Drug Evaluation and Classification: Review of the Program and Opportunities for Enhancement

February 2019

607 14th Street, NW, Suite 201 | Washington, DC 20005 | 202-638-5944

AAA Foundation for Traffic Safety

Title

Drug Evaluation and Classification: Review of the Program and Opportunities for Enhancement *(February 2019)*

Authors

Douglas J. Beirness[1] and Amy J. Porath[2]

[1]*Beirness & Associates, Inc.*
[2]*Canadian Centre on Substance Use and Addiction*

Foreword

Drugged driving has a detrimental impact on traffic safety and presents a significant challenge to law enforcement. Unlike alcohol, it is more difficult to accurately assess drivers for impairment due to other drugs. The Drug Evaluation and Classification Program (DECP) was developed to help law enforcement officers combat the problem; however, the program is not without its challenges.

This report offers a summary of information regarding the DECP and identifies opportunities for improvements that could serve to strengthen the program. Materials presented in this report should be a useful reference for the traffic safety community.

C. Y. David Yang, Ph.D.

Executive Director
AAA Foundation for Traffic Safety

About the Sponsor

AAA Foundation for Traffic Safety
607 14th Street, NW, Suite 201
Washington, D.C. 20005
202-638-5944
www.aaafoundation.org

Founded in 1947, the AAA Foundation for Traffic Safety in Washington, D.C., is a not-for-profit, publicly supported charitable research and education organization dedicated to saving lives by preventing traffic crashes and reducing injuries when crashes occur. Funding for this report was provided by voluntary contributions from AAA/CAA and their affiliated motor clubs, individual members, AAA-affiliated insurance companies, and other organizations or sources.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

ARIDE	Advanced Roadside Impaired Driving Enforcement
BAC	Blood alcohol concentration
CIT	Compulsory Impairment Test

CNS	Central nervous system
CTI	Clinical Test for Impairment
DEC	Drug Evaluation and Classification
DECP	Drug Evaluation and Classification Program
DRE	Drug Recognition Expert
FDA	Food and Drug Administration
FIT	Field Impairment Test
FTN	Finger to Nose test
HGN	Horizontal Gaze Nystagmus test
IACP	International Association of Chiefs of Police
ICADTS	International Council on Alcohol, Drugs and Traffic Safety
LAPD	Los Angeles Police Department
LOC	Lack of convergence
MRB	Modified Romberg Balance
NHTSA	National Highway Traffic Safety Administration
NIDA	National Institute on Drug Abuse
OLS	One Leg Stand test
SFST	Standard Field Sobriety Tests
VGN	Vertical Gaze Nystagmus test
WAT	Walk and Turn test

Introduction

The use of psychoactive drugs by drivers poses a risk to traffic safety and presents a significant challenge to law enforcement. Whereas the Standardized Field Sobriety Tests (SFST) and breath testing technology have become invaluable tools in enforcement efforts to combat alcohol-impaired driving, determining driver impairment by drugs has proven to be considerably more complex. Foremost, there are numerous types of drugs, many of which have effects that differ dramatically from those of alcohol. In addition, unlike alcohol, most drugs cannot be readily measured in breath at the side of the road. The lack of tools and procedures to adequately assess drivers for impairment due to drugs other than alcohol was a serious gap in law enforcement efforts to remove these high-risk drivers from the road.

The Drug Evaluation and Classification Program (DECP) was developed to help fill this void. Since its inception in the late 1970s, the DECP has been adopted by every state as well as Canada and has served as the basis for similar programs in other countries. Despite its widespread use, the program is not without its challenges. The program involves a detailed assessment protocol that requires ample training. Only a select group of officers have the opportunity and desire to acquire the specialized skills required to become proficient in the techniques of the DECP. The results of the assessment are not always readily accepted by the courts and are subject to numerous challenges. Nevertheless, the DECP continues to grow and has become a central component of the response to drug-impaired driving.

The purpose of this report is to examine the evidence supporting the DECP and to identify opportunities for potential improvements that would serve to strengthen the program.

The specific objectives of the project were to:

- Conduct a review of the literature on the DECP to identify the strengths and limitations of the program.
- Investigate and identify opportunities and make recommendations to bolster the DECP program, the DECP assessment protocol, the DECP training, and the DECP certification and re-certification process.
- Determine the approach and potential obstacles that may be involved in implementing recommended improvements to the DECP.

The report begins with a review of the evidence supporting the various components of the DECP as well as a review of the evidence pertaining to its validity and effectiveness.

The second section describes practices for assessing suspected drug-impaired drivers in selected countries around the world to determine if there are alternative or additional approaches that could be adopted.

New technologies that could be introduced into the DECP to assist with measurement and improve efficiency are discussed in the third section.

Key informant interviews were also conducted with individuals who have a connection to, or involvement with, the DECP to gather their thoughts and insights into the program as well as suggestions for improvement. The results are presented in the fourth section.

The fifth section examines data from a large number of DECP evaluations to examine differences in the various indicators from the assessment according to the type of drug ingested. In addition, the data are used to determine if there are key indicators included in the DECP assessment that can help identify the category of drug used.

The report concludes with a summary and a series of suggestions for enhancing the DECP.

Literature Review

History and Overview of the DECP[1]

The origins of the DECP date back to the 1970s when police officers in Los Angeles noted the relatively high rate at which drivers arrested for impaired driving were being released because they were shown to have a low blood alcohol concentration (BAC) or had not been drinking at all. In response to the growing recognition of the need for a method to assess impairment by drugs other than, or in addition to, alcohol, officers with the Los Angeles Police Department (LAPD) consulted with experts in medicine, toxicology, and behavioral psychology to pool their collective knowledge about the signs, symptoms and behavioral effects associated with the ingestion of various drugs. The objective was to develop a standardized procedure to assist in the identification of drivers suspected of being under the influence of drugs.

The result of their efforts was a drug assessment protocol that was officially recognized by the LAPD in 1979. It incorporated interviews, behavioral tests, and measurements of vital signs and other clinical indicators that can be affected by psychoactive substances. The unique aspect of the protocol was the integration of the various tests and measurements into a comprehensive procedure to assess impairment and identify the category of drugs most likely to be the cause of impairment.

The first Drug Recognition Expert[2] (DRE) school was held in Los Angeles in 1980. Instructors included physicians, behavioral researchers and other scientists. The program eventually captured the attention of the National Highway Traffic Safety Administration (NHTSA), which worked with the LAPD to formalize the program into what is now known as the DECP. Pilot programs were set up in three states in 1987, and programs were added in three more states in 1988.

The resultant assessment protocol is a systematic and standardized procedure that involves a series of interviews and observations, as well as psychomotor tests and measurements of vital signs and clinical indicators, followed by a toxicological evaluation. The assessment consists of the following 12 steps (International Association of Chiefs of Police, 2015):

1. Breath alcohol test: A breath test is conducted to rule out alcohol as a cause of impairment or determine if alcohol is contributing to the observed signs and symptoms.

2. Interview of the arresting officer: The officer who made the arrest is questioned to gather information about the traffic stop and the suspect's behavior and demeanor, as well as other pertinent information that might be relevant to the assessment.

[1] The information in this section on the development of the DECP was compiled from the DECP website (DECP.org), the DRE Course manual (IACP 2015), personal interviews, and presentations on the history of the DECP at the 17th Annual IACP Training Conference on Drugs, Alcohol and Impaired Driving (July 20, 2011).

[2] Some jurisdictions prefer to use the word "evaluator" or "technician" as a substitute for "expert."

3. Preliminary examination: The subject is asked about existing medical conditions or injuries and examined to look for any evidence of a medical condition that requires immediate medical assistance (e.g., equivalence of pupil size, equal eye tracking). The first of three pulse measurements is taken.

4. Examination of the eyes: The examining officer assesses horizontal gaze nystagmus (HGN), vertical gaze nystagmus (VGN) and lack of convergence (LOC).

5. Divided attention psychophysical tests: Four tests are conducted – Walk and Turn (WAT), One Leg Stand (OLS), Modified Romberg Balance (MRB), and Finger to Nose (FTN).

6. Examination of vital signs: Blood pressure and body temperature are measured. The second measurement of pulse is also taken.

7. Dark room examinations: Pupil size is measured in room light, direct light, and near total darkness. Pupil reaction to light and rebound dilation are also assessed.

8. Examination of muscle tone: The subject's arm muscles are examined to assess whether muscles are rigid or flaccid.

9. Examination for injection sites: The subject is examined for evidence of recent injection.

10. Suspect's statements and other observations: Further questioning of the subject is conducted, including the use of drugs.

11. Opinion of the evaluator: The evaluator reviews the findings from the evaluation and forms an opinion of the category (or categories) of drug[3] responsible for the signs and symptoms observed during the evaluation.

12. Toxicological examination: A specimen of bodily fluid (blood, urine or oral fluid) is collected and sent to a toxicology laboratory for analysis of drug content.

The measurements and observations taken during a DEC evaluation are recorded by the DRE on what is commonly referred to as a drug influence evaluation "face sheet." An example of a face sheet is provided in Appendix A. A narrative report is also prepared by the evaluating officer that summarizes the results of the evaluation.

The purposes of the procedure are: (1) to provide the officer with the necessary evidence to determine whether or not the suspect is impaired, (2) to determine whether the observed impairment is due to drugs rather than a medical condition and, (3) to determine which category (or categories) of drugs might be responsible for the observed impairment (IACP, 2015; see also Porath-Waller, Beirness, & Beasley, 2009 and Logan, Kacinko, & Beirness,

[3] A drug is defined as any substance that when taken into the body can impair the ability of the person to operate a vehicle safely. Drugs are divided into seven categories – central nervous system (CNS) depressants, inhalants, dissociative anesthetics, cannabis, CNS stimulants, hallucinogens, and narcotic analgesics. These categories reflect the commonalities in the signs and symptoms associated with various substances and are not necessarily intended to reflect pharmacological properties. (See Page, 2007, for a detailed discussion of the drug categories.)

2016). The results of the tests, when corroborated by toxicological evidence of drug use, are generally deemed to provide sufficient evidence to proceed with drug-impaired driving charges.

The DEC training program involves 72 hours of instruction in the techniques of a drug influence evaluation followed by field certification that involves the evaluation of at least 12 subjects believed to be impaired by drugs other than alcohol. Candidates must also complete a comprehensive examination. Those who are successful can be certified as a DRE. To maintain certification, DREs must complete a minimum of four evaluations and attend a minimum of eight hours of approved recertification training every two years.

The DECP is supported by the National Highway Traffic Safety Administration (NHTSA) and is governed by the Highway Safety Committee of the International Association of Chiefs of Police (IACP). In 1992, a set of minimum standards was adopted specifying the requirements for training, certification and recertification of DREs and DRE instructors, and standards for agency participation. A Technical Advisory Panel has been formed to assist the Highway Safety Committee on specific matters pertaining to the DECP, including the curriculum, training, and technical aspects of the program.

In 2016, there were 8,277 certified DREs in the United States, plus another 607 in Canada, though some may not have been active. Officers from the United Kingdom, Germany, Australia and China have also been trained as DREs (DECP, 2017).

Vital Signs and Clinical Indicators

The DEC evaluation includes measurements of the subject's pulse, blood pressure, and body temperature. Psychoactive drugs can affect the physiological mechanisms responsible for these systems by either mimicking the actions of neurotransmitters in the autonomic nervous system (agonistic action) or by blocking the action of neurotransmitters (antagonistic action) (Julien, Advokat, & Comaty, 2008). Some drugs have excitatory effects while others have inhibitory effects, causing the physiological systems to react differently.

Other clinical indicators such as observations and measurements are made of the eyes and muscle tone. Drugs can affect the mechanisms underlying these systems as well, causing diverse effects. Each drug category has a relatively distinct pattern of potential effects on these indicators. Measurements of vital signs and other clinical indicators can provide valuable clues as to the type of substance that might be responsible for the observed effects.

A summary of the typical effects of drugs on vital signs and other clinical indicators used in the DECP is presented in the form of a table known as the DRE Matrix (Appendix B). The expected effect of each of the seven drug categories on nine indicators is listed in the table. Comparing the results from the evaluation with those listed in the matrix should be a guide in identifying the category of drug the suspect most likely ingested.

In the course of the development of the DECP, the effects of drugs on vital signs and other clinical indicators were gathered from common medical knowledge and experience. However, a complete accounting of the sources of this information was not apparent or readily available.

There are, in fact, numerous sources of information on the potential effects of various types of drugs. In addition to many medical and pharmacology texts, there are online resources available from the National Institute on Drug Abuse and the Food and Drug Administration.[4] The Diagnostic and Statistical Manual of Mental Disorders also describes the criteria used to diagnose intoxication by various types of drugs (American Psychiatric Association, 2013). In addition, limited information on the effects of particular pharmaceutical products can be found in drug product monographs, many of which are available online. Product monographs can provide a wealth of information about the pharmacokinetics and pharmacodynamics of specific drugs at therapeutic levels and the findings from clinical trials that are required in the drug approval process. Product monographs also provide information on additional effects of the drug, which can include effects on heart rate, blood pressure, and body temperature. Information on adverse reactions ("side effects") from drug trials are also reported. This information reflects therapeutic doses provided to patients with a particular condition, and therefore offers little or no information regarding potential abuse by healthy subjects.

In reviewing the information on drug effects from these sources, drug effects on vital signs such as heart rate and blood pressure are typically only reported if there was a notable effect. Where no effect is reported, it is not clear whether the drug had no effect or it was simply not measured.

As a means to substantiate the drug effects listed in the matrix, a number of sources of such information were consulted. The effects on vital signs for each drug category in the matrix were matched with the references. The results are presented in a series of tables in Appendix C. The complete citation for each of the sources listed is included in the reference list.

In reviewing the literature on clinical drug indicators, there were cases in which no corroborating reference was found for a sign or symptom and/or there were contradictions between the matrix and sources on the effects of some drugs. For example, whereas many sources indicate that cannabis causes an increase in blood pressure (Julien et al., 2008; Leikin & Paloucek 2007), others (e.g., Korsmeyer & Kranzler, 2009) indicate the opposite, suggesting that observed low blood pressure in cannabis users may be a result of orthostatic or postural hypotension (i.e., a drop in blood pressure associated with standing up from a lying or seated position). Such apparent contradictions do not necessarily invalidate the vital signs/clinical indicators noted in the DECP Matrix. Drug effects can be variable and not all effects are always evident in every subject who has used a particular category of drug. The drug categories used in the DECP often include numerous substances, not all of which have exactly the same effects or the same intensity of effects. The absence of an indicator, or even an opposite effect, can sometimes be observed. Such effects may be related to the dose of the drug ingested, frequency of use, time since ingestion, tolerance to the drug, interactions with other drugs ingested, health conditions, and individual differences. In the field, observed effects may diverge from those in clinical research due to unrestricted dosing.

[4] The websites for NIDA and the FDA are www.drugabuse.gov and www.fda.gov/drugsatfda/Drugs, respectively.

In compiling this information, the reference materials rarely cite original research studies from which the information was derived. Textbooks and other source materials typically report drug effects on vital signs as "facts" without direct reference to the original research, let alone a description of the methods employed. Hence, it is not known whether these "facts" are the result of a collection of case reports or double-blind drug administration studies, whether the subjects were healthy volunteers or patients, the dose of drug administered, or the concentration of drug at the time of measurement.

The evidence presented in support of the vital signs and clinical indicators presented in this report is not intended to be definitive. The intention was to examine the basis for the DECP Matrix in the medical literature. For the most part, the information was gleaned from pharmacology and medical sources. The strongest evidence supporting the drug effects listed in the DECP Matrix would involve a comprehensive systematic review of the literature on each of the vital signs and indicators for each of the seven drug categories. Such an exercise, though worthwhile, was beyond the scope of this project.

It should also be noted that the DECP Matrix is a "category-based" summary and reference guide. It was never intended as a definitive source of information on drug effects. Drug effects depend on a number of factors such as the specific substance ingested, the dose, pharmacokinetics, time since ingestion, tolerance, and individual differences. The matrix was designed as a general guide for those who have studied and been certified in the DECP. The specific drugs and substances within the various drug categories are quite varied and differ somewhat in their associated physical and behavioral effects. The matrix cannot address all the subtle details and differences among specific drugs. Rather, it is intended to be a quick reference guide for the examiner to help refresh their memory of particular drug effects and the decision process, which were studied extensively in the training program.

Psychophysical/Divided Attention Tests

The three tests that comprise the Standardized Field Sobriety Test (SFST) – the Walk and Turn (WAT), One Leg Stand (OLS) and Horizontal Gaze Nystagmus (HGN) (Tharp et al., 1981) – are embedded in the DECP protocol. The DECP protocol also includes two additional psychophysical tests — Modified Romberg Balance (MRB) and Finger to Nose (FTN). These two tests were among the original set of 10 tests examined by Burns and Moskowitz (1977) in the development of the SFST. Each of these tests is briefly described below[5].

Horizontal Gaze Nystagmus (HGN) test

HGN is an involuntary jerking of the eye that occurs as the eyes gaze from side to side. During the HGN test, the individual is instructed to follow an object (such as a pen or finger) with their gaze as it is moved at a steady pace slowly and horizontally from side to side. The officer assesses three indicators of HGN (referred to as "clues") in each eye, for a total of six possible clues.

One Leg Stand (OLS) test

[5] Portions of this section are drawn from Porath-Waller & Beirness (2014).

In this test, the individual is instructed to stand with his or her arms at their sides, raise one foot approximately six inches off the ground, and count aloud from 1,001 (i.e., 1001, 1002, 1003, 1004, etc.) until told to stop. The evaluator uses a timer to ensure the test duration is 30 seconds. The test is performed on each leg.

Walk and Turn (WAT) test

In the WAT test, subjects are instructed to place their right foot in front of their left foot touching heel to toe, place their arms at their sides, and not to begin until told to do so. Subjects are to take nine steps, heel-to-toe, along a straight line. After taking nine steps, subjects are instructed to turn by leaving the lead foot on the ground and taking a series of small steps with the other foot until facing the opposite direction. Subjects are then to take nine steps back along the line in the same heel-to-toe manner.

Modified Romberg Balance (MRB) test

Subjects are instructed to stand with their feet together, arms at their sides, head tilted slightly back and eyes closed. When told to begin, subjects are instructed to remain in that position until the subject believes 30 seconds has elapsed and then open their eyes and lower their head. The officer assesses the amount of front-to-back and side-to-side sway displayed by the subject during the test as well as the actual amount of elapsed time. This test assesses postural balance and alterations in the perception of time (International Association of Chiefs of Police, 2015).

Finger to Nose test

Subjects are instructed to stand with their feet together, hands at their side, palms facing forward, index fingers extended, head tilted slightly back and eyes closed. When instructed, they are to raise the indicated hand and touch the tip of their nose with the tip of their finger and then return the hand to their side. There are six trials, three with each hand. Officers are to note any evidence of body sway as well as eyelid and body tremors.

Psychophysical/Divided Attention Tests to Assess Impairment by Alcohol

Considerable research has been conducted on the SFST for alcohol (e.g., Tharp et al., 1981; Burns & Anderson, 1995; Stuster, 1997; Stuster & Burns, 1998). These studies typically had police officers administer the SFST to drivers and use the results to decide whether to arrest the driver or not. Arrest decisions were then compared with a measure of BAC. The primary measure was overall accuracy — i.e., the percentage of subjects who were correctly identified by the officer as either impaired or not. For example, Stuster et al. (1998) reported that officers correctly identified drivers with BACs over or under .08% in 91% of cases based on their performance on the SFST. Overall accuracy included cases that were correctly identified as having a BAC over .08% (known as the sensitivity of the test) plus cases correctly identified with a BAC under .08% (known as specificity).

A disadvantage of overall accuracy as a measure of test performance is that in situations where either impaired or unimpaired drivers predominate, it can provide an incomplete and possibly misleading measure of the validity of the test. If either the sensitivity or specificity are low, the overall accuracy may nevertheless be high. It is only in cases where the

prevalence of impaired drivers in the tested population is close to 50%, or in cases where sensitivity and specificity are almost equal, where accuracy will closely approximate both sensitivity and specificity (Alberg, Park, Hager, Brock & Diener-West, 2004).

To illustrate, Stuster et al. (1998) indicated that overall accuracy of the SFST was 91%. Although sensitivity and specificity were not reported, these measures can be calculated from the data tables in the report. Overall, the combined battery of three tests that comprise the SFST had a sensitivity of 98% but specificity was 71%. For HGN, the sensitivity was 98% and the specificity was 63%; for the WAT, the sensitivity was 92% and the specificity was 47%; for the OLS, the sensitivity was 92% and the specificity was 59%.

The three tests of the SFST have proven to be of considerable value in the enforcement of alcohol-impaired driving laws and have been widely implemented in the United States, across Canada, and in parts of Australia. In addition, individual components of the SFST have also been incorporated into the field impairment testing procedures used in many other countries (see section on international practices).

Psychophysical/Divided Attention Tests to Assess Impairment by Drugs Other than Alcohol

Research on the SFST has been instrumental in establishing a foundation for the use of behavioral tests for impairment in drivers. The inclusion of the three tests of the SFST plus FTN and MRB in the DEC protocol has contributed to the perception and use of these procedures as general tests of impairment. However, the SFST was developed and validated as a test of alcohol impairment and has not been systematically validated as a test of drug-induced impairment. In fact, relatively few studies have examined these tests as a means to assess the impairing effects of other psychoactive substances (see Bramness et al., 2003; Smink et al., 2008; Brookoff et al., 1994; Silber et al., 2005; Downey et al., 2012, Papafotiou et al., 2005a and 2005b; Bosker et al, 2012; Logan et al., 2016; Porath-Waller & Beirness, 2014). Rather than relating performance on psychophysical testing with specific drugs, the DEC program first assesses impairment, and then determines the cause of the impairment.

Experimental studies provide limited evidence on the effects of various drugs on commonly employed tests of impairment. Drug administration studies can be challenging and are ethically constrained in the types of drugs and doses that can be administered. Nevertheless, these studies illustrate that several types of psychoactive substances can have adverse effects on performance of these tasks. Further research is needed with different types of drugs and doses to document the nature and extent of the effects of various substances. Research using novel methods and incorporating a broad range of subject characteristics would also help expand the base of knowledge in this area.

Limitations of the Research on Psychophysical/Divided Attention Tests

At first glance, there would appear to be some disagreement in the evidence on the effect of drugs on the psychophysical/divided attention tasks that are included in the DEC protocol. However, in laboratory studies, due to ethical and safety considerations, the dose of drug administered may be well below that which individuals might choose to self-administer. In studies that sampled individuals in naturalistic settings, including those using the results

of DEC evaluations on suspected impaired drivers, unspecified doses of drugs were self-administered. The disadvantage of the former approach is that lower doses provide limited understanding of the effectiveness of the SFST and other psychophysical tests in detecting drivers who have ingested considerably larger doses of drugs. A limitation of the latter approach is that drug levels may not have been tested or reported and that the suspects may have been under the influence of multiple drugs. Hence, it is not possible to determine the sensitivity of these tests in identifying impairment at known drug levels.

Another prominent factor to consider in evaluating the impact of drugs on psychophysical test performance is the action of various types of substances on the brain. Depending on mechanism of action in the brain, different substances can be expected to have different effects on cognitive and psychomotor performance. For example, HGN is a good indicator of alcohol use, depressants (e.g., Smink et al., 2008) and dissociative anesthetics (e.g., Cheng et al., 2007) but is not typically affected by cannabis (e.g., Papafotiou et al., 2005a, b).

The relative absence of data on drug-free performance on psychophysical/divided attention tests limits the ability to evaluate specificity. It is important to know the extent of normal, drug-free variability in the performance of these tests to better understand the influence of psychoactive substances. In this context, Rubenzer (2008) has indicated that there is also a need to explicitly investigate how performance on the SFST is related to age, sex, medical and psychiatric conditions, race, drug tolerance, and other potentially relevant characteristics within the population.

Some authors have expressed concern about inconsistencies in the administration of the SFST that could reduce its effectiveness. For example, Barone and Crampton (2005) noted that during the HGN test, the most common errors made by officers were moving the object from side to side an incorrect number of times, failing to move the object with the correct timing, and failing to properly estimate a 45-degree angle. Rubenzer (2008) noted several possible sources of interrater disagreement in the administration of HGN. These included difficulty in accurately estimating an angle of 45 degrees, judging when nystagmus has occurred, and difficulty in administering the assessment of smooth pursuit with the correct motion and uniform velocity. The author suggested that these difficulties indicated a need for further training or the use of instruments to aid in administering the test. It is not known, however, how these differences in administration affect scoring or the ultimate decision of the officer in terms of the subject's impairment.

Key Indicator Studies

In the course of a DEC evaluation, the officer will collect in excess of 100 pieces of information. The officer must then attempt to assemble and integrate this information to develop an opinion about impairment and the category (or categories) of drug(s) most likely responsible for the observations. The extent of information available is too vast to reasonably expect a person to consider every piece of data in rendering a decision. Hence, it has been suggested that DREs might rely on only one or two "pivotal" signs and symptoms to guide their decision concerning drug category while ignoring others, even if contradictory to their judgment (Shinar and Schechtman 2005).

Several studies have also examined the data elements collected in the course of a DRE assessment in an attempt to identify the best set of predictors of impairment by the various types of drugs (Heishman et al., 1996; 1998; Shinar & Schechtman, 2005; Schechtman &

Shinar, 2005; Porath-Waller et al., 2009; Porath-Waller & Beirness, 2010). These studies serve to validate the signs and symptoms of drug use collected as part of DECP evaluations. In addition, the identification of key signs and symptoms of drug use suggests that there are key indicators collected during a DECP evaluation that can be pivotal in the determination of drug category. In the context of the overall evaluation results, these indicators should be assigned greater weight in the process of determining the most likely drug category (or categories) involved. The current DECP provides no differential weighting to the different indicators.

In any event, these studies suggest that it might be possible to develop a more efficient means of analyzing and weighting combinations of signs and symptoms associated with various drugs to provide guidance in the prediction of drug categories. Whether this is best accomplished through more extensive training in the role of key indicators or through the development of an algorithm that could be applied to the data collected remains to be determined.

The Accuracy of DECP Evaluations

The measurements and observations of vital signs, clinical indicators, and psychophysical/divided attention tests, along with the interviews collected during DEC evaluations, provide officers with a broad spectrum of evidence upon which to base their opinion about a subject's impairment and the category (or categories) of drugs most likely to be the cause of the observed symptoms. When the DRE's opinion of drug category is corroborated by toxicological evidence, the evaluation is generally sufficient to proceed with drug-impaired driving charges[6]. A key indicator of the validity of the DECP would be the degree of correspondence between the officer's opinion of the category of drug and the results of toxicological tests of bodily fluid.

There are two general types of studies evaluating the accuracy of DECP evaluations — experimental laboratory and field studies. These two research approaches differ from each other in several ways but together provide a broader assessment of the accuracy of DECP evaluations than either alone. Laboratory studies provide researchers the opportunity to control various factors, including the type(s) and dose(s) of drugs ingested, the time elapsed between drug administration and behavioral assessment, and the variables collected from the assessment procedure. From a research perspective, using the same group of DREs and volunteers repeatedly over several sessions reduces variance attributable to individual differences among subjects and DREs. However, this approach reduces the inherent variability associated with differences in the skills of individual DREs and the range of drugs, doses administered, and drug combinations self-administered by drivers. Consequently, it is possible that their estimates of the accuracy of the DECP are higher than can be achieved under normal field conditions. Laboratory studies can also include a placebo condition in which volunteers are given an inactive substance to control for the effects associated with the expectation of receiving an active drug and to assess subjects who are not under the influence of any substance. A "double blind" procedure, in which neither the volunteer nor the DRE doing the evaluation are aware of what drug has been ingested, is a means to reduce the likelihood of bias associated with the subject or evaluator being aware of the type and amount of drug administered. Restricting the questioning of

[6] It is not essential that the toxicology results match the DRE opinion to proceed with charges.

volunteers is used to eliminate the potential influence introduced by admissions of drug use. Procedures employing these restrictions provide a rigorous test of the psychophysical assessment rather than a test of the complete DECP procedure in the actual context and circumstance of an arrest.

Field studies of the DECP involve a retrospective review of DEC evaluations conducted in an enforcement setting. In these cases, the DREs' interview of the arresting officer and the conversation (and often admission) of the suspect are already incorporated into officers' conclusions. Furthermore, the officers are generally aware of the prevailing drugs of choice in their environment, thus providing them with valid prior probabilities for the different drug categories. Judgments of suspected drug use by people arrested for an impaired driving offense are compared with the results of toxicological tests for the presence of drugs. Whereas a key feature of experimental studies is the degree of control the researcher can exert over the situation, such controls are not possible in field settings. This can both help and hinder the observed accuracy with which DREs identify impairment and the drug responsible. For example, whereas laboratory studies use volunteers who have been administered known quantities of one specific substance (or a placebo), field studies involve evaluations of drivers who may have self-administered unknown quantities of one or more psychoactive substances. These drivers may have medical, physical or mental health issues that can mimic or interact with the effects of drugs. The task of the DRE is to determine whether the driver is impaired and whether the impairment is a result of drug use, and if so, to identify the category (or categories) of drug(s) most likely responsible for the symptoms observed.

Importantly, compared with laboratory studies, the quantities and type(s) of drug(s) ingested by suspected impaired drivers can be considerably larger. Because of the more profound effects, higher doses are easier to detect. However, laboratory studies do not typically examine polydrug scenarios whereas, in the real world, drugs are often used in combination with other drugs and/or alcohol. Concurrent use of more than one substance can mask some symptoms and enhance others, creating challenges for identifying the substances involved.

It should also be noted that experimental studies typically restrict the time allowed for the DRE to conduct an examination of the subject and often do not allow the DRE to conduct an interview with the subject. The DECP training emphasizes that officers take into consideration the “totality” of the situation and the evidence from the assessment, including an interview with the subject and information gathered by the arresting officer.

A small number of laboratory investigations have been conducted to assess the validity of some of the components of the DECP to identify the influence of various types of drugs based on the effects observed (see Beirness, LeCavalier & Singhal, 2007, for a critical review). The experimental laboratory studies indicated that officers trained in the DECP are generally able to detect impairment in subjects who have been administered drugs. These studies, however, do not present strong support for the accuracy with which they can identify the particular class(es) of the drugs involved. Although for some drug categories the classification rate is better than chance, the miss rate is high, and the false alarm rate is also high. The lack of distinctive clinical and psychomotor symptoms associated with the relatively low doses of some of the drugs administered in the studies likely played a role in the findings. The fact that some drugs are detected with greater accuracy than others

suggests that the effects of these substances are more prominently manifested in the symptomology assessed by the DEC procedure. In real-world situations, the doses of drugs ingested are unknown and may be significantly higher than those ethically permitted in a laboratory setting. It would be expected that with higher doses, the accuracy with which DREs can detect drug impairment and identify the category (or categories) of drugs responsible for the impairment would be greater.

Similarly, few field studies have been conducted to provide a real-world test of the DEC procedure (for a critical review, see Beirness, et al., 2007; also see Beirness et al., 2009; Porath-Waller et al., 2009; Porath-Waller & Beirness, 2010). Studies conducted with data collected in enforcement settings generally report higher overall accuracy than laboratory studies. This is most likely the result of the additional information provided by or available during a complete 12-step evaluation. Although some might argue that this latter approach does not provide a pure assessment of the DECP, others would suggest that it is more reflective of the real-world implementation of the program. Either way, together the two types of studies provide evidence of the accuracy of the DECP to further its use in efforts to remove drug-impaired drivers from the roads.

Discussion

The research studies supporting the DECP are not perfect — a fact that should be evident from the discussion in previous sections. They are subject to numerous practical and ethical constraints that limit the conclusions and generalizability to other situations and populations. It should also be acknowledged that the purpose of the DEC protocol is not to diagnose a medical condition but, rather, to provide evidence in a criminal investigation. The standards, methods, and ultimate consequences of the two processes are very different.

The experimental and field evaluations of the DECP provide different perspectives on the accuracy of the program. Overall, experimental laboratory studies do not present strong support for the accuracy with which officers trained in the DECP can detect and identify the particular class(es) of drugs ingested.

The demand characteristics are markedly different in field studies and in experimental studies. In experimental studies, the situation can compel evaluators to make a drug call even in situations when they were not confident about their opinion. In fact, Bigelow et al. (1985) instructed DREs to provide an opinion about the suspected drug category even if they weren't as confident as they would normally be in an enforcement situation. In the Heishman et al. (1996) study, when the subject was deemed "not impaired," DREs could still indicate the class of drug they believed the subject had ingested. In field studies, the judgment by the DREs may also be influenced by other factors such as the act of referral to the DRE by the arresting officer and awareness of the prevailing drugs in their local area, among others.

Restricting the ability of the examiner to question the participant about drug use facilitates an assessment of the validity of the psychophysical signs and symptoms of drug use, unbiased by admissions of the suspect. The experimental controls used to enhance methodological rigor in laboratory investigations, however, create an environment that substantially differs from that in the field where DECP procedures are employed. This brings to light several important considerations. First, DECP assessments made in an enforcement context have the benefit of information obtained by the initial investigating

officer and from suspect interviews. The mere fact that an individual has been presented for assessment by a DRE serves as an indication that the individual has been detained for suspicion of driving under the influence of drugs other than, or in addition to, alcohol. In many cases, there may be physical evidence of drug use and/or the individual may have performed poorly on a field sobriety test. Often, the suspect will confess to drug use during the interview process or upon presentation of the evidence obtained during the assessment. Although Smith et al. (2002) found evidence of the veracity of such statements, suspects may not reveal all aspects of their drug use such as the amount or the number of drugs ingested. Reliance on such confessions is not recommended (International Association of Chiefs of Police, 2015).

Many other methodological concerns can be identified to underscore the limitations of research in this area, both in the laboratory and in the field. The experience of the DREs involved, the restricted elements of the DECP evaluation, the type of fluid specimen collected and tested, the detection thresholds for toxicological analysis, the elapsed time between arrest and specimen draw, and the selection and exclusion procedures for cases are among the issues that can influence the generalizability of studies of the DECP. An additional issue that limits comparison and synthesis of studies assessing the DECP is inconsistent reporting of statistics as measures of validity. The statistic most commonly reported is the percentage of DRE classifications that are correct, either in terms of identifying drug-positive drivers or the specific drug category involved. The classification rate combines sensitivity and specificity, which some studies do not report or provide sufficient data to calculate. Classification rates may be misleading and conceal very low sensitivity or specificity, and thus not accurately reflect the utility of the DRE assessment.

Despite the limitations, officers trained in the DECP are able to discern impairment and specify the category of drug responsible with a degree of accuracy that, for some drug categories, exceeds chance, and in some cases reaches a rather high level. The judgments of DREs concerning drug use should be corroborated by toxicology in most cases and false positive cases should be minimal. The field studies to date largely demonstrate this to be the case. However, a substantial proportion of drug-positive cases may be missed — or misspecified. It is possible that these cases did not manifest observable symptoms of drug use, possibly because of waning drug influence in the interval between identification as a suspected impaired driver and assessment by a DRE. In enforcement settings, the number of drug-negative cases would be expected to be minimal. Nevertheless, it is important that when these cases are presented, that they be identified with a high degree of accuracy. In addition, field studies most likely underestimate the number of false negative cases. This is because drivers who may have ingested drugs but do not display outward signs and symptoms of drug use are unlikely to be subjected to a DEC evaluation.

In conclusion, the literature provides mixed evidence of the accuracy of the DECP. While the procedures have been designed to be detailed and objective, conducting the assessment effectively requires extensive training and experience. There is also a subjective element of the process, by which the officer must determine the extent to which an observation surpasses the standard of “average” or “normal” as well as the interpretation of statements made by the subject and the arresting officer. Continuing efforts are needed to help identify and reduce the opportunities for errors to creep into the process. There also remains room for improvement in the consistency, accuracy, and efficiency of the DECP. Finally, as a general note, the research on the DECP is limited and much of it is dated (i.e., published

prior to 2000). As the program evolves, contemporary research is needed to assess the more mature DECP.

International Practices

As has been the case in North America, many countries around the world have been working to adapt their well-established practices for assessing alcohol impairment in drivers to deal more effectively with the different effects of other (or additional) psychoactive substances. This section examines some of the practices used in other countries to detect and assess drivers suspected of being under the influence of drugs other than alcohol. The objectives were to gain an appreciation for, and understanding of, the approaches used in other countries and to investigate other, or additional procedures, tests, or techniques that could potentially be adopted to enhance the DECP.

Methods

A search of the literature was conducted using a number of common databases such as PsychINFO, PubMed, Safety Lit, and Pub Med. In addition, a search of the Proceedings of the International Conference on Alcohol, Drugs and Traffic Safety (ICADTS) was conducted[7]. Search terms included sobriety tests, assessment of drug effects, drug-driver assessment, impaired driver assessment, detecting drug drivers, police drug-driver programs, drug-driver evaluation, and variations of these terms.

Websites of national transportation, public safety and/or police agencies were also searched for descriptions of programs or procedures used by law enforcement in the assessment of suspected drug-impaired drivers.

Professional colleagues in several countries were also contacted to obtain detailed information about procedures and practices pertaining to the assessment of suspected drug-impaired drivers.

Results

The amount of information available on drug-impaired driving assessment practices was often limited. However, detailed information was obtained from seven jurisdictions around the world. These are described in the following pages.

Australia

The state of Victoria in Australia uses a two-part process for investigating drug-impaired driving. Together, the two pieces involve a progressive evidence-gathering process to determine whether the driver is impaired and the cause of impairment. The first part, the Roadside Impairment Assessment (RIA), involves basic investigation skills such as observations of the vehicle in motion, interview with the driver, and observations of the driver's appearance and behavior. A template document is used to help ensure the observations are made in a standardized manner. These observations are recorded on a form to ensure a standardized presentation. The data collected are used to form the basis of an opinion about suspected drug impairment. There is a six-hour training program for the RIA.

[7] Available at *icadtsinternational.com*.

An evidential breath test for alcohol is conducted to establish the extent to which alcohol may be involved. This is followed by the Standard Impairment Assessment (SIA). This assessment is a structured and systematic assessment conducted by a specially trained officer in a controlled setting such as a police station. The SIA involves a standard series of questions and observation followed by physical tests of impairment. The assessment is videotaped to demonstrate that the procedure was performed properly.

The physical tests are based on the SFST. The three validated tests — i.e., HGN, WAT, and OLS — are scored so as to identify impairment equivalent to a BAC of .05% (Stuster & Burns, 1998). There is also a 30-second time estimation test (similar to the estimation of 30 seconds during the Modified Romberg Balance test used in the DEC protocol) and the Finger to Nose test. The final step is the collection of an oral fluid sample for analysis of drug content.

The SIA is a 32-hour competency-based program and requires reassessment every 12 months. A self-paced CD instructional program containing six modules — the Human Body, Drugs and the Body, Testing for Drugs, Drug Categories and Effects, the Roadside Impairment Assessment and the Standard Impairment Assessment — is available to officers.

All police officers receive training in the RIA procedure. A selected group of officers are trained in the SIA procedure.

Finland

Drivers in in Finland are subject to mandatory alcohol and drug testing without suspicion. There is zero tolerance for many drugs and an impairment standard also exists. The police use a Standardized Field Sobriety Observation Sheet to record observations of the driver, driving behavior and interactions with the driver. The observations include an assessment of the subject's eyes (e.g., pupil size, reaction to light, nystagmus, and redness), appearance, speech, general behavior, and balance. No psychomotor tests are performed by the police officer on site.

An assessment by a physician is also performed under controlled conditions. This includes tests of walking, turning, balance, motor coordination (finger-to-finger test), pupil size, reaction to light and nystagmus. The physician concludes with a statement about the degree of functional disorder observed and the suspected cause of the disorder (drugs and/or alcohol, medication, disease, or injury).

Germany

The approach to drug-impaired driving in Germany involves observations by the officer at roadside followed by more in-depth testing if deemed appropriate. Officers assess driving style, reactions, physical signs, appearance, speech, response to questions, orientation, mood, eye condition, and pupil size. A breath test for alcohol can be performed. Officers use a checklist to note their observations. A point system is used to score observations and assist in the decision to proceed with an administrative or criminal charge. A negative alcohol test combined with noted irregularities of speech, standing, walking, motor coordination and/or suspicious odors can lead to more in-depth testing.

A training program was developed for the police based on the DECP, with noted adjustments for legal and technical differences between countries. This program consists of a one-week program for drug experts and a separate program of two half days for all police officers.

The Netherlands

The Netherlands is currently in the process of implementing legal limits for psychoactive drugs. The limits are to be based on behavioral impairment beyond which drugs affect the ability to drive, comparable to that associated with a level of alcohol of .05%. Police are being trained to identify clinical signs of impairment. On the basis of these signs and symptoms, drivers can then be required to provide a sample of oral fluid to be screened for drugs (SWOV, 2015).

The signs and symptoms used by the police to determine whether a driver is under the influence of a drug include:

- runny nose/sniffling;
- dry mouth;
- jaw tension;
- droopy eyelids;
- watery or bloodshot eyes;
- eyelid tremors;
- pupil size;
- reaction to light;
- unsteady on one's feet;
- uncontrolled movements;
- drowsy appearance;
- hyperactivity/aggression;
- thick/slurred speech; and,
- grinding teeth.

The extent of training in the recognition of signs and symptoms of drug use is yet to be determined.

New Zealand

In New Zealand, police officers first conduct a roadside assessment of drivers to determine whether there is "good cause to suspect" that a driver has consumed drugs. The assessment includes observations of the vehicle in motion and while stopping. Once the vehicle is stopped, the officer makes note of the driver's appearance and behavior, including condition of the eyes, speech, balance, and evidence of consumption. A breath test may also be performed.

On the basis of this assessment, if the officer has good cause to suspect that the driver has consumed drugs, he or she can require the driver to undergo a Compulsory Impairment Test (CIT) to assess impairment due to drugs. The CIT must be performed by an officer who has been trained in the procedure. The CIT involves an assessment of the eyes, including pupil size, reaction to light, lack of convergence, HGN and VGN. This is followed by the WAT and OLS tests. These tests are administered and scored as in the SFST. Unsatisfactory performance (i.e., impairment) on the CIT leads to a demand for a blood test to determine drug content. The sequence of procedures to be followed along with the instructions for all tests and observations to be made are contained on an eight-page form that is completed by the police officer. The driver has the right to consult with legal counsel before answering questions or performing tests; however, refusal is an offense.

The Drugged Driver Impairment Training course is designed to train and certify officers in the CIT. The course is eight hours. It includes instruction on the nature of drugged driving, the law, how to establish "good cause to suspect," and the tests of the CIT. There is a 90-minute period allocated to practicing the administration of the tests.

Norway

Norwegian drivers suspected of drug-impaired driving are brought to the police station for examination by a physician. A blood sample is drawn for toxicological analysis. The physician then conducts the Clinical Test for Impairment (CTI). The CTI consists of a series of tests and observations that include alertness, eye signs, vestibular function, motor coordination, pulse, physical signs of drug use, and appearance. Impairment is determined by the overall judgment of the physician (Bramness, Khiabani & Mørland, 2010; Bramness, Skurtveit & Mørland, 2003).

Although several of the tests appear to resemble those that comprise the SFST and/or those in the DEC protocol, the procedures and scoring are not identical. For example, the Romberg Balance test is performed on one leg with arms stretched out for five seconds. The Finger to Nose test is performed with the arms stretched out to the side, rather than held at the subject's side. In many cases, the tests do not have a standard scoring system but the signs and symptoms of impairment are observations and conclusions based on the physician's judgment.

United Kingdom

Police in the U.K. are trained in the use of the Field Impairment Test (FIT) in situations when drug use is suspected. There are two parts to the training — field impairment testing and drug influence recognition. Training is a minimum two-day course.

Field impairment testing involves eye exams (e.g., examination of pupils, HGN, Romberg Balance, WAT, OLS, and FTN tests). The officer is also required to note driving behavior, record observations of the driver's behavior, and conduct a roadside breath test prior to administering the FIT. The tests are administered at roadside or at the police station if the officer deems it to be appropriate. The officer is required to read the instructions for all tests to the suspect. This helps ensure consistency in administration of the tests.

Drug influence recognition training provides the officer with information on the influence of drugs as well as common signs and symptoms of various classes of drugs. Although listed as separate programs, it is recommended that the courses be taken together.

As of March 2015, if on the basis of the assessment of the driver there is reason to believe the driver is affected by drugs, the officer can require the driver to provide a sample of oral fluid for screening at roadside. A positive oral fluid screen and/or suspected drug impairment provides the officer with grounds to take the driver to the police station for examination by a certified health care practitioner (either a custody nurse[8] or a forensic physician). The health care practitioner will conduct a medical assessment as well as an assessment of drug influence, which may include the FIT, to establish whether the suspect has "a condition that might be due to the use of a drug." The health care provider does not have to make an opinion about impairment. If circumstances warrant (i.e., the suspect is deemed to be impaired by the consumption of a drug), the health care practitioner will also collect a blood sample for toxicological testing.

It is recommended (but not mandatory) that certified health care providers attend the FIT training so as to ensure they are familiar with its administration. The Faculty of Forensic and Legal Medicine has recommended standards for certified health care providers who conduct these assessments.

The tests that comprise the FIT are well known to those familiar with the SFST and the DECP. The addition of the Romberg Balance and Finger to Nose tests to the three tests of the SFST, plus training in the signs and symptoms of drug use, make it look very similar to the Advanced Roadside Impaired Driving Enforcement (ARIDE) program. There is, however, nothing in the materials provided that suggest the HGN, WAT and OLS tests are scored in the same manner as in the SFST. A judgment of impairment is left to the officer. The training programs are minimal compared with those offered by the IACP. The involvement of a certified health care professional in conducting a medical assessment establishes that the observed signs and behaviors are not a result of a medical condition and facilitates the collection of a blood sample.

Discussion

The scan of international programs to assess drug impairment in drivers was not intended to be exhaustive or necessarily representative of programs and procedures around the

[8] Some larger police detachments have a nurse who attends to individuals who have been arrested and confined in holding cells (referred to as "custody suites") at the police station, reducing reliance on a forensic physician. Having a custody nurse on site also eliminates the time delay for the physician to arrive at the police station.

world. The jurisdictions and programs included are intended to be informative and illustrative rather than exemplary.

There is considerable variation among countries in the procedures used to detect, assess, and gather evidence of drug impairment. Some countries use interviews and observations of the driver's appearance and outward behavior at the roadside, which can lead to oral fluid drug screening and/or possibly arrest. The use of psychomotor tests of impairment at roadside is less common in other countries than in North America. The absence of roadside testing necessitates that a suspect be taken to the police station for testing. Although this allows for more in-depth testing under controlled conditions, it limits the amount of screening that occurs at roadside. Hence, officers may only identify those with obvious impairment for further testing, allowing those with more subtle indicators of impairment to continue driving.

Some countries require more in-depth assessments to be conducted and interpreted by physicians or other health care professionals. Although often recommended, it is not evident that physicians and other health care professionals are required to attend a training program to become proficient in the administration of the tests and procedures necessary to render an opinion of the suspect's impairment. If the suspect is deemed impaired, the physician or health care professional can proceed with the collection of a blood sample for analysis. This obviates the need to transport the suspect to a medical facility and wait for an authorized person to draw a blood sample, thereby reducing the time elapsed between arrest and sample collection. Because of the rapid elimination of some drugs from the body, this shorter interval provides greater accuracy in estimating the drug concentration at the time of driving.

The common factor across most countries is the use of some combination of the same set of psychomotor/divided attention tests to assess drug influence — i.e., HGN, WAT, OLS, FTN and MRB as used in the DEC protocol. In fact, several countries indicate that they adopted (or adapted) the tests from the United States and rely on the American evidence of their validity. Widespread use does not increase the validity of these tests but does serve to increase confidence in the value of these procedures to assess impairment by drugs.

It should be noted, however, that in adopting some of the psychomotor/divided attention tests from the SFST/DEC protocol, the tests may not necessarily be performed in the same manner. For example, the Finger to Nose tests used in Norway have the subject extend his or her arms out to the side. In the DEC protocol, the starting position has subjects place their arms at their side. Other countries may not necessarily use the standardized scoring of clues to determine impairment. Some countries use the same set of clues as in the SFST or DECP but may not necessarily use the same criterion for impairment. Interpretation of test performance is often based on the opinion of the examiner.

Where information was provided about the extent and type of training provided to police officers on drug influence assessment, differences were evident among countries. The length of training varied from eight to 32 hours, considerably shorter than the 72 hours for the DECP training. At least one program is available as a self-paced CD.

The difference in the hours of instruction can be attributable, at least in part, to the fact that programs are not necessarily comparable in scope. In some cases, the goal is to familiarize enforcement personnel with the law and the process required to collect the

necessary evidence to warrant further investigation or support an impaired driving charge. Others have a second level of more intensive instruction to train officers to collect the evidence — i.e., to administer the tests. Some training programs provide limited or no practice on the administration of the psychomotor/divided attention tasks.

In several European countries, there is considerable reliance on observations and questioning of the suspect. In the United States, this type of evidence has sparked opposition from critics of the DECP because it provides the officers with clues about drug use that are not necessarily related to test performance and impairment. In some cases, the purpose of this approach is to determine whether there is reason to suspect the driver has ingested a psychoactive substance that would be sufficient to warrant further testing; in others, the interviews and observations of the suspect provide valuable clues about possible drug use. In the DECP, this type of evidence contributes to the totality of the situation, providing context and other important information essential to understanding the circumstances.

As noted above, psychomotor testing is not necessarily performed in the field as it often is in the United States (i.e., SFST). Rather, such testing is conducted under more controlled conditions by another person, either a physician or an officer who has undergone special training to conduct the tests, interpret the results, and form an opinion about impairment. In some countries, the collection of vital signs, clinical indicators, psychophysical/divided attention tests and interviews begins to resemble the DEC protocol. This evidence, when combined with toxicological analysis of a blood sample that indicates the presence of one or more drugs, is generally sufficient to proceed with impaired driving charges.

In conclusion, despite international differences in the legal requirements and traditions that dictate how impaired driving investigations are conducted, there are numerous similarities in the general process and types of tests that are performed. This includes a two-stage assessment process that involves an initial roadside assessment followed by a more intensive and detailed assessment at the police station. The types of tests administered are often similar but may differ in the scoring and interpretation. The use of physicians and/or other health care practitioners to conduct the second stage of the assessment is common in European countries. This approach provides greater validity in terms of ruling out medical conditions that could be responsible for the observations. The major drawbacks are the potential for a lengthy time delay waiting for a physician to arrive at the police station to conduct the assessment and the cost of the physician's services. On the other hand, if the physician determines that the suspect is impaired by a drug, a blood specimen can be drawn immediately on site. Ultimately, the approach is dictated by legislative requirements and legal precedents, which vary from country to country.

New Technologies

Many of the tests included in the DEC protocol are assessed through observations and/or measurements made by the evaluating officer. Some of the measurements are made by observation (e.g., body sway); others are assessed using standard medical devices (e.g., thermometer, sphygmomanometer). The tools and techniques used to take measurements of vital signs are relatively simple and require standardized techniques and methods that, if not adhered to stringently, have the potential to affect the measurements. New technologies are available to assess/measure performance and/or clinical indicators in a manner that could save time, enhance precision, increase reliability, and reduce measurement error. This section examines some of these technologies and assesses their potential to streamline and/or enhance the accuracy of various measurements made during a DECP evaluation.

Method

The 12 steps of the DECP were examined for opportunities for improving efficiency and accuracy. Discussions with DRE instructors, observations of DRE certification sessions, and reviews of completed drug-influence face sheets helped to identify areas for potential enhancement. These included the measurement of blood pressure, eye indicators, and Romberg Balance. While the measurement of body temperature with electronic digital thermometers did not appear to present an issue, the relatively small variations in body temperature associated with different drug types that are noted in the fifth section (Data Analysis) prompted an investigation into other technologies that might be more accurate. Finally, the use of computer tablet applications to replace pencil and paper for recording the evaluation was investigated as a potential means to improve efficiency.

Internet searches were conducted to identify devices that employ newer technologies to measure vital signs. A search of the medical literature (using PubMed) was conducted for evidence pertaining to the accuracy of new technologies.

Through regular contacts with the DECP, we were aware that the Institute for Traffic Safety Management and Research (ITSMR) at the State University of New York at Albany had developed a computer tablet application for recording DECP evaluations. A meeting was held with the New York State DRE Coordinator and personnel at ITSMR to learn more about the tablet application.

Results

Computer tablet applications

The DECP drug influence evaluation face sheet is a familiar feature of the program. The face sheet provides a one-page template for the officer to complete as the evaluation progresses. (See Appendix A for an example of a drug influence evaluation face sheet[9].) It includes the questions to be asked of the subject along with space and/or check boxes to record the responses, graphics for recording the observations and scoring of tests performed, and space for making notes and diagrams. When completed, the face sheet provides a comprehensive picture of the subject's responses and performance on each

[9] The face sheet is modified from time to time to incorporate changes to the protocol and can also vary among police departments.

component of the protocol. The face sheet is typically completed manually by the evaluating officer using pencil and paper.

In an effort to improve the efficiency and accuracy of recording and reporting of DECP evaluations, the New York DECP worked with the Institute for Traffic Safety Management and Research (ITSMR) at the State University of New York at Albany to develop an electronic data entry, reporting, and management system. The system has two primary components, a web-based application and a tablet application. The tablet app is used by DREs in the field to complete, record, and submit their evaluations and narrative reports. Toxicology results can be entered subsequently when they are received from the lab.

The application also allows the data from an evaluation to be uploaded directly to the DRE National Tracking System (NTS) (DECP, 2017). The NTS is a national database for the exclusive use of IACP and NHTSA to assist in monitoring the program. Data uploads to the NTS database save the officer the time and effort of undertaking this task manually. This function also opens the door for expanding the NTS to include more data elements that would enhance the opportunity for better reporting and research. It is estimated that only about two-thirds of DREs enter their evaluations into the NTS (DECP, 2017). Automated uploads of data would help to ensure that the NTS is as complete as possible.

The tablet application accommodates real-time data entry by the DRE, streamlining the entire process and improving the overall efficiency of the evaluation and reporting process. The application allows the DRE to collect all fields required from the DEC face sheet, including the ability to draw images associated with the Walk and Turn test, the Finger to Nose test, lack of convergence, and One Leg Stand. The data can be output in a standard face sheet format.

The tablet application is an easy-to-use guided program that runs on Android, Windows and IOS operating systems. There is a four-hour training program on the use of the tablet app provided by ITSMR. The application effectively replaces the paper-and-pencil approach to recording and brings the evaluation process into the digital age with the promise of improved efficiency and accuracy of the evaluation and reporting processes. The application is flexible and can be tailored to meet the requirements of individual states. Several screen pages can be used to facilitate the evaluation and recording. Prompts for missing information can be added. There is also the capability to add photos and video to supplement the evaluation. The evaluation results are saved to the tablet and uploaded to a secure server, and can be accessed online via a secure website.

The system allows DECP coordinators to monitor the program in real time without having to wait for DREs to submit hard copy reports. Coordinators can also use the system to generate periodic reports on the number of evaluations performed by drug category, date and time, officer, and the accuracy of drug calls.

Tablets can vary in price from about $350 to $500 including a protective case. There is no requirement for police services to have consistency in the type of devices used so long as they are compatible with the application.

The application is owned and managed by ITSMR at the State University of New York at Albany. Agencies sign a licensing agreement that includes training and support. The cost of the agreement varies according to the number of DREs. For example, as of July 2017, the

cost for the Year 1 license for a police department with one to 25 DREs is $23,000. This drops to $13,500 for Year 2. For a service with 351 to 500 DREs, the costs are $110,000 for Year 1 and $50,000 for Year 2.

In addition to New York, several other states have implemented the tablet-based application, including Vermont, West Virginia, Ohio, Indiana and Massachusetts. The introduction of the tablet has been well-received, particularly by the younger DREs. Some of the more experienced DREs appear reluctant to abandon the traditional paper and pencil approach. In New York, DREs who are issued a tablet are required to perform four enforcement evaluations in a year. Not surprisingly, among the states with the greatest increases in DRE evaluations in 2016 were those that implemented the tablet application (Hayes, 2017). If this is a general effect associated with the use of this technology across all police services, the cost would be well-justified.

Force Plates for Assessing Balance

The Romberg Balance test requires the officer to observe the subject and assess the extent of postural sway in both the front-to-back and side-to-side planes. The measurement is typically a judgment of the distance from the center position in these two directions. This is a relatively simple approach to measuring sway that has limited reliability and validity.

Over the past 100 years, a variety of approaches have been used to measure postural sway (Stevens & Tomlinson, 1971). Today, many medical, physical therapy, and geriatric settings utilize force plates to assess postural sway. Force plates, which consist of pressure sensors embedded in a small platform, have been used for many years in other fields (medicine, kinesiology, physiotherapy) to measure various aspects of balance. The subject simply stands on the force plate to perform the Romberg Balance test. The test would proceed in the usual manner. As the subject attempts to maintain balance, the pressure exerted on various parts of the feet changes. These pressure changes are automatically captured by the sensors and translated by the software into various measures of balance, including maximum deviations from center, front-to-back and side-to-side sway, area of sway, frequency of movements, etc. These devices have considerable potential to improve reliability and overall accuracy of the measurement of balance.

The number of devices required would depend on the size of the department and the number of evaluations performed. In many cases, one device per station where evaluations are performed would be sufficient. There are many such devices on the market, many of which are capable of much more sophisticated measurements than would be required for this particular application. The specifications for the device would have to be examined and experts consulted to find the type of force plate that would best meet the needs of the DECP. The cost of force plates can vary dramatically depending on the system and the requirements of the situation. It is estimated that the cost of implementation of force plates for a DECP would be $1,500 to $2,500. The most significant costs associated with the introduction of force plates to assess postural sway would involve determining the specific metrics to be used, developing procedures for use of such a device, the establishment of normative data, and officer training. In addition, force plates would require periodic maintenance and calibration checks.

Ocular recorders

The reactions and movements of the eyes are susceptible to the influence of psychoactive substances and provide prime indicators of drug influence (Kosnoski, Yolton, Citek et al., 1998; Griffiths, Marshall & Richens, 1984; Willetts, 1969). Eye movements and reactions are particularly valuable to drug influence evaluations because they are not under conscious control and vary according to drug category. The challenges associated with the use of eye reactions and movements as indicators of drug influence are most often related to their acceptance by the court as evidence. Those who have not studied or been trained in the assessment of eye reactions and movements often have difficulty understanding these measures and how they are affected by drug use. A video taken during the evaluation process can provide a powerful demonstration.

Ocular recording devices consist of a small video camera and monitor along with a chin rest or positioning system to help ensure the subject's eyes are correctly aligned with the camera. The devices record eye movements and reactions such as HGN, VGN, pupil size, reaction to light, rebound dilation, and lack of convergence during SFST and DECP evaluations. Such devices are capable of recording the examiner's instructions and video evidence of eye reactions that can be reviewed by DRE instructors/coordinators and presented in court. These recordings supplement the officer's description of the effects observed.

Ocular recording devices have been designed specifically for DECP evaluations. They are housed in a portable hand-held instrument that rests against the subject's forehead. They are equipped with a pupilometer to measure pupil size, two switchable lights directed at each eye to assess reaction to light, an adaptor to block outside light, and infrared LEDs for operation in near total darkness. Audio and video are stored on a removable Secure Digital (SD) card. Each evaluation can be stored on a separate SD card and copied to a computer hard drive if necessary.

These devices have the potential to streamline and enhance a DECP evaluation and the presentation of evidence in court. The audio/video recording may provide valuable evidence on HGN, VGN, pupil size, rebound dilation, or reaction to light that can be difficult for laypersons to comprehend.

Ocular recording devices are not essential to complete an evaluation. DREs are still required to conduct the eye examinations according to standards; the device merely provides a means to record and facilitate the eye measurements and movements. These devices have a price tag in the $4,000 to $5,000 range. Training in use of the device is not extensive and can be completed in a few hours. A department could have a device available at stations where DECP evaluations are performed. Requirements and costs of periodic maintenance were not available.

Automatic blood pressure measurement

As part of the DEC training, officers are taught to measure blood pressure with a sphygmomanometer and stethoscope. This is a manual procedure that requires the officer to read and record the pressure from a dial at the correct moments. Although not difficult to

learn, the technique is an acquired skill that requires practice. In addition, observing and recording the readings from an analogue dial are subject to error.

Conventional manual measurement is no longer considered to be the best method for evaluating blood pressure (Myers, 2014). Automated blood pressure monitors are currently in use in hospitals and physicians' offices; similar models are available for personal use at home. Automated systems increase accuracy, reduce the probability of situational increases in blood pressures (e.g., "white coat" syndrome), and minimize observer bias (Myers, Kaczorowski, Dawes & Godwin, 2014). These devices also facilitate the recording of multiple blood pressure measurements and can provide an average value of the readings. Many of these devices measure and display pulse at the same time.

Automatic blood pressure measurement devices vary in quality and price. Simple models available for home use can be found for less than $100; higher-end models used in medical facilities can be closer to $1,000. Although most such devices are portable, they contain components that can be sensitive to wear and tear and would be best left (or fixed) in a single location. Periodic service and calibration checks are required to maintain proper functioning and accuracy. The selection of a device for use would require professional medical advice.

The implementation of automatic blood pressure measurement devices would not eliminate the need to provide instruction on the theory of blood pressure and the effects of psychoactive drugs.

Automated body temperature measurement

Body temperature measurements are not difficult to take with a mercury-in-glass thermometer. Accurate measurement requires that the thermometer be correctly placed under the tongue, an appropriate time be allowed for the device to attain equilibrium with the body temperature, and that the thermometer be read and recorded accurately. Concerns about the health and safety risks of such devices (e.g., glass breakage, mercury poisoning) have resulted in the replacement of these types of thermometers in many institutional and professional settings (Blumenthal, 1992).

Electronic digital systems for measuring body temperature have been available for many years. Battery operated oral digital thermometers often include an auditory signal to indicate when the peak temperature has been achieved. This feature helps ensure the thermometer is kept in place until a stable peak temperature is attained. These thermometers also provide a digital reading that virtually eliminates reading errors. Such devices are relatively inexpensive ($20-$40) and require little training.

Body temperature can also be assessed by measuring thermal radiation from the tympanic membrane in the ear with a hand-held probe. This type of thermometer provides an easy means of obtaining a digital reading of body temperature within a few seconds. This technique is currently being used in medical settings and some models are available for home use. A variety of models are available in a range of prices from approximately $25 to $125. Training would be minimal. The selection of a device for use would require professional medical advice.

Studies that have compared temperature readings of different types of thermometers with core body temperature (measured by pulmonary artery catheter) have shown that electronic digital thermometers gave readings that were from 0.198 degrees Celsius below core temperature to 0.188 C above core temperature (Giuliano, Scott, Elliot & Giuliano, 1999; Schmitz, Bair, Falk & Levine, 1995). Tympanic thermometer readings varied by less 0.1C of core temperature and were not significantly different. Dowding (2002) reported variations in electronic digital temperature readings according to the person using the device, suggesting that the use of these devices requires greater training in their use. Tympanic thermometer readers not only did not differ from actual body temperature but showed no differences in the readings obtained by different users.

The research indicates that tympanic thermometers are the most accurate reliable approach for the measurement of body temperature. Although the absolute variation in readings among the different types of thermometers is relatively small, the changes in body temperature associated with drug use are also small. Greater measurement accuracy can help determine whether body temperature differs from average or varies over the course of the evaluation.

Discussion

The implementation of new technologies has the potential to enhance the accuracy of the measurements of vital signs and clinical indicators and improve the efficiency of the overall DECP evaluation. Automated measurements of body temperature and blood pressure are relatively easy and inexpensive additions that require little in the way of additional training. Ocular recording devices provide a means to document and facilitate the understanding and acceptance of the effects of drugs on eye signs and movements that are critical elements of DECP evaluations.

The implementation of force plates presents a major change to the assessment of balance. Although the procedure for conducting the Romberg Balance test would not change, the metrics and their interpretation could involve substantial modification. This would require considerable preliminary work to establish normative data and identify the point at which performance is deemed to vary significantly from the norm and reflect impairment. This would be a longer-term initiative that involves assessing the potential benefits against the financial implications.

The replacement of the traditional pencil-and-paper approach to the recording of DECP evaluations with a computerized tablet application would likely greatly improve the efficiency of recording and reporting of evaluations. The use of tablets also appears to increase the number of evaluations performed.

Technological innovations could bring the DECP into the digital age. In addition to improved efficiency and accuracy, the use of new technologies may enhance the perception of the program as being modern, scientific and credible. The implementation of these technologies, however, is a process that will require consultation with experts in the selection of products as well as the development of standardized procedures and training programs. Pilot programs can be used to provide guidance on how best to utilize new procedures.

The costs associated with implementing changes can be substantial. Acquiring new equipment is the initial investment. Additional costs are associated with regular maintenance and calibration where necessary. More significant are the costs associated with the development of training programs and the retraining of existing DREs in the use of the new equipment.

Key Informant Interviews

Interviews were conducted with individuals involved with, or closely connected to, the DECP in some capacity. These key informants were deemed to have considerable knowledge about various aspects of the program, including its operation, the specifics of the tests and procedures, the strengths and weaknesses of the program, the training program, and/or the requirements for certification. Their experience with the program may also have led them to develop opinions and suggestions on the types of changes/adjustments that could be made to enhance the efficiency and effectiveness of the program.

A discussion guide consisting of a series of questions and themes to be discussed was developed to guide the interviews (Appendix D). The purpose of the guide was to facilitate discussion and was not intended to limit discussion to specific topics or to be a questionnaire to be completed. The objective of the interviews was to capture the thoughts and opinions of a diverse group of stakeholders regarding the perceived strengths and limitations of the DECP along with suggestions for improvement. Any and all aspects of the DECP were open for discussion — e.g., technical issues, legal challenges, administrative issues, training, recertification, and the interaction with other programs such as Advanced Roadside Impaired Driving Enforcement (ARIDE), Standardized Field Sobriety Test (SFST), and oral fluid drug screening.

A list of potential key informants was developed that included DRE state/provincial coordinators, program managers, DREs, toxicologists, prosecutors, defense counsel, and researchers. Thirty-one interviews were conducted by the investigators. Half the interviews were conducted in person, the others by telephone. Interviews lasted anywhere from 30 minutes to more than two hours. Respondents were assured of anonymity and confidentiality and that neither their names nor organizational affiliations would be reported or in any way associated with specific comments. Notes were taken; however, interviews were not recorded. The information and comments collected during the interviews were reviewed and common themes within each of the broad areas were identified. Individual statements deemed noteworthy were flagged as well.

Just over three-quarters of those interviewed were active DREs with experience in the program ranging from one to more than 25 years, with the mean being 12 years. More than half of the DREs included were state/regional coordinators and program managers. The remainder of those who participated in the interviews were prosecutors, defense counsel, and toxicologists.

Strengths of the Program

A common comment was that the DECP was deemed to be the best method available for responding to the problem of drug-impaired driving. It was also noted that the DECP is the foundational component in a more comprehensive system of drug-impaired driving enforcement that includes both the SFST and ARIDE programs. The latter two programs are not only key elements for detection and identification of potential impaired drivers, but they can also serve to develop interest and expertise in this area of law enforcement that can help recruit officers for the DECP training.

It was repeatedly stated that the major strength of the DECP is the systematic and standardized nature of the assessment. The consistency with which all DREs conduct the evaluation provides credibility. The evaluation is not a random or arbitrary collection of tests; rather, the tests were specifically selected to help identify the effects of various types of substances. Many viewed the clinical signs as the most important elements of the evaluation because these are physiological reactions to the presence of drugs that cannot be controlled by conscious effort on the part of the subject. The psychophysical/divided attention tests are important to demonstrate the adverse physical and cognitive effects of drugs that can be linked to driving ability.

Continuing education and recertification were also viewed as elements that serve to strengthen the program. The Annual IACP Training Conference on Drugs, Alcohol, and Impaired Driving, at which attendees are credited with professional development hours, was given high marks by many of those interviewed. DREs are required to take advantage of educational and professional development events and opportunities where available to advance their knowledge and improve their skill sets.

Although sometimes seen as being slow to approve and implement changes, the Technical Advisory Panel (TAP) makes ongoing improvements to the program to add value and help keep it current with new knowledge, advanced techniques, and improved tools. For example, recent changes have included the introduction of ultraviolet penlights for use in assessing pupil size in near total darkness, modifications of the standards for average pupil size, and the change in the assessment of hippus to an observation of pupillary unrest. Continual improvement is an essential component of a strong program.

While the DECP's primary purpose lies in the investigation of drug-impaired driving cases, many interviewees indicated that the DECP training was valuable in many other areas of law enforcement that involve people who are under the influence of alcohol and/or drugs and may present a danger to themselves, the officer, or others. The ability to recognize the signs and symptoms of drug use can be of tremendous value in assessing situations such as domestic violence or where personal injury is involved.

A DRE's expertise on drugs and their effects is also utilized in educational settings. DREs are often active in school and community settings. Giving presentations to students, parents, teachers and others about the specific acute and chronic effects of different types of drugs adds tremendous value and credibility in these settings. Ancillary programs such as DITEP (Drug Impairment Training for Educational Professionals), in which DREs serve as instructors, provide training to teachers and school administrators as well as those who deal with youth in other settings to help them recognize the signs and symptoms of drug use so that they might be in a better position to deal with those involved in a safe and professional manner. For example, probation settings and situations involving injury are other areas where knowledge of drug effects can be beneficial to those at the scene. In fact, emergency medical technicians have benefited from DECP training in drug use and effects.

The growing legalization of cannabis and the opioid crisis were noted in various contexts throughout the interviews. In many cases, respondents noted the media attention to these issues often highlighted the DECP and its beneficial role and value in responding the problems created. The publicity was generally seen as serving to raise public awareness of the DECP and enhance public perceptions of the DECP as a credible and dependable law

enforcement program. The media attention also helped highlight the need for an expansion in the number of officers trained as DREs.

Program Oversight

The purpose of oversight is to maintain program fidelity by ensuring that those involved in the application and management of the DECP uphold the standards and integrity of the program. Oversight of the program involves supervision and management and can occur at various levels — test administration, evaluation review, the local department, and at the state and national levels.

The majority of respondents expressed a strong level of satisfaction with the degree of program oversight. As with any large program, variations in procedures can begin to creep into everyday practices. In the absence of appropriate correction from those in positions of authority, even minor variations can become detrimental. Oversight begins with monitoring the quality and consistency with which evaluations are performed. Instructors and coordinators need to spot potential issues in evaluations and bring them to the attention of the DRE for immediate remediation. The potential for procedures to "drift" was noted as a potential problem when oversight becomes lax. If there is an area that could benefit from greater oversight, it is at the level of the evaluation, ensuring the fidelity of test administration.

It was often expressed that the states had a good working relationship with the IACP. IACP program managers were available and willing to assist with technical and management issues when requested. IACP plays a crucial role in overseeing the operation of the program and ensuring its integrity.

There were a couple of instances where concern was expressed about the apparent disconnect between local coordinators/managers who were not trained in the DECP and DREs who were active in conducting evaluations. This would appear to be an issue restricted to particular areas or regions. Respondents felt strongly that DRE coordinators/program managers should be experienced DREs who understand and support the program.

The TAP was viewed as another level of oversight to the program. Any changes to the procedures or interpretations of test results must be vetted and approved by an expert panel to ensure that any modifications to the program have merit and that all DREs are informed of these changes through their state and local program coordinator.

Limitations of the Program

Interviews with defense counsel questioned the scientific basis of the DECP, referring to it as "voodoo" or "junk science". At the same time, however, it was acknowledged that there is a need to do something about the increasing drug-impaired driving problem and that the DECP is the best program available to deal with it.

One of the most frequently mentioned areas limiting the DECP involved recruitment, training, and attrition. The DECP training is not for every police officer. The training is challenging and demanding. It requires commitment and dedication and a great deal of hard work. Not everyone succeeds. Not surprisingly, it attracts many of the best officers,

who are also most likely to be successful in other areas of law enforcement and are often promoted. These promotions can remove them from active service as a DRE. In the absence of ongoing recruitment into the program, the number of DREs decreases.

In addition, it was noted that some DREs are not particularly active in terms of the number of evaluations completed. There can be several reasons for this, including reassignment and/or promotion. Nevertheless, the number of DREs in a department can be viewed as an indicator of the capacity to deal with drug-impaired driving. Officers who are not active (referred to as "paper" DREs) enhance the perception of capacity but may, in fact, be limiting the ability for others to be trained.

Recruiting officers into the DECP can be challenging as well. The DECP training is well-known as a tough course and not everyone is successful. Officers need to demonstrate not only the interest but an active involvement in impaired driving enforcement to be considered. This should involve proficiency in the SFST and ARIDE training. These courses are often viewed as the prerequisites for DECP training. Some suggested that before being accepted into the DECP training, officers should commit to remaining active for a minimum number of years (e.g., five years) and/or until a specified number of evaluations have been completed.

Several of those who participated in the interviews noted that as departments seek to increase the number of DREs, there needs to be an emphasis on quality over quantity. Potential recruits for the DECP training should be assessed for interest, motivation, ability, and commitment prior to being accepted into the training program.

There was some concern expressed about the cost of the training. Taking an officer out of service for two weeks to be trained is a significant expense. In smaller communities, officers typically have to travel to attend training. When certification cannot be completed locally, it is often conducted over a series of consecutive days in Phoenix or Jacksonville, Florida. Although there is an overall efficiency to this approach, there is a cost associated with travel to one of these locations. The public perception of sending police officers to popular "vacation" destinations needs to be countered with information about the reality of the number of hours and the extent of effort required for training and certification.

The DECP training itself was the subject of numerous comments. It is a very intense course that involves a great deal of difficult material that has to be learned in a short period of time. There are also skills to be acquired that require practice to establish proficiency. It was suggested that the course, currently two-weeks, be lengthened or perhaps broken into two parts.

Certification is also an intense process. At times the goal seems to be to get the required number of evaluations done rather than learning something in the process. While acknowledging the efficiency of conducting certifications in a concentrated period of time in locations with access to drug-using populations, several people indicated that this approach was not conducive to optimal learning. Spreading the certification process over a longer period of time would provide greater opportunity for interactions with instructors and would be beneficial to skill development. The logistical and operational challenges associated with other approaches to certification were noted. For example, having DRE trainees conduct evaluations for certification during regular police shifts would most likely be given low priority. Identifying an individual who has ingested drugs and obtaining the

individual's permission to participate in an evaluation for training purposes would have to be balanced against ongoing police duties. In addition, this must be coordinated with the availability of a DRE instructor to observe the evaluation. The inherent inefficiencies of this approach have the potential to extend the certification process over a long period of time. Alternatively, conducting certifications in two separate sessions (rather than one) would provide greater opportunity for review and study but would add considerably to the costs associated with time and travel.

One of the most frequently reported limitations of the DECP involved the court system. The involvement of a well-informed and motivated prosecutor was seen as a critical element in a successful case. The evidence provided by a DRE is often not well understood by the courts. It is subject to legal challenges and interpretations that don't necessarily fall in favor of the prosecution. Juries can have a difficult time comprehending the significance of various pieces of evidence, including the clinical indicators (e.g., the eye examinations) and the results of toxicology. It was reported that prosecutors are often not well-versed in the DECP and their knowledge of drugs and drug-impaired driving in particular is limited. It also cannot be assumed that the judiciary has knowledge of the DECP and the types of evidence provided by the 12-step evaluation. Drug-impaired driving cases are seen as challenging and there are few prosecutors who are prepared to accept the challenge.

It was also noted that DREs are not always well-prepared to testify in court. Drug-impaired driving trials can involve complex and detailed issues. DREs are required to present key evidence in a straightforward and convincing manner. This can be improved with experience but additional training would also prove beneficial.

On the other hand, the important roles played by prosecutors and toxicologists in supporting the DECP were acknowledged. Several respondents noted that prosecutors and toxicologists who have attended the DECP training or have otherwise acquired advanced knowledge about the program have a better understanding of the process and are generally able to present testimony that is supportive of the DRE findings.

The Technical Advisory Panel (TAP) was viewed as providing expertise and credibility to the DECP. However, as noted earlier, TAP is often perceived as slow to implement program modifications and improvements. In addition, some of the changes are made without adequate explanation to DREs.

Potential Enhancements

Interview participants provided many suggestions for potential enhancements to the DECP. Many of these ideas were directly related to the limitations of the program and for the most part represent obvious corrections or improvements for the issues noted. Other suggestions were unrelated to any expressed limitations of the program but appeared to be generated in the spirit of enhancing some aspect of the DECP.

One of the most frequently mentioned potential enhancements was the need for new, updated validation studies. The original validation studies are viewed as dated, yet there are no newer studies to rely on. Defense counsels have become well-versed on the problems associated with the original validation studies and exploit the limitations relentlessly. Newer studies could correct for some of the shortcomings of the older ones by providing stronger, more contemporary evidence. Additional research on some of the individual

components of the evaluation was also suggested as a means to strengthen and update the evidence base for the program. For example, although the Finger to Nose test is included in the DECP protocol, research is needed to establish a standard scoring system with validated clues. Research should also evaluate relative changes in pupil size under different lighting conditions, and there needs to be greater understanding of the incidence and circumstances surrounding observations of HGN following cannabis use.

Another area for improvement that was repeatedly mentioned was providing more opportunities for local workshops to build on the knowledge and skills of DREs. The national conference was viewed as a tremendous learning experience but typically limited funds are available and only a few DREs from a given department are able to attend. A suggestion was made to hold regional workshops or conferences that would be accessible to a greater number of DREs. These events could include some of the same topics as the national conference but might also include local issues such as changing patterns of drug use, state legislation and the implications for enforcement, and discussions of specific local cases.

There were several comments related to training. Some would like to see the training extended over a longer period of time, possibly dividing it into two parts separated by a week or two to allow trainees time to study and prepare for the next session. Others thought that some parts of the training program were covered in greater depth than necessary and the key information was sometimes lost in the excessive detail.

Encouraging the use of experts and professionals as instructors who are not DREs was mentioned as a way to improve training. Some DRE instructors are not comfortable with certain topics that might be better left to experts in the particular field. For example, prosecutors, toxicologists, and ophthalmologists could provide instruction on topics in their area of expertise. The inclusion of these subject matter experts could elevate the level of training, add some variety to the training, and provide a valuable resource for addressing questions and furthering discussion.

New Technologies

Opinions were mixed when interviewees were specifically asked about the potential introduction of new technologies. Some older, more experienced respondents were of the opinion that the traditional way of doing things worked just fine and didn't need to be changed. There was also concern expressed about the cost of new technologies, the need for changes to the training program, retraining existing DREs, the requirement for calibration checks, potential court challenges, and additional defense counsel requests for disclosure of maintenance records on new equipment.

On the other hand, other (often younger) respondents were open to the idea of bringing new technology into the evaluation process. In particular, tympanic thermometers and automatic blood pressure devices were frequently cited as new equipment that could be implemented to improve the accuracy and efficiency of the evaluation. The replacement of pencil and paper with tablets to record the evaluation was mentioned by several respondents. They perceived this to be the wave of the future, providing additional options for improving the recording and reporting of evaluations.

There was also considerable discussion of the implementation of oral fluid drug screening and/or testing as a new technology that could be introduced. At times it wasn't clear whether the respondent was referring to oral fluid screening, which provides an indication of the presence of certain drugs, or oral fluid drug testing, which requires a sample of oral fluid to be tested at a toxicology laboratory as a potential alternative to blood or urine testing to provide evidence of the category of drug ingested.

Oral fluid drug screening is being pilot tested in some jurisdictions to help identify drivers who may have a specific substance(s) in their system.[10] A swab or collection device is either wiped along the tongue or placed in the mouth for a brief period and then analyzed on-site to indicate the presence of a number of common substances. A positive result provides an indication of the presence of one or more substances and could lead to the subject being required to undergo evaluation by a DRE. The use of oral fluid screening has the potential to add efficiency to the initial investigation of suspected impaired driving by providing objective evidence of drug use at roadside, prior to arrest. Oral fluid screening would not necessarily be part of the 12-step evaluation.

Oral fluid testing for drugs involves the collection of a specimen that is submitted to a toxicology lab for analysis. The major advantage of using oral fluid as the sample medium following evaluation by a DRE is that it can be readily and easily collected at the time of the evaluations, eliminating the delay in getting a blood sample drawn. Oral fluid is also superior to urine as a test medium in that it typically detects active drug products rather than metabolites and better reflects recent drug use (Bosker & Huestis, 2009). Using oral fluid would, however, need to be assessed by TAP with expert advice from the toxicology community to ensure the validity and acceptability of the results.

There was also mention of moving Step 12 of the evaluation (i.e., the collection of the bodily fluid specimen for analysis of drug content) to Step 1. The rationale for the suggestion is to obtain a specimen as soon as possible to limit the inevitable decrease in drug levels associated with the time lag between arrest and the collection of the specimen following the evaluation. It was indicated that there is often sufficient evidence of driving and/or behavior from the arrest to justify a warrant for the specimen.

Discussion

Despite the numerous critiques and expressed limitations of the DECP, the overall impression from the interviews is that the program has a strong foundation. It was obvious that those involved in the DECP are passionate about the program and their work. Their commitment and dedication come through in everything they say about the program, including their comments about perceived limitations and suggestions for improvement.

Drug-impaired driving is a challenging issue from many perspectives. The DECP is considered by those interviewed to be the premier approach for dealing with this problem. Even as new detection technologies such as oral fluid drug screening become increasingly available, interviewees anticipate the DECP will continue to be needed to assess suspected impaired drivers to provide detailed evidence of the nature and extent of drug effects. In

[10] In the interest of clarity, supplemental information on oral fluid screening and testing has been added by the authors.

addition, as the legalization of cannabis continues to spread across jurisdictions, the need for a strong DECP program may become increasingly more evident.

Those closely involved in the DECP provided evidence of the considerable support for the program and expressed their commitment to ensure it remains strong and vibrant. They provided numerous suggestions for the continued development and enhancement of the program. Evolution of the techniques, methods, and tools used in the evaluation was generally viewed as a means to help keep the program current and relevant.

Data Analysis

A fundamental premise of the DECP is that drug use and impairment can be identified through a systematic assessment of individuals using a standard set of tests, measurements and observations. This premise is based on the understanding that different types of psychoactive drugs affect brain functioning in different ways and thereby produce distinct patterns of signs and symptoms that vary by drug type (Julian et al., 2008). The extent of the observed effects will also vary between and within individuals according to factors such as the dose administered, pharmacokinetics of the specific substance, elapsed time from administration, extent of prior use, existing medical conditions, and concomitant use of other substances. Nevertheless, research on the validity and accuracy of the DECP indicates that the systematic assessment of suspected drug users can be used to identify the type of drug ingested (Beirness et al., 2009).

The common signs and symptoms of each of the seven categories of drugs are summarized in the DECP Matrix (Appendix B). As noted previously, the matrix is only intended to be a quick reference guide of some of the effects associated with broad categories of drugs. Even a cursory glance at the matrix reveals that in many cases, specific indicators are common across two or more categories of drugs. For example, horizontal gaze nystagmus is associated with the use of CNS depressants, inhalants and dissociative anaesthetics but not other categories; a slow pulse is typically an effect related to the use of CNS depressants and narcotic analgesics but not other categories. In many cases, the indicator is expected to be "normal" or not present. Hence, the pattern of signs and symptoms is an important aspect of the evaluation that allows the officer to suspect a particular drug category.

During the course of a DEC evaluation, well over 100 separate pieces of information are collected, including vital signs, eye signs, interview responses, and psychophysical/divided attention tests. Combined with the toxicological confirmation of the category of drug ingested, these data provide a rich source of information about the signs, symptoms and impairing effects of various types of substances. This section examines the data from a large number of DECP evaluations to illustrate the effects of different drug categories.

Database of DECP Evaluations

Over the past several years, a database of 5,920 DECP evaluations has been created. The initial set of 1,349 evaluations was provided by the Canadian DRE Coordinator to assess the degree of correspondence between the opinion of the evaluator on the category of drug ingested by the suspect and the results of the toxicological analysis of the specimen collected (Beirness et al., 2009). All data elements from the drug-influence evaluation face sheets, the narrative reports prepared by the evaluating officer, and toxicology reports were subsequently coded and entered into a database to examine other research questions (Porath-Waller & Beirness, 2009). Additional evaluations were collected as part of a project for NHTSA and a project examining cannabis use by drivers (Logan et al., 2016). This latter project included 301 evaluations involving only cannabis cases that had blood toxicology provided by IACP[11] along with evaluations provided by DRE coordinators in 15 states[12] and

[11] These evaluations were provided by C. Hayes from IACP and were used in a separate project (Hartman et al., 2016).
[12] The states were: AZ, CA, CO, IN, MD, MI, MN, MT, NC, NM, OK, PA, TX, WA and WI.

Canada. In addition to cases involving each of the seven drug categories, the database includes cases involving several common drug combinations (e.g., opioids and stimulants, depressants and stimulants). Each case includes the results of all clinical and psychophysical tests administered as part of the DEC evaluation, the opinion of the evaluating officer as to the drug category (or categories) responsible for the observations, and the results of toxicological tests for the presence of drugs. No personal and potential identifying information (e.g., suspect and officer names, location, date of birth) was included in the database. The cases in the database do not represent a random sample of DEC evaluations. The drug categories with sufficient cases for analysis were central nervous system (CNS) depressants (n = 741), CNS stimulants (n = 762), narcotic analgesics (n = 461), and cannabis (n = 1,605).

The database also includes a set of 437 drug-free cases. This latter set of cases represents a collection of evaluations performed on volunteers, typically as part of certification sessions or following public education seminars, workshops or conference sessions about the DECP. Some people who have not been using drugs might display indicators that resemble those associated with drug use and/or not perform well on some of the tasks included in the evaluation. Such findings might simply be attributable to individual differences or possibly other physical/medical conditions that could cause vital signs to vary or result in poor psychomotor performance. The inclusion of a set of drug-free evaluations in the database allows comparison of the performance of subjects who have ingested a substance that can adversely affect their ability to operate a vehicle with that of drug-free subjects.

The sample of drug-free cases does not represent a random or representative sample. The average age of the drug-free comparison group (M = 33.4, SD = 11.6) was just over a year older than the drug-positive cases (M = 32.1, SD = 12.6). Although the difference is statistically significant ($F(1, 3612) = 4.19$, $p = .04$), the absolute magnitude of the difference is unlikely to be of consequence. Males were considerably more prevalent than females among both the drug-positive cases (72.9%) and the drug-free comparison group (82.9%) (χ^2 (1, N=3951) = 20.0, p<.001).

This section examines the individual indicators from the set of DECP evaluations according to drug category and explores patterns of indicators that distinguish various drug categories.

Results

Signs and symptoms of drug use

The first series of analyses examined the individual signs and symptoms of different categories of drugs. These bivariate analyses on the vital signs and eye indicators would be expected to be consistent with the effects noted in the DECP Matrix.

A series of bivariate analyses were performed to examine differences in various drug signs and symptoms according to drug category. Only the single drug categories of CNS depressants, CNS stimulants, narcotic analgesics, and cannabis had sufficient cases to be included. The no-drug cases were included for comparison. The analyses examined the effects listed in the matrix as well as several other of the indicators assessed during a DEC evaluation.

Figure 1 illustrates the differences between drug categories on horizontal gaze nystagmus (HGN). HGN was determined by showing four or more clues of nystagmus (i.e., lack of smooth pursuit, distinct and sustained nystagmus at maximum deviation, nystagmus prior to 45 degrees in both eyes). Subjects who had ingested CNS depressants were most likely to show HGN. Each of the three validated clues of HGN showed a similar pattern of effects. HGN is known to be an effect of alcohol (Burns and Moskowitz, 1977) and other depressant drugs (American Psychiatric Association, 2013). HGN occurs naturally in a small proportion of the population even in the absence of these substances (IACP, 2016).

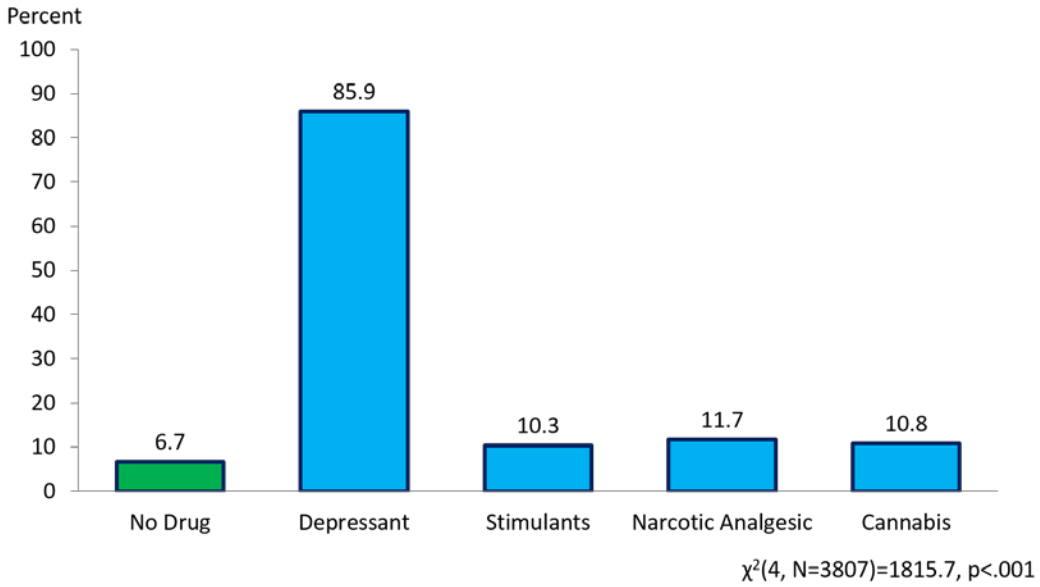


Figure 1. Horizontal Gaze Nystagmus (4+ Clues) by Drug Category

Figure 2 displays the percentage of subjects who displayed vertical gaze nystagmus (VGN) in each of the drug groups. VGN is typically only evident in cases that display HGN and a high dose of the drug is present. As expected, depressant cases were most likely to display VGN.

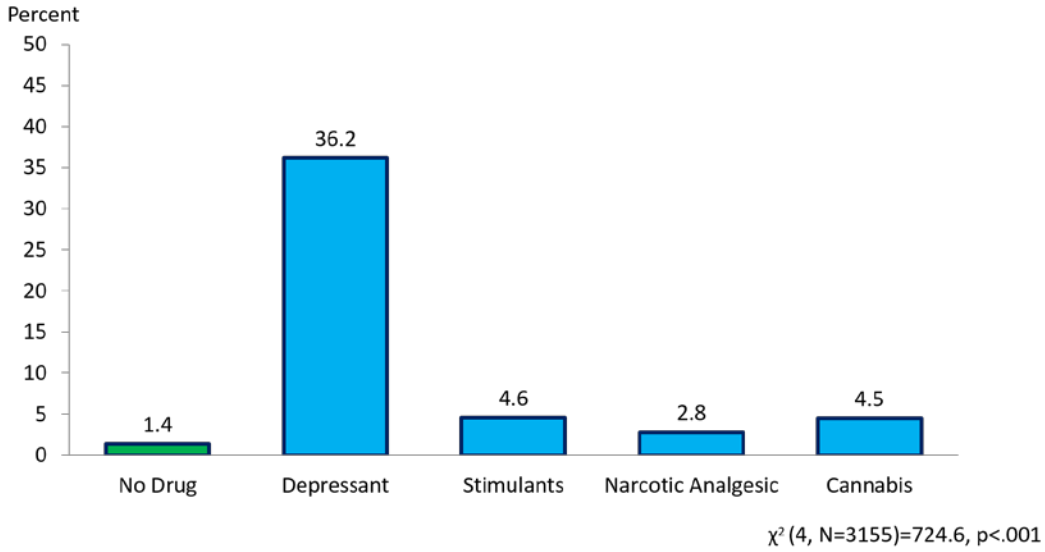


Figure 2. Vertical Gaze Nystagmus by Drug Category

Figure 3 shows the percentage of cases that displayed lack of convergence according to drug group. Lack of convergence is the inability of a person's eyes to converge, or "cross," as the person attempts to focus on a stimulus as it is moved slowly toward the bridge of the subject's nose. The inability to cross one's eyes is not uncommon in the general population. However, as evident in Figure 3, the use of some types of drugs (CNS depressants, cannabis) interferes with the ability to converge one's eyes.

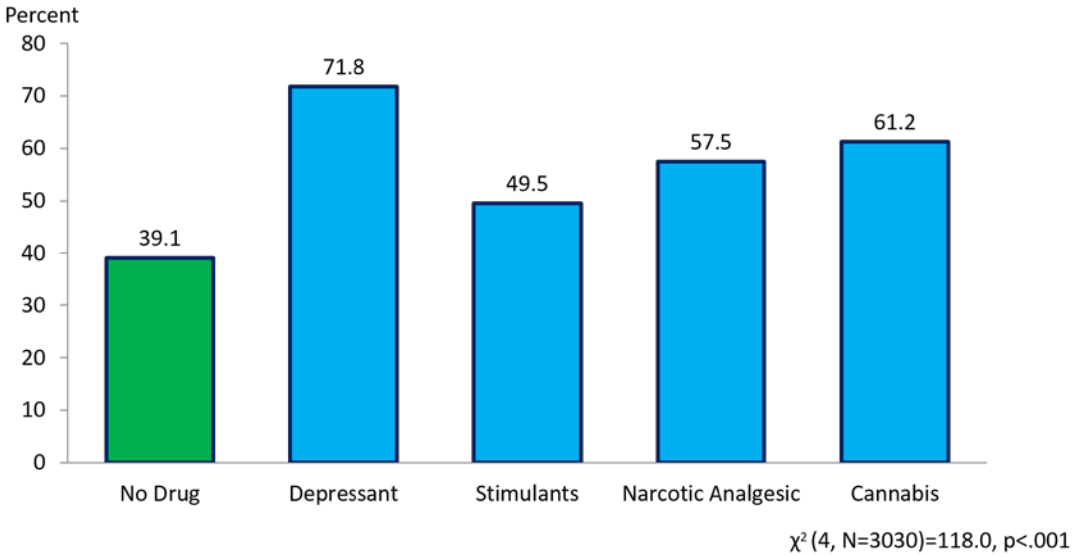


Figure 3. Lack of Convergence by Drug Category

Pupil size under various light conditions is also an indicator that can be used to help identify various drug categories. Figures 4, 5 and 6 show the average pupil size according to drug category under three different light conditions — room light, direct light, and near total darkness. Of these three light conditions, room light is the most variable condition under which to observe pupil size because the evaluator typically has little control over it. Nevertheless, Figure 4 shows significant differences in average pupil size between the drug

groups under room light, with cannabis having the largest average pupil size and narcotic analgesics showing the smallest.

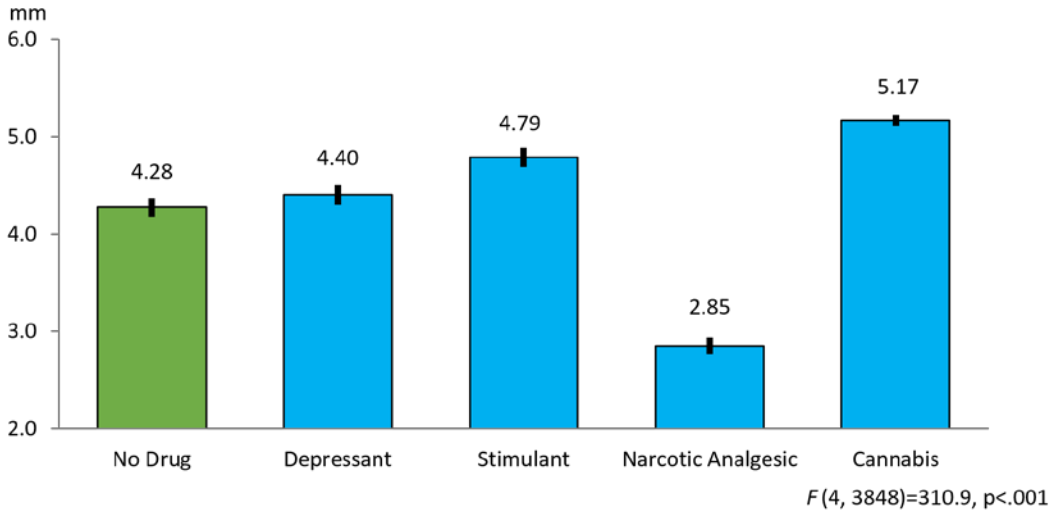


Figure 4. Pupil Size in Room Light by Drug Category

By shining a light directly into the subject's eyes, the light condition is more controlled and one would anticipate that the pupils would constrict under this condition. Figure 5 shows considerable differences in the average pupil size in direct light across the various drug groups. The cannabis group has the largest average pupil size and the narcotic analgesic group has the smallest.

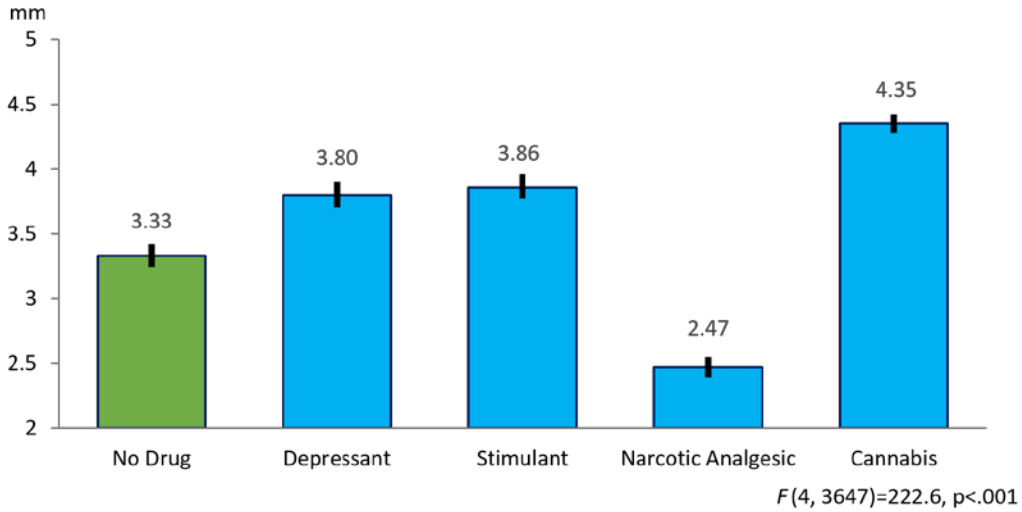


Figure 5. Pupil Size in Direct Light by Drug Category

The evaluator also has greater control in the near total darkness condition, especially since the introduction of the small ultraviolet lights to examine pupils under dark conditions. As evident in Figure 6, the largest mean pupil size was again evident in the cannabis group and the narcotic analgesic group was the smallest.

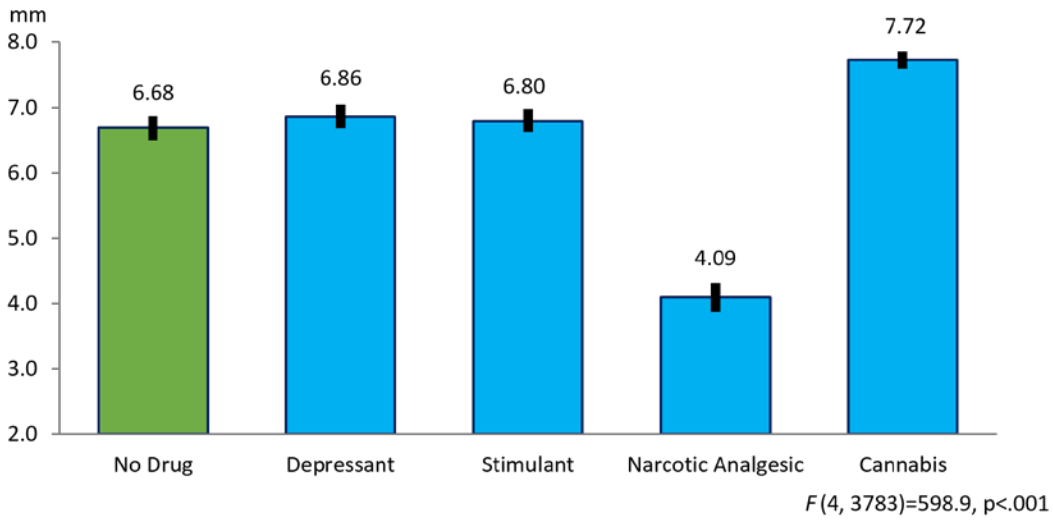


Figure 6. Pupil Size in Near Total Darkness by Drug Category

Typically, pupils will constrict quickly with the introduction of a light and return to normal relatively quickly when the light is withdrawn. The speed with which this reaction occurs can be another indicator of drug presence. It is a subjective judgment on the part of the evaluating officer and requires experience to distinguish among the three categories of speed of reaction — slow, little to no reaction, or normal. Figure 7 shows clear differences in the speed with which a subject's pupils react to light according to the category of drug ingested. Whereas those under the influence of CNS depressants or stimulants often show slow reaction to light, those who are positive for narcotic analgesics show little or no reaction. Cannabis generally has little effect on reaction to light.

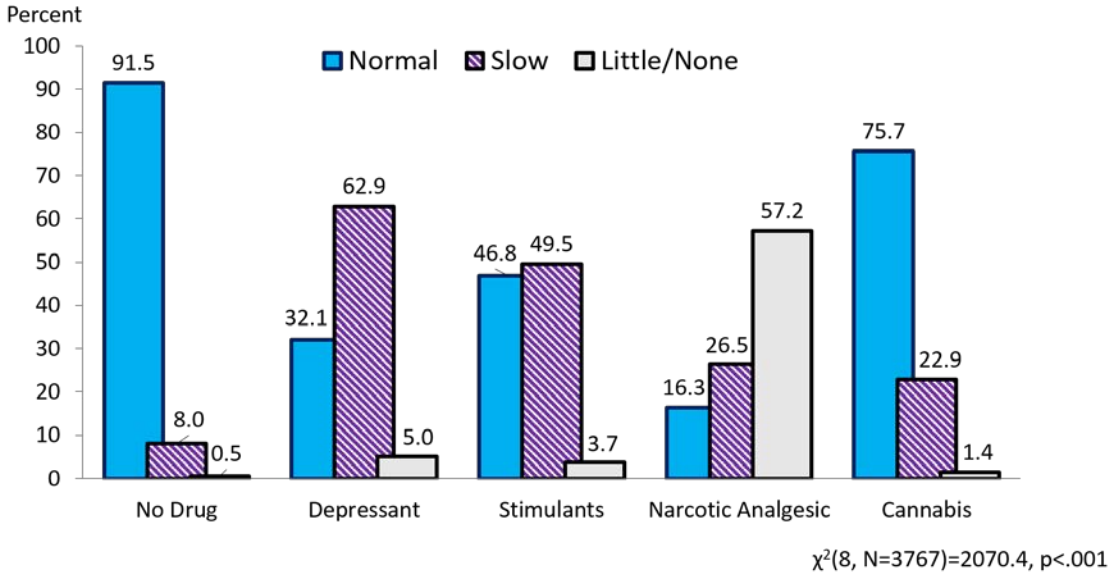


Figure 7. Reaction to Light by Drug Category

In the presence of direct light, the pupils will normally constrict and remain constricted until the light is withdrawn. In some cases, the pupil will initially constrict and then start to enlarge (dilate), sometimes in a pulsing manner, while the light is still being shone into

the eyes, an effect known as rebound dilation. Figure 8 displays the percentage of subjects in each drug category who displayed rebound dilation, which is most commonly observed among subjects who test positive for cannabis.

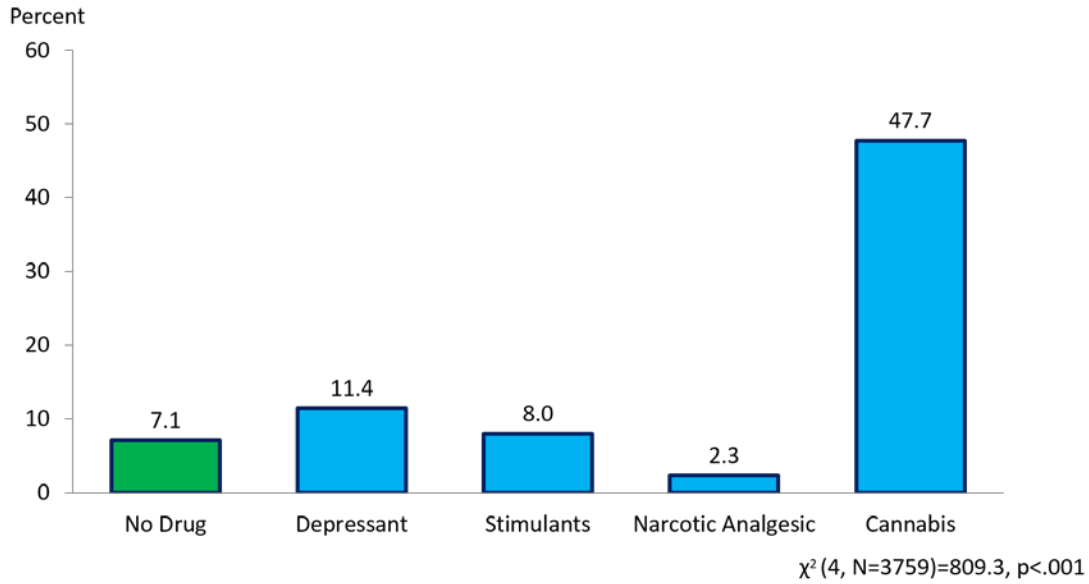


Figure 8. Rebound Dilation by Drug Category

In addition to effects on the eyes, drugs can also have effects on the cardiovascular system, including heart rate and blood pressure. Figure 9 shows the average pulse rate (beats per minute) according to drug category. Pulse is measured three times throughout the DEC evaluation. In Figure 9, these three measurements were averaged for each subject and then averaged within drug category. Although the absolute differences between drug categories are relatively small, they are statistically significant. Stimulants and cannabis are associated with elevated pulse. The narcotic analgesics group had an average lower pulse than the cannabis and stimulant groups but there was little difference with the no-drug group. Examination of the three separate pulse measurements revealed a similar pattern.

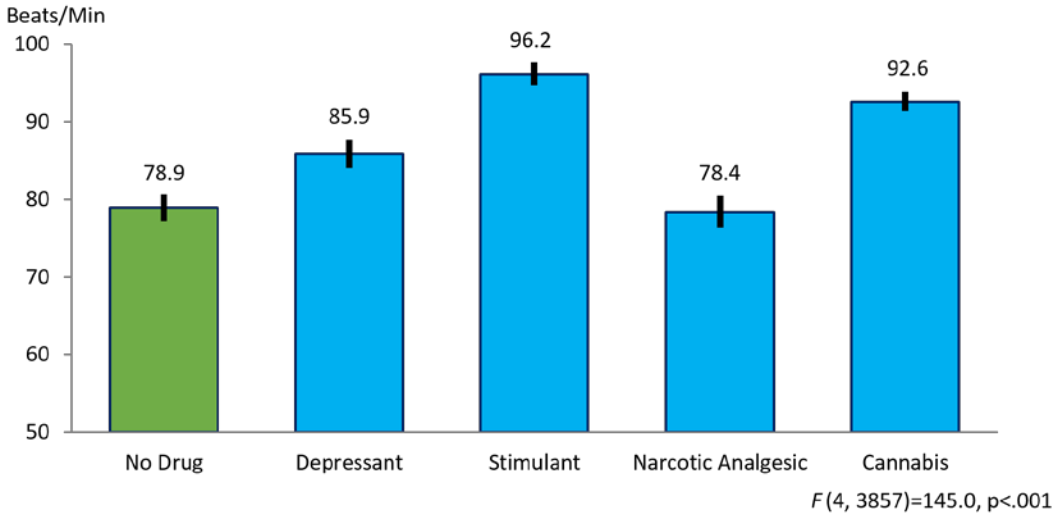


Figure 9. Mean Pulse by Drug Category

Blood pressure was assessed as either normal, high or low by comparing systolic and diastolic blood pressure against age norms.[13] If either systolic or diastolic pressure was above the normal range for the person's age, blood pressure was considered to be high; if either systolic or diastolic pressure was below the normal range for the person's age, it was deemed to be low. Otherwise, blood pressure was considered to be normal. Figure 10 illustrates the differences in the proportion of subjects in the three ranges of blood pressure according to drug category. Subjects who tested positive for cannabis showed the highest proportion in the high blood pressure category followed by those in the stimulant category. Low blood pressure was most common among those who had ingested narcotic analgesics, followed by CNS depressants.

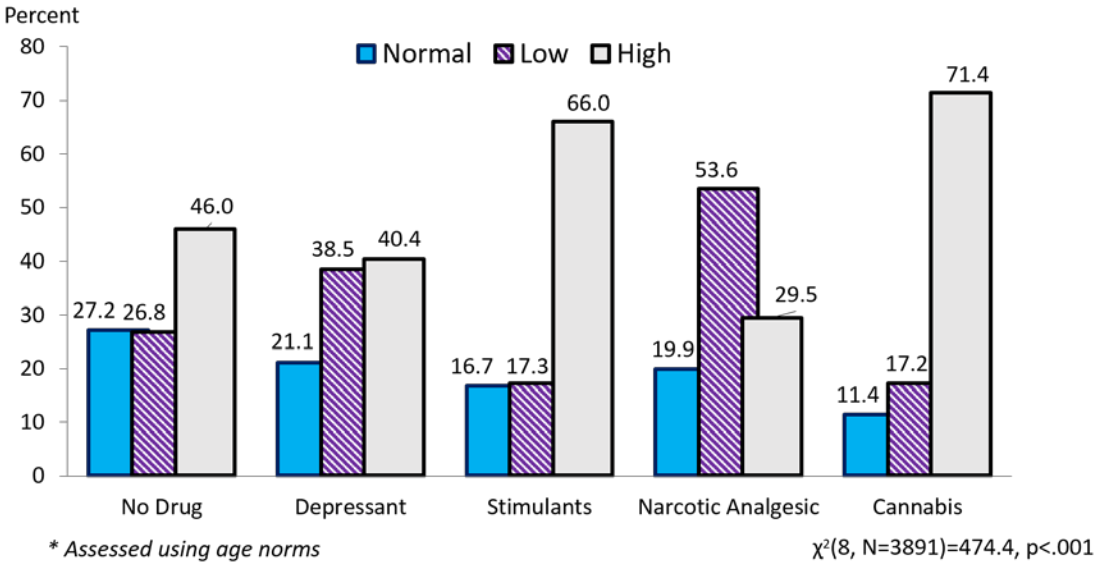


Figure 10. Blood Pressure Groups by Drug Category*

Figure 11 shows the average body temperature of subjects within each of the drug groups. Differences across the drug groups are relatively small but statistically significant. It is also of interest to note that the mean body temperature in each group is below 37 degrees Celsius, but most are within the normal range (i.e., 36.5 – 37.5). The small range of average temperatures within drug groups would suggest that body temperature would have limited ability to easily discriminate among drug groups on an individual case basis.

[13] The table of blood pressure age norms (including minimum and maximum values) was obtained from *www.idealbloodpressureinfo.com*.

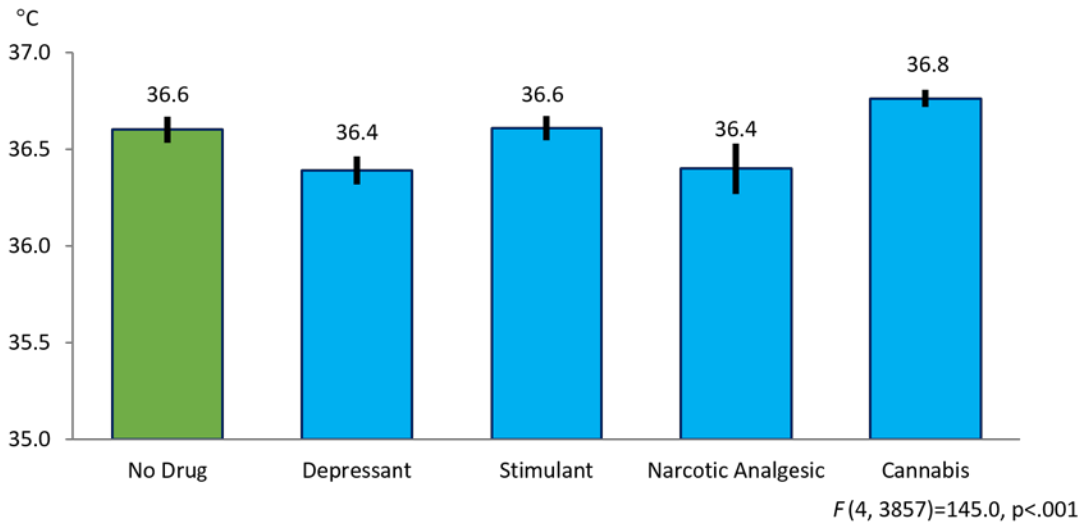


Figure 11. Body Temperature by Drug Category

Droopy eyelids are a symptom of drug use often noted by evaluating officers. Figure 12 shows the percentage of subjects in each drug category who displayed droopy eyelids. The symptom occurs most frequently among those who tested positive for narcotic analgesics and depressants. Almost half of cannabis positive subjects also displayed droopy eyelids.

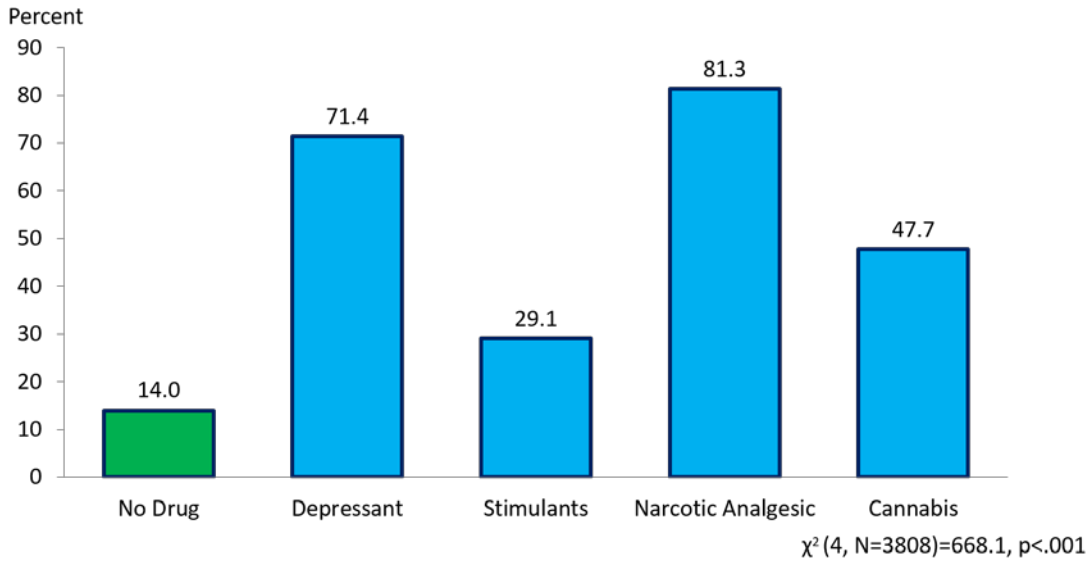


Figure 12. Droopy Eyelids by Drug Category

Figure 13 shows the percentage of subjects in each drug category who displayed reddened conjunctiva – a general reddening of the white part of the eye. This symptom is clearly most commonly associated with cannabis use, with 40% of those in the cannabis group displaying this symptom.

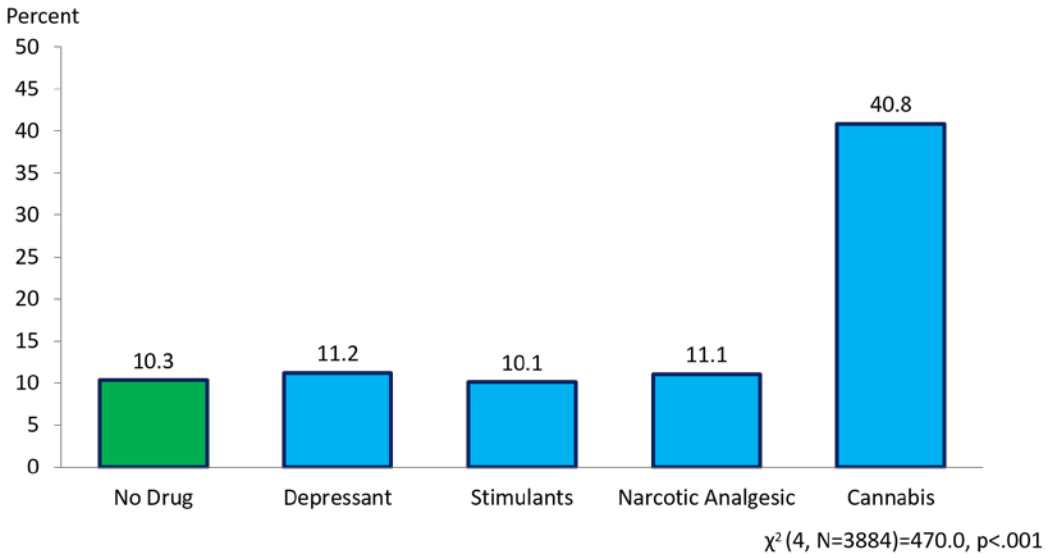


Figure 13. Reddened Conjunctiva by Drug Category

Figure 14 shows the proportion of subjects in each drug category who displayed eyelid tremors. Eyelid tremors are often observed during performance of the Modified Romberg Balance test and/or the Finger to Nose test when subjects are instructed to close their eyes and tip their head back slightly. Eyelid tremors were most commonly observed among those who were positive for cannabis, followed by those positive for stimulants.

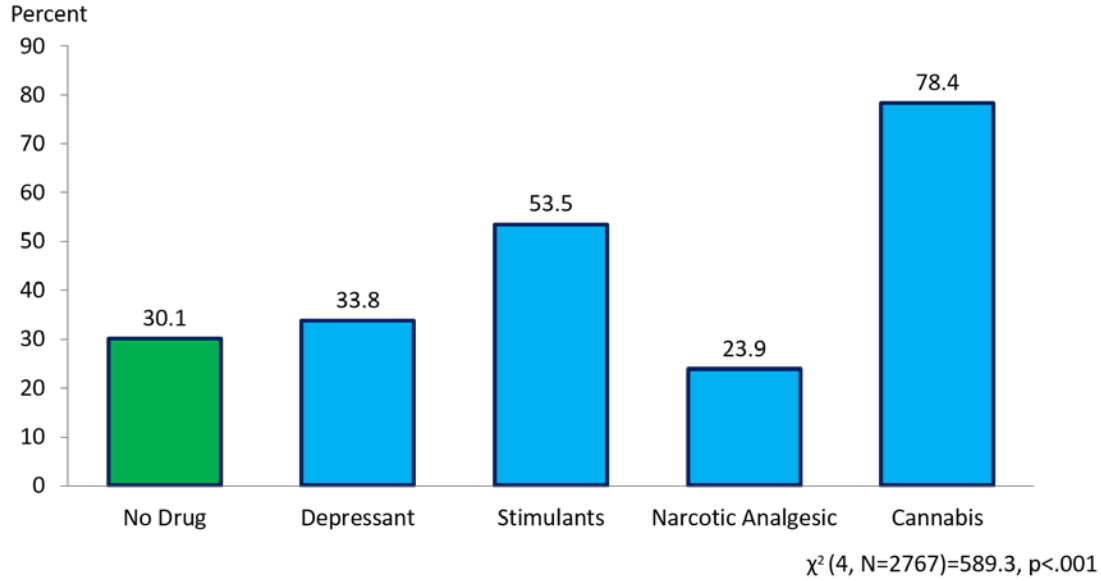


Figure 14. Eyelid Tremors by Drug Category

The bivariate analyses illustrate clear differences in many of the signs and symptoms assessed in a DECP evaluation according to the type of drug ingested. In most cases, the signs and symptoms associated with the drug categories examined were consistent with those listed in the DECP Matrix. In reviewing the results, it is apparent that some signs and symptoms are most prominent in those who have ingested a particular type of drug. For example, HGN was most prevalent among those who had ingested depressants,

constricted pupils in room light was most evident in those who had ingested narcotic analgesics, and rebound dilation was associated with cannabis use. Other signs were prominent in more than one drug category. For example, pupillary dilation and elevated blood pressure were evident in both stimulant and cannabis users; slow reaction to light was most common among those who had used depressants or stimulants. At the same time, the analyses revealed that while users of a particular substance are more likely to display a particular sign or symptom, not all users of that substance display the sign or symptom. Hence, reliance on a single sign or symptom to determine the category of drugs involved could easily lead to errors. Determining the category of drugs involved requires DREs to review all elements of an evaluation to identify patterns of signs and symptoms.

Patterns of Drug Signs and Symptoms

The data were also used to examine the patterns of signs and symptoms associated with particular drug categories. The overall purpose of this series of analyses was to identify the signs and symptoms that were the strongest predictors of a given drug category. A series of binary logistic regression analyses was conducted on the set of DECP cases to determine the prediction of four categories of drugs (CNS stimulants, cannabis, CNS depressants and narcotic analgesics) from the drug-related signs and symptoms assessed during an evaluation. Separate analyses were conducted for each of the four drug categories. Signs and symptoms that were included in the model for all analyses were grouped into three conceptual groups:

1. *Clinical indicators* (i.e., systolic blood pressure, body temperature, mean pulse rate, muscle tone, body tremors and leg tremors).

2. *Performance on the psychophysical tests* (i.e., total number of clues on the HGN test, total sway during the MRB test, and number of misses on the FTN test).

3. *Appearance and physiological response of the eyes* (i.e., condition of the eyes [e.g., normal, reddening of the conjunctiva, bloodshot, watery], condition of eyelids [e.g., normal, droopy], eyelid tremors, convergence, pupil size in room light, darkness and direct light, rebound dilation, and reaction to light). The different dimensions of eye condition were coded individually; with the exception of "normal", the categories were not mutually exclusive.

These three categories reflect logical groups of indicators. The clinical indicators are primarily vital signs of bodily functions. They are used to provide clues to the type of substance ingested. The psychophysical tests are measures of balance, coordination, and divided attention typically indicative of impairment. HGN is included in this group of indicators because it is used as part of the SFST as an indicator of impairment. The appearance and physiological response of the eyes provide a unique set of indicators associated with the use of various types of drugs. Separating the indicators into these three groups also reduced the number of variables in each analysis, thereby increasing the power of the analysis and reducing the extent of common variance among indicators.

Signs and symptoms that were not statistically significant at the bivariate level were excluded from the final model (i.e., number of clues on the OLS test, estimate of 30 seconds on the MRB test, number of clues on the WAT test, diastolic blood pressure). Two drug-related signs and symptoms were also excluded from the final model because their initial

inclusion violated the statistical assumption of adequacy of expected frequencies (i.e., vertical gaze nystagmus, pupillary unrest). That is, more than 20% of cells had an expected frequency of less than five. When this assumption is violated, statistical power is attenuated and it restricts the goodness-of-fit criteria used to evaluate the model (Tabachnick & Fidell, 2007).

Prediction of CNS Stimulants from Drug-Related Signs and Symptoms in DECP Evaluations

Prediction of CNS stimulants from clinical indicators

A binary logistic regression analysis was conducted to determine which clinical indicators from the overall model distinguished the CNS stimulant drug category from the no-drug category (i.e., the reference group). Results indicated that the set of clinical indicators significantly distinguished the CNS stimulant cases from the no-drug cases, $\chi^2(7, N = 1{,}199) = 40.20$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the CNS stimulant drug category compared with the no-drug category are displayed in Table 1. Using a Bonferroni correction ($p < .007$) to control for Type I error, only two of the clinical indicators significantly contributed to the prediction of CNS stimulants: average pulse rate and muscle tone (Table 1). The results showed that the odds of a suspected drug-impaired driver having used CNS stimulants increase by 7% for every one unit increase in average pulse rate. Findings also revealed that suspected drug-impaired drivers who consumed CNS stimulants were more likely than those who did not consume any drugs to have a rigid muscle tone compared to a normal one.

Table 1. Prediction of Drug Category From Clinical Indicators Among DEC Evaluations: CNS Stimulants vs. No-Drug Cases

Clinical Indicators	B	SE	Wald χ^2 Test	OR	95% CI for OR
Mean pulse rate	.06	.02	13.96	1.07*	1.03, 1.10
Absence vs. presence of leg tremors	-.30	.64	.22	.74	.21, 2.59
Absence vs. presence of body tremors	-.25	.56	.20	.78	.26, 2.35
Systolic blood pressure	-.01	.01	.91	.99	.97, 1.01
Body temperature (in C)	-.81	.36	5.11	.45	.22, .90
Flaccid vs. normal muscle tone	1.79	.80	5.10	6.01	1.27, 28.55
Rigid vs. normal muscle tone	2.33	.68	11.78	10.22*	2.71, 38.54

*$p < .0071$

Prediction of CNS stimulants from performance on psychophysical tests

A binary logistic regression analysis was also conducted to determine which psychophysical tests from the overall model distinguished the CNS stimulant drug category from the no-drug category. The findings showed that the set of tests significantly distinguished the CNS stimulant cases from the no-drug cases, $\chi^2(3, N = 1{,}199) = 107.86$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the CNS stimulant drug category compared with the no-drug category are displayed in Table 2. Using a Bonferroni correction ($p < .0167$) to control for Type I error, each of the three psychophysical tests significantly contributed to the prediction of CNS stimulants: number of clues on the HGN test, total sway on the MRB test, and number of misses on the FTN test (Table 2). The results showed that the odds of a suspected drug-impaired driver having used CNS stimulants increase by 19% for every additional clue recorded on the HGN test. Findings also revealed that the odds of a suspected drug-impaired driver having used CNS stimulants increase by 34% for every one unit increase in total sway on the MRB test. Finally, the results indicated that the odds of a suspected drug-impaired driver having used CNS stimulants increase by 16% for every additional miss recorded on the FTN test.

Table 2. Prediction of Drug Category From Psychophysical Tests Among DEC Evaluations: CNS Stimulants vs. No-Drug Cases

Clinical Indicators	B	SE	Wald $\chi 2$ Test	OR	95% CI for OR
Number of clues on Horizontal Gaze Nystagmus test	.17	.06	9.53	1.19*	1.06, 1.32
Total sway on Modified Romberg Balance test	.29	.05	39.11	1.34*	1.22, 1.47
Number of misses on Finger to Nose test	.15	.03	22.75	1.16*	1.09, 1.23

*$p < .0167$

Prediction of CNS stimulants from appearance and physiological response of the eyes

A binary logistic regression analysis was then conducted to determine which signs and symptoms from the overall model related to the appearance and physiological response of the eyes distinguished the CNS stimulant drug category from the no-drug category. Results indicated that this set of signs and symptoms significantly distinguished the CNS stimulant cases from the no-drug cases, $\chi^2(13, N = 1{,}199) = 350.72$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the CNS stimulant drug category compared with the no-drug category are displayed in Table 3. Using a Bonferroni correction ($p < .0038$) to control for Type I error, only three of the signs and symptoms significantly contributed to the prediction of CNS stimulants: pupil size in direct light, reaction to light, and condition of the eyelids (Table 3). The results showed that the odds of a suspected drug-impaired driver having used CNS stimulants increase by 75% for every one-unit increase in pupil size in direct light. Findings also revealed that suspected drug-impaired drivers who consumed CNS stimulants were more likely than those who did not consume any drugs to have a slow compared with a normal reaction to light. Finally, the results indicated that suspected drug-impaired drivers who used CNS stimulants were more likely to have normal vs. droopy eyelids.

Table 3. Prediction of Drug Category From Appearance and Physiological Response of the Eyes Among DEC Evaluations: CNS Stimulants vs. No-Drug Cases

Signs and Symptoms	B	SE	Wald $\chi 2$ Test	OR	95% CI for OR
Condition of the eyes: normal	-1.07	.49	4.81	.34	.13, .89
Condition of the eyes: reddening of the conjunctiva	-.38	.49	.60	.68	.26, 1.79
Condition of the eyes: bloodshot	.49	.42	1.38	1.63	.72, 3.70
Condition of the eyes: watery	.34	.37	.87	1.41	.69, 2.89
Convergence (present, absent)	.28	.27	1.07	1.32	.78, 2.25
Presence of eyelid tremors (yes, no)	.74	.28	6.95	2.10	1.21, 3.66
Pupil size in room light	-.05	.15	.09	.96	.71, 1.28
Pupil size in darkness	-.24	.14	2.88	.79	.60, 1.04
Pupil size in direct light	.56	.16	12.64	1.75*	1.28, 2.37
Rebound dilation	-.34	.53	.42	.71	.25, 2.01
Slow vs. normal reaction to light	2.89	.32	80.57	17.90*	9.54, 33.61
Little to none vs. normal reaction to light	2.82	1.13	6.24	16.70	1.83, 152.17
Eyelids (normal, droopy)	.96	.32	9.06	2.61*	1.40, 4.88

*$p < .0038$

Prediction of Cannabis From Drug-Related Signs and Symptoms in DEC Evaluations

Prediction of cannabis from clinical indicators

A binary logistic regression analysis was conducted to determine which clinical indicators from the overall model distinguished the cannabis drug category from the no-drug category. Results indicated that the set of clinical indicators significantly distinguished the cannabis cases from the no-drug cases, $\chi^2(7, N = 2{,}042) = 17.84$, $p < .01$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the cannabis drug category compared with the no-drug category are displayed in Table 4. Using a Bonferroni correction ($p < .007$) to control for Type I error,

none of the clinical indicators significantly contributed to the prediction of cannabis (Table 4).

Table 4. Prediction of Drug Category From Clinical Indicators Among DEC Evaluations: Cannabis vs. No-Drug Cases

Clinical Indicators	**B**	**SE**	**Wald χ^2 Test**	**OR**	**95% CI for OR**
Mean pulse rate	.03	.01	4.77	1.03	1.00, 1.06
Absence vs. presence of leg tremors	1.21	.56	4.75	3.37	1.13, 10.02
Absence vs. presence of body tremors	.19	.48	.16	1.21	.47, 3.12
Systolic blood pressure	-.03	.01	5.11	.97	.95, 1.00
Body temperature (in C)	.22	.36	.39	1.25	.62, 2.53
Flaccid vs. normal muscle tone	1.12	.76	2.20	3.08	.70, 13.55
Rigid vs. normal muscle tone	.15	.64	.05	1.16	.33, 4.11

$^{*}p < .007$

Prediction of cannabis from performance on psychophysical tests

A binary logistic regression analysis was also conducted to determine the psychophysical tests from the overall model that distinguished the cannabis drug category from the no-drug category. Findings showed that the set of tests significantly distinguished the cannabis cases from the no-drug cases, $\chi^2(3, N = 2{,}042) = 382.38$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the cannabis drug category compared with the no-drug category are displayed in Table 5. Using a Bonferroni correction ($p < .0167$) to control for Type I error, two of the three psychophysical tests significantly contributed to the prediction of cannabis: total sway on the MRB test and number of misses on the FTN test (Table 5). The results showed that the odds of a suspected drug-impaired driver having used cannabis increase by 55% for every one unit increase in total sway on the MRB test. Findings also revealed that the odds of a suspected drug-impaired driver having used cannabis increase by 50% for every additional miss recorded on the FTN test.

Table 5. Prediction of Drug Category From Psychophysical Tests Among DEC Evaluations: Cannabis vs. No-Drug Cases

Clinical Indicators	B	SE	Wald $\chi 2$ Test	OR	95% CI for OR
Number of clues on Horizontal Gaze Nystagmus test	.09	.05	2.77	1.09	.98, 1.21
Total sway on Modified Romberg Balance test	.44	.05	79.25	1.55*	1.41, 1.71
Number of misses on Finger to Nose test	.40	.03	187.39	1.50*	1.41, 1.59

*$p < .0167$

Prediction of cannabis from appearance and physiological response of the eyes

A binary logistic regression analysis was conducted to determine which signs and symptoms from the overall model related to the appearance and physiological response of the eyes distinguished the cannabis drug category from the no-drug category. Results indicated that this set of signs and symptoms significantly distinguished the cannabis cases from the no-drug cases, $\chi^2(13, N = 2{,}042) = 840.57$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms of the eyes for the cannabis drug category compared to the no-drug category are displayed in Table 6. Using a Bonferroni correction ($p < .0038$) to control for Type I error, several signs and symptoms significantly contributed to the prediction of cannabis, including the condition of the eyes, convergence, eyelid tremors, pupil size in direct light, rebound dilation, and condition of the eyelids (Table 6). Findings revealed that suspected drug-impaired drivers who used cannabis were more likely than those who did not consume drugs to have normal eyes or reddening of the conjunctiva. Such individuals were also more likely to experience eyelid tremors and rebound dilation, and have droopy eyelids. The results from the analysis also showed that the odds of suspected drug-impaired drivers having used cannabis increase by 70% for every one-unit increase in pupil size in direct light. Finally, the findings indicated that suspected drug-impaired drivers who consumed cannabis were more likely than subjects who did not consume any drugs to have a slow or little to no reaction to light.

Table 6. Prediction of Drug Category From Appearance and Physiological Response of the Eyes Among DEC Evaluations: Cannabis vs. No-Drug Cases

Signs and Symptoms	B	SE	Wald $\chi 2$ Test	OR	95% CI for OR
Condition of the eyes: normal	1.35	.42	10.40	3.85*	1.70, 8.74
Condition of the eyes: reddening of the conjunctiva	1.52	.37	16.73	4.59*	2.21, 9.51
Condition of the eyes: bloodshot	1.00	.35	8.03	2.73	1.36, 5.46
Condition of the eyes: watery	.02	.28	.01	1.02	.59, 1.77
Convergence (present, absent)	.59	.22	7.12	1.81	1.17, 2.80
Presence of eyelid tremors (yes, no)	1.20	.23	28.08	3.31*	2.13, 5.15
Pupil size in room light	.09	.13	.49	1.09	.85, 1.40
Pupil size in darkness	.12	.12	1.00	1.12	.90, 1.41
Pupil size in direct light	.53	.14	14.79	1.70*	1.30, 2.22
Rebound dilation	2.09	.35	36.45	8.08*	4.10, 15.93
Slow vs. normal reaction to light	1.31	.34	14.54	3.72*	1.89, 7.30
Little to none vs. normal reaction to light	3.88	1.20	10.40	48.38*	4.58, 511.41
Eyelids (normal, droopy)	1.13	.26	19.16	3.08*	1.86, 5.11

*$p < .0038$

Prediction of CNS Depressants From Drug-Related Signs and Symptoms in DEC Evaluations

Prediction of CNS depressants from clinical indicators

A binary logistic regression analysis was conducted to determine which clinical indicators from the overall model distinguished the CNS depressant drug category from the no-drug category. Results indicated that the set of clinical indicators significantly distinguished the CNS depressant cases from the no-drug cases, $\chi^2(7, N = 1{,}178) = 49.85$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the CNS depressant drug category compared to the no-drug category are displayed in Table 7. Using a Bonferroni correction ($p < .007$) to control for Type I error, only two of the clinical indicators significantly contributed to the prediction of

CNS depressants: systolic blood pressure and muscle tone (Table 7). The results showed that the odds of suspected drug-impaired drivers having used CNS depressants decrease by 4% for each one-unit increase in systolic blood pressure. Findings also revealed that suspected drug-impaired drivers who consumed CNS depressants were more likely than those who did not consume any drugs to have a flaccid compared to a normal muscle tone.

Table 7. Prediction of Drug Category From Clinical Indicators Among DEC Evaluations: CNS Depressants vs. No-Drug Cases

Clinical Indicators	**B**	**SE**	**Wald χ^2 Test**	**OR**	**95% CI for OR**
Mean pulse rate	.03	.01	3.85	1.03	1.00, 1.05
Absence vs. presence of leg tremors	.12	.61	.04	1.13	.34, 3.72
Absence vs. presence of body tremors	-.21	.55	.15	.81	.28, 2.38
Systolic blood pressure	-.04	.01	11.99	.96*	.94, .98
Body temperature (in C)	-.78	.39	4.11	.46	.22, .98
Flaccid vs. normal muscle tone	2.74	.76	13.00	15.44*	3.49, 68.36
Rigid vs. normal muscle tone	.49	.69	.52	1.64	.43, 6.32

*$p < .007$

Prediction of CNS depressants from performance on psychophysical tests

A binary logistic regression analysis was also conducted to determine which psychophysical tests from the overall model distinguished the CNS depressant drug category from the no-drug category. The findings showed that the set of tests significantly distinguished the CNS depressant cases from the no-drug cases, $\chi^2(3, N = 1{,}178) = 960.03$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the CNS depressant drug category compared with the no-drug category are displayed in Table 8. Using a Bonferroni correction ($p < .0167$) to control for Type I error, each of the three psychophysical tests significantly contributed to the prediction of CNS depressants: total number of clues on the HGN test, total sway on the MRB test, and number of misses on the FTN test (Table 8). The results showed that the odds of suspected drug-impaired drivers having used CNS depressants increase by 146% for every additional clue recorded on the HGN test. Findings also revealed that odds of suspected drug-impaired drivers having used CNS depressants increase by 45% for every one-unit increase in total sway on the MRB test. Finally, the results indicated that the odds of suspected drug-impaired drivers having used CNS depressants increase by 31% for every additional miss recorded on the FTN test.

Table 8. Prediction of Drug Category From Psychophysical Tests Among DEC Evaluations: CNS Depressants vs. No-Drug Cases

Clinical Indicators	B	SE	Wald $\chi 2$ Test	OR	95% CI for OR
Number of clues on Horizontal Gaze Nystagmus test	.90	.06	230.10	2.46*	2.19, 2.77
Total sway on Modified Romberg Balance test	.37	.07	32.01	1.45*	1.27, 1.64
Number of misses on Finger to Nose test	.27	.06	23.18	1.31*	1.17, 1.46

*$p < .0167$

Prediction of CNS depressants from appearance and physiological response of the eyes

A binary logistic regression analysis was conducted to determine which signs and symptoms from the overall model related to the appearance and physiological response of the eyes distinguished the CNS depressant drug category from the no-drug category. Results indicated that this set of signs and symptoms significantly distinguished the CNS depressant cases from the no-drug cases, $\chi^2(13, N = 1{,}178) = 642.96$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the CNS depressant drug category compared with the no-drug category are displayed in Table 9. Using a Bonferroni correction ($p < .0038$) to control for Type I error, only three of the signs and symptoms significantly contributed to the prediction of CNS depressants: convergence, reaction to light, and condition of the eyelids (Table 9). The results showed that suspected drug-impaired drivers who used CNS depressants were more likely to experience lack of convergence compared with those who did not consume any drugs. Findings also revealed that suspected drug-impaired drivers who consumed CNS depressants were more likely than those who did not consume any drugs to have a slow or little to no reaction to light. Suspected drug-impaired drivers who used CNS depressants were also more likely to have droopy eyelids.

Table 9. Prediction of Drug Category From Appearance and Physiological Response of the Eyes Among DEC Evaluations: CNS Depressants vs. No-Drug Cases

Signs and Symptoms	B	SE	Wald $\chi 2$ Test	OR	95% CI for OR
Condition of the eyes: normal	-.82	.46	3.23	.44	.18, 1.08
Condition of the eyes: reddening of the conjunctiva	-1.08	.46	5.49	.34	.14, .84
Condition of the eyes: bloodshot	-.15	.40	.14	.86	.39, 1.88
Condition of the eyes: watery	.47	.36	1.74	1.60	.80, 3.23
Convergence (present, absent)	1.91	.27	50.73	6.78*	4.00, 11.48
Presence of eyelid tremors (yes, no)	.35	.27	1.75	1.42	.84, 2.40
Pupil size in room light	-.10	.14	.55	.90	.69, 1.18
Pupil size in darkness	.11	.13	.68	1.11	.87, 1.43
Pupil size in direct light	.35	.15	5.55	1.42	1.06, 1.90
Rebound dilation	.43	.45	.92	1.53	.64, 3.68
Slow vs. normal reaction to light	2.89	.31	87.27	18.02*	9.83, 33.06
Little to none vs. normal reaction to light	4.15	1.28	10.49	63.63*	5.15, 785.77
Eyelids (normal, droopy)	2.36	.27	77.09	10.55*	6.24, 17.86

*$p < .0038$

Prediction of Narcotic Analgesics From Drug-Related Signs and Symptoms in DEC Evaluations

Prediction of narcotic analgesics from clinical indicators

A binary logistic regression analysis was conducted to determine which clinical indicators from the overall model distinguished the narcotic analgesics drug category from the no-drug category. Results indicated that the set of clinical indicators significantly distinguished the narcotic analgesics cases from the no-drug cases, $\chi^2(7, N = 898) = 54.66$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the narcotic analgesics drug category compared with the no-drug category are presented in Table 10. Using a Bonferroni correction ($p < .007$) to control for Type I error, only two of the clinical indicators significantly contributed to the prediction

of narcotic analgesics: systolic blood pressure and muscle tone (Table 10). The results showed that the odds of suspected drug-impaired drivers having used narcotic analgesics decrease by 4% for each one-unit increase in systolic blood pressure. Findings also revealed that suspected drug-impaired drivers who consumed narcotic analgesics were more likely than those who did not consume any drugs to have a flaccid compared with a normal muscle tone.

Table 10. Prediction of Drug Category From Clinical Indicators Among DEC Evaluations: Narcotic Analgesics vs. No-Drug Cases

Clinical Indicators	**B**	**SE**	**Wald χ^2 Test**	**OR**	**95% CI for OR**
Mean pulse rate	.002	.01	.03	1.00	.98, 1.03
Absence vs. presence of leg tremors	.11	.70	.03	1.12	.28, 4.45
Absence vs. presence of body tremors	-.33	.55	.28	.72	.21, 2.42
Systolic blood pressure	-.04	.01	7.73	.96*	.94, .99
Body temperature (in C)	-.12	.62	.10	.88	.42, 1.87
Flaccid vs. normal muscle tone	3.45	.77	19.87	31.56*	6.92, 143.99
Rigid vs. normal muscle tone	1.03	.84	1.50	2.79	.54, 14.36

*$p < .007$

Prediction of narcotic analgesics from performance on psychophysical tests

A binary logistic regression analysis was also conducted to determine which psychophysical tests from the overall model distinguished the narcotic analgesics drug category from the no-drug category. The findings showed that the set of tests significantly distinguished the narcotic analgesics cases from the no-drug cases, $\chi^2(3, N = 898) = 245.60$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the narcotic analgesics category compared with the no-drug category are displayed in Table 11. Using a Bonferroni correction ($p < .0167$) to control for Type I error, two of the three psychophysical tests significantly contributed to the prediction of narcotic analgesics: total sway on the MRB test and number of misses on the FTN test (Table 11). The results showed that the odds of a suspected drug-impaired driver having used narcotic analgesics increase by 70% for every one-unit increase in total sway on the MRB test. Findings also indicated that the odds of a suspected drug-impaired driver having used narcotic analgesics increase by 23% for every additional miss recorded on the FTN test.

Table 11. Prediction of Drug Category From Psychophysical Tests Among DEC Evaluations: Narcotic Analgesics vs. No-Drug Cases

Clinical Indicators	B	SE	Wald $\chi 2$ Test	OR	95% CI for OR
Number of clues on Horizontal Gaze Nystagmus test	.11	.06	3.22	1.11	.99, 1.25
Total sway on Modified Romberg Balance test	.53	.05	102.55	1.70*	1.54, 1.88
Number of misses on Finger to Nose test	.21	.04	29.08	1.23*	1.14, 1.33

*$p < .0167$

Prediction of narcotic analgesics from appearance and physiological response of the eyes

A binary logistic regression analysis was conducted to determine which signs and symptoms from the overall model related to the appearance and physiological response of the eyes distinguished the narcotic analgesics drug category from the no-drug category. Results indicated that this set of signs and symptoms significantly distinguished the narcotic analgesics cases from the no-drug cases, $\chi^2(13, N = 898) = 599.54$, $p < .0001$.

The regression coefficients, chi-square tests, odds ratios and 95% confidence intervals for the signs and symptoms for the narcotic analgesics drug category compared with the no-drug category are displayed in Table 12. Using a Bonferroni correction ($p < .0038$) to control for Type I error, only three of the signs and symptoms significantly contributed to the prediction of narcotic analgesics: pupil size in darkness reaction to light, and condition of the eyelids (Table 12). The findings revealed that the odds of a suspected drug-impaired driver having used narcotic analgesics decrease by 55% for every one unit increase in pupil size in darkness. Suspected drug-impaired drivers who consumed narcotic analgesics were more likely than those who did not consume any drugs to have a slow or little to no reaction to light. Finally, suspected drug-impaired drivers who used narcotic analgesics were also more likely to have droopy eyelids.

Table 12. Prediction of Drug Category From Appearance and Physiological Response of the Eyes Among DEC Evaluations: Narcotic Analgesics vs. No-Drug Cases

Signs and Symptoms	B	SE	Wald $\chi 2$ Test	OR	95% CI for OR
Condition of the eyes: normal	-.17	.80	.04	.85	.18, 4.04
Condition of the eyes: reddening of the conjunctiva	.33	.73	.20	1.39	.34, 5.71
Condition of the eyes: bloodshot	.81	.70	1.35	2.25	.57, 8.88
Condition of the eyes: watery	.26	.66	.15	1.30	.36, 4.71
Convergence (present, absent)	.74	.47	2.49	2.09	.84, 5.20
Presence of eyelid tremors (yes, no)	.38	.51	.56	1.46	.54, 3.95
Pupil size in room light	-.76	.32	5.89	.47	.25, .86
Pupil size in darkness	-.79	.22	12.41	.45*	.29, .70
Pupil size in direct light	.37	.30	1.51	1.44	.81, 2.58
Rebound dilation	-.91	1.29	.50	.40	.03, 5.02
Slow vs. normal reaction to light	3.24	.56	33.11	25.56*	8.48, 77.11
Little to none vs. normal reaction to light	4.20	1.12	14.11	66.97*	7.47, 600.81
Eyelids (normal, droopy)	2.50	.48	27.53	12.16*	4.78, 30.91

*$p < .0038$

Discussion

The data from DEC evaluations provided an opportunity to examine the signs and symptoms associated with different categories of drugs as presented by individuals who were suspected of drug-impaired driving. This approach differs from drug administration studies in which subjects are provided with a specific dose of a given drug and the effects are monitored and measured in a controlled environment as the concentration of the drug waxes and wanes. DEC evaluations are typically conducted on individuals who have ingested an unknown amount and type of one or more psychoactive substance(s) over an unspecified time period. In addition, the quantity of substance(s) ingested by impaired driving suspects can be considerably larger than could be ethically administered in a clinical or laboratory setting.

The set of DECP evaluations does not represent a random sample of evaluations. The criteria for inclusion required a match between the drug category indicated in the toxicology report and the officer's opinion of the drug category ingested by the suspect. Cases that lacked agreement between the toxicology and the officer's opinion were excluded. Overall agreement between the officer's opinion and the toxicology has been reported to be between 85%-95% (Beirness et al., 2009; IACP, 2016). The reasons for disagreement can be diverse, including unusual drug effects, masking of symptoms by other substances, polydrug use, medical/physical conditions that mimic drug symptoms, or simply an incorrect call by the DRE. Nevertheless, the exclusion of cases where there is disagreement between the officer's opinion and the toxicology raises questions about the excluded cases and the implications for the representativeness of the data. If these cases were the result of a lack of strong indicators of drug use, their inclusion might have served to weaken the observed relationships.

The database includes a limited number of evaluations involving hallucinogens (six), inhalants (22), and dissociative anaesthetics (27). These substances are less frequently encountered in traffic enforcement situations and the small number of evaluations involving these types of substances precluded their inclusion in the analysis.

The initial set of analyses of the DEC evaluation data examined differences in signs and symptoms of drug use across drug categories. In most cases, the general signs and symptoms were consistent with those listed in the DECP Matrix. It was clear, however, that the range of effects within a given drug category can be large and not all suspects in a given drug category displayed the same effects or the same degree of effects. For example, the DECP Matrix indicates that reaction to light is expected to be slow among those who have ingested depressants or stimulants, and normal among those who had used cannabis. Subjects positive for narcotic analgesics are expected to show little or no reaction to light. However, the data also indicate that about 23% of cannabis users showed slow reaction to light. Stimulant users were almost equally likely to show normal or slow reaction to light. At the very least, these findings serve to confirm the general effects of different categories of drugs as noted in the matrix, while at the same time illustrating the variability of drug effects.

The data analysis confirmed that the effects of drugs on body temperature differ significantly by drug category. However, the magnitude of the differences in body temperature were extremely small, with the mean values across drug categories differing by less than 0.5 degrees Celsius (i.e., 36.4 to 36.8). Despite the significant differences in mean body temperature among drug categories, the limited range of values would suggest that body temperature might not be a good discriminator of drug category. The multivariate analyses revealed this to be the case. In no drug category was body temperature shown to be a significant predictor.

In the univariate analysis, blood pressure was assessed as either high, low or normal when compared with established norms for people according to age. This approach revealed very clear differences among drug groups that were mostly consistent with the DECP Matrix. It was noted that almost half (46%) of the subjects in the no-drug group were assessed as having blood pressure above the range for their age group. This might suggest a general population issue with elevated blood pressure or may reflect a situational increase in blood pressure associated with measurement process, sometimes referred to as "white coat

hypertension" or "white coat syndrome". While the former situation would be an issue of concern for public health, the latter situation could possibly be attenuated by having DREs measure blood pressure on two occasions over the course of the evaluation.

Multivariate analyses were conducted to identify a prominent set of signs and symptoms assessed during a DEC evaluation for predicting each of the four major drug categories examined. Given that more than 100 individual pieces of information are collected and recorded during a DEC evaluation, the signs and symptoms of drug use were grouped into three conceptual groups — i.e., clinical indicators, performance on the psychophysical tests, and eye indicators. The analysis identifies those indicators that best distinguish a given drug category from drug-free cases after accounting for the common or shared variance across indicators. For example, the best set of signs and symptoms that distinguished subjects who tested positive for cannabis compared with drug-free subjects were: total sway on the Modified Romberg Balance test, number of misses on the Finger to Nose test, reddened conjunctiva, presence of eyelid tremors, presence of leg tremors, pupil size in direct light, rebound dilation, slow or little reaction to light, and droopy eyelids. This set of signs and symptoms represents the most prominent set of indicators across a large group of people who had ingested unknown quantities of cannabis. These symptoms can be used to begin the process of reviewing the results of an evaluation, using the results of this set of indicators to provide an initial indication of possible drug category (or categories). A DRE would still be expected to consider the data from the other tests and observations made during the course of a complete DEC evaluation to support the opinion. It should be recognized, however, that individual results can vary. Hence, this best set of indicators should never exclude consideration of other signs and symptoms assessed during the DEC evaluation and the totality of the situation. Also, it is important to note that many predictor variables included in the multivariate analyses were treated as continuous variables, assuming a linear relationship with the prediction of drug category. It is possible that nonlinear models or the use of meaningful categorizations or thresholds for the predictor variables could yield different patterns of outcomes. That said, in the absence of information to guide the selection of predefined groups or cutoff points, the linear assumption is the most parsimonious.

The ability to identify sets of signs and symptoms of various categories of drugs opens the possibility of developing algorithms to predict the category of drug from the data collected during a DEC evaluation. Assigning weights to the various measurements could be used to assess the probability that a particular category of drug was responsible for the observations. This type of approach could facilitate the evaluator in forming an opinion of category of drug ingested.

Conclusion

The DECP was developed out of a growing need for an effective procedure for law enforcement officers to assess drivers suspected of driving while under the influence of psychoactive substances other than alcohol. Existing knowledge about drug effects on vital signs, responses of the eyes, and psychomotor performance was used to establish a set of tests and procedures to assess indicators that could be affected by the use of various types of substances. The resultant 12-step procedure has proliferated and there are currently more than 8,000 certified DREs throughout the United States and Canada.

The basis for the effects of drugs on vital signs such as blood pressure and heart rate can be found in the medical literature, which is generally consistent with the drug effects listed in the DECP Matrix. Where reports of conflicting effects exist, results could be due to variations in drug dose or the effects of a specific substance rather than all drugs within a category. In cases where literature addressing the effect on a sign or indicator was not identified, the DECP Matrix typically indicates the expected effect is that the sign or indicator is either not present or within normal range for drug-free subjects.

Experimental studies examining the DECP provide limited evidence of the ability of the DECP protocol to identify various categories of drugs on the basis of clinical indicators and commonly employed tests of impairment. Experimental laboratory studies indicate that officers trained in the DECP are generally able to detect impairment in subjects who have been administered drugs. These types of studies, however, do not present strong support for the accuracy with which they can identify the particular class of the drug administered or correctly identify drug-free subjects. Drug administration studies of the type employed can be challenging and are ethically constrained in the types of drugs and doses that can be administered. The lack of distinctive clinical and psychomotor symptoms associated with relatively low doses of some of the drugs administered in the studies likely played a role in the findings. The doses of drugs ingested by suspected drug-impaired drivers are unknown and may be significantly higher than those ethically permitted in a laboratory setting. It would be expected that with higher doses, there would be higher rates of accuracy with which DREs can detect drug impairment and be able to identify the category (or categories) of drugs ingested.

Studies of the accuracy of the DECP conducted in enforcement settings report higher overall accuracy than laboratory studies. This is most likely the result of the additional information provided by and available during a complete 12-step evaluation. Taken together, the results of the two types of studies provide evidence of the validity of the DECP to further its use in efforts to remove drug-impaired drivers from the roads.

Further research with different types of drugs and doses is needed to document the nature and extent of the effects of various substances to strengthen the evidence. Research using novel methods and incorporating a broad range of subject characteristics would also help expand the base of knowledge in this area.

Relative to many other police training programs, the DECP training is longer and more intensive, which can lead to greater costs. A DECP assessment also takes considerably longer to administer than an SFST or a breath test for alcohol, and is not as definitive given drug-impaired driving is a more challenging issue than alcohol-impaired driving from many

perspectives. Nevertheless, the DECP provides a tool for dealing with drug-impaired drivers that is regarded by stakeholders as the premier approach for dealing with this problem based on the data and research evidence available today. Even as new detection technologies such as oral fluid drug screening become increasingly available, stakeholders anticipate the DECP will continue to be needed to assess suspected impaired drivers to provide detailed evidence of the nature and extent of drug effects. In addition, as the legalization of cannabis continues to spread across jurisdictions, the need for a strong DECP program may become increasingly more evident.

Suggestions for Program Enhancement

A primary objective of the comprehensive review of the DECP was to identify areas where potential modifications could be implemented to enhance the efficiency and effectiveness of the program. The literature review, analysis of data from DEC evaluations, review of international programs, and key informant interviews all identified areas of the DECP that could be modified to improve the program. This section presents a list of potential modifications to the DECP. It should be noted, however, that a suggestion for changing some aspect of the DECP does not necessarily imply that a problem exists. Rather, the following suggestions are made primarily for the purposes of enhancing the DECP.

The suggestions listed here are presented in three groups according to proposed time frame for implementation — short term, intermediate term, and longer term. Each of the suggestions listed below includes a brief description of how the change would serve to enhance some aspect of the DECP. In presenting this list of potential enhancements, it is recognized that each must be reviewed and evaluated by the DECP Technical Advisory Panel.

Short term

1. Encourage the implementation of computer tablets for use by DREs to record the results of evaluations. Several states have implemented the tablet application developed by ITSMR in New York. The efficiencies associated with this technology include the digital recording of evaluations, simplifying the production of narrative reports, sending evaluations electronically for review by coordinators, and inserting toxicology results quickly and easily. The application can also format the results of an evaluation into traditional print format for those who prefer (and possibly require) a paper copy of the report.

Initial indications suggest the introduction of tablets is associated with an increase in the number of evaluations performed. This could be an effect associated with the novelty of new equipment or could be tied to a requirement to perform a specified number of evaluations to maintain the use of the tablet.

The costs associated with hardware, software, training and maintenance are not insignificant and need to be weighed against improved efficiency and a potential increase in the number of evaluations performed.

2. Examine the potential value associated with the use of equipment that provides automated measurement of (a) body temperature and (b) blood pressure and pulse. Measuring body temperature is neither difficult nor complicated. Automated electronic oral

thermometers are currently in use in many programs. However, in light of the findings from the database of evaluations that show extremely small differences in body temperature among drug categories, an increase in accuracy of temperature measurement would be a positive step. To this end, thermometers that measure body temperature from the tympanic membrane have been shown to provide somewhat more accurate readings than other types of thermometers. The cost of implementing tympanic thermometers is relatively small and training is minimal.

The measurement of blood pressure with a sphygmomanometer and stethoscope is not particularly difficult but does require training and practice to ensure accuracy. Automatic blood pressure measurement devices have been available for many years. Many different models are available for use in medical settings as well as for personal use at home. Their accuracy and reliability appear to be very good. Some models have built-in self-calibration checks. Many models also have the option of multiple sequential measurements. These devices typically measure pulse at the same time. The use of automated blood pressure (and pulse) measurement would serve to enhance efficiency and accuracy.

Not every DRE would need to be issued an automatic blood pressure device. A limited number of devices could be located at the station for shared use. The cost of acquiring the equipment, training and maintenance would be minimal.

3. Allow the use of ocular recording devices. Ocular recording devices have been available for several years. Such devices provide a clear image of the eye and eye movements and facilitate the assessment of pupil size under various light conditions, reaction to light, HGN, and rebound dilation. The ability to record the assessment of eye indicators would allow DREs and instructors the opportunity to review the evidence at some later point in time. The video of the eye assessment would also provide valuable evidence in court.

There is an initial cost of several thousand dollars for the equipment. Not every DRE requires their own but devices can be shared within a police service. Training in the use of the device is not extensive and can be done locally in small groups as required.

The use of ocular recording devices could remain optional. Such devices provide an enhancement to the assessment and help clarify the evidence through video recording and opportunity to review. The absence of the device should not detract from a DECP assessment.

Even in the absence of implementation of ocular recording devices for DECP evaluations, the use of these devices could be beneficial for training purposes. Having an ocular recording device available during training could help illustrate the assessment of eye indicators.

4. Revise the drug influence evaluation face sheet to include a place for the recording of the presence or absence of eyelid and leg tremors. DREs are trained to watch for eyelid and leg tremors during the evaluation, which may be evident while the subject is performing the OLS, MRB and FTN tests. The DRE is to note the presence of tremors by making a note on the face sheet. If such a note is made, it is clear the DRE observed the presence of eyelid and/or leg tremors. If no such notation is present, it is not clear whether the DRE did not observe tremors, failed to note their presence, or simply neglected to watch for tremors.

There is currently no check box or designated space where the presence or absence of eyelid and leg tremors are recorded. Including check boxes for the presence or absence of eyelid and leg tremors on the face sheet would clarify the information.

Although in principle, this is a relatively simple modification, the challenge in implementing it is finding space on the existing crowded face sheet. The issue with space, however, could be easily resolved with the implementation of computer tablets (see suggestion No. 1).

5. Add a second measurement of blood pressure. During the DECP evaluation, pulse is measured on three separate occasions. This is to help ensure that stress, fear, and/or nervousness that might be affecting pulse at the beginning of the evaluation would have eased over the course of the evaluation. Hence, a better indication of pulse would be obtained by the time of the third measurement. In addition, differences in heart rate over the course of an evaluation can be an indication of the use of multiple drugs.

A similar argument can be made for increasing the number of times blood pressure is measured during an evaluation. Adding an additional measurement of blood pressure would help alleviate concerns about situational hypertension.

The increased time required for a second blood pressure measurement could be offset with the implementation of automatic blood pressure devices (see suggestion No. 2).

Intermediate term

6. Develop standard scoring systems for the Finger to Nose test and the Modified Romberg Balance test. The FTN and MRB tests were included among the tests examined by Burns and Moskowitz (1977) in studies leading to the development of the Standardized Field Sobriety Test (SFST). These two tests performed well as measures of impairment due to alcohol but were not among the top three tests examined. These two tests were, however, included among the psychophysical/divided attention tests in the DECP protocol. Although various measurements are taken and errors recorded, there are neither standard scoring systems nor validated clues for either test.

Studies should be undertaken to develop standard systems of scoring for these two tests. Subsequent analyses would be able to identify a series of threshold values ("clues") that are indicative of drug use in a manner similar to that used in the scoring of the three tests of the SFST.

A standardized system of scoring would facilitate the reporting and interpretation of the results of these two tests.

7. Encourage the use of subject matter experts to instruct at DRE training schools. DRE instructors are expected to be able to teach all sessions of the DRE training. Some of the topics, however, are complex (e.g., physiology, toxicology, legal issues, case and court preparation). Not all instructors are comfortable teaching these topics and answering questions about them. The use of subject matter experts as guest instructors for specific

topics could enhance the training by allowing these experts to share their wealth of knowledge and understanding of some of the complex issues involved in drugs and driving.

The costs associated with guest instructors would depend on the individual and whether or not travel is required. Wherever possible, appropriate local experts with the necessary skills and expertise in particular areas who are willing to spend a few hours to assist could most likely be found at minimal cost.

8. Consider the inclusion of measuring respiration rate as a vital sign. Respiration rate is a vital sign that can be affected by the use of some substances. For example, narcotic analgesics decrease respiration rate (Julian et al., 2008; Schuckit, 2006), whereas stimulants increase respiration rate (American Psychiatric Association, 2013; Julian et al., 2008). The inclusion of respiration rate should be investigated for its ability to enhance the prediction of drug category by DREs.

Respiration rate is readily measured by watching the rise and fall of the chest as the person inhales and exhales. Training would be minimal as would be the time required to add this indicator to the evaluation.

Longer term

9. Examine the use of force plates for measuring balance/body sway. The assessment of body sway during the MRB is done without the aid of any measurement device. The extent of movement in the front-back and side-side planes are estimated by the evaluator. There is a tendency to record the magnitude of sway in whole numbers, suggesting a lack of precision. The approach to measurement also assumes that people sway in two dimensions that are perpendicular to each other.

Greater accuracy in the measurement of body sway could be achieved with the use of force plates. These devices measure changes in the pressure applied to the feet, have been available for many years, and offer a variety of measures to assess sway. Research on the use of force plates should be undertaken to determine how these devices can be integrated into the assessment and the best measures of postural sway/balance.

The greatest costs of implementing force plates would be associated with the research required to develop standardized measures and procedures for interpreting the results. This, however, would not necessarily be borne by individual departments. The acquisition of the equipment and having the system programmed for use in the DECP would cost a few thousand dollars. Training and equipment maintenance of the system are not anticipated to be extensive.

10. Investigate the use of oral fluid to test for the presence of drugs. The final step in a DECP evaluation is the collection of a sample of bodily fluid for toxicological analysis of drug content. This is typically either blood or urine. Blood is the preferred medium because tests most often reveal the presence (and concentration) of the primary drug(s) rather than metabolites. The limitation associated with the collection of the blood sample is that legislation typically requires that blood draws be performed under the supervision of a physician. This can introduce substantial delays between the evaluation and sample collection. During this time, drug levels can decrease significantly.

Urine samples can be collected without medical supervision immediately following the evaluation. The drawback is that urine samples typically contain drug metabolites rather than the parent drug. The presence of metabolites indicates the use of the primary drug at some previous point in time that can be long before driving.

Oral fluid samples can be collected quickly and unobtrusively without special facilities or professional training. Toxicology testing of oral fluid samples generally reveals the presence of the primary psychoactive drug rather than metabolites and provides a better indicator of recent drug use than urine. Some drugs, however, are more difficult to detect in oral fluid than others.

The use of oral fluid as a medium to test for drugs could help reduce the time lag between the evaluation and the sample collection, strengthening the connection between the drug(s) and the symptoms observed.

With the assistance of toxicologists, the use of oral fluid as a medium in which to test for the presence of drugs should be investigated. A list of limitations and caveats associated with oral fluid drug testing should form part of the investigation. In addition, it would be necessary to consider any actual or perceived requirements of the relevant legislation and/or the courts to ensure oral fluid will be accepted.

11. Encourage and support new validation studies. The original DECP validation studies (e.g., Bigelow et al. 1985; Heishman et al., 1996, 1998) were conducted many years ago and are considered by many to be dated. New studies are needed to address some of the limitations of these previous studies to strengthen the evidence base for the DECP.

12. Incorporate a probationary period following completion of training and certifications during which a candidate must review a specified number (e.g., five) of enforcement evaluations under the supervision of an instructor. The number of evaluations performed by a DRE candidate as part of the certification requirement is minimal (i.e., 6). Six other evaluations are observed and scored. In many cases, these evaluations are performed on volunteers, not drivers suspected of impaired driving. Although officers selected for DECP training are often those with an interest and experience in impaired driving enforcement, conducting an evaluation on suspected drug-impaired drivers can be more challenging than evaluations conducted on volunteers.

A probationary period during which newly trained DREs conduct a series of evaluations under the supervision of an instructor would help to ensure that DREs are competent and comfortable with their new skills. Full certification would be on the recommendation of the supervisor in the officer's home department.

References

Alberg, A.J., Park, J.W., Hager, B.W., Brock, M.V. & Diener-West, M. (2004). The use of "overall accuracy" to evaluate the validity of screening or diagnostic tests. *Journal of General Internal Medicine*, 19(5), 460-465.

American Psychiatric Association. (2013) *Diagnostic and Statistical Manual of Mental Disorders* (5th Edition). Arlington, VA: American Psychiatric Publishing.

Barone, P. T., & Crampton, J. S. (2005). Do "Standardized" Field Sobriety Tests Reliably Predict Intoxication? Knowledge for drunk driving litigation. Michigan Bar Journal, 84(7), 22.

Beirness, D. J., Beasley, E., & LeCavalier, J. (2009). The accuracy of evaluations by Drug Recognition Experts in Canada. *Canadian Society of Forensic Science Journal,* 42(1), 75-79.

Beirness, D. J., LeCavalier, J. & Singhal, D. (2007). Evaluation of the Drug Evaluation and Classification Program: A Critical Review of the Evidence, Traffic Injury Prevention, 8:4, 368-376.

Bigelow, G.E., Bickel, W.E., Roache, J.D. Liebson, I.A. and Nowowieski, P. (1985) Identifying Types of Drug Intoxication: Laboratory Evaluation of the Subject Examination Procedure. DOT HS 806 753. Washington, D.C.: National Highway Traffic Safety Administration.

Blumenthal, I. (1992). Should we ban the mercury thermometer? *Journal of the Royal Society of Medicine*, 85,553-555.

Bosker, W. M. & Huestis, M.A. (2009). Oral fluid testing for drugs of abuse. Clinical Chemistry, 55 (11), 1910-1931.

Bosker, W. M., Theunissen, E. L., Conen, S., Kuypers, K. P. C., Jeffery, W. K., Walls, H. C., ... & Ramaekers, J. G. (2012). A placebo-controlled study to assess Standardized Field Sobriety Tests performance during alcohol and cannabis intoxication in heavy cannabis users and accuracy of point of collection testing devices for detecting THC in oral fluid. Psychopharmacology, 223(4), 439-446.

Bramness, J. G., Khiabani, H. Z., & Mørland, J. (2010). Impairment due to cannabis and ethanol: clinical signs and additive effects. Addiction, 105(6), 1080-1087.

Bramness, J. G., Skurtveit, S., & Mørland, J. (2003). Testing for benzodiazepine inebriation — relationship between benzodiazepine concentration and simple clinical tests for impairment in a sample of drugged drivers. European Journal of Clinical Pharmacology, 59(8-9), 593-601.

Brookoff, D., Cook, C. S., Williams, C., & Mann, C. S. (1994). Testing reckless drivers for cocaine and marijuana. New England Journal of Medicine, 331(8), 518-522.

Brunton, L.L., Lazo, J.S. & Parker, K.L. (2006). *Goodman and Gilman's The pharmacological basis of therapeutics* (11th edition). New York: McGraw-Hill.

Burns M., & Anderson E.W. (1995). A Colorado validation study of the Standardized Field Sobriety Tests (SFST) battery. Washington, D.C.: National Highway Traffic Safety Administration.

Burns, M., & Moskowitz, H. (1977). Psychophysical tests for DWI arrest (No. DOT-HS-802-424). Washington, D.C.: National Highway Traffic Safety Administration.

Cheng, W. C., Ng, K. M., Chan, K. K., Mok, V. K., & Cheung, B. K. (2007). Roadside detection of impairment under the influence of ketamine — evaluation of ketamine impairment symptoms with reference to its concentration in oral fluid and urine. *Forensic Science International*, 170(1), 51-58.

Cruz, S. L., Rivera-García, M. T., & Woodward, J. J. (2014). Review of toluene action: clinical evidence, animal studies and molecular targets. Journal of Drug and Alcohol Research, 3, 235840. http://doi.org/10.4303/jdar/235840

Dowding, D. (2002). An investigation into the accuracy of different types of thermometers. *Nursing Times*, 1 October, 2002.

Downey, L.A., King, R., Papafotiou, K., Swann, P., Ogden, E., & Stough, C. (2012). Examining the effect of dl-3, 4-methylenedioxymethamphetamine (MDMA) and methamphetamine on the standardized field sobriety tests. Forensic Science International, 220(1), e33-e36.

Drug Evaluation and Classification Program (DECP) (2017). 2016 Annual Report of the IACP Drug Evaluation and Classification Program. Arlington, VA: International Association of Chiefs of Police.

Giuliano, K.K., Scott, S.S., Elliot, S. & Giuliano, A.J. (1999). Temperature measurement in critically ill orally intubated adults: a comparison of pulmonary artery core, tympanic and oral methods. *Critical Care Medicine*, 27, 2188-2193.

Griffiths, A.N., Marshall, R.W. & Richens, A. (1984). Saccadic eye movement analysis as a measure of drug effects on human psychomotor performance. *British Journal of Clinical Pharmacology*, 18, 73S-82S.

Hanson, G. & Venturelli, P.J. (2000). *Drugs and Society* (5th Edition). Toronto Ontario: Jones and Bartlett Publishers.

Hayes, C. (2017). Safeguarding Our "High"ways. Combating Drugged Driving. Plenary presentation at the Minnesota Towards Zero Deaths Conference, St. Paul, MN. October 2017. Available at: www.minnesotatzd.org/events/conference/2017/documents/presentations/Drugged_driving.pdf

Hartman, R.L., Richman, J.E., Hayes, C.E., and Huestis, M.A. (2016). Drug Recognition Expert (DRE) examination characteristics of cannabis impairment. *Accident Analysis & Prevention,* 92, 219-29.

Heishman, S. J., Singleton, E. G., & Crouch, D. J. (1996). Laboratory validation study of drug evaluation and classification program: ethanol, cocaine, and marijuana. *Journal of Analytical Toxicology*, 20(6), 468-483.

Heishman, S.J., Singleton, E.G. and Crouch, D.J. (1998). Laboratory validation study of drug evaluation and classification program: Alprazolam, d-amphetamine, codeine, and marijuana. Journal of Analytical Toxicology, 22, 503-514.

Inaba, D.S. & Cohen, W.E. (2014). *Uppers, Downers, and All Arounders. Physical and Mental Effects of Psychoactive Drugs*. (8th Edition). Medford Oregon: CNS Productions Inc.

International Association of Chiefs of Police. (2015). Drug Recognition Expert Course. Participant Manual (Revised October, 2015). Arlington, VA: International Association of Chiefs of Police and National Highway Transportation Safety Administration.

Julien, R.M. Advokat, C.D. & Comaty, J.E. (2008). *A primer of drug action. A comprehensive guide to the actions, uses and side effects of psychoactive drugs* (11th edition). New York: Worth Publishers.

Khiabani, H.Z., Mørland, J. & Bramness, J.G. (2008). Frequency and irregularity of heart rate in drives suspected of driving under the influence of cannabis. European Journal of Internal Medicine, 19, 608-612.

Korsmeyer, P. and Kranzler, H.R. (Eds.) (2009). *Encyclopedia of Drugs, Alcohol and Addictive Behavior* (2nd Edition). New York: Macmillan Reference USA.

Kosnoski, E.M., Yolton, R.L., Citek, K., Hayes, C.E. & Evans, R.B. (1998). The Drug Evaluation and Classification Program: Using ocular and other signs to detect drug intoxications. *Journal of the American Optometric Association*, 69(4), 211-227.

Leikin, J.B. & Paloucek, F.P. (2007). *Poisoning and Toxicology Handbook*. (4th Edition). New York: CRC Press.

Logan, B., Kacinko, S.L. & Beirness, D.J. (2016). *An Evaluation of Data from Drivers Arrested for Driving Under the Influence in Relation to Per se Limits for Cannabis.* Washington, D.C.: AAA Foundation for Traffic Safety.

McKim, W.A. & Hancock, S. (2013). *Drugs and Behavior. An Introduction to Behavioral Pharmacology* (7th Edition), Toronto, Ontario: Pearson.

Myers, M.G. (2014). Replacing manual sphygmomanometers with automated blood pressure measurement in routine clinical practice. *Clinical and Experimental Pharmacology and Physiology*, 41,46-53.

Myers, M.G., Kaczorowski, J., Dawes, M., & Godwin, M. (2014). Automated office blood pressure measurement in primary care. *Canadian Family Physician*, 60, 127-132.

Owusu-Bempah, A. (2014). Cannabis impaired driving: An evaluation of current modes of detection. *Canadian Journal of Criminology and Criminal Justice*, 56(2), 219-240.

Page, T.E. (2007). The classification of drugs by category. In M. Burns (Ed.) *Medical-Legal Aspects of Drugs*. Second Edition. Tucson, AZ: Lawyers and Judges Publishing Company, Inc.

Papafotiou, K., Carter, J. D., & Stough, C. (2005a). The relationship between performance on the standardised field sobriety tests, driving performance and the level of Δ9-tetrahydrocannabinol (THC) in blood. *Forensic Science International,* 155(2), 172-178.

Papafotiou, K., Carter, J. D., & Stough, C. (2005b). An evaluation of the sensitivity of the Standardised Field Sobriety Tests (SFSTs) to detect impairment due to marijuana intoxication. *Psychopharmacology*, 180(1), 107-114.

Peragallo, J., Biousse, V. & Newman, N.J. (2013). Ocular manifestations of drug and alcohol abuse. *Current Opinion in Ophthalmology*, 24(6), 566-573.

Porath-Waller, A. J., & Beirness, D. J. (2010). Simplifying the process for identifying drug combinations by drug recognition experts. *Traffic Injury Prevention*, 11(5), 453-459.

Porath-Waller, A. J., & Beirness, D. J. (2014). An examination of the validity of the standardized field sobriety test in detecting drug impairment using data from the drug evaluation and classification program. *Traffic Injury Prevention,* 15(2), 125-131.

Porath-Waller, A. J., Beirness, D. J., & Beasley, E. E. (2009). Toward a more parsimonious approach to drug recognition expert evaluations. *Traffic Injury Prevention*, 10(6), 513-518.

Rome, E.S. (2001). It's a rave new world: Rave culture and illicit drug use in the young. *Cleveland Clinic Journal of Medicine*, 68(6), 541-550.

Rubenzer, S. J. (2008). The Standardized Field Sobriety Tests: a review of scientific and legal issues. *Law and Human Behavior*, 32(4), 293.

Schechtman, E. & Shinar, D. (2005) Modeling drug detection and diagnosis with the 'drug evaluation and classification program.' *Accident Analysis & Prevention,* Vol. 37, pp 852-861.

Schmitz, T., Bair, N., Falk, M. & Levine, C. (1995). A comparison of five methods of temperature measurement in febrile intensive care patients. *American Journal of Critical Care*, 4, 286-292.

Schuckit, M.A. (2006). *Drugs and alcohol abuse: A clinical guide to diagnosis and treatment*. New York: Springer.

Shinar, D., & Schechtman, E. (2005). Drug identification performance on the basis of observable signs and symptoms. *Accident Analysis & Prevention,* 37(5), 843-851.

Silber, B. Y., Papafotiou, K., Croft, R. J., & Stough, C. K. (2005). An evaluation of the sensitivity of the standardised field sobriety tests to detect the presence of amphetamine. Psychopharmacology, 182(1), 153-159.

Smink, B. E., Lusthof, K. J., De Gier, J. J., Uges, D. R., & Egberts, A. C. (2008). The relation between the blood benzodiazepine concentration and performance in suspected impaired drivers. Journal of Forensic and Legal Medicine, 15(8), 483-488.

Smith, J.A.; Hayes, C.E., Yolton, R.L.; Rutledge, D.A., & Citek, K. (2002). Drug recognition expert evaluations made using limited data. *Forensic Science International,* Vol. 130(2/3), 167-173.

Stevens, D.L. & Tomlinson, G.E. (1971). Measurement of human postural sway. *Proceedings of the Royal Society of Medicine*, 64, 653-655.

Stuster, J. (1997). The Detection of DWI at BACs Below 0.10. DOT-HS-808-654. Washington, D.C.: National Highway Traffic Safety Administration.

Stuster, J. & Burns, M.(1998). Validation of the Standardized Field Sobriety Test Battery at BACs Below 0.10 Percent. DOT-HS-808-839. Washington, D.C.: National Highway Traffic Safety Administration.

SWOV (2015). *The use of drugs and medicines behind the wheel*. SWOV Fact sheet, July 2015, The Hague.

Tabachnick, B.G., & Fidell, L.S. (2007). Using Multivariate Statistics (5th ed.). Boston, MA: Allyn and Bacon.

Tharp, V., Burns, M., & Moskowitz, H. (1981). Development and field test of psychophysical tests for DWI arrest (No. DOT-HS-805-864). National Highway Traffic Safety Administration.

Willetts, G.S. (1969). Ocular side-effects of drugs. *British Journal of Ophthalmometry*, 53,252-262.

Appendix A: DECP Face Sheet

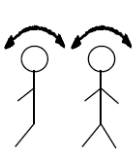
DRUG INFLUENCE EVALUATION

Evaluator	DRE # / Rolling Log #		Evaluator's Agency
Recorder/Witness	Crash: ☐ None ☐ Fatal ☐ Injury ☐ Property		Arresting Officer's Agency
Arrestee's Name (Last, First, Middle)	Date of Birth	Sex / Race	Arresting Officer (Name, ID#)
Date Examined / Time /Location	Breath Results: / Results:	Test Refused ☐ / Instrument #:	Chemical Test: Urine ☐ Blood ☐ / Test or tests refused ☐

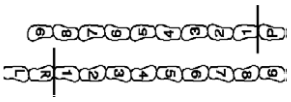
Miranda Warning Given ☐ Yes ☐ No / Given By:	What have you eaten today? When?	What have you been drinking? How much?	Time of last drink?	
Time now/ Actual	When did you last sleep? How long	Are you sick or injured? ☐ Yes ☐ No	Are you diabetic or epileptic? ☐ Yes ☐ No	
Do you take insulin? ☐ Yes ☐ No	Do you have any physical defects? ☐ Yes ☐ No	Are you under the care of a doctor or dentist? ☐ Yes ☐ No		
Are you taking any medication or drugs? ☐ Yes ☐ No	Attitude:	Coordination:		
Speech:	Breath Odor:	Face:		
Corrective Lenses: ☐ None ☐ Glasses ☐ Contacts, if so ☐ Hard ☐ Soft	Eyes: ☐ Reddened Conjunctiva ☐ Normal ☐ Bloodshot ☐ Watery	Blindness: ☐ None ☐ Left ☐ Right	Tracking: ☐ Equal ☐ Unequal	
Pupil Size: ☐ Equal ☐ Unequal (explain)	Vertical Nystagmus ☐ Yes ☐ No	Able to follow stimulus ☐ Yes ☐ No	Eyelids ☐ Normal ☐ Droopy	

Pulse and time 1. ____ / ____ 2. ____ / ____ 3. ____ / ____	HGN	Right Eye	Left Eye	Convergence Right eye Left eye	ONE LEG STAND
	Lack of Smooth Pursuit				
	Maximum Deviation				
	Angle of Onset				

Romberg Balance



Walk and turn test



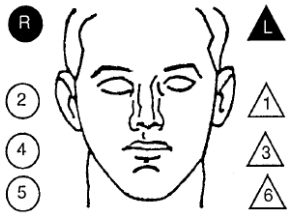
	1st Nine	2nd Nine
Cannot keep balance		
Starts too soon		
Stops walking		
Misses heel-toe		
Steps off line		
Raises arms		
Actual steps taken		

ONE LEG STAND

L R
☐ ☐ Sways while balancing
☐ ☐ Uses arms to balance
☐ ☐ Hopping
☐ ☐ Puts foot down

Internal clock ____estimated as 30 seconds	Describe Turn	Cannot do test (explain)	Type of footwear:

Draw lines to spots touched



PUPIL SIZE	Room light 2.5 – 5.0	Darkness 5.0 – 8.5	Direct 2.0 – 4.5
Left Eye			
Right Eye			

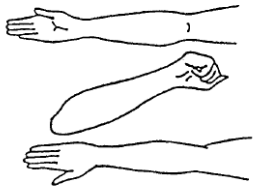
Nasal area:

Oral cavity:

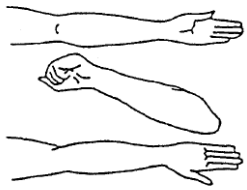
REBOUND DILATION ☐ Yes ☐ No

REACTION TO LIGHT:

RIGHT ARM



LEFT ARM



Blood pressure	Temperature

Muscle tone:
☐ Near Normal ☐ Flaccid ☐ Rigid
Comments:

What drugs or medications have you been using?	How much?	Time of use?	Where were the drugs used? (Location)
Date / Time of arrest:	Time DRE was notified:	Evaluation start time: / Evaluation completion time:	Precinct/Station:
Officer's Signature:	DRE #	Reviewed/approved by / date:	

Opinion of Evaluator:
☐ Rule Out ☐ Alcohol ☐ CNS Stimulant ☐ Dissociative Anesthetic ☐ Inhalant
☐ Medical ☐ CNS Depressant ☐ Hallucinogen ☐ Narcotic Analgesic ☐ Cannabis

Revised. 05/2013

Appendix B: DECP Matrix

	CNS Depressants	Inhalants	Dissociative Anesthetics	Cannabis	CNS Stimulants	Hallucinogens	Narcotic Analgesics
HGN	Yes	Yes	Yes	No	No	No	No
Vertical Nystagmus	Yes (high dose)	Yes (high dose)	Yes	No	No	No	No
Non-convergence	Yes	Yes	Yes	Yes	No	No	No
Pupil Size	Normal[1]	Normal[4]	Normal	Dilated/Normal	Dilated	Dilated	Constricted
Reaction to Light	Slow	Slow	Normal	Normal	Slow	Normal[3]	Little or None Visible
Pulse Rate	Down[2]	Up	Up	Up	Up	Up	Down
Blood Pressure	Down	Up/Down[5]	Up	Up	Up	Up	Down
Body Temp	Normal	Up/Down/ Normal	Up	Normal	Up	Up	Down
Muscle Tone	Flaccid	Flaccid	Rigid	Normal	Rigid	Rigid	Flaccid

1. Soma and Quaaludes usually dilate pupils
2. Quaaludes and alcohol may elevate
3. Certain psychedelic amphetamines cause slowing
4. Normal but may dilate.
5. Down with anesthetic gases; Up with volatile solvents

Appendix C: Vital Signs/Clinical Indicators from the DECP Matrix with References

Each of the following tables presents the vital signs and clinical indicators examined as part of a DECP evaluation as noted in the DECP Matrix for each of the seven drug categories. The left column indicates the signs/symptoms examined. Any additional signs/symptoms are listed at the bottom of the table.

The second column (labeled DECP Matrix Expected Effect) lists the expected effect of the drug for each indicator, per the DECP Matrix (see Legend below).

The third column lists references for the effect(s). References reporting an effect contradictory to that listed in the DECP Matrix are marked with the appropriate legend symbol.

It is important to note that the information and related references presented in this table are not intended to be comprehensive or definitive. The intent was to illustrate a basis for the drug effects utilized by the DECP in the medical/scientific literature. The references listed are primarily pharmacology and/or medical texts, compendiums, and summaries that list or refer to specific drug effects, including both original studies and reviews of other studies. For the most part, the specific drug effects noted are typically listed as "facts" without reference to the original studies. Hence, the effects noted are taken at face value without critical assessment. Places where studies show inconclusive or divergent outcomes are indicative of an area where more research may be merited.

Legend

✓	**Sign or indicator is present**
X	**Sign or indicator is not present**
↑	**Sign or indicator is elevated**
↓	**Sign or indicator is reduced**
Normal	**Sign or indicator is within the average range for drug-free subjects**

CNS Depressants

Signs/symptoms	DECP Matrix Expected Effect	References
Nystagmus[14]	✓	American Psychiatric Association (2013) Peragallo et al. (2013)
Non-convergence	✓	
Pupil Size	Normal	↓ (high dose Zolpidem) Leikin & Paloucek (2007)
Reaction to Light	↓	Schuckit (2006)
Pulse Rate	↓	American Psychiatric Association (2013) Leikin & Paloucek (2007) Rome (2001) Little to no effect - McKim & Hancock (2013) (benzodiazepines)
Blood Pressure	↓	American Psychiatric Association (2013) Leikin & Paloucek (2007) Little to no effect benzodiazepines - McKim & Hancock (2013)
Body Temp	Normal	↓ Schuckit (2006)
Muscle Tone	Flaccid	Brunton, Lazo & Parker (2006) Julien et al. (2008) McKim & Hancock (2013) Schuckit (2006) Stedman (1990)
Additional Effects		
Respiration	↓	American Psychiatric Association (2013) McKim & Hancock (2013) Schuckit (2006)
Ptosis (Droopy Eyelids)	✓	Peragallo et al. (2013)

[14] The literature does not necessarily distinguish between horizontal and vertical gaze nystagmus. Vertical gaze nystagmus may be evident in cases of high doses of drugs that induce horizontal gaze nystagmus. The exception is dissociative anaesthetics. Vertical gaze nystagmus may be evident even at low doses of these drugs.

Inhalants

Signs/symptoms	DECP Matrix Expected Effect	References
Nystagmus	✓	American Psychiatric Association (2013) Kosnoski et al. (1998) Schuckit (2006)
Non-convergence	✓	Kosnoski et al. (1998)
Pupil Size	Normal	
Reaction to Light	↓	
Pulse Rate	↑	American Psychiatric Association (2013) Leikin & Paloucek (2007)
Blood Pressure	↓ ↑	Brunton, Lazo & Parker (2006) Leikin & Paloucek (2007)
Body Temp	↓ / ↑ / Normal	↓ Cruz et al. (2014) ↑ National Library of Medicine
Muscle Tone	Flaccid	American Psychiatric Association (2013)
Additional Effects		
Tremors	✓	American Psychiatric Association (2013)
Respiration Depression	✓	Schuckit (2006)

Dissociative Anesthetics

Signs/symptoms	DECP Matrix Expected Effect	References
Nystagmus	✓	American Psychiatric Association (2013) Leikin & Paloucek (2007) Hanson & Venturelli (2000) (high dose) Brunton et al. (2006) Rome (2001) Peragallo et al. (2013)
Non-convergence	✓	
Pupil Size	Normal	Dilation - Brunton et al. (2006) Leikin & Paloucek (2007)
Reaction to Light	Normal	
Pulse Rate	↑	Leikin & Paloucek (2007) Julien et al. (2008) Brunton et al. (2006) Schuckit (2006) Hanson & Venturelli (2000)
Blood Pressure	↑	American Psychiatric Association (2013) Leikin & Paloucek (2007) Hanson & Venturelli (2000) Julien et al. (2008) Brunton et al. (2006) Rome (2001)
Body Temp	↑	American Psychiatric Association (2013) Hanson & Venturelli (2000) Julien et al. (2008) Leikin & Paloucek (2007) Schuckit (2006) (↑possible) American Psychiatric Association (2013)
Muscle Tone	Rigid	American Psychiatric Association (2013) Brunton et al. (2006) Leikin & Paloucek (2007)
Additional Effects		
Tremors	✓	Julien et al. (2008) Leikin & Paloucek (2007)
Respiratory Depression	✓	Brunton et al. (2006) Leikin & Paloucek (2007) Rome (2001)

Cannabis

Signs/symptoms	DECP Matrix Expected Effect	References
Nystagmus	X	Possible: Peragallo et al. (2013) Schuckit (2006)
Non-convergence	X	Kosnoski et al. (1998)
Pupil Size	Dilated /Normal	Dilated: Leikin & Paloucek (2007) Normal: Stafford (1993)
Reaction to Light	Normal	
Pulse Rate	↑	Korsmeyer & Kranzler (2009) Inaba & Cohen (2014) Julien et al. (2008) Khiabani et al (2008) Leikin & Paloucek (2007) McKim & Hancock (2013) Schuckit (2006) Stafford (1993)
Blood Pressure	↑	Julien et al. (2008) Leikin & Paloucek (2007) Rome (2001) ↓ (supine or orthostatic) Korsmeyer & Kranzler (2009) ↓ (supine or orthostatic) Inaba & Cohen (2014)
Body Temp	Normal	↑ Ashton (1999) ↓ Julien et al. (2008)
Muscle Tone	Normal	
Additional Effects		
Tremors	✓	Ashton (1999) Julien et al. (2008) Leikin & Paloucek (2007) Schuckit (2006)
Reddened Conjunctiva (Conjunctival Injection)	✓	American Psychiatric Association (2013) Ashton (1999) Julien et al. (2008) Korsmeyer & Kranzler (2009) Leikin & Paloucek (2007) McKim & Hancock (2013) Peragallo et al. (2013) Rome (2001) Schuckit (2006)
Droopy eyelids (Ptosis)	✓	McKim & Hancock (2013)

CNS Stimulants

Signs/symptoms	DECP Matrix Expected Effect	References
Nystagmus	X	
Non-convergence	X	
Pupil Size	Dilated	American Psychiatric Association (2013) Hanson & Venturelli (2000) Julien et al. (2008) Leikin & Paloucek (2007) Peragallo et al. (2013) Rome (2001)
Reaction to Light	Slow	Kosnoski et al. (1998) Spotts & Spotts (1980)
Pulse Rate	↑	Brunton et al. (2006) Hanson & Venturelli (2000) Julien et al. (2008) Korsmeyer & Kranzler (2009) Leikin & Paloucek (2007) McKim & Hancock (2013) Schuckit (2006) ↑↓American Psychiatric Association (2013) ↑↓Rome (2001)
Blood Pressure	↑	Brunton et al. (2006) Hanson & Venturelli (2000) Julien et al. (2008) Korsmeyer & Kranzler (2009) Leikin & Paloucek (2007) McKim & Hancock (2013) Schuckit (2006) ↑↓American Psychiatric Association (2013) ↑↓Rome (2001)
Body Temp	↑	Brunton et al. (2006) Hanson & Venturelli (2000) Julien et al. (2008) Leikin & Paloucek (2007)
Muscle Tone	Rigid	Schuckit (2006)
Additional Effects		
Tremors/Spasms (Dystonia)	✓	Hanson & Venturelli (2000) Julien et al. (2008) Leikin & Paloucek (2007) Rome (2001) Schuckit (2006)
Respiration	↑	American Psychiatric Association (2013) Julien et al. (2008)

		Korsmeyer & Kranzler (2009) **Hanson & Venturelli (2000)** **Brunton et al. (2006)**
Arrhythmias	✓	**Brunton et al. (2006)** **Hanson & Venturelli (2000)** **Korsmeyer & Kranzler (2009)** **Rome (2001)** **Spotts & Spotts (1980)**

Hallucinogens

Signs/symptoms	DECP Matrix Expected Effect	References
Nystagmus	X	
Non-convergence	X	
Pupil Size	Dilated	American Psychiatric Association (2013) Brunton et al. (2006) Hanson & Venturelli (2000) Inaba & Cohen (2014) Julien et al. (2008) Leikin & Paloucek (2007) McKim & Hancock (2013) Peragallo et al. (2013) Rome (2001) Schuckit (2006)
Reaction to Light	Normal	Brunton et al. (2006)
Pulse Rate	↑	Brunton et al. (2006) Inaba & Cohen (2014) Julien et al. (2008) Leikin & Paloucek (2007) McKim & Hancock (2013) Schuckit (2006) Constricted: Leikin & Paloucek (2007) (mescaline) Hanson & Venturelli (2000) (mescaline)
Blood Pressure	↑	Julien et al. (2008) Leikin & Paloucek (2007) Inaba & Cohen (2014) Schuckit (2006) Brunton et al. (2006) Hanson & Venturelli (2000) Rome (2001)
Body Temp	↑	Korsmeyer & Kranzler (2009) Hanson & Venturelli (2000) Inaba & Cohen (2014) Julien et al. (2008) Leikin & Paloucek (2007) McKim & Hancock (2013) Schuckit (2006)
Muscle Tone	Rigid	Julien et al. (2008) Leikin & Paloucek (2007) Brunton et al. (2006) Inaba & Cohen (2014) Hanson & Venturelli (2000)

Additional Effects		
Dry Mouth/Thirst	✓	**<u>MDMA</u>: Hanson & Venturelli (2000)** **Inaba & Cohen (2014)** **Julien et al. (2008)** **Leikin & Paloucek (2007)** **Schuckit (2006)**
Tremors	✓	**American Psychiatric Association (2013)** **Julien et al. (2008)** **Leikin & Paloucek (2007)** **Schuckit (2006)**

Narcotic Analgesics

Signs/symptoms	DECP Matrix Expected Effect	References
Nystagmus	X	
Non-convergence	X	
Pupil Size	Constricted	American Psychiatric Association (2013) Brunton et al. (2006) Hanson & Venturelli (2000) Inaba & Cohen (2014) Julien et al. (2008) Leikin & Paloucek (2007) McKim & Hancock (2013) Peragallo et al. (2013) Rome (2001) Dilated: (meperidine) Schuckit (2006)
Reaction to Light	Little/None	
Pulse Rate	↓	Brunton et al. (2006) Inaba & Cohen (2014) Leikin & Paloucek (2007)
Blood Pressure	↓	Brunton et al. (2006) Korsmeyer & Kranzler (2009) Hanson & Venturelli (2000) Inaba & Cohen (2014) Julien et al. (2008) Leikin & Paloucek (2007) McKim & Hancock (2013) Rome (2001)
Body Temp	↓	Julien et al. (2008)
Muscle Tone	Flaccid	Leikin & Paloucek (2007) Schuckit (2006)
Additional Effects		
Respiratory Depression	✓	Julien et al. (2008) Hanson & Venturelli (2000) Schuckit (2006)
"On the nod"	✓	American Psychiatric Association (2013) Hanson & Venturelli (2000) Inaba & Cohen (2014)
Myoclonus (Muscle spasms)	✓	Leikin & Paloucek (2007)

Appendix D: Key Informant Discussion Guide

Background

Beirness and Associates is conducting a study for the AAA Foundation for Traffic Safety to review the Drug Evaluation and Classification Program (DECP) and examine options for enhancing the program. This includes all aspects of the program — recruitment, training, certification, the tests, procedures, and administration of the program. To accomplish this objective, we are inviting a number of people who have knowledge of, interest in, and/or experience with the DECP to share their thoughts and opinions about the strengths and limitations of the DECP.

Interviews are confidential and all comments will remain anonymous. Nothing you say will be linked to you or your agency or department. No names will be used in the reporting of results.

We anticipate that each interview will last up to an hour.

Topic Areas for Discussion

Involvement in DECP

- Role, extent of involvement, specific areas of involvement, years involved, number of evaluations conducted
- Recruitment/entrée into DECP
- Training: initial and continuing
- Amount (or percentage) of time devoted to DEC activities
- Specific activities

Overall Impressions of the DECP

- Importance of the DECP in impaired driving enforcement
- Comprehensiveness/consistency of the training
- Management/oversight of the program (agency and/or state/provincial coordination)
- Funding for the program
- Impact of the program on your career path as an officer

Strengths of the Program

- What do you see as the major strengths of the program?
  (What makes it a good program? e.g., recruitment, training, continuing education, supporting research, prosecution, toxicology, management, funding)
- How can we build on those strengths to make the program better?
- What are the most important aspects of a DRE evaluation?
- Is there value of the DECP beyond impaired driving?

- Perceptions of the most important/effective parts of the DECP

Limitations of the Program

- What would you say are the weak points of the program?
  (e.g., recruitment, training, continuing education, supporting research, prosecution, toxicology, management support, funding)
- What is the impact of these weaknesses?
- To what extent do these limitations affect your ability to do your job as a DRE?
- Are there parts of the DECP evaluation that you feel aren't very relevant or useful?

Suggestion for Enhancing the Program

- What areas/aspects of the program could be improved/updated/revised? What would be needed to make it better?
- Validating FTN, MRB?
- Updated normative data for clinical and psychophysical indicators?
- Ongoing training, educational updates, regional conferences/workshops, case reports, research updates, new or emerging drug alerts?
- Are there efficiencies that could be introduced in training/operations/management?

New Technologies

- Are you aware of technologies that could be introduced to assist and/or enhance the efficiency/accuracy of evaluations?
- Would you consider using assisting technologies?
- What would be involved in introducing new technologies?
- Who should approve these technologies?
- What would be the drawbacks/limiting factors?

Future of DECP

- What do you see in the future for the DECP?
- Are there legislative changes that could better support DREs and enhance the program?