

Crying Wolf

What never before published data proves about Standardized Field Sobriety Tests

By Greg Kane, M.D.

I recently reviewed never before published data from the latest, most up to date Standardized Field Sobriety Test validation research put out by the National Highway Traffic Safety Administration. The study is commonly called the 1998 San Diego study.¹ What I discovered may surprise you.²

To a first approximation the SFST works this way: the test says everyone is guilty; the officer ignores the test and arrests or releases people according to his unstandardized gut instincts.

DUI defense attorneys often counter SFST evidence by attacking the way the SFST was performed in this particular defendant's case. The DUI defense bar complains about the un-peer reviewed science-for-hire used to support claims of SFST accuracy.

The San Diego study's raw data suggests a new defense. The science has been done. **The science proves SFSTs do not work.** The science proves that if juries rely on the SFST to decide the guilt of drivers charged with DWAI at the current 0.05% level, they will wrongly convict ninety-three percent of the innocent drivers who go to trial.

Field Sobriety Tests

Field sobriety tests are imagined to be accurate, objective measures of blood alcohol concentration.³ Suspect drivers do two coordination exercises (One Leg Stand, Walk And Turn), and an officer checks their eyes for jerkiness (Horizon-

Figure 1. What the San Diego SFST Validation Study Keeps Secret

1 a. This data was reported:
Officer BAC guesstimates.¹²

| Measured BAC | | Officer's BAC Estimate 0.08% | | | % correct |
|--------------|----|--|------------|------|-----------|
| | | < 0.08 | ≥ 0.08 | miss | |
| > 0.08 | 4 | 210 | 214 | 98% | |
| < 0.08 | 59 | 24 | 11 | 71% | |
| Totals → | | 63 | 234 | 327 | |
| Correct → | | 94% | 90% | | |
| | | Innocent driver conviction rate 29% | | | |

4 high-BAC drivers were guesstimated innocent
210 high-BAC drivers were guesstimated guilty
59 innocent drivers were guesstimated innocent
24 innocent drivers were guesstimated guilty

1 b. This data was kept secret.
How the SFST actually performed.

| Measured BAC | | SFST 0.08% | | | % correct |
|--------------|----|--|------------|------|-----------|
| | | < 0.08 | ≥ 0.08 | miss | |
| > 0.08 | 0 | 213 | 210 | 100% | |
| < 0.08 | 24 | 59 | 11 | 29% | |
| Totals → | | 24 | 272 | 296 | |
| Correct → | | 100% | 78% | | |
| | | Innocent driver conviction rate 71% | | | |

0 high-BAC drivers passed the SFST
213 high-BAC drivers failed the SFST
24 innocent drivers passed the SFST
59 innocent drivers failed the SFST

Arrest Accuracy

In this study group the "accuracy" of officer's unstandardized BAC guesstimates was 90%. We've seen before the chief statistical utility of this number is to fool people into thinking a test works.⁶ NHTSA validation studies can and do (and in this study did) inflate this so called "accuracy" by skewing the mix of sober and impaired drivers they choose to study. Even so, the "accuracy" of SFST's ≥0.08% answer is only 78%. Even with the study group heavily skewed, SFSTs could not be made to look useful. The NHTSA could not report the true SFST results and still claim the SFST is useful. The San Diego SFST validation study does not reveal its SFST results.

tal Gaze Nystagmus). If they fail any one component, the standardized SFST interpretation criteria are clear: their blood alcohol concentration is above 0.10%. Or 0.08%. Or 0.05%. Or 0.04% - whatever level the NHTSA wishes to "validate" in this particular study.

Some SFST apologists claim the tests are "valid" scientific predictors of blood

alcohol concentration. Others say they identify impairment. SFST admissibility at trial varies by state, but in some jurisdictions failed FSTs are used to convict drivers whose BACs were in the legal range. Drivers proven to have BACs of zero may be charged with impairment by some other drug. The government's theory: **they failed the SFST, they must have been impaired by something.**

How did FSTs get to be this important? The foundation of field sobriety tests' forensic use⁴ lies in three un-peer reviewed "validation studies" paid for by the NHTSA in the 1990s.⁵ In the Colorado, Florida, and San Diego studies police on patrol duty administered FSTs to drivers suspected of alcohol impairment. "Using SFSTs" the officers estimated each driver's BAC (San Diego) and/or made decisions to arrest or release suspects (Colorado, Florida). These validation studies compared the officers' BAC estimates / arrest decisions with drivers' actual BACs, misapplied statistics,⁶ and pronounced success.

The San Diego study puts it this way: "Decision analyses found that officers' estimates of whether a motorist's BAC was above or below 0.08 or 0.04 percent were extremely accurate."⁷

The Data

I wanted to see SFST validation studies' raw data, so I asked the NHTSA and the Southern California Research Institute (the contractor, as I read the reports, for the Colorado and Florida studies) to release copies. No dice.

I found researcher Dr. Mike Hlastala,⁸ who sent me a Microsoft Excel file of the data set for the 1998 San Diego study. He originally got the file from

the NHTSA via a FOIA request. Later I got a second Excel file from helpful Dr. Jack Stuster, principal author of the San Diego study. The two data sets were identical.

I've put this never before published data online. You can review the records and download your own copy at: FieldSobrietyTest.info/raw.html.

Crying Wolf

With the official data on my PC, I looked inside the latest, most up to date NHTSA SFST validation study. Here's what the data shows.

The SFST cries wolf.

When drivers are impaired, the SFST cries "impaired." When drivers are not impaired, the SFST still cries "impaired."

Pause for a moment to take in the bigness of that deal. Police and courts don't use SFSTs just to identify impaired drivers, they also use the test, or imagine they do, to identify and release innocent drivers. But a test that cries wolf can't do that. The SFST cannot possibly do what the government says it does.

In the old fable, when folks in the village heard the shepherd cry "Wolf!" they couldn't tell whether their sheep were being attacked or not, because the shepherd boy always cried "wolf."

Nowadays folks in the courthouse can't tell whether a driver was impaired or not, because the SFST always cries "impaired."

Always cries "impaired"? OK, I exaggerate — by one percentage point. In the San Diego validation study, using the standardized interpretation criteria for a 0.04% BAC, ninety-nine percent of everyone given the SFST failed.

296 drivers took the SFST
292 failed — 99%.
4 passed — 1%

Twenty-nine innocent people took the SFST. Twenty-seven failed—ninety-three percent. On innocent people the SFST cries wolf ninety-three times out of one hundred. On innocent people, the accuracy is seven percent. **Seven percent!**

So when a jury hears that a driver failed an SFST, how can they tell whether the driver was really impaired, or whether the driver was just one of those ninety-three percent of innocent drivers who also fail the SFST? They can't. If juries rely on the SFST to decide the guilt of drivers charged with DWAI at the current 0.05% level, they will wrongly convict ninety-three percent of the innocent drivers who go to trial.

And yet the NHTSA claims: "Using only the standardized 3-test battery (Walk-and-Turn, One-Leg Stand, Horizontal Gaze Nystagmus), officers seldom erred when they decided to arrest a driver. Breath or blood specimens confirmed that 93% of the arrested drivers were above 0.05% BAC."⁹

You got that? The NHTSA has a way to make you believe a test with an innocent driver false conviction rate of ninety-three percent does the opposite — gives the correct answer ninety-three percent of the time. Howdo they do that?

Part of how they do that is to use statistical tricks. The way to validate a shepherd who shouts "Wolf!" three times a day is to stand him beside a pack of wolves. That way it looks like he's really spotting wolves, instead of just randomly crying "Wolf!"

Figure 1c. The SFST at 0.04% BAC

The SFST has two sets of interpretation rules. One is imagined to target a BAC of 0.08%. Or 0.10%. Or 0.05%. The other is imagined to target 0.04%. Here's what the unpublished raw data proves about the SFST at the 0.04% level.

The NHTSA's misleading "accuracy" statistic, the go-to statistic the agency uses to validate the SFST in every validation study, including this one, looks pretty good: 91%.¹² But here, at the low 0.04% BAC, the NHTSA's study report didn't mention this favorable looking "accuracy." Why?

Look at the accuracy of the SFST on innocent drivers: 7%. That's not a typo. Seven percent. On innocent drivers, the SFST gives the wrong answer 93% of the time! The NHTSA couldn't tell you about the favorable looking (but misleading) 91% accuracy without letting you see the real 93% inaccuracy. The San Diego SFST validation study does not reveal its SFST results.

| | | SFST 0.04% | | | % correct |
|--------------|--------|---------------------------------|--------|-----|-----------|
| | | < 0.04 | ≥ 0.04 | | |
| Measured BAC | ≥ 0.04 | 2 | 286 | 287 | 99% |
| | < 0.04 | 2 | 27 | 29 | 7% |
| Totals = | | 4 | 313 | 317 | |
| Correct = | | 91% | | | |
| | | Innocent driver conviction rate | | 93% | |

Instead of studying the SFST with what scientists call a "random sample" of drivers, NHTSA validation studies skew the groups of drivers they study. They load up on drunks. Study groups skewed to drunks inflate the accuracies these studies "discover." Skewed samples make it look like the SFST is really spotting impairment, instead of just crying "impaired" for most people tested.

Earlier articles in this series deconstruct the statistics.¹⁰

Magician's Misdirection

The other part of how the NHTSA makes you think ninety-three wrong is ninety-three right is to fool you with a

magician's misdirection.

SFST validation studies gather data on and report **officer decisions**. They also gather data on SFST accuracy, but they keep those results secret. Unless you know what to look for, you'll probably miss the distinction. This is an SFST study; you figure you're seeing the SFST's accuracy. You're not. You're seeing the (statistically enhanced) accuracy of the police officers' guesstimate. The inaccuracy of the SFST stays hidden because the accuracy of the SFST itself is never released.

It's as if the agency did a shepherd validation study by having the village policeman drive out and check for

wolves himself — and the study reported the officer's performance as if it were the shepherd's, thus keeping the shepherd's inaccuracy secret. "When the shepherd was identifying wolves, the officer's decisions were 93% accurate."

Look at Figure 1a. This is accuracy information released in NHTSA's SFST report. But it does not reflect the SFST. It reflects officers' estimates. In this group of drivers, officers' estimates were ninety percent accurate.

Now look at Figure 1b. Compiled from never before published data the NHTSA claims to have lost, **this** is the SFST's accuracy. The SFST predictions were only 78 percent accurate.¹¹

Figure 2 OFFICERS DID NOT USE SFST TO GUESSTIMATE BAC

The unpublished raw data from the San Diego SFST validation study proves police officers could not have used SFST results to identify impaired drivers. Let me ask you to estimate driver BACs. Your only information: real SFST scores from the latest, most up to date SFST validation study. Let's see if it's possible to duplicate the study officers' results. I've selected all the drivers in the study with the following SFST score: each failed the HGN test at the 0.08% BAC level, and each passed both the OLS and WAT tests. Please estimate each driver's BAC.

How many different BACs did you predict for these thirteen drivers? What criteria did you use to pick each BAC? Where did those criteria come from? How did you know they would work? Second question: It turns out seven of these thirteen drivers were innocent. Please identify the innocent drivers.

Now let's see how San Diego police officers did. Remember, these drivers had **identical** SFST scores.¹⁴ According to standardized SFST interpretation criteria detailed in the study's report, each driver should have had a BAC estimate of "**? 0.08.**" Instead, officers somehow came up with these results: 0.03, 0.05, 0.06, 0.06, 0.06, 0.07, 0.08, 0.09, 0.12, 0.12, 0.14, 0.14, 0.19. That's nine different BAC guesstimates, from 0.03% all the way up to 0.19%. The high guesstimate was more than six times the low, for identical SFST scores! One SFST score; nine different BACs. How did officers do that? What criteria did they use? How did they know those criteria would work? Officers did not, could not had they wanted to, rely on these identical SFST scores to come up with nine different BAC guesstimates. In the San Diego SFST validation study, officers **did** SFSTs, they did not use SFSTs.

What's more, instead of the SFST's standardized interpretation results - BAC **<0.08** or **≥0.08** - officers were somehow able to guesstimate BAC levels to 1 part in 100. There are no standardized SFST interpretation criteria for estimating BAC to 1 part in 100. Officers did not, could not had they wanted to, use these identical SFST scores to come up with their nuanced, 1 part in 100, BAC scores. In the San Diego SFST validation study, officers **did** SFSTs, they did not use SFSTs.

| Case | HGN | OLS | WAT | SFST prediction | Your BAC estimate |
|------|-----|-----|-----|-----------------|-------------------|
| 29 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 34 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 54 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 56 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 62 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 76 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 114 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 134 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 166 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 176 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 189 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 232 | > 4 | < 2 | < 2 | ≥ 0.08% | |
| 280 | > 4 | < 2 | < 2 | ≥ 0.08% | |

Remember how you had no way to pick which of these drivers were innocent? The study officers had a way. They knew almost exactly which SFST results to throw out. The SFST said every one of these thirteen drivers was guilty. The SFST cried wolf. The SFST was wrong. Seven were innocent. Guess what, officers correctly identified five of those innocent drivers as having low BACs. How'd they do that? How did officers know which SFSTs to throw out? What criteria did they use? How did they know those criteria would work? You couldn't do it. Neither can I. Because it can't be done, not with the SFST. Officers must have used some method other than the SFST to determine driver BAC levels. The guys doing the validation study knew the SFST doesn't work, so they just ignored the results! That's what the validation study proves.

Figure 3. OFFICER'S IGNORED THE SFST

Over and over SFST study officers ignored SFST results and systematically violated standardized SFST interpretation criteria. They knew the test didn't work. They ignored it.

Figure 3a SFST 0.08%

| | | | |
|--------------|--------|---------------------------------|-----------------------------------|
| | | SFST 0.08% | |
| | | < 0.08 | ≥ 0.08 |
| Measured BAC | ≥ 0.08 | 0 ignored: 0 ignored: n/a | 213 ignored: 4 ignored: 2% |
| | < 0.08 | 24 ignored: 0 ignored: 0% | 59 ignored: 36 ignored: 66% |
| | | ignored TRUE | 2% |
| | | ignored FALSE | 59% |

The boxes in 3a show you how often officers ignored the SFST. The top number in each box (0, 213, 24, 59) is the driver counts from Figure 1b. The middle number, from a driver by driver analysis of the unpublished raw data, is a count of how many of those SFST results officers ignored. The bottom number turns that count into a percentage: 4 is 2% of 213, etc. When the SFST gave the correct answer, officers rejected the test only 2% of the time.

When the SFST gave the wrong answer, officers rejected the test a whopping 59% of the time! How'd they do that? How did officers know which SFSTs to ignore and which SFSTs to accept? The chance of this distribution of answers happening by chance, at random, is tiny. Officers must have had some other way to determine driver impairment.

The pattern of officers ignoring the SFST repeats across the entire study.

Figure 3b

| Component test | SFST rejection rate | |
|---------------------------------|---------------------|---------------|
| | True answers | False answers |
| Horizontal Gaze Nystagmus 0.04% | 2% | 54% |
| Horizontal Gaze Nystagmus 0.08% | 4% | 49% |
| One Leg Stand | 4% | 62% |
| Walk and Turn | 3% | 63% |
| SFST 0.08% | 2% | 59% |
| SFST 0.04% | 1% | 48% |

The officer's ninety percent accuracy is touted in the study's final report. The SFST's seventy-eight percent accuracy is not included at all. **The SFST's accuracy is kept out of SFST validation study reports.** If the NHTSA told you the SFST is only seventy-eight percent accurate, you wouldn't believe the test was worthwhile.

Look at Figure 1c. Compiled from never before published data the agency claims to have lost, this is the SFST's accuracy performance at the 0.04%

BAC level. If the NHTSA released the raw SFST study data, people could see that on innocent people at the current legal limit the SFST gives the wrong answer ninety-three percent of the time. No one would believe a test like that was worthwhile.

Releasing the SFST validation data would prove the test is not "valid." The SFST's accuracy is kept out of SFST validation study reports. The agency claims the data is lost.

Figure 4a: Show me the graph

This is a graph of never-before-published data from the NHTSA's 1998 San Diego SFST validation study. Data was gathered by seven experienced alcohol enforcement officers given personal SFST training by Dr. Marcelline Burns,¹⁵ (effectively the inventor of the SFST). Each point on the graph represents one driver in the study: total FST score on the x-axis, Blood Alcohol Concentration on the y-axis. Drivers who failed the FST at the 0.04% BAC level are represented by solid dots [●]. Drivers who passed the SFST at that level are represented by open squares [□].

296 drivers performed at least one of the three components of the SFST.

292 failed – 99%

4 passed – 1%

Three of these four drivers were tested by a single officer. The fourth was tested only with OLS.

In other words, seven highly experienced alcohol enforcement officers,

personally trained by Dr. Burns, patrolled a major US city for more than five months, stopping and assessing hundreds of motorists. And in all those months, in all those hundreds of tests, only one officer ever completed even a single SFST that came back "non-impaired" at the 0.04% BAC level. NHTSA science proves that for six of seven highly experienced DUI patrol officers, every single driver who is able to take the SFST fails the SFST.

At BAC 0.04% six of seven officers did SFSTs that failed every driver who could take the test. Their accuracy on innocent drivers was zero percent. Zero percent! See if you can spot the magician's misdirection in the study report's description of these facts: "Officers' estimates of whether a motorist's BAC was above 0.04 percent but lower than 0.08 percent were accurate in 94 percent of the decisions to arrest and in 80 percent of cases overall."¹⁶

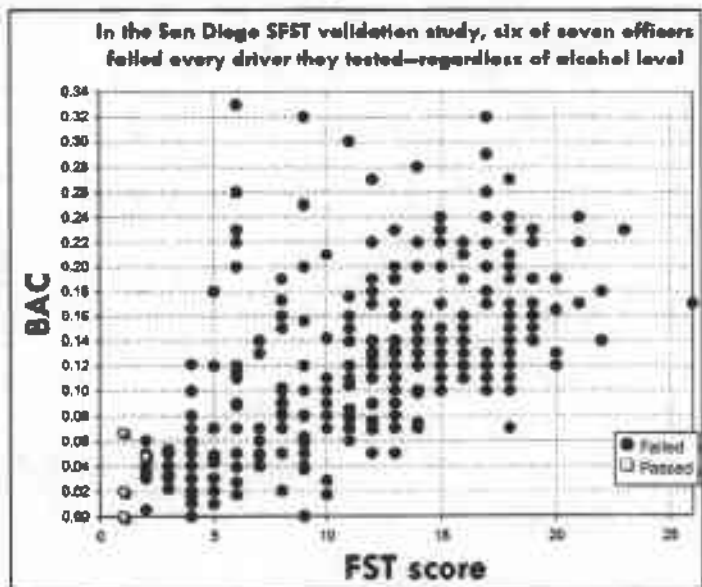
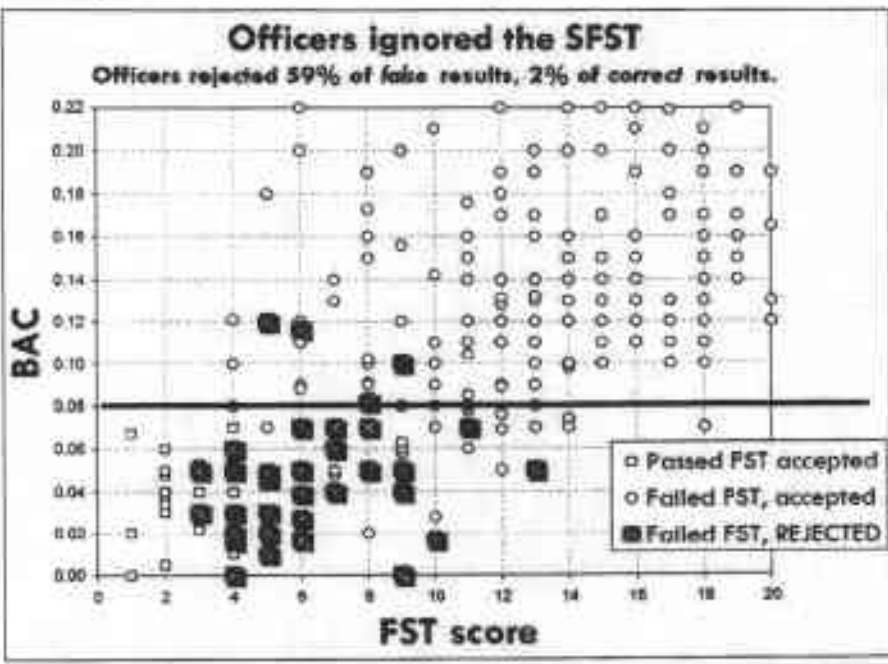


Figure 4d. Officers systematically ignored the SFST BAC 0.08%

This data from the San Diego validation study shows where the SFST's mistakes are (below the black line), and which of those SFST mistakes study officers corrected by ignoring the SFST results (black boxes).

The dark horizontal line marks the then legal BAC limit, 0.08%. The open circles below the black line and black boxes below the black line are SFST mistakes—59 open circles (some stacked on each other). 35 of the circles have been converted to black boxes (some stacked), representing drivers whose mistaken SFST result, "guilty," was ignored by the officer in favor of the officer's unstandardized gut instinct, "innocent."

If the NHTSA had reported the actual SFST results, people would have known that the accuracy of the SFST on innocent people is 29%. Instead, study officers were allowed to ignore the SFST, effectively correcting the test's mistakes. The NHTSA then kept the actual SFST results secret and reported only the officer-corrected numbers.



SFSTs Did Not Spot Impairment

All right, the SFST isn't perfect, but at least it helps officers spot impaired drivers, right? No. The answer is No.

There's fancy math to prove this, but common sense will do. Think about it this way: to identify impairment a test must also identify non-impairment—otherwise it can't tell one from the other. The raw data prove the SFST cried "impaired" for ninety-nine percent of all the people tested. The SFST can't tell the difference between impaired and not impaired. You cannot use the SFST to tell which drivers were and which were not impaired.

Try it yourself. Figure 2 asks you to use real SFST results to estimate BAC levels for real drivers. Do the exercise. Answer the questions. Did San Diego study officers use the SFST to guide their BAC estimates? They couldn't have. It's not possible.

The SFST cannot have been how study officers identified impairment.

Study Officer's Ignore the SFST

The SFST cries wolf a lot. For innocent drivers, it cries "impaired" ninety-three percent of the time. On innocent people, the test gives the wrong answer ninety-three percent of the time.

To understand how study officers dealt with the SFST when it cried wolf,

Figure 4b. INNOCENT drivers, BAC 0.04%

These data from the San Diego validation study shows SFST performance on innocent drivers at BAC 0.04%.

29 innocent drivers took the SFST. 2 passed. 27 failed—93%.

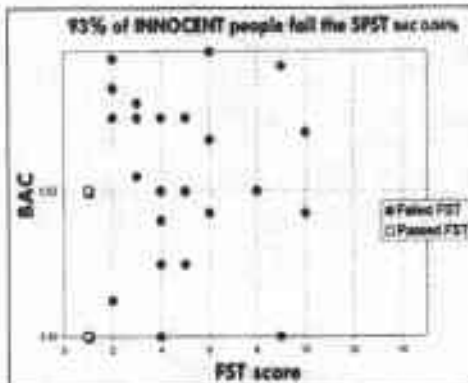
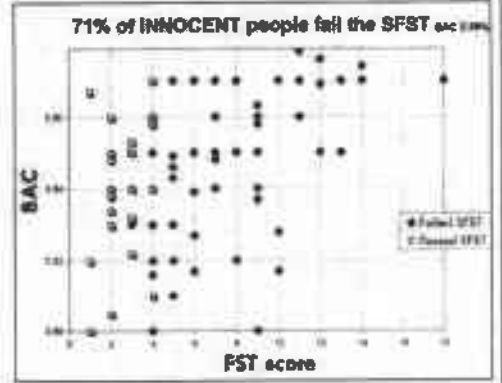


Figure 4c. INNOCENT drivers, BAC 0.08%

This data from the San Diego validation study shows SFST performance on innocent drivers at BAC 0.08%.

83 innocent drivers took the SFST. 24 passed. 59 failed—71%.



pretend you're the village policeman, out in the pasture checking for wolves. You look around. Nary a beastie. The shepherd kid cries "Wolf!" What do you do? Do you grab your radio and call in the wolf SWAT squad? Of course not. The shepherd **always** cries wolf. You can't trust him. You don't trust him. You've just checked for yourself; you know there's no wolf. You ignore the shepherd.

That's how police in validation studies deal with the SFST.

Figure 3 keeps track of how often study officers reversed the SFST's decision. At the then legal 0.08% BAC limit the SFST mistakenly identified fifty-nine innocent people as having high BACs. Officers knew better. Officers simply ignored thirty-five of those SFST results. When the SFST gave the wrong answer, fifty-nine percent of the time study officers simply ignored that answer.

To a first approximation the SFST works this way: the test says everyone is guilty; the officer ignores the test and releases people he knows are innocent. The guys doing the validation study knew the SFST doesn't work, so they just ignored the results! That's what the validation study proves.

Greg Kane, MD owns Med-Mal Experts, Inc., (www.medmalExperts.com) an Englewood based consulting firm that reviews medical malpractice claims and refers attorneys across America to physician expert witnesses in all specialties. 303-741-0993 FST@medmalExperts.com

© Greg Kane, M.D. 2008

Endnotes

- 1 Jack Stuster and Marcelline Burns, *Validation of the Standardized Field Sobriety Test Battery At BACs Below 0.10 % 1998*.
- 2 This article is about deconstructing the logic and mathematics of the NHTSA's SFST validation theory. Nothing in this article is a statement about the knowledge or intentions of the NHTSA or it's contractors or any SFST researchers.
- 3 The only appropriate criterion measure to assess the accuracy of SFSTs is BAC [blood alcohol concentration]. Measures of impairment are irrelevant because performance of the SFSTs must be correlated with BAC level, rather than driving performance. BAC provides an objective and reliable measure that states have recognized as presumptive and/or per se evidence of impairment, depending on the statute.
Jack Stuster and Marcelline Burns, *Validation of the Standardized Field Sobriety Test Battery At BACs Below 0.10 Percent, 1998*, p.10.
- 4 NHTSA *DWI Detection and Standardized Field Sobriety Testing, Student Manual, 2004 Edition, Page VIII-2*; Stuster, *supra* n. 3 at 3 has a nice roundup of the peripheral SFST studies.
- 5 Marcelline Burns and Ellen Anderson, *A Colorado Validation Study of the Standardized Field Sobriety Test (SFST) Battery, 1995*; Marcelline Burns, *A Florida Validation Study of the Standardized Field Sobriety Test (S. F. S. T.) Battery, 1997*; Stuster *supra*.
- 6 Greg Kane, Percentages of Percentages, Why "Validation" Studies Fail to Validate, TRIAL TALK, Aug./Sept. 2006, p. 33, available at: www.FieldSobrietyTest.info; Greg Kane, Field Sobriety Tests: How Basic Science Proves They Have Little Power to Tell Impaired From Sober, TRIAL TALK, Oct./Nov. 2006, p. 57, available at: www.FieldSobrietyTest.info; Greg Kane, *Field Sobriety Tests: Two Statistical Tricks Let NHTSA Contractors Prove Any Roadside Sobriety Test Is "Extremely Accurate"* TRIAL TALK, April/May 2007, p. 35, available at www.FieldSobrietyTest.info.
- 7 Stuster, *supra* n. 3 at iii.
- 8 Hlastala, Polissar and Oberman, *Statistical Evaluation of Standardized Field Sobriety Tests*, J. OF FORENSIC SCIENCE, May 2005, vol. 50 n. 3.
- 9 Burns & Anderson, *supra* n. 5, *Technical Summary*.
- 10 See Greg Kane, M.D., *Field Sobriety Tests: Percentages of Percentages, Why "Validation" Studies Fail to Validate*, TRIAL TALK, August/September 2006 at 31; Greg Kane, M.D., *Field Sobriety Tests: How Basic Science Proves They Have Little Power to Tell Impaired From Sober*, TRIAL TALK, October/November 2006 at 57; Greg Kane, M.D., *Field Sobriety Tests: Two Statistical Tricks Let NHTSA Contractors Prove Any Roadside Sobriety Test Is "Extremely Accurate,"* TRIAL TALK, April/May 2007 at 33.
- 11 Don't be fooled by the officers' high 90% "accuracy." The articles cited in note 6 above explain why this accuracy is a statistical trick.
- 12 Data in Figure 1a from Stuster, *supra*, at 18, fig. 4. Data in Figures 1b and 1c is from the previously unpublished raw data set for this same study. A copy of the complete dataset is available free at: www.FieldSobrietyTest.info/raw.html.
- 13 Kane *supra*, n. 6. More statistical shenanigans: the *lower* the BAC limit, the *higher* the phony accuracy becomes. This is a property of the statistic, not of the test itself.
- 14 I'd like to give you the drivers' raw HTN, OLS and WAT scores, but doing that would violate standardized SFST interpretation criteria listed in the validation study. Violating those criteria would invalidate the test:

IT IS NECESSARY TO EMPHASIZE THIS VALIDATION APPLIES 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.

NHTSA *DWI Detection and Standardized Field Sobriety Testing, Student Manual, 2004 Edition, page VIII*.
- 15 Stuster, *supra* n. 3 at 13.
- 16 *Id.* at iii.