Evaluation of the Field Certification Process for Drug Recognition Experts (DREs)

Dr. David Shinar, Dr. Robert B. Voas, Elizabeth A. Langston and A. Scott Tippetts
National Public Services Research Institute
Landover, MD USA

The Drug Evaluation and Classification Program

In areas of the United States where drug use is wide-spread, traffic enforcement officers frequently encounter drivers under the influence of drugs. These drivers often show signs of impairment similar to those associated with alcohol, but have a low or zero Blood Alcohol Concentration (BAC). In the past, although the impairment caused by drug use may have been apparent, without training in drug recognition, police officers were often not able to make an arrest which could be successfully prosecuted. During the 1970s, the Los Angeles Police Department pioneered an effort to evaluate persons for drug use. A systematic evaluation was developed which allowed the officers to make a judgement of drug impairment and the drug category involved based on physical and behavioral signs. With the results of this evaluation, the officer could establish probable cause to justify requiring a blood or urine sample from the suspect for toxicological analysis.

Because the program was designed primarily for traffic enforcement purposes, emphasis was placed on the detection of drugs that have the potential to impair driving. Drugs were grouped into seven categories based on the similarity of their behavioral and physiological effects upon the user. The seven categories of drugs include; Central Nervous System (CNS) Depressants, CNS Stimulants, Narcotic Analgesics, Cannabis, Phencyclidine (PCP), Inhalants, and Hallucinogens. Although alcohol is technically a CNS Depressant, alcohol is considered separately from the Depressant category because Blood Alcohol Concentration (BAC) can easily be assessed using breath testing devices. In most jurisdictions, a person suspected of driving while under the influence of drugs who also has a BAC over the legal limit will be prosecuted solely on the alcohol charge, eliminating the need for a drug examination.

The Drug Evaluation and Classification (DEC) program is supported by the National Highway Traffic Safety Administration (NHTSA) and is expanding to...
police departments throughout the United States. A standardized curriculum and criteria for certification have been developed to foster national acceptance of this program.

**DEC Training Programs**

The DEC Training Program consists of 72 hours of classroom training, followed by field training in which evaluations of from 12 to 15 actual suspects are supervised by a NHTSA-Certified Instructor. After a trainee has completed the course work, passed a series of written tests, completed the required number of field evaluations and passed a proficiency test, he or she is eligible for certification as a Drug Recognition Expert (DRE).

The field training takes place immediately after the trainee has passed the final written test following the classroom portion of the training. Each trainee is required to complete a minimum of 12 evaluations of actual suspects under the supervision of an instructor. The trainees work in groups of from two to four persons to complete each evaluation. One trainee administers the test, another records the data, and the remainder act as observers. Each trainee must be the "hands-on" test administrator at least six times to be certified. At the end of the training, the trainees must also have correctly identified four of the seven drug categories.

This study, funded by the NHTSA, was conducted to evaluate the learning of the trainees during the field certification phase of the training and the accuracy with which they could evaluate drug-impaired subjects. Trainees in DRE training programs in eleven sites across the United States participated in this study. Classes consisted of from 17 to 35 trainees and data were received from 199 trainees. The classes followed the standard NHTSA curriculum but, in addition to the normal tests required for certification as a Drug Recognition Expert, the trainees in this study participated in a special testing procedure for experimental purposes.

**Study Objective**

Three aspects of trainee performance of the field certification phase were evaluated — the ability to (1) administer the test, (2) correctly record the results of the test, and (3) correctly interpret the results of the test. This paper deals with the third of these, evaluating the ability to interpret the results. This task, which we will call "diagnosis," requires that the DRE first determine whether the suspect is impaired and then determine the category of drug used by the impaired suspect by evaluating the physical evidence garnered by the examination. In the DEC field training program the trainee is taught to recognize symptoms associated with the seven drug categories. The trainee's progress in accurately judging impairment and identifying the correct drug category was evaluated.
METHODOLOGY

The observations for each test conducted during the 30- to 45-minute DRE examination is recorded on a Drug Influence Evaluation (DIE) Form. This form is used to record physical signs such as blood pressure, pulse, and pupil dilation; results of psychomotor tests such as the Walk-And-Turn (WAT) and the One-Leg-Stand (OLS); signs of the suspect's demeanor such as speech, appearance, and attitude; and physical signs of drug use such as puncture marks, or debris in the oral or nasal cavity. If these test observations are correctly recorded, sufficient information from the examination is captured on the DIE Form that a trained Drug Recognition Expert can make a valid diagnosis without observing the suspect directly.

Measuring the trainee’s proficiency in judging the correct drug category required a standardized measure across all eleven sites and 199 trainees. To meet this need, a set of 16 DIE Forms ("exemplars") were adapted from the records of actual cases where a DRE made a diagnosis which was confirmed by toxicological examination. Each trainee analyzed one of these test forms and recorded whether the subject was impaired and identified the impairing drugs following each supervised evaluation. This provided a measure of student progress in drug category identification through the field training program.

The ability of the trainees measured by this test should not be confused with a Certified DRE’s ability to make judgements in a real-life setting. Actual cases differ from this exercise in the following ways: Test forms were selected to include those with only one category of drugs, whereas 45% to 72% of the DRE cases in department files involve more than one drug category. The selected test forms had an even distribution of the seven drug categories, whereas in real life four drug categories; stimulants, narcotics, cannabis, and depressants predominate. The other three categories; PCP, hallucinogens, and inhalants, are rarely seen. This special set of cases was selected for testing purposes — all cases that were especially easy or difficult were eliminated. In arrest situations Officers frequently gain additional information from the circumstances of the arrest (i.e. drugs found in the suspect’s possession) or the suspect’s admission of drug use during or after the examination. This exemplar exercise measures only the ability of the trainee to correctly interpret the indices provided on the Drug Influence Evaluation form. It does not intend to measure the accuracy of the complete drug apprehension process. Thus, the absolute performance levels were expected to be lower than those obtained in "real-life."

Administering the Exercise

The 16 special DIE test forms were randomly ordered to insure that each form (and each drug category) appeared roughly the same number of times for each evaluation. Each trainee was given a packet of forms at the start of the field certification phase, which started as soon as the class was completed. During field certification, each trainee was required to complete at least 12 to 20 evaluations of impaired suspects before being approved for certification. Because the purpose of this exercise was to measure the incremental improvement
of the trainees as they progressed through the field certification process, they were asked to complete one of the exemplar forms after each actual evaluation they completed. The trainees completed these forms on their own with no feedback, and without referring to any reference materials. Completing the DIE test form exercise required about five minutes.

**ANALYTIC METHOD**

Due to the nature of the data used in this study, defining accuracy becomes a complex issue. There are eight conditions represented (with actual condition defined as the toxicological laboratory's finding), and thus eight diagnoses that can be made: one for each of the seven drug types, and one special "no drug" condition representing alcohol-only suspects. The exemplars in this study included two cases of each of the seven drug categories and two cases of alcohol alone, for a combined total of 16 exemplars. In order to more closely approximate reality, DREs were not made aware of the fact that each exemplar represented only one of the seven drug categories (or none at all). Since in actual field situations more than one drug may be present, the exemplar exercise allowed DREs to name as many categories as they thought to be necessary.

The simplest method of assessing accuracy (and perhaps the most commonly used) is to calculate the overall "hit" rate, or correct classifications as a percentage of all diagnoses. Any other type of diagnosis is considered a "miss" of the drug actually represented. While straightforward and easy to interpret, it is inappropriate to this task for various reasons. First, naming more than one drug increases the probability of a correct identification by chance alone, thereby inflating the hit rate as a percentage of all diagnoses. This creates the potential for strategic "over-diagnosing" unless some penalty is introduced for the additional diagnosed categories.

For a theoretically meaningful mathematical treatment of the data, a precise model of the nature of the task was required. The model chosen for this analysis was to view the DRE performance in terms of seven present/absent decisions per exemplar, then use Signal Detection Theory to calculate sensitivity and bias scores for each of these seven drug categories. By looking at the subset of all cannabis diagnoses, for example, the data fall neatly into a 2x2 matrix, where one factor is the chemically-determined drug condition (present/absent) and the other is the DRE-diagnosed drug condition. The hit rate is the proportion of chemically-detected cases that are also diagnosed-present (the remainder being the misses). The false positive rate is the proportion of chemically-absent cases that are diagnosed-present (the remainder being correct rejections). Using these proportions, sensitivity and bias scores derived from Signal Detection Theory were calculated for each drug type (i.e., each of the seven decisions).

These measures of sensitivity and bias can be calculated for each stage in the evaluation sequence, and then analyzed using conventional statistical procedures to assess changes over time. Because these measures were based on aggregate data, thereby reducing the number of data points, the power of the test is greatly reduced (in terms of $N$) as opposed to using each exemplar as a
separate case. However, this is somewhat offset by the greater stability represented by each aggregate point, since the high variance associated with the unaggregated dichotomous data is eliminated.

RESULTS

A learning effect could be manifested in two ways. The primary learning expected is that the DRE gains increased ability to recognize patterns of signs associated with a drug category; this learning is termed accuracy (or sensitivity). A second type of "learning" occurs if the DRE becomes less conservative in his or her reporting threshold, reporting a drug presence if there is any evidence of the patterns of signs associated with the drug. This "learning" is strategic and is independent of accuracy. Thus, correctly "hitting" more drugs may be a learned bias change, which would increase both "true" and "false" positives. Both types of "learning" were tested.

At each trial stage (1 through 15), the A's and the $\beta$s for each of the seven drug types were calculated. The sensitivity for the trainees was initially .893 (or 89% of the signal and noise distributions correctly classified), and increased to .923. In terms of the area left for improvement (1-.893, or .107) the learning over 15 trials resulted in reducing misclassification by .030 over .107 or 28.1%, as is shown by the solid regression line in Figure 1. This increase in sensitivity is statistically significant ($r = .615, p = .007$).

**Figure 1**

Change in Sensitivity

![Figure 1](image-url)

raw scores = composite (mean) of 7 drug types

1551
As shown in Figure 2, bias throughout the evaluation sequence remained well above zero, indicating that when behavioral and physiological conditions were borderline, the DRE's tended to give the suspects the benefit of the doubt (as opposed to asserting drug-impairment). A slight trend toward using a more "aggressive" criterion was not significant (p = .096).

This is especially interesting given the policy of what constitutes a "correct" assessment. In general, only one drug needs to be confirmed by toxicological exam even when two or more drugs are diagnosed present. One might expect DREs to acquire the strategy of naming two or three drugs when impairment is evident in order to increase their chances of meeting the minimum policy criterion of correctly identifying at least one. Even if we were to consider the trend shown in Figure 2 to be significant, the DREs, nevertheless, would be committing a false positive diagnosis only one-fourth as often as they miss a drug that is actually present.

**Figure 2**

*Change in Bias*

![Figure 2 Graph](image)

raw scores = composite (mean) of 7 drug types