# Alcohol-Impairment Tests for DWI Arrests

Marcelline Burns and Herbert Moskowitz, Southern California Research Institute, Los Angeles

**A study undertaken to develop an improved battery of field tests of driver sobriety is discussed. In most states, 0.10 percent is the blood alcohol concentration at which a driver is presumed to be driving while intoxicated, but the mean blood alcohol concentration of arrested drivers is closer to 0.17 percent. This reflects the difficulty of the police officer's task. Tests for determining whether a motorist is driving while intoxicated and whether an arrest should be made must be sensitive to alcohol impairment and meet the severe constraints imposed by limited time and the characteristics of the roadside environment. Six types of tests were selected for evaluation based on the literature, field observations, and pilot studies. In a laboratory study, 10 police officers administered the tests to 238 participants at 0-0.15 percent levels of blood alcohol concentration. Based on analysis of the data, a battery of three tests was selected: one-leg stand, walk and turn, and alcohol gaze nystagmus. By using the police officers' scores for these three tests, it was possible to correctly classify 83 percent of the participants as above or below 0.10 percent blood alcohol concentration.**

In spite of a variety of efforts aimed at reducing and controlling the problem of drinking drivers, nationwide traffic accident statistics still show a high proportion of alcohol-related fatalities and injuries. Currently, the major approach to dealing with the alcohol-impaired driver is deterrence, a process that begins with the police officer in the field.

It appears that traffic-enforcement officers have difficulty detecting and arresting drinking drivers [commonly referred to as driving while intoxicated (DWI) or driving under the influence (DUI)]. Evidence abounds that only a small proportion of alcohol-impaired drivers are detected and arrested by the police. Estimates of the probability of being arrested while driving under the influence of alcohol vary from 1 in 200 to 1 in 2000.

Additional evidence of problems in the execution of this police function can be found in the blood alcohol concentration or blood alcohol content (BAC) distribution of DWIs. In the United States, the mean BAC of individuals arrested for driving while under the influence is approximately 0.17 percent. Yet in most states the BAC at which a driver is presumed to be alcohol impaired is 0.10 percent. This problem is also reflected in the discrepancy between the BAC distribution of persons on the road and those arrested for driving while intoxicated. A California roadside survey found that, even though 67 percent of all drivers whose BAC was 0.10 percent or higher had a BAC between 0.10 and 0.14 percent (1), only 26 percent of the drivers arrested as DWIs had BACs in that range. Thus, although the BACs of a majority of alcohol-impaired drivers are below 0.15 percent, 73 percent of arrested DWIs were at or above that BAC level.

The low probability of arrest for drinking drivers and the alcohol levels at which they actually are arrested reflect at least two areas of difficulty. First, police officers on traffic patrol must detect drinking drivers by observing traffic and noting vehicle-handling errors that may or may not be the effect of alcohol. Then they must assess the stopped driver at the roadside. Is the person impaired by alcohol? What is the probable BAC? Should the driver be arrested or released? Usually, police officers are reluctant to arrest unless they are very sure that a chemical test will produce a BAC reading of 0.10 percent or higher. It is a waste of time and effort to transport and test a suspect who cannot be booked, and law-enforcement personnel generally prefer to avoid practices that may bring charges of harassment and generate poor police-community relations.

Typically, the conditions for roadside evaluation of alcohol impairment are less than optimal. The officer must make a decision within a short period of time in an environment that may complicate the evaluation. Conditions such as lighting, noise, space, terrain, and onlookers vary widely, and the officer's assessment process must be adapted accordingly. Further, the driver's degree of alcohol impairment may be masked or confounded by drinking history, physical condition, age, and other variables about which there may be little or no immediate information. The consequence of these various areas of difficulty is that a large proportion of drivers who have BACs in the 0.10-0.15 percent range are not detected or are not arrested. The result is an underrepresentation of low BACs among arrested drivers.

The study described here was undertaken to examine and improve procedures for behavioral testing at the roadside after a driver has been stopped. It is believed that the police officer's use of improved behavioral tests of sobriety to assist in making a decision to arrest or release the DWI suspect will result in both an increased arrest rate and a BAC distribution for those arrested that more closely corresponds to the on-the-road BAC distribution.

In states that have a "per se" law, the BAC test provides sufficient evidence for a DWI conviction. Most states, however, are operating under a "presumptive" law; that is, a measured BAC of 0.10 percent or higher is presumptive only, and the arresting officer is required to give additional evidence of impairment, from both observed driving performance and behavioral tests. The most extensively used types of behavioral tests examine balance, coordination, and speech, but the specific tests and the procedures for administering them vary widely among law-enforcement agencies and individual officers.

The objective of this study was to develop and standardize an improved battery of behavioral tests for use by police officers in assessing DWI suspects at the roadside and providing behavioral evidence of impairment for court proceedings.

## SELECTION OF TESTS

Roadside tests of alcohol impairment must discriminate between levels of impairment under a wide variety of circumstances. The first effort in this study was to locate tests that appeared to be potentially suitable.

An important criterion for the candidate tests was that the performance results be quantifiable. It was also considered essential that test variance be small in relation to the alcohol effect so that differences in individual performance would not obscure alcohol-related impairment. Tests were sought that would be sensitive to alcohol beginning at 0.05 percent BAC and would yield scores that correlate well with BACs of 0.05-0.30 percent. The tests should be short and easy to administer without specialized apparatus or equipment. In addition, the standardized administration and scoring methods to be developed should not be so complex that it would require

difficult or lengthy training of police officers to use the tests in the field. Another important criterion in the selection process was that the test battery should be credible and acceptable to DWI suspects, law-enforcement personnel, and the judiciary.

A literature search was undertaken to locate potentially suitable behavioral measures. In addition, the application of currently used tests was observed by individuals who rode along with city and state police officers and sheriff's deputies in the cities and rural areas of five states. The 16 types of tests described below were selected for evaluation:

1. Alcohol gaze nystagmus (AGN)—The DWI suspect is asked to cover one eye and follow the movement of a small light or object with the other eye without changing the position of the head. The light is moved slowly to points that require 30° and 40° lateral deviation of the open eye. The eye is observed for a jerking movement that occurs in the presence of alcohol. The test is repeated with the other eye. The deviation required of the eye should not exceed approximately 40°. More extreme angles of deviation can result in the jerking movement even in the absence of alcohol (2).
2. Coin pickup—Three coins (or chips or matches) are placed on the floor. The individual is instructed to stand in one location, pick up the coins one at a time, and hand them to the examiner.
3. Color naming (attention diagnostic method, modified)—A card presents the numbers 10-59, in random order, in four colors, by row. The suspect is instructed to find a sequence of 10 numbers, beginning with some designated number, and to report the color of each. The verbal response, for example, might be, "10-blue, 11-white, 12-yellow, 13-red" and so on. The response measure is the time it takes the individual to report the colors of 10 numbers.
4. Finger count—The DWI suspect is instructed to touch each finger in succession, counting aloud. The examiner demonstrates: "Watch what I do. 1-2-3-4-5-5-4-3-2-1."
5. Finger to nose—The suspect stands erect with eyes closed, arms extended horizontally. Instructions are to touch the nose with the index finger, alternating right and left hands as instructed. A demonstration is given by the examiner.
6. Grip strength—The individual is instructed to squeeze as hard as possible a dynamometer of the type shaped like a pistol grip with grooves for each finger. This instrument measures the force exerted in isometric contraction.
7. Letter cancellation—The suspect is required to cancel a given letter every time it appears in a paragraph of text. The test lasts 30 s.
8. One-leg stand—The suspect is instructed to stand for 30 s with one leg held straight, slightly above the floor, in a forward direction. The eyes are to remain open.
9. Romberg (balance)—The suspect is instructed to stand with feet together, head tipped back, eyes closed, and arms at side. The position is demonstrated. The examiner observes anterior-posterior sway during a 45-s trial.
10. Serial performance—The device for the serial-performance test consists of a small box on the face of which are mounted five toggle switches and a small red light bulb. The box is presented to the individual with all switches in the center position. The instructions are to move the switches to locate the correct sequence of up-down positions. The red bulb lights up when the problem is solved. The response measure is the time to solution.
11. Subtraction, addition, counting backwards—The suspect is instructed to subtract 3 from 102 and to repeat the subtraction until the result equals some specified number (or to add 3 continuously in the same manner). The same general instructions may be given for counting backwards.
12. Tapping rate—The suspect is instructed to tap a telegraph key as rapidly as possible. The number of taps are recorded by electronic counter during a 10-s trial.
13. Tongue twisters—The suspect is asked to repeat as rapidly as possible such words as "methodist", "episcopal", and "sophisticated statistics".
14. Two-point tactile discrimination—The forearm or back of hand of the suspect is touched lightly and simultaneously with two pinpoints. The suspect's eyes are closed. The test begins with no separation between the two points, and the suspect is asked, "How many places am I touching your arm?" Trials are repeated with increasing separation between the pinpoints. The response measure is the first separation to which the person responds, "Two."
15. Tracing—The suspect is required to trace a pathway (maze) on a piece of paper. Three 20-s trials are given.
16. Walk and turn, heel to toe—The suspect is instructed to walk a straight line, touching heel to toe each step for nine steps, and then to turn and return along the same line in the same manner. A demonstration is given.

## FIRST LABORATORY EVALUATION

The 16 tests listed above were evaluated in the laboratory by using nine pilot subjects and BACs ranging from 0 to 0.10 percent. Ten of the tests were identified as being unsuitable for roadside sobriety testing. The primary reasons for discarding tests were that (a) the results showed great variability among individuals and could not be interpreted in the single case without individual baseline data or (b) the test required special devices and conditions that could not be adapted for roadside use.

The following six tests appeared to merit further study: AGN, finger count, finger to nose, one-leg stand, tracing, and walk and turn.

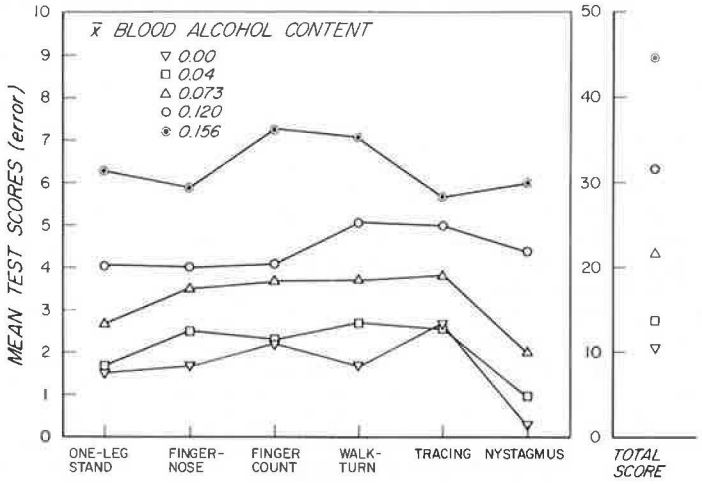
## SECOND LABORATORY EVALUATION

A large-scale laboratory study was undertaken to evaluate the six tests chosen for further research. Procedures for administering the tests, including instructions, demonstrations, and scoring, were established. To evaluate the tests and procedures as they would be used by law-enforcement personnel, 10 officers were recruited from city police, sheriff's office, and highway patrol personnel. These officer-examiners administered the test battery and made "arrest-or-release" decisions about test subjects in laboratory sessions.

Some of the officers who participated in the study had been assigned to specialized DWI units and had made thousands of DWI arrests; others had only minimal experience with drunk drivers. Before serving as examiners, all officers attended training sessions in which they learned how to administer the test battery in a standardized manner, including giving instructions and demonstrations as well as scoring and interpreting the driver's performance.

The 238 participating subjects—168 men and 70 women—were recruited through employment offices and newspaper advertisements. Their ages ranged from 20 to 71 years and averaged 26.5 years. They reported drinking habits that characterized 62 of them as light

Figure 1. Mean test scores of study participants by BAC.




drinkers, 86 as moderate drinkers, and 90 as heavy drinkers (3). These subjects were given three drinks of vodka and orange juice over a 1.5-h period, in doses based on body weight, to produce BACs from 0 to 0.15 percent. The distribution of participants by drinking category and measured BAC is given below:

| BAC (%) | Light Drinkers | Moderate Drinkers | Heavy Drinkers |
|---|---|---|---|
| 0 | 26 | 27 | 26 |
| 0.05 | 36 | 16 | 3 |
| 0.075 | | 6 | 7 |
| 0.10 | | 37 | 13 |
| 0.15 | | | 41 |
| Total | 62 | 86 | 90 |

BACs were measured by taking breath samples with a gas chromatograph intoximeter. Neither the participants nor the examiners knew the amount of alcohol or the measured BAC for any individual. The examiners had no contact with the participants prior to testing; their observations of an individual's general behavior, appearance, and speech were roughly equivalent to such observations at roadside. Typical interrogation and interaction with the laboratory "DWI suspect" was permitted. The examiner administered the test battery and scored the individual's performance of each test on a 1-10 scale—the score increasing as a function of error and poor performance—and then recorded a decision as to whether the individual should be arrested or released.

## RESULTS AND DISCUSSION

The police officers' scores showed all of the tests to be sensitive to the impairing effects of alcohol. The participants' scores correlated closely with BAC ($p < 0.01$). As Figure 1 shows, there was a consistent increase in mean error score with increase in mean BAC. However, the important question for the objectives of the study was whether the test results help the police officer in deciding the individual case. That is, do the tests discriminate between alcohol-impaired individuals and those who have not consumed alcohol and thus facilitate the DWI arrest decision?

The question of primary interest, then, is whether the officers were able to make the correct arrest-or-release decisions. The data show that the officers did correctly decide to arrest 84 percent of the persons who had BACs higher than 0.10 percent and to release 73 percent of those who had BACs lower than 0.10 percent. It is important to note that, although the officers were instructed to score test performance on the scale of 1-10 error points, no particular score or cutoff was designated as a criterion. The arrest decisions were based solely on the examiners' own interpretations and judgments of test scores and performance. Overall, the arrest-or-release decisions were correct for 76 percent of the participants, based on an arrest criterion of 0.10 percent BAC.

The officers indicated that they would have arrested 47 participants whose BACs were lower than 0.10 percent. It is believed that, since police officers are usually reluctant to make false arrests, this high "false-alarm" rate was an artifact of the laboratory setting, in which the officers were less conservative about making arrests than they would be in the field.

Further analysis of the scoring records revealed a related and particularly interesting phenomenon. It appears that the officers made arrest decisions "as though" 0.08 percent, and not 0.10 percent, were the legal BAC limit. It is not suggested that the officers consciously shifted the arrest criterion to the lower BAC. Rather, we believe that they consistently observed impairment at 0.08 percent and in their decision-making process simply equated observed impairment with arrest. Obviously, if the BAC criterion is 0.10 percent, this results in a high rate of false arrests. In states that have a 0.08 percent criterion, such as Utah, or in Europe, these cases would not be false arrests.

## SELECTION OF FINAL TEST BATTERY

The police officer in the field is confronted with the individual driver and must make a decision to arrest or release. If the officer arrests, he or she may later be required to present as evidence in court proceedings a report of the driver behavior that led to the decision to arrest. The battery of sobriety tests has value for the officer only if it (a) facilitates the immediate decision to arrest or release and (b) enables him or her to give credible and convincing testimony in court.

The laboratory evaluation described here demonstrated that the six chosen tests were useful to the police officers in determining the individual's level of intoxication. The officers' scoring of the tests correlated with BAC as follows:

| Test | Correlation with BAC |
|---|---|
| Nystagmus | 0.67 |
| Finger to nose | 0.47 |
| Finger count | 0.31 |
| One-leg stand | 0.48 |

| Test | Correlation with BAC |
|---|---|
| Tracing | 0.44 |
| Walk and turn | 0.55 |

However, since administration and scoring of a six-test battery can require as much as 15 min, a longer time than that usually allotted for a roadside examination, it was necessary to select some subset of the tests as a final test battery that would meet real-world time constraints. This was done by using stepwise discriminant analysis. The discriminant model derives linear functions of the test-battery scores and makes the maximum separation of the BAC groups.

The analysis was carried out by using the BMDP7M analysis program from BioMedical Computer Programs (4). This program computes the set of linear classification functions by choosing variables in a stepwise manner. At each step the variable with the highest standard F-statistic is chosen. By using specified prior probabilities and pooled within-group variances, group classification functions are obtained.

The discriminant analysis led to the selection of three tests: AGN, one-leg stand, and walk and turn. The total score derived from these three measures appears to be the best predictor of BAC. By using the officers' scores for these three tests alone, more than 83 percent of the participants in the laboratory evaluation were correctly classified. Note that, although the officers had administered six tests and had made their own observations of general behavior, appearance, and speech, their decisions were correct for only 76 percent of the participants. This indicates that using the criterion scores from the analysis would greatly improve the rate of correct arrest-or-release decisions.

It is of considerable interest to examine the cases that were incorrectly classified by the discriminant analysis. Eight misclassifications involved participants whose BACs were lower than 0.08 percent but whose test scores placed them in the >0.10 percent group. Six of these individuals were light drinkers, and it can be reasonably assumed that the misclassification demonstrates their lack of tolerance for alcohol. On the other hand, 10 people who had BACs greater than 0.12 percent were classified in the <0.10 percent group. As might be expected, all were heavy drinkers whose drinking practices appear to have led to the development of a chronic, partial tolerance to some of the impairing effects of alcohol. It appears that the tests are measuring true impairment, which varies for a given BAC as a result of the tolerance or lack of tolerance to alcohol produced by the individual's drinking practices.

The discriminant analysis confirms that a few individuals may perform in a manner that is not congruent with their BAC level but is usually explainable in terms of a tolerance effect. Long-term, heavy drinkers are likely to be less behaviorally impaired than the individual who is a light to moderate consumer of alcohol. An infrequent or inexperienced drinker may show impairment at relatively low levels of alcohol. These persons can present a problem for sobriety testing, as do drivers whose BACs fall within a narrow ±0.02 percent range around 0.10 percent. It is hoped, however, that the experienced officer has developed observation skills that enable him or her to use such factors as age, appearance, and locale as additional clues to the suspect's drinking habits. The three-test battery, together with the information derived from observation and interrogation, will serve to minimize the number of incorrect arrest-or-release decisions made at the roadside.

## CONCLUSIONS

The laboratory evaluation of behavioral tests of alcohol impairment led to the selection and recommendation of three tests for a sobriety test battery: the one-leg stand, walk and turn, and AGN. The data indicate that, if balance and walking are examined and the eyes are checked for the jerking nystagmus movement, and if these tests are administered and the results observed and evaluated by a police officer trained for the task, the officer will have as much information as can be obtained routinely and quickly from behavioral tests at the roadside.

These conclusions are in close agreement with those of the only other known study of similar scope and methodology (5, 6). Based on the examination records of 495 Finnish drivers, Pentillä and others concluded that an optimal test battery should include measures of walking, balance, and nystagmus. Their longer battery also included counting and picking up small objects.

Note that both groups of investigators found nystagmus to be a particularly sensitive and valuable measure. Further study of AGN is in progress at Southern California Research Institute.

## REFERENCES

1. G. A. Beitel, M. D. Sharp, and W. D. Glauz. Probability of Arrest While Driving Under the Influence of Alcohol. Journal Studies on Alcohol, Vol. 36, No. 1, 1975.
2. G. Aschan. Different Types of Alcohol Nystagmus. Acta Oto-Laryngologica, Supplement 140, 1958, pp. 69-78.
3. D. Cahalan, I. H. Cisin, and H. Crossley. American Drinking Practices. College and University Press, New Haven, CT, 1969.
4. W. J. Dixon, ed. BioMedical Computer Programs. Univ. of California Press, Berkeley, 1973.
5. A. Pentillä, M. Tenhu, and M. Kataja. Clinical Examination for Intoxication in Cases of Suspected Drunken Driving. Statistical and Research Bureau, TALJA, Helsinki, Finland, 1971.
6. A. Pentillä, M. Tenhu, and M. Kataja. Examination of Alcohol Intoxication in Cases of Suspected Drunken Drivers—II. Statistical and Research Bureau, TALJA, Helsinki, Finland, 1971.

*Publication of this paper sponsored by Committee on Vehicle User Characteristics.*