has required, at times, a reexamination of the admissibility of evidence that long has been admitted under the *Frye* test, which may result in exclusion of evidence that for years routinely has been admitted. See, e.g., *United States v. Llera Plaza*, 179 F.Supp.2d 523 (E.D.Pa.2002) (excluding aspects of evidence of latent fingerprint identification evidence on the basis of *Daubert/Kumho Tire* and Rule 702 analysis). As lawyers and courts become fully aware of the relatively recent additional requirements of *Kumho Tire* and revised Rule 702, this process of reexamination can be expected to continue. It may mean, in a very real sense, that "everything old is new again" with respect to some scientific and technical evidentiary matters long considered settled. Alarmists may see this as undesirable,

envisioning courtrooms populated by mad scientists in white lab coats and overzealous judges in black robes, busily undoing established precedent. The more probable outcome is that judges, lawyers and expert witnesses will have to learn to be comfortable refocusing their thinking about the building blocks of what truly makes evidence that is beyond the knowledge and experience of lay persons *555 useful to them in resolving disputes. The beneficiaries of this new approach will be the jurors that have to decide increasingly complex cases. *Daubert*, *Kumho Tire*, and now Rule 702 have given us our marching orders, and it is up to the participants in the litigation process to get in step.

C. Applying *Daubert/Kumho Tire* and Rule 702 in this Case

Many of the state cases debate whether SFST evidence is "scientific" or "novel science," and therefore subject to *Frye* analysis in the first instance.⁴² Under the Federal Rules of Evidence, this debate is irrelevant, as newly revised Rule 702 and the *Daubert/Kumho Tire* cases require the same analysis for any evidence that is to be offered under Rule 702. Thus, if the SFSTs in this case are being offered as direct evidence of intoxication or impairment, they then become cloaked in a scientific or technical aura, and the factors articulated in *Daubert/Kumho Tire* and Rule 702 must be evaluated by the district court under Rule 104(a) before such evidence may be admitted.⁴³

With regards to the HGN test, from the testimony before me, the materials submitted for my review by counsel, my review of all of the state cases decided to date, and many of the articles cited in those cases, it cannot be disputed that there is a sufficient factual basis to support the causal connection between observable exaggerated horizontal gaze nystagmus in a suspect's eye and the ingestion of alcohol by that person. This connection is so well established that it is appropriate to be judicially noted under Rule 201.⁴⁴ That being said, however, it must quickly be added that there also are many other causes of nystagmus that are unrelated to alcohol consumption. The *Schultz* court identified thirty-eight possible causes of *556 nystagmus,⁴⁵ and, in his testimony, Colonel Rabin agreed that most of the *Schultz* factors did, or possibly could, cause nystagmus in humans. Thus, the detectable presence of exaggerated HGN in a driver clearly is circumstantial, not direct, evidence of alcohol