MEASUREMENT OF BLOOD ALCOHOL CONCENTRATION IN MICHIGAN DRIVERS:
A Review of the Literature

Paul L. Olson

Prepared for
The Office of Highway Safety Planning
State of Michigan
Highway Safety Project MTR-85-033A

December 1985

UMTRI The University of Michigan Transportation Research Institute
A literature review was conducted to assess the value of two means of screening DUIL suspects, i.e., a new field sobriety screening battery incorporating gaze nystagmus, and portable breath testers (PBT’s) employing fuel-cell technology. The hope is that the wider use of these techniques would improve detection of DUIL offenders, especially at lower levels of BAC, and reduce the incidence of false positive identifications.

The evaluations of the field sobriety battery indicate that it is very effective in reducing false positives. How well it improves the identification of DUIL offenders, particularly those at BAC’s close to 0.10%, is less clear. The investigations reviewed report results ranging from no improvement to very great improvements. Significant methodological problems with some of the studies make it difficult to draw firm conclusions from the data.

The evaluations of PBT’s also suggest that improvements can be realized on the two criteria mentioned above. However, only one study has examined the use of PBT’s for evidential purposes. That study was small in scope and employed only one person to do the testing.

While the bulk of available evidence indicates that both of the BAC screening techniques reviewed have merit, further work is indicated to try to reduce some of the problems and limitations noted.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>The Michigan Field Sobriety Test (MFST)</td>
<td>5</td>
</tr>
<tr>
<td>Introduction</td>
<td>5</td>
</tr>
<tr>
<td>Development</td>
<td>6</td>
</tr>
<tr>
<td>Evaluations Reported in the Scientific Literature</td>
<td>7</td>
</tr>
<tr>
<td>Evaluations Not Reported in the Scientific Literature</td>
<td>13</td>
</tr>
<tr>
<td>Conclusions</td>
<td>15</td>
</tr>
<tr>
<td>Portable Breath Testers</td>
<td>16</td>
</tr>
<tr>
<td>Introduction</td>
<td>16</td>
</tr>
<tr>
<td>Evaluation — General</td>
<td>17</td>
</tr>
<tr>
<td>Evaluation — Fuel Cell Instruments</td>
<td>21</td>
</tr>
<tr>
<td>Conclusions</td>
<td>25</td>
</tr>
</tbody>
</table>

REFERENCES ................................................. 26
LIST OF TABLES

1. Percent in Each BAC Category for Drivers Arrested by Various Procedures .................................................. 9

2. Percent of Subjects Who Would Have Been Detained Based on Typical Field Sobriety Tests Compared with Nystagmus Test ................................................. 10

3. Percent of Subjects Who Would Have Been Detained Based on Nystagmus Test as a Function of Level of Training ......................................................... 10

4. Percent of Subjects Who Would Have Been Detained Based on Standard Tests as Compared to Experimental Field Sobriety Test Battery ............................. 11

5. Officers' Estimate Compared With Actual BAC .................................................. 12

6. Percent of Subjects in Each Dose Category Who Were Classified as "Should be Detained" as a Function of Police Agency .................................................. 12

7. Comparison of Results of Mesa, Arizona, Field Test of Alco-Sensor III .................. 23
LIST OF FIGURES

1. Relationship between BAC and the probability of accident involvement ... 2

2. BAC distributions of two groups: Roadside-survey drivers and arrested drivers. ................................................ 3

3. Correlation between results with Alcotest and corresponding blood-alcohol levels. .............................................. 19
**Introduction**

Driving by persons under the influence of alcohol remains one of the most significant problems in motor vehicle safety. The use of alcohol is linked with about half of all motor vehicle deaths, hundreds of thousands of injuries, and billions of dollars in annual costs.

Efforts to reduce the traffic toll attributed to alcohol have taken two general directions. One is education, i.e., trying in a variety of ways to bring home to the motoring public the dangers of driving after drinking. The other approach is enforcement, i.e., arrest, punishment, and/or rehabilitation of persons who drive while under the influence of alcohol.

Theoretically, it would be possible to have an alcohol enforcement system that relied on direct observation of and court testimony concerning the driving capabilities of the suspect. In a way this would be the fairest approach, since the effects of alcohol vary from person to person and it is the effect of alcohol on driving performance that is the ultimate concern. However, it is not a very practical system, introducing a great deal of subjectivity into the judicial process, and probably making it much more difficult to obtain convictions.

Because of this problem, the usual method has been to set a maximum blood alcohol concentration (BAC) at and above which the driver is assumed to be intoxicated. The limits are established based on various factors, one of which is research relating BAC to the probability of accident involvement. An example (from Borkenstein, 1964) is given in Figure 1. The figure indicates that the probability of being involved in an accident begins to increase at a BAC of about 0.05%, and is about three times higher than when sober at 0.08% and about six times higher at 0.10%. Most states set the maximum BAC at either 0.08% or 0.10%. Michigan uses 0.10%.

Despite many years of effort, very little has been achieved in terms of reducing the incidence of drunk driving. Among the possible reasons for this is that the likelihood of being arrested while driving under the influence of alcohol is very low. Some indication of how low has been provided by Beitel, Sharp, and Glauz (1975), who compared survey data on BAC levels with arrest rates in a small area. A summary of their data is provided in Figure 2. The investigators found that the probability of arrest was about 1 in 200 for a BAC equal to or greater than 0.10%, and about 1 in 50 for a BAC in the range from 0.20% to 0.24%. These rates are disappointingly low, but are probably higher than normal, since the area was being patrolled by specially trained Alcohol Safety Action Project (ASAP) teams, who were looking for alcohol offenders in particular.
Figure 1. Relationship between BAC and the probability of accident involvement. (From: Borkenstein, 1964.)
Figure 2. BAC distributions of two groups: Roadside-survey drivers and arrested drivers. (From: Beitel et al., 1975.)
Three things must happen before a drunk driving arrest can occur. First, the car must pass within the field of view of a police officer. Second, during the time it can be observed, there must be something about its operation that will rouse the suspicions of the officer. Third, with the suspect driver stopped, the officer must make a decision, normally based on so-called “field sobriety tests,” whether the individual is likely enough under the influence to make an arrest.

Determining whether a vehicle operator is under the influence of alcohol, based on a brief observation of driving behavior, is difficult. A systematic study of this problem has been reported by Harris et al. (1980), who estimated the probability of the BAC’s being in excess of 0.10% for several observable behaviors (e.g., speeding, erratic lane changing, running a stop sign).

Based on this work, a “Drunk Driver Detection Guide” was prepared and evaluated at 10 law enforcement agencies throughout the United States. Officers were trained in use of the guide and their DUIL arrest rates for a three-month test period were compared with average arrest rates for the previous twelve months. A statistically significant increase of 12% was noted during the test period. However, a quarter-by-quarter comparison of arrest rates in the year preceding the test period shows a steady increase, so the use of an average as the basis for comparison is questionable.

When confronted with a driver who may have been drinking, the officer must decide if an arrest is warranted. Again, the task is not easy. If the BAC is not much in excess of the legal maximum, or even if it is and the suspect is an experienced drinker, behavioral signs such as unsteadiness and slurred speech may not be of such magnitude as to be obvious. To aid in the arrest decision, research has been conducted to develop an improved battery of field sobriety tests. In addition, portable breath-testing devices have become available that make possible accurate, objective determination of BAC. The primary purpose of this report is to review the research on both of these topics as a necessary prelude to research that will be conducted in the State of Michigan.
The Michigan Field Sobriety Test (MFST)

Introduction

The MFST consists of three tests that have been selected to meet certain criteria, i.e.:

a. They are capable of discriminating between persons who are sober and those who are not.

b. They can be scored objectively.

c. They are easy to administer under field conditions.

The three tests are called: Horizontal Gaze Nystagmus, One-Leg Stand, and Walk and Turn. A brief description of each follows.

Horizontal Gaze Nystagmus. This test is a relative newcomer to field sobriety testing. With head held steady, the suspect is asked to fixate an object (e.g., a pencil) as it is moved around to the side of his/her head. In a normal individual the eyes will often begin an involuntary, rapid, back-and-forth movement as they approach maximum excursion. The rapid motion is called nystagmus. Nystagmus occurs under a number of conditions in humans. Because of the way this test is conducted the result is called horizontal gaze nystagmus. The interesting thing about it is that the angle at which nystagmus begins in this test is affected by alcohol, and it is possible for an officer to estimate BAC by noting the angle at which nystagmus is first observed.

One-Leg Stand. This is a simple test, variations of which have been used in field sobriety testing for years. The suspect is asked to stand on one foot, with the other leg held stiffly out in front, eyes on the extended foot, for 30 seconds.

Walk and Turn. This is another test that has been commonly used, in various forms, for many years. The suspect is required to take nine heel-to-toe steps along a line, turn and return to the start point in the same way.

The last two tests have a “divided attention” feature. That is, the suspect assumes the start position, which requires maintaining his/her balance, and then must listen to the instructions. An inebriated individual may be able to maintain balance rather well, but only by giving a great deal of conscious attention to the task. If their conscious attention is required to listen to the instructions, they likely will either experience difficulty in maintaining balance, or fail to comprehend the instructions.
Investigators have long known that alcohol affects the ability to process visually acquired information. Much research has been directed toward understanding why it does so (e.g., Bernstein et al., 1965; Flom et al., 1976; Goldberg, 1963; Moskowitz, Ziedman, and Sharma, 1976; Schroeder et al., 1972; Ziedman, Moskowitz, and Nieman, 1980). Most of these investigators used direct measures of information processing. Some noted the effect of alcohol on eye movements, but assumed it was a possible cause of the reduced visual capabilities they were concerned with.

It was realized some time ago that alcohol affected the nystagmus phenomenon. For example, Aschan (1958) refers to studies relating various forms of nystagmus to BAC that go back as far as 1934. Wilkinson, Kime, and Purnell (1974) reported consistent changes in horizontal gaze nystagmus with alcohol dose. Lehti (1976) may have been the first to actually calculate a correlation (−0.76) between BAC and the onset of nystagmus. This work apparently led to the idea of using gaze nystagmus as an objective measure of BAC.

Development

Recognizing the difficulties of assessing BAC in the field, the National Highway Traffic Safety Administration (NHTSA) sponsored research to try to develop improved measures. As a first step a contract was given to the Southern California Research Institute in 1975. The objectives were: (1) to conduct an evaluation of tests in current use, (2) develop more sensitive and reliable measures, and (3) standardize test administration procedures. The final report was issued in June of 1977 (Burns and Moskowitz, 1977).

In this study the authors selected ten common field sobriety tests for evaluation. Ten experienced police officers then administered these tests to a total of 238 volunteers. The subjects were classified as light, moderate or heavy drinkers. They were dosed to BAC's ranging from 0 to 0.15%, with only the heavy drinkers getting the maximum levels.

All of the field sobriety tests were found to be sensitive to BAC to some degree. An analysis was then conducted to find a limited combination of tests that gave the best discrimination. Based on this work, the authors recommended a combination of three tests that they felt gave as much information as was practical at roadside. The three tests were the One-Leg Stand, Walk-and-Turn, and Horizontal Gaze Nystagmus.

A follow-on contract was awarded to the same agency in 1978. The objectives were: (1) to standardize the administration and scoring procedures, and (2) conduct laboratory and field evaluations of the new test battery (Tharp, Burns, and Moskowitz, 1981).

The laboratory phase was conducted in much the same way as the 1977 study. Ten experienced police officers administered the test battery to a total of 297 volunteers with
BAC's ranging from 0 to 0.18%. Their primary task was to determine whether each subject's BAC was above or below 0.10%. The officers were able to correctly classify about two-thirds of the subjects whose actual BAC was above 0.10%, and about 90% of those whose BAC was less than 0.10%. Overall, the officers made 81% of these classifications correctly.

In the field study, data were collected on more than 3,000 police stops during a three-month period. The following results were reported:

a. The percentage of persons stopped who were subsequently arrested increased from 6.3 to 7.6, comparing pre- and post-training rates. This is a 20% increase. However, small changes in either of these figures would make a large change in the percent difference.

b. The percentage of persons stopped who had a BAC of more than 0.10% and were subsequently arrested increased from 61.9 to 69.2. Overall, 93.6% of the officers' decisions were correct before training on the new battery, and 94.1% were correct after training. Less than 2% of arrests were of persons subsequently found to have a BAC under 0.10% (false positive).

c. The mean BAC's of persons arrested did not change. However, there were some problems in the study that may have biased the data. In particular, there were a large number of refusals to take chemical tests.

d. There was some evidence that the officers were able to make more accurate estimates of BAC after training.

In sum, the results of the study were encouraging, and indicated that the use of the three-test battery would improve discrimination of drivers who were over the legal limit.

The studies that have just been described were concerned with the development and initial evaluation of the tests that are used in the MFST. The next section will discuss other research concerned with the value of these tests.

**Evaluations Reported in the Scientific Literature**

The development of the new field sobriety test battery by the California group generated a great deal of interest. The tests have been made part of the recommended procedures of a number of jurisdictions (e.g., Augsburger and Good, undated; Studdard, 1984). Several further studies have been reported. One of the first of these was by Anderson, Schweitz, and Snyder (1983). Anderson et al. set up a field study using police in four jurisdictions (District of Columbia, Arlington, and the Maryland and North Carolina
The officers were given a one-day training session and asked to supply information on stops and arrests for alcohol over the next three-month period. Their principal results are given in Table 1, which compares arrest records using PBT's (type[s] unspecified), with the new field sobriety test battery and "normal procedures." The data presented in Table 1 indicate that the ability to discriminate drunk from sober drivers using the field sobriety battery is nearly as good as with PBT's, and about 50% better than with normal procedures. Moreover, the number of arrests of persons having BAC's equal to or less than 0.04% was much lower with the new battery than with conventional procedures.

The results reported by Anderson et al. suggest that the new field sobriety battery is a far better discriminator than "standard procedures," and about as good as a portable breath analysis instrument. However, as the authors acknowledge, the study design had some defects. One is that the officers were not randomly assigned to different groups, and differences in outcomes may be attributable to selection and assignment bias. The research report contains no description of the "standard procedures" that were used. Thus it is not possible to determine whether the same results could be expected elsewhere.

The evaluation data developed by Moskowitz and his coworkers indicate that gaze nystagmus is the most sensitive test in the battery. Compton (1984) carried out a controlled field study to measure how well police officers could judge BAC, based solely on the use of gaze nystagmus. In this study volunteer subjects were dosed to BAC's ranging from 0 to about 0.15% and drove their own cars down a closed street to a checkpoint, where they were stopped and evaluated by police officers. The subjects remained in their cars, and the test was administered through the open driver's window.

Table 2 compares results obtained with the nystagmus test and a "typical" checkpoint procedure. It is not clear what this "typical" procedure was. Apparently it involved nothing more than trying to determine whether there was alcohol on the driver's breath, and noting how he/she responded to one or two questions. While the difference between the procedures was not very great for drivers above 0.10%, there was a substantial difference in the incidence of false positives.

Two levels of training were used for the participating officers. Those that were "fully trained" had been given 16 hours of instruction and had some field experience in the use of gaze nystagmus. Other officers were given three hours of training and had no field experience prior to the test. The performance of the two groups is compared in Table 3. The results indicate that the essential elements of the procedure are quickly acquired.
<table>
<thead>
<tr>
<th>Procedure</th>
<th>False Positive 0 - .04%</th>
<th>Difficult to Assess Depends on Other Data .05 - .09%</th>
<th>Arrest Supported by BAC Data .10%+</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal procedure using PBT (D.C. control)</td>
<td>0</td>
<td>10</td>
<td>89</td>
<td>164</td>
</tr>
<tr>
<td>Sobriety test battery and PBT (D.C., MD &amp; Arlington)</td>
<td>2</td>
<td>8</td>
<td>90</td>
<td>581</td>
</tr>
<tr>
<td>Sobriety test battery, No PBT (NC); Arrest indicated by 2 test combined decision rule only</td>
<td>4</td>
<td>11</td>
<td>86</td>
<td>279</td>
</tr>
<tr>
<td>Sobriety test battery No PBT (NC) Officer arrest only</td>
<td>4</td>
<td>12</td>
<td>83</td>
<td>289</td>
</tr>
<tr>
<td>Normal procedures, No PBT (NC)</td>
<td>26</td>
<td>15</td>
<td>59</td>
<td>309</td>
</tr>
</tbody>
</table>

Some rows do not add to 100 due to rounding
From: Anderson et al., 1983
TABLE 2

PERCENT OF SUBJECTS WHO WOULD HAVE BEEN DETAINED BASED ON TYPICAL FIELD SOBRIETY TESTS COMPARED WITH NYSTAGMUS TEST

<table>
<thead>
<tr>
<th>BAC</th>
<th>Typical Test</th>
<th>Nystagmus Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>47</td>
<td>15</td>
</tr>
<tr>
<td>0.05-0.09</td>
<td>87</td>
<td>64</td>
</tr>
<tr>
<td>0.10-0.15</td>
<td>87</td>
<td>95</td>
</tr>
</tbody>
</table>

From: Compton, 1984

TABLE 3

PERCENT OF SUBJECTS WHO WOULD HAVE BEEN DETAINED BASED ON NYSTAGMUS TEST AS A FUNCTION OF LEVEL OF TRAINING

<table>
<thead>
<tr>
<th>BAC</th>
<th>Fully Trained Officers</th>
<th>Briefly Trained Officers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>0.05-0.09</td>
<td>60</td>
<td>68</td>
</tr>
<tr>
<td>0.10-0.15</td>
<td>100</td>
<td>89</td>
</tr>
</tbody>
</table>

From: Compton, 1984

Compton (1985a) has also carried out a more extensive study in which gaze nystagmus was combined with three other tests (driver behavior during approach and stopping, driver appearance, and a divided attention task). All of these tests could be administered without the driver having to leave the car. Hence they were suitable for checkpoint use. These tests were compared with a "typical" procedure in which the officer observed the subject and engaged him/her in brief conversation (to smell the breath). It is not clear from the report how the results from the various tests were weighted to arrive at a decision.
In this study volunteer subjects were dosed to various BAC's and drove their own cars down a closed street through three police checkpoints manned by officers from different agencies. Over half of the subjects were sober; the rest were dosed to levels up to 0.15%.

The principal findings of the study are presented in Table 4. The data here are very similar to those presented in Table 2. It appears that the main benefit of the new test battery is in reducing the number of false positive identifications.

**TABLE 4**

<table>
<thead>
<tr>
<th>BAC</th>
<th>Standard Tests</th>
<th>New Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>47</td>
<td>16</td>
</tr>
<tr>
<td>0.05-0.09</td>
<td>87</td>
<td>61</td>
</tr>
<tr>
<td>0.10-0.15</td>
<td>87</td>
<td>95</td>
</tr>
</tbody>
</table>

From: Compton, 1985a.

Table 5 compares the actual and estimated BAC's using the combined test battery. At the highest and lowest levels of the subjects' BAC the classifications were more accurate than in the middle range (0.05% to 0.09%). The correlation derived from these data was -0.64. This compares favorably with the -0.76 reported by Lehti (1976) referred to earlier.

Table 6 compares the performance between the various agencies involved in the test. The major difference is in the percentage of false positives reported by Agency 2. The investigator could offer no explanation for this discrepancy.

The studies reported by Compton are interesting and indicate improved ability to discriminate between drunk and sober drivers using gaze nystagmus and other tests. The problem with laboratory-type studies like this lies in the possibility that the participating officers, knowing that a high percentage of approaching drivers will have been drinking, and that their performance is being monitored and compared with actual BAC doses, may
TABLE 5
OFFICERS' ESTIMATE COMPARED WITH ACTUAL BAC

<table>
<thead>
<tr>
<th>BAC</th>
<th>Officers' Estimate of BAC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-0.04</td>
</tr>
<tr>
<td>0-0.04</td>
<td>69%</td>
</tr>
<tr>
<td>0.05-0.09</td>
<td>20%</td>
</tr>
<tr>
<td>0.10-0.15</td>
<td>5%</td>
</tr>
</tbody>
</table>

From: Compton, 1985a.

TABLE 6
PERCENT OF SUBJECTS IN EACH DOSE CATEGORY WHO WERE CLASSIFIED AS "SHOULD BE DETAINED" AS A FUNCTION OF POLICE AGENCY

| BAC       | Agency
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0-0.04</td>
<td>18</td>
</tr>
<tr>
<td>0.05-0.09</td>
<td>77</td>
</tr>
<tr>
<td>0.10-0.15</td>
<td>91</td>
</tr>
</tbody>
</table>

From: Compton, 1985a.

behave differently than they would under normal circumstances. Hence, a more realistic evaluation is desirable.

Pierce (1984) has reported an evaluation of the MFST under actual operating conditions, using the Detroit Freeway Patrol of the Michigan State Police. Officers were trained in the use of the MFST and instructed to use it, together with any other tests they wished, in identifying drivers under the influence of alcohol and drugs. Comparisons were made between arrests during the test period and the previous two years, both for the post under investigation and other posts in the same district.
The findings were that DU11 arrests increased by 44% at the study post, compared with a district-wide average increase of 10%. However, there was a great deal of variability in arrest rates among the nine posts in the district. Three other posts showed increases as well, and two of them had larger increases than the study post. Thus, although the results are encouraging, there can be no certainty that the reported differences are real.

During the test period arrests for driving under the influence of drugs (DUID) increased from 3 to 9. While this represents a 300% improvement, the numbers are too small to be meaningful. It was also noted that the mean BAC of persons arrested went from 0.18% in the two years prior to the test to 0.16% during the test period. The statewide average was 0.16% in the year preceding the test.

It was also reported that, based on subsequent breath testing, 98.4% of the suspects were correctly classified as having a BAC of 0.10% or more. However, given the mean BAC of 0.16%, it is possible that relatively few suspects were near 0.10%. More meaningful perhaps is a comparison of estimated and actual BAC's. Pierce reports that 34.4% of suspects were estimated to have a BAC the same as subsequently measured, 73% were within 0.02%, and 86.5% were within 0.03%. The 34.4% figure is somewhat misleading, because the upper limit for estimates was 0.20%, and all of these were considered as "exact" if the subsequent breath test was 0.20% or greater. Since the mean BAC was 0.16%, it is probable that a significant number of the suspects tested 0.20% or higher. Even so, the results indicate good performance.

Evaluations Not Reported in the Scientific Literature

Through personal contact with officials in several states other evaluations of the field sobriety test battery incorporating gaze nystagmus have been uncovered. Written reports of the work to be described are not available, although some were referred to as "in preparation." In part due to the lack of written documentation, it is difficult to assess the scientific adequacy of these studies.

Ohio. Ohio State Police have been using the tests incorporated into the MFST for about 1.5 years. They are training eight classes of 30 officers each year in use of the battery. When the officers return to the field after training they are encouraged to use the new tests. However, they are allowed to continue to use any other tests that they feel comfortable with in addition to the new battery.
No evaluations have been run of the accuracy with which officers can estimate BAC, or of changes in arrest rates. However, a comparison has been made of changes in mean BAC's of arrested individuals prior to and after training. The mean BAC's of persons arrested by the State Police in Ohio has been about 0.16%. After training the mean dropped to about 0.14%.

The Department has enjoyed success with the new battery as a basis for probable cause in the Ohio courts. Officers have been allowed to testify about gaze nystagmus. The judges generally do not require that an expert testify. There is a recent District Court of Appeals decision that affirms gaze nystagmus as a legitimate means of screening for probable cause. Some of the success that the Department has enjoyed in court may be attributable to seminars on the new field sobriety battery that they have conducted for judges and prosecutors.

**Illinois.** The Illinois State Police have also been using the MFST test battery for about 1.5 years. From the very start they decided to conduct an evaluation of the effectiveness of the tests. A number of officers who had been on desk duty for some time were selected to participate in the study. They were divided into two groups of 13 each. One group received the training, the other did not. Their DUlL arrests were monitored for the next 12 months. During that period the trained officers made 92 arrests where a subsequent breath test was taken. The average BAC was 0.155%. In the meantime the control group of officers (reduced to nine due to transfers and promotions) made 51 arrests where a subsequent breath test was taken. The average BAC was 0.192%. As it happened, the state-wide average BAC in Illinois at that time was about 0.19%. It is also reported that, as other groups of officers have been trained, their average BAC's have dropped to about 0.15%.

Each officer in Illinois maintains a log in which his/her estimate of the suspect's BAC is entered, along with the results of the subsequent Breathalyzer test. These logs have proven useful in court because, if the officer's testimony about the driver's BAC is challenged, the log can be produced to show how accurate his/her estimates have been on previous occasions.

Illinois currently has one case in which testimony has been allowed concerning gaze nystagmus. They also have a recent supreme court ruling allowing the use of road blocks as a drunk-driving countermeasure.

**Maryland.** The Maryland State Police had a contract with the National Highway Traffic Safety Administration to train a nationwide cadre of officers in the use of the new
field sobriety battery. They have also trained their own officers. They are in the process of conducting an evaluation of its effectiveness at the present time.

In the Maryland evaluation each commander in the state was asked to submit a list of ten officers selected on any basis, but not necessarily on DUIL arrests. Five of these were selected from each list, for a total of 126. These officers were put through the field sobriety training program, and 94 were finally certified. The performance of this group of officers on DUIL arrests was compared with other state troopers in 1984. It was found that the 94 trained officers made an average of 31.5 arrests each during that period, while the average for the other officers was 20.2 each. The trained troopers represented 13.7% of the total force, and made 19.9% of the DUIL arrests. The data analysis of this study is not complete at the present time. However, the investigators believe that the trained troopers are making more arrests of suspects at the 0.10%–0.15% levels.

Maryland reports good results in the court system. Judges have allowed officers to testify about gaze nystagmus. They have a recent appellate court decision affirming the use of gaze nystagmus as well.

The work carried out by these states is very interesting and continues to point toward improved DUIL arrest performance with the MFST battery. The change in average BAC's in Ohio compares well with that reported by Pierce in Michigan. However, the change in average BAC's in Illinois is almost too good to believe. It may be that other factors, such as motivation, are biasing the results.

**Conclusions**

The MFST is based on research sponsored by the National Highway Traffic Safety Administration, and is intended to improve the identification of persons operating a motor vehicle under the influence of alcohol or drugs. The available research on the topic indicates that the test battery outperforms other field sobriety test procedures by a significant margin.

Including the work done by the original developers of the MFST battery, six evaluations of it have been reported in the technical literature. In addition, we have uncovered three other studies, not in the technical literature. Three of these have been "laboratory" tests in the sense that all aspects were under close control by the experimenter and the "suspects" were paid subjects deliberately dosed to specific BAC's. Such studies can, and in this case do, show the potential value of the procedure. However, as has been pointed out, the motivational aspects of such a test may alter the attitudes of
the participating officers and yield a result that bears little resemblance to what would happen in the real world.

Six of the studies reviewed in this paper, including the three evaluations conducted by state agencies, were field evaluations, intended to overcome the shortcomings just described. It is very difficult to design and carry out a study with sufficient controls to yield reliable results under such conditions. In reading the published reports, it is apparent that the investigator's experienced many problems, or were forced to make compromises that may have affected the results. Thus, while the data consistently indicate that the MFST is an improvement over present procedures, the degree to which its use would actually increase the likelihood of arrests of individuals under the influence, particularly those at lower BAC's, is not clear. A more extensive testing program, with better control over the relevant variables, is indicated.

**Portable Breath Testers (PBT's)**

**Introduction**

If alcohol concentration in the blood is the basis for determining impairment, then it has to be possible to measure it with reasonable accuracy. BAC can be measured in a blood or urine sample. However, drawing blood is offensive to some people, and requires a qualified technician. A urine sample may or may not be available when needed. At best these tests require some laboratory equipment, persons who know how to operate it, and take time. These are all significant disadvantages in DUIL enforcement.

One very good possibility is to measure the alcohol concentration in the suspect's breath. As Goldberg and Havard (1968) have pointed out, alcohol passes from the arterial blood to the air in the lungs. It has been shown experimentally that the alcohol contained in 1 millimeter of blood is in equilibrium with 2.1 liters of alveolar air at 98.4 degrees F. This relationship makes it practical to measure BAC indirectly, but with accuracy, using the suspect's breath. The challenge is to find a way or ways of doing it that is accurate, easy to use, and reasonable in cost.

Two basic types of breath-testing device have developed. One is designed primarily for preliminary screening by officers in the field, the other for use in a central location. In recent years some of the devices have been used in both settings. Schmidt (1980) has listed and described the various types of breath-testing devices. These are as follows:

**Test tubes.** This equipment is designed solely for quick, field evaluation. It relies on the change in color of a chemical after a certain volume of breath is exhaled through the tube.
**Oxidation-photometric.** The Borkenstein Breathalyzer relies on this method. Alcohol in the exhaled breath passes through a chemical solution, changing its transmissivity. This change is then measured with a light source and photocell.

**Electrochemical cell.** The Alco-Sensor III, under consideration for use by the State of Michigan, uses this principle. Alcohol in exhaled breath is oxidized in a fuel cell by allowing an electric current to flow from the anode to the cathode. The current-time product is proportional to the amount of alcohol being oxidized.

**Semiconductor sensors.** Metal oxide semiconductors are used in these devices. They show a significant increase in conductivity in the presence of reducing gases.

**Gas-chromatograph.** A breath sample is introduced into a stream of carrier gas and fed into a heated column filled with material that has different retention times for the different constituents of the gas sample. The gas components are analyzed according to the time needed to move through the column.

**Infrared absorption.** The breath sample is introduced into an optical cell and the transmission of an infrared beam is measured. Changes in transmission are proportional to the concentration of alcohol in the sample.

The technical literature contains a number of studies in which available devices have been subjected to evaluation. Many of these have been conducted in laboratories, under carefully controlled conditions. However, some have been field evaluations, attempting to approximate actual conditions of use. The next section of this report will review a number of these studies. Initially, the review will consider evaluations of portable breath testers (PBT's) other than fuel-cell devices like the Alco-Sensor III, because an understanding of the methodology employed can be instructive. Finally, studies of fuel-cell instruments will be covered, including the Alco-Sensor III.

**Evaluation – General**

One of the earliest field evaluations of breath-testing equipment was carried out in Sweden and reported by Bjerver, Andreasson, and Bonnichsen (1965). They were working with a test tube device called “Alcotest.” Alcotest does allow for some gradations in estimated BAC, based on the distance up the tube that a color change can be noted. The investigators sent questionnaires out to a number of police stations, asking that they be filled in and returned, along with the legal blood or urine sample, each time the Alcotest was used. The latter were analyzed according to the usual procedures and the results compared with the BAC estimated from the Alcotest. A total of 1,148 cases were reviewed.
The results are shown in Figure 3. The vertical line at 0.5 (corresponding to 0.05%) represents the legal limit of interest, and the horizontal line its estimate based on the Alcotest. While no correlation is reported, it appears low, and it is apparent that a substantial fraction of the cases fell in the upper left and lower right quadrants of the figure (false positives and false negatives, respectively).

A device called the ALERT (Alcohol Level Evaluation Roadside Tester) was given an extensive evaluation in Canada (Picton and Bowthorpe, 1977). A total of 34 of the devices were distributed to the 10 RCMP locations in Alberta that had arrested the greatest number of impaired drivers during 1975. During the test period a total of 1,722 drivers were evaluated. The results show an increase of 75% in alcohol arrests and a decline in mean BAC from 0.160% to 0.153% at the test locations. This compares with an increase in arrests of 24% and a decline in mean BAC from 0.175% to 0.167% at the control locations. Thus, the ALERT seemed not to affect average BAC, but may have been a factor in increasing DUIL arrests.

Studies of this type have a basic problem in that it is not possible to have a clean experimental and control group arrangement. One group will get the new equipment, the other will not. The equipment itself may have a positive or negative effect on the members of the experimental group. Some attention should be given this problem by the experimenter to minimize its possible impact on the outcome. Attitude surveys and special training are possibilities. In this case the investigators may have exacerbated the problem by deliberately giving the ALERT's to the police units with the best alcohol-arrest records. This is unfortunate, because it is possible that selecting these officers as the experimental group motivated them to try even harder.

Emerson et al. (1980) were concerned with the evaluation of instruments that could be used at central locations. What is apparently the same study has been reported by Isaacs et al. (1980) and Birch (1980). The investigators purchased four units each of three instruments:

1. CMI Intoxilyzer 4011A
2. Intoximeters Inc. Gas Chromatograph Intoximeter MK IV (GCI)
3. Smith and Wesson Breathalyzer 1000

Twelve locations were selected and each one had one of the instruments for two months in the six-month test period. Results from the instruments were compared with BAC's as determined by chemical analysis of blood samples. All three instruments were found to perform well under the conditions of the test. Unfortunately, the authors do not
Figure 3. Correlation between results with Alcotest and corresponding blood-alcohol levels. (From: Bjerver et al., 1965.)
present correlations. An inspection of the scatter plots suggests correlations of 0.95 or better. However, the authors, particularly Isaacs et al., point out that the results in these field evaluations were not as accurate as in controlled laboratory tests run earlier. They feel that much of the difficulty may be attributable to some suspects not blowing hard enough or long enough to provide an adequate sample of alveolar air.

Three other "studies" of breath-analysis instruments were uncovered in the literature review (Grambow et al., 1980; Kitagawa, Yoneda and Nakajima, 1980; Tsukamoto, 1980). While each of these was concerned with a different device, the reports were very similar in that they:

a. Were primarily concerned with describing the virtues of a particular instrument.

b. Presented data indicating very high correlations between results obtained with their devices and some other measure of BAC (not always specified), but no information about how these tests were conducted.

Recently a passive alcohol sensor (PAS) has become available. Intended solely for quick, unobtrusive screening of suspects seated in their cars, the PAS looks like a flashlight. In use it is held in front of the suspect’s mouth while the officer engages him/her in brief conversation. A fan draws exhaled breath into the device where it is analyzed. Two evaluations have been reported. In Compton’s study (1985b) subjects were dosed to various BAC’s and drove their own cars down a closed street through police check points. The results indicate that 10% of the sober subjects gave a positive response, as did 75% of subjects in the range of 0.05% to 0.09%, and 94% of subjects in the range of 0.10% to 0.15%.

Jones and Lund’s study (1985) of the PAS was conducted in much the same way as that of Compton, except that a comparison was made between officers with and without the PAS. They report that 68% of drivers with BAC’s of 0.10% or more and 45% of drivers with BAC’s of 0.05%–0.099% were detected using the PAS, compared to 45% and 24%, respectively, using conventional methods.

The data from these two studies indicate that the PAS can greatly improve performance in detecting drivers under the influence of alcohol. However, a comparison with data on the use of horizontal gaze nystagmus and other field sobriety tests that do not require instrumentation (see Tables 2 and 4) suggests that the PAS may be no better than they are. It is also apparent that the "conventional methods" used in Jones and Lund’s
study with PAS were less effective than those used in the work reported by Compton (1984, 1985a), and summarized in Tables 2 and 4.

**Evaluation – Fuel Cell Instruments**

Breath-test devices based on fuel cell technology have become popular in recent years because they can be made very small, are potentially quite accurate, easy to use, and relatively inexpensive. There have been a number of evaluations reported.

One of the earliest of these was by Jacobs and Goodson (1973), who carried out an evaluation of the fuel cell for the Alco-Sensor. Their conclusions were as follows:

a. When calibrated to 0.10%, accuracy can be expected to be within plus or minus 0.01%, with a standard deviation of 0.005%

b. With an actual BAC of 0.052% the Alco-Sensor would read 0.043% to 0.063% 95% of the time. With an actual BAC of 0.103% it would read 0.088% to 0.107% 95% of the time. And, with an actual BAC of 0.157% it would read 0.134% to 0.153% 95% of the time. In short, 95% of the readings at various BAC’s will be within 0.02%, indicating a high degree of reliability. However, the device had a tendency to underestimate the BAC at higher levels.

c. Obtaining a positive reading on a nondrinker is “virtually impossible.” The chance of obtaining a negative reading on a person who had been drinking is 1 in 10,000.

Extensive testing of another instrument, the “Alcometer,” has been reported by Jones, Jones, and Williams (1977) and Jones and Goldberg (1978). While the test methods and numbers of subjects are not clear, it is apparent that these were laboratory tests. In both investigations the instruments performed very well, according to the investigator.

However, Jones (1976) also conducted a laboratory evaluation of an instrument called ASD (Alcohol Screening Device). He described it as having high precision but low accuracy. By this he means it produced the same BAC estimate from an identical sample on repeated occasions, but it gave positive readings when there was no alcohol present, and its readings did not correspond closely to actual BAC on many tests.

One of the most ambitious efforts to evaluate PBT’s under field conditions has been reported by Bishop, Goransson, and Oates (1977). Four devices were tested, the Alcohalt, Alcosensor, Alert, and BAT III. The Alcosensor used here appeared similar to the Alco-Sensor III being considered for use in Michigan, but it was actually a “go- or no-go” device rather than one that could give quantitative readings.
Numbers of each instrument were purchased and distributed to police agencies in six states. Each state was to have established experimental and control groups, distributed the PBT's, seen to their maintenance, and collected data to be forwarded to the investigators. However, serious problems were encountered. There was poor cooperation from a number of the states. As a consequence much data were lost and many control groups were set up late or not at all. Some of the instruments experienced operational problems, with the result that they were dropped entirely from service in some states. Because of these problems, the value of the study is somewhat limited. It was found that the number of arrests for DUIL increased 31.8% for agencies using the PBT's, and 23.1% for the control agencies. Mean BAC's were not different, however, being 0.181% and 0.179% for the experimental and control groups, respectively. The investigators feel this may be misleading, and point out that the three states that most closely followed the research plan all reported lower mean BAC's for the experimental than for the control group.

A question that often confronts a hospital emergency room staff is whether a given patient had been drinking. It is important to know, because the answer affects diagnosis and treatment. Gibb et al. (1984) decided to try the Alco-Sensor III as a means for checking when there was doubt. They classified patients as "cooperative" or "uncooperative," based on whether they were capable of blowing into the instrument enough to give a sample of alveolar air. Comparisons were made with blood alcohol tests. The correlation between Alco-Sensor readings and blood tests for cooperative patients was 0.96. For uncooperative patients it was 0.72. Even samples obtained nasally from uncooperative patients correlated moderately well with actual BAC (r = 0.69). The authors conclude that the Alco-Sensor III is sufficiently accurate to be of use in rapidly assessing BAC in this setting, even when the patient is unable to cooperate fully.

A small-scale laboratory and field evaluation of the Alco-Sensor III was carried out by the Mesa, Arizona, Police Department (Watts, 1984). In the laboratory phase five readings were taken at two-minute intervals at each of two BAC levels, using a simulator. All of the readings were within 0.001% of the actual concentration. Tests were also run after the unit was kept in a refrigerator at 4 degrees C for an hour and in an oven at 40 degrees C, also for an hour. The simulator solution was 0.10% in each case. The Alco-Sensor III read low by an average of about 0.015% when cold and was within 0.003% when warm.

In the field phase the unit was given to officers and used for screening of suspects. Those who tested positive were then brought to the station and tested again, using an
Intoxilyzer Model 4011AS. A total of 13 tests were conducted. The results are given in Table 7. The correspondence is generally good, although in two cases there is a difference of about 0.03%. Where there are differences, the Intoxilyzer always read higher, even though there was a time delay of 20 minutes or more between the two readings, with the Alco-Sensor always being used first. Watts suggests that the differences in readings that do exist may be attributable to the time lapse. However, unless all the suspects were measured during the ascending phase of the BAC, this is not likely.

**TABLE 7**

**COMPARISON OF RESULTS OF MESA, ARIZONA, FIELD TEST OF ALCO-SENSOR III**

<table>
<thead>
<tr>
<th>Alco-Sensor III</th>
<th>Intoxilyzer</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.202</td>
<td>0.220</td>
</tr>
<tr>
<td>0.258</td>
<td>0.160</td>
</tr>
<tr>
<td>0.136</td>
<td>0.140</td>
</tr>
<tr>
<td>0.140</td>
<td>0.170</td>
</tr>
<tr>
<td>0.138</td>
<td>0.150</td>
</tr>
<tr>
<td>0.104</td>
<td>0.124</td>
</tr>
<tr>
<td>0.076</td>
<td>0.079</td>
</tr>
<tr>
<td>0.104</td>
<td>0.104</td>
</tr>
<tr>
<td>0.122</td>
<td>0.124</td>
</tr>
<tr>
<td>0.127</td>
<td>0.127</td>
</tr>
<tr>
<td>0.129</td>
<td>0.160</td>
</tr>
<tr>
<td>0.080</td>
<td>0.080</td>
</tr>
<tr>
<td>0.185</td>
<td>0.190</td>
</tr>
</tbody>
</table>

From: Watts, 1984

The State of Maryland conducted an extensive field evaluation of the Alco-Sensor (Yohman and Ebare, undated) to: (1) determine whether it would increase the number of arrests at lower BAC's, (2) determine whether it would reduce the number of arrests of persons not legally intoxicated, and (3) to evaluate its acceptance and use by officers. A total of 405 units were used in the test, which ran for one year. The units were calibrated weekly at first, then monthly when that was found adequate. They were distributed to six state police barracks and used in routine field screening of suspects. Individuals found to be in excess of 0.08% were arrested and transported to the barracks for a blood or Breathalyzer test.

A total of 4,730 tests are included in the data. Results are summarized as follows:
a. The correlation between the Alco-Sensor readings and follow-up tests was 0.79.

b. It could not be established that use of the Alco-Sensor increased the number of arrests at BAC's in the range of 0.08% to 0.12%, although there was some indication that it might.

c. The number of arrests of persons below 0.08% was reduced significantly.

d. There were great differences in use levels of the device by participating officers. About 20% of the officers accounted for 70% of its use.

The Maryland study is the most ambitious and realistic reported to date. It indicates that PBT's can be expected to experience some problems in acceptance with individual officers, and they may not produce much or any change in the arrests of persons at low BAC's. Further, the correlation between readings taken using the PBT by officers in the field and later tests run in the station will likely be much lower than suggested by laboratory investigations.

McDonough (undated) has reported an evaluation of the Also-Sensor III for evidential use under both laboratory and field conditions in Idaho. The laboratory test was run to assess the accuracy of the unit at various BAC's and under various temperatures and other conditions. The unit performed well under all the tests, except for radio frequency interference (RFI), confirming work reported by the National Bureau of Standards (1983).

It is not clear from the report how the field test was carried out. A total of 170 persons were screened using the Alco-Sensor III, the Intoximeter 3000, and blood tests. Apparently, all of the tests were conducted by the author. The results indicate that 30% of the Alco-Sensor readings were within 0.001% of the readings on the Intoximeter. A correlation of 0.99 is reported.

McDonough refers to the Yohman and Ebare study, arguing that her work indicates better accuracy for the Alco-Sensor, and pointing out that the tests in the Maryland study were run by many state troopers. This is true, and the Maryland results probably underestimate the correlation that would be obtained when the Alco-Sensor III is used by highly practiced officers for evidential purposes in a central setting. However, McDonough's results, based on tests run by a single skilled individual, probably overestimate the correlation. The question is the degree to which each study is in error.
Conclusions

The research reviewed in this section suggests that PBT's, such as the Alco-Sensor III, are accurate, reliable, and easy to use. Most of the research has been concerned with the use of these devices by officers in the field. This is useful, but the hope is that it can be used for evidential purposes in the State of Michigan. Only one published study has dealt with that issue (McDonough, undated), and her work was of limited scope and used only one person to administer the tests. This leaves open the question of how well the instrument would function under more typical conditions, with a number of officers doing the screening, different kinds of time pressure, etc. Further research should be undertaken to address this issue.
REFERENCES


Augsburger, A. and Good, G. Use of alcohol gaze nystagmus as a part of roadside sobriety testing. The Ohio State University, College of Optometry. Columbus, Ohio.


Compton, R.P. Preliminary field evaluation of a passive alcohol sensor used to screen drivers at a sobriety checkpoint. Research notes. U.S. Department of


Watts, V. A comparative study on two breath screening devices, the Alcosensor III and the Alcotest. Mesa Police Department, Crime Laboratory, Mesa, Arizona, 1984.

