Traffic Injury Prevention (2014) 15, 125–131

ISSN: 1538-9588 print / 1538-957X online
DOI: 10.1080/15389588.2013.800638

Taylor & Francis
Taylor & Francis Group

An Examination of the Validity of the Standardized Field Sobriety Test in Detecting Drug Impairment Using Data from the Drug Evaluation and Classification Program

AMY J. PORATH-WALLER and DOUGLAS J. BEIRNESS

Canadian Centre on Substance Abuse, Ottawa, Ontario, Canada

Received 4 February 2013, Accepted 25 April 2013

Objective: The purpose of this study is to assess the validity of the 3 components of the Standardized Field Sobriety Test (SFST), including the Horizontal Gaze Nystagmus (HGN), One Leg Stand (OLS), and Walk and Turn (WAT) tests, in identifying impairment among suspected drug-impaired drivers using data recorded during drug evaluation and classification (DEC) evaluations.

Methods: Data from 2142 completed DEC evaluations of central nervous system (CNS) stimulants, CNS depressants, narcotic analgesics, cannabis, or no drugs were analyzed using multinomial logistic regression.

Results: All drug categories were significantly associated with impaired performance. On the HGN, users of CNS depressants were significantly more likely to experience lack of smooth pursuit and distinct nystagmus at maximum deviation compared to non-drug users. On the OLS, users of all drug classes were significantly more likely to sway while balancing and use their arms to maintain balance but significantly less likely to hop compared to drug-free cases. Users of CNS depressants, CNS stimulants, and narcotic analgesics were significantly more likely to put their raised foot down during the test. On the WAT, users of CNS depressants, CNS stimulants, and narcotic analgesics were significantly less likely to keep their balance while listening to test instructions compared to those who had not used drugs. Users of CNS depressants were less likely to touch heel-to-toe while walking, whereas individuals who had used narcotic analgesics were less likely to take the correct number of steps.

Conclusions: These findings provide support for the use of the SFST as a screening tool for law enforcement to identify impairment in persons who have used CNS stimulants, CNS depressants, cannabis, or narcotic analgesics.

Keywords: Standardized Field Sobriety Test (SFST), impairment testing, impaired driving, drugged driving, Drug Evaluation and Classification (DEC) Program

Introduction

Police officers are often faced with deciding whether or not a suspected impaired driver should be arrested and taken to the station for an evidential breath or blood test. Even after the development and widespread implementation of small, handheld breath alcohol testing devices, police officers often required evidence of behavioral and/or cognitive impairment to assist in arrest decisions. In addition, this evidence was often critical in court to show that the suspect was adversely affected by alcohol.

Over the past several decades, a wide variety of coordination tests have been used by police officers to determine whether or not a driver was impaired by alcohol. These included such tasks as reciting the alphabet, counting backwards by 3s, walking a straight line, touching finger to nose, standing steadiness, and picking up coins. These are primarily tests of balance, coordination, and speech, skills that were believed to be adversely affected by alcohol in a dose-related manner (Burns and Moskowitz 1977). However, the procedures for administering and interpreting these tests varied widely. Most of these tests also lacked any type of scientific evidence of validity and reliability.

The search for a short battery of tests to detect driver impairment that could be utilized and scored at roadside began in earnest during the 1970s. The National Highway Traffic Safety Administration set out to develop a standardized battery of tests that would identify impaired drivers in a valid and reliable manner. The researchers initially identified a set of 16 potentially suitable tasks that had been described in the research literature as being sensitive to the effects of alcohol. Through an initial review and pilot study, the number of tests was reduced to 10—One Leg Stand, Finger to Nose, Finger Count, Walk and Turn, Tracing Mazes, Nystagmus, Romberg Body Sway, Subtraction, Letter Cancellation, and Backward Counting. All tests were found to be correlated with blood alcohol concentration (BAC); that is, greater performance deficits were evident as the BAC of the person increased. Further statistical analysis identified the 3 best tests

Address correspondence to Amy J. Porath-Waller, PhD, Canadian Centre on Substance Abuse, 75 Albert Street, Suite 500, Ottawa, ON, Canada K1P 5E7. E-mail: aporath-waller@ccsa.ca

Laboratory studies such as those described previously are of considerable value in terms of validating the overall SFST procedure and identifying specific components of this battery and/or patterns of performance on these tests. Building on this limited literature, the objective of the present study was to examine data from the components of the SFST that are recorded during DEC evaluations as a means to assess the validity of the SFST in identifying impairment among suspected drug-impaired drivers. In addition to assessing the validity of the SFST for detecting impairment due to cannabis and central nervous system (CNS) stimulants, the current work will also examine the drug classes of narcotic analgesics and CNS depressants.

Methods

Sample

Data from 2142 DEC evaluations conducted across Canada involving a single drug category that were conducted during 1995–2009 were used in this study. As part of the DEC Program administration in Canada, all evaluations conducted by drug recognition experts (DREs) and the corresponding toxicology reports are routinely sent to the national coordinator at the Royal Canadian Mounted Police (RCMP) Headquarters in Ottawa. The DEC evaluations were made available for analysis as part of a larger evaluation of the DEC program. These DEC reports contain all of the data collected during the evaluation of suspects as well as the opinion of the DRE about the category of drug involved. The DEC evaluations were supplemented by the results of the toxicological tests performed on the bodily fluid sample collected from suspects at the conclusion of the DEC examination. Recent work by Beirness and colleagues (2009) reported an overall accuracy rate of 95 percent for these evaluations, as determined by a match between the drug category noted by the evaluator and that in the toxicology report. All personal identifying information (i.e., suspect's name, file number, etc.) was removed from the DEC evaluations by the RCMP prior to their receipt by the investigators. The data from the DEC evaluations were entered by the investigators into a database that was saved on a password-protected computer.

Four classes of drugs were represented in this set of evaluations: CNS stimulants ($n = 852$), CNS depressants ($n = 135$), narcotic analgesics ($n = 312$), and cannabis ($n = 703$). There were also 140 "no-drug" cases whereby the opinion of the evaluator was that the suspect was not under the influence of any drug and no drug was found as a result of toxicological analysis of the bodily fluid sample provided. Both of these criteria had to be met in order to be classified as a no-drug case.

Standardized Field Sobriety Test

The 3 tests of the SFST (HGN, OLS, and WAT) are imbedded in the 12-step protocol of the DEC evaluation. Data from the DEC evaluations on the 3 tests that comprise the SFST battery were analyzed for their potential association with the 4 drug categories. These 3 tests are briefly summarized below, as are the specific signs observed during the tests that were included in the current analysis.

Horizontal Gaze Nystagmus Test: HGN is an involuntary jerking of the eye that occurs naturally as the eyes gaze to the side. During the HGN test, the eyes of an individual are observed as the individual follows a slowly moving object, such as a pen, horizontally with his or her eyes as it is moved from side to side. The officer separately observes the left and right eye for 3 signs: lack of smooth pursuit (present, absent); distinct nystagmus at maximum deviation (present, absent); and nystagmus onset before 45° (present, absent). Research has shown that 88 percent of individuals who present 4 or more clues, between the 2 eyes, on this test will have likely have a BAC of 80 mg/dL or greater (Stuster and Burns 1998). In the current analysis, signs from only the left eye were used for the sake of parsimony.

One Leg Stand Test: In this test, the individual is instructed to stand with one foot approximately 15 cm off the ground and count aloud from 1000 (1000, 1001, 1002, etc.) for 30 seconds. There are 4 signs from the OLS test that are scored; that is, swaying while balancing on one leg; using arms to maintain balance; hopping during test; and putting the raised foot down. Research has indicated that 83 percent of individuals who exhibit 2 or more indicators in the performance of this test will have a BAC of 80 mg/dL or greater (Stuster and Burns 1998). In the current analysis, the signs for both legs were summed and averaged to provide an overall indication of impairment becaus it was not possible to determine the "preferred leg" of the suspect.

Walk and Turn Test: In the WAT test, the participant is directed to take 9 steps, heel-to-toe, along a straight line. After taking the 9 steps, the participant must turn on one foot and return in the same manner in the opposite direction. There are 8 signs of impairment that can be observed during this test; that is, could not keep balance while listening to the test instructions; started the test before the instructions were completed; stopped walking during the test; did not touch heel-to-toe while walking; stepped off the line; used arms to maintain balance; took the incorrect number of steps; and turned improperly (not as demonstrated). Research has shown that 79 percent of individuals who exhibit 2 or more indicators in the performance of this test will have a BAC of 80 mg/dL or greater (Stuster and Burns 1998). In the current analysis, the signs for the first 9 steps and the second 9 steps were summed to provide an overall indication of impairment. The improper turn (i.e., whether the suspect performed the turn correctly or incorrectly) from this test was not assessed in the current study due to data coding limitations.

Data Analysis

A series of multinomial logistic regression analyses was performed to assess the prediction of drug category from the various signs observed during the SFST battery (Tabachnick and Fidell 2007). Separate analyses were conducted for each of the 3 components of the SFST (i.e., HGN, OLS, and WAT tests). Multinomial logistic regression allows the prediction

Table 1. Contribution of signs from the HGN test in predicting drug category

Signs	χ^2 to Remove	*df*
Lack of smooth pursuit	32.53*	4
Distinct nystagmus at maximum deviation	100.82*	4
Nystagmus onset before 45°	5.84	4

*$P < .0167$.

of an outcome variable that has more than 2 categories from a set of predictor variables that may be continuous, discrete, dichotomous, or a mix. Classification rates for the outcome categories were also calculated as part of the analyses because they provide an estimate of the success of the model in correctly predicting the outcome category for cases for which the outcome is known (Tabachnick and Fidell 2007).

Results

Prediction of Drug Category From Performance on Horizontal Gaze Nystagmus Test

A multinomial logistic regression analysis was performed on the set of DEC cases to assess the prediction of drug category from performance on the HGN test. Results indicated that the set of 3 signs from the HGN test significantly distinguished the 4 drug categories of CNS stimulants, CNS depressants, narcotic analgesics, and cannabis from the no-drug cases, $\chi^2 (12, N = 2142) = 442.65$, $P < .0001$. The classification rate for these drug categories was 42.2 percent; that is, less than half of all cases were correctly classified based on the inclusion of these 3 signs from the HGN test. The classification rate was 94.6 percent for CNS stimulants, 70.1 percent for CNS depressants, 0 percent for narcotic analgesics, and 1 percent for cannabis. Table 1 shows the unique contribution of the individual predictors to the overall multinomial logistic regression model by comparing models with and without each predictor. Using a Bonferonni correction ($P < .0167$) to control for type I error, 2 of the 3 signs significantly contributed to the prediction of drug category: lack of smooth pursuit and distinct nystagmus at maximum deviation (Table 1).

As a follow-up to the overall multinomial logistic regression analysis, a series of binary logistic regression analyses was conducted to determine the specific signs from the HGN test that distinguished each of the 4 drug categories from the no-drug cases (reference group). The regression coefficients, chi-square tests, odds ratios (ORs), and 95 percent confidence intervals for these analyses appear in Table 2. The ORs indicate whether there is an increased or decreased likelihood of the signs being associated with the particular drug category as compared to the no-drug group; ORs greater than 1 reflect an increased likelihood, whereas ORs less than 1 reflect a decreased likelihood (in some instances the ORs have been flipped to avoid stating double negatives and ease interpretation for the reader). The results indicated that users of CNS depressants were significantly more likely to experience lack of smooth pursuit and distinct nystagmus at maximum deviation compared to individuals who were not positive for drug use.

Prediction of Drug Category From Performance on One Leg Stand Test

A separate multinomial logistic regression analysis was performed to predict drug category from performance on the OLS test and the results showed that all 4 signs from this psychophysical test significantly distinguished the 4 DEC drug categories from the no-drug cases, $\chi^2 (16, N = 2142) = 305.79$, $P < .0001$ (Table 3). Based on this set of 4 signs, 43.6 percent of all cases were correctly classified, with classification being the highest for CNS stimulants (59.9%), followed by cannabis (55.4%) and narcotic analgesics (10.6%). No CNS depressant cases were correctly classified based on these signs from the OLS test.

To examine the specific signs from the OLS test that distinguished the 4 drug categories from the no-drug cases, a series of binary logistic regression analyses was performed. As shown in Table 4, users of all 4 drug categories were significantly more likely to sway while balancing on one leg or use their arms to maintain balance during the OLS test compared

Table 2. Prediction of drug category vs. no drug use from signs observed during the HGN test (*$p < .0167$)

Signs	*B*	*SE*	Wald's χ^2 test	OR	95% Confidence interval for OR
CNS depressants					
Lack of smooth pursuit	−1.76	0.46	14.67*	0.17	0.07, 0.42
Distinct nystagmus at maximum deviation	−2.77	0.52	28.39*	0.06	0.02, 0.17
Nystagmus onset before 45°	−0.75	0.70	1.13	0.48	0.12, 1.88
CNS stimulants					
Lack of smooth pursuit	−0.35	0.39	0.79	0.71	0.33, 1.52
Distinct nystagmus at maximum deviation	−0.08	0.46	0.03	0.92	0.37, 2.28
Nystagmus onset before 45°	−0.02	0.70	0.001	0.98	0.25, 3.88
Narcotic analgesics					
Lack of smooth pursuit	−0.73	0.42	3.07	0.48	0.21, 1.09
Distinct nystagmus at maximum deviation	1.22	0.59	4.34	3.39	1.07, 10.67
Nystagmus onset before 45°	−0.57	0.79	0.53	0.56	0.12, 2.63
Cannabis					
Lack of smooth pursuit	−0.12	0.40	0.10	0.88	0.41, 1.93
Distinct nystagmus at maximum deviation	−0.15	0.47	0.10	0.86	0.35, 2.15
Nystagmus onset before 45°	−0.24	0.71	0.12	0.78	0.20, 3.12

Table 3. Contribution of signs from the OLS test in predicting drug category (*$p < .0125$)

Signs	χ^2 to Remove	*df*
Swayed while balancing on one leg	22.74*	4
Used arms to maintain balance	28.89*	4
Hopped during test to maintain balance	19.57*	4
Put raised foot down	65.59*	4

to individuals who had not used drugs. Users of CNS depressants, CNS stimulants, and narcotic analgesics were also significantly more likely to put their raised foot down during the test. In contrast, the drug users across all 4 drug categories were less likely to hop during the OLS test to maintain their balance compared to those who had not used drugs.

Prediction of Drug Category From Performance on the Walk and Turn Test

To predict drug category from performance on the WAT test, a multinomial logistic regression analysis was conducted and the findings revealed that the set of 7 signs from this test significantly distinguished the 4 drug categories from the no-drug cases, $\chi^2\,(28, N = 2142) = 273.89$, $P < .0001$. An overall classification rate of 42.8 percent was calculated based on these 7 signs. Classification was found to be highest for CNS stimulants (72.2%), followed by cannabis (39.7%), CNS depressants (9%), and narcotic analgesics (3.5%). Three signs from the WAT test that significantly predicted drug category were could not keep balance while listening to instructions, did not touch heel-to-toe while walking, and taking an incorrect number of steps during the test ($P < .0071$; Table 5).

The specific signs from the WAT test that distinguished the 4 drug categories from the no-drug cases were examined and the results are shown in Table 6. The findings revealed that users of CNS depressants, CNS stimulants, and narcotic analgesics were significantly less likely to keep their balance while listening to the test instructions compared to individuals who were not impaired by drugs. In addition, users of CNS depressants were less likely to touch heel-to-toe while walking, whereas individuals who had used narcotic analgesics were less likely to take the correct number of steps during the WAT test.

Table 5. Contribution of signs from the WAT test in predicting drug category (*$p < .0071$)

Signs	χ^2 to Remove	*df*
Could not keep balance while listening to the test instructions	75.48*	4
Started the test before the instructions were completed	6.57	4
Stopped walking during the test	6.09	4
Did not touch heel-to-toe while walking	35.37*	4
Stepped off the line	6.90	4
Used arms to maintain balance	6.81	4
Number of steps taken (correct, incorrect)	17.13*	4

Discussion

The results of the present study show that CNS depressants, CNS stimulants, narcotic analgesics, and cannabis are significantly associated with impairment on the SFST, with prediction being highest for CNS stimulants. The pattern of signs on the various tests of the SFST varied by drug category, which provides support for the validity of using the SFST to identify persons who are impaired by drugs other than alcohol.

Consistent with Bosker and colleagues (2012), the current investigation found that cannabis adversely affected performance on the OLS test but not the WAT and HGN tests. These results, however, contrast with those reported by Papafotiou et al. (2005a), who noted that cannabis was related to impairment on all 3 tests of the SFST battery. According to the DEC Program, cannabis is not one of the drugs that produces HGN. It is possible that the HGN displayed by participants in Papafotiou et al.'s (2005a) study may have occurred because they consumed drugs other than cannabis. In their report, Papafotiou and colleagues noted that the subject's blood

Table 4. Prediction of drug category vs. no drug use from signs observed during the OLS test (*$p < .0125$)

Signs	*B*	*SE*	Wald's χ^2 test	OR	95% Confidence interval for OR
CNS depressants					
Swayed while balancing on one leg	0.28	0.11	6.96*	1.32	1.07, 1.62
Used arms to maintain balance	0.35	0.10	12.04*	1.42	1.16, 1.73
Hopped during test to maintain balance	−0.46	0.11	17.80*	0.63	0.51, 0.78
Put raised foot down	0.37	0.09	16.32*	1.45	1.21, 1.74
CNS stimulants					
Swayed while balancing on one leg	0.24	0.09	6.90*	1.27	1.06, 1.51
Used arms to maintain balance	0.37	0.09	16.71*	1.44	1.21, 1.72
Hopped during test to maintain balance	−0.45	0.10	20.92*	0.64	0.53, 0.77
Put raised foot down	0.26	0.08	9.51*	1.30	1.10, 1.53
Narcotic analgesics					
Swayed while balancing on one leg	0.38	0.10	16.35*	1.47	1.22, 1.77
Used arms to maintain balance	0.37	.09	15.57*	1.45	1.20, 1.73
Hopped during test to maintain balance	−0.42	0.10	17.55*	0.65	0.54, 0.80
Put raised foot down	0.24	0.09	7.67*	1.28	1.07, 1.52
Cannabis					
Swayed while balancing on one leg	0.29	0.09	9.93*	1.33	1.11, 1.59
Used arms to maintain balance	0.26	0.09	8.26*	1.30	1.09, 1.55
Hopped during test to maintain balance	−0.37	0.10	14.23*	0.69	0.57, 0.84
Put raised foot down	0.04	0.09	0.18	1.04	0.88, 1.23

Table 6. Prediction of drug category vs. no drug use from signs observed during the WAT test (*$p < .0071$)

Signs	*B*	*SE*	Wald's χ^2 test	OR	95% Confidence interval for OR
CNS depressants					
Could not keep balance while listening to the test instructions	0.54	0.13	18.53*	1.72	1.34, 2.20
Started the test before the instructions were completed	0.13	0.23	0.31	1.14	0.72, 1.80
Stopped walking during the test	−0.02	0.10	0.05	0.98	0.80, 1.20
Did not touch heel-to-toe while walking	0.22	0.05	17.30*	1.24	1.12, 1.38
Stepped off the line	0.04	0.09	0.20	1.04	0.87, 1.25
Used arms to maintain balance	0.03	0.03	1.23	1.03	0.98, 1.09
Number of steps taken	−0.60	0.30	3.91	0.55	0.30, 1.00
CNS stimulants					
Could not keep balance while listening to the test instructions	0.49	0.11	18.94*	1.63	1.31, 2.03
Started the test before the instructions were completed	0.29	0.21	1.86	1.34	0.88, 2.03
Stopped walking during the test	0.05	0.09	0.24	1.05	0.88, 1.25
Did not touch heel-to-toe while walking	0.08	0.05	2.65	1.08	0.98, 1.19
Stepped off the line	−0.07	0.08	0.77	0.93	0.79, 1.10
Used arms to maintain balance	0.001	0.004	0.06	1.00	0.99, 1.01
Number of steps taken	−0.50	0.24	4.47	0.60	0.38, 0.96
Narcotic analgesics					
Could not keep balance while listening to the test instructions	0.64	0.12	29.95*	1.89	1.51, 2.38
Started the test before the instructions were completed	0.32	0.22	2.15	1.38	0.90, 2.13
Stopped walking during the test	−0.07	0.10	0.43	0.94	0.77, 1.14
Did not touch heel-to-toe while walking	0.09	0.05	3.08	1.10	0.99, 1.21
Stepped off the line	0.01	0.09	0.01	1.01	0.85, 1.20
Used arms to maintain balance	0.04	0.02	3.74	1.04	1.00, 1.08
Number of steps taken	−0.88	0.26	11.44*	0.42	0.25, 0.69
Cannabis					
Could not keep balance while listening to the test instructions	0.24	0.11	4.22	1.26	1.01, 1.58
Started the test before the instructions were completed	0.15	0.22	0.48	1.16	0.76, 1.78
Stopped walking during the test	0.04	0.09	0.18	1.04	0.87, 1.25
Did not touch heel-to-toe while walking	0.08	0.05	2.48	1.08	0.98, 1.19
Stepped off the line	−0.08	0.09	0.78	0.93	0.78, 1.10
Used arms to maintain balance	0.002	0.004	0.28	1.00	0.99, 1.01
Number of steps taken	−0.34	0.24	1.93	0.72	0.45, 1.15

samples were only tested for THC. Papafotiou et al. (2005a) also documented that cannabis was significantly related to impaired performance on the WAT test, a finding not evident in the current study. In reconciling these differing results, it is possible that they may be the result of differences in cannabis use history. In the study by Papafotiou et al. (2005a), the reported frequency of cannabis use of the participants varied from once a week to once every 2 to 6 months. In contrast, the present study was based on DEC evaluations conducted on suspected drug-impaired drivers who had self-administered drugs in doses that would be expected to exceed those that are ethically allowed in laboratory settings. Previous research has shown that heavy cannabis users develop tolerance to the impairing effects of THC on neurocognitive measures (Hart et al. 2001; Ramaekers et al. 2011). It is conceivable that the cannabis users in the current study developed tolerance to the impairing effects of THC as well, which may have affected their performance on the WAT test. Although cannabis users in the current work did not exhibit performance deficits on the WAT test, they did present such deficits on the OLS test. In accounting for these seemingly contradictory results, it is possible that the OLS may be too sensitive for determining drug use and that many individuals may not have very good balance even when they are not under the influence of drugs (Jackson et al. 2000). This highlights the need for normative data to evaluate the performance of individuals on the SFST battery who are not impaired by drugs.

Contrary to previous research (Downey et al. 2012; Silber et al. 2005), the present study found that CNS stimulants were significantly associated with impaired performance on the WAT and OLS. The apparent discrepancy in these results is most likely a consequence of different doses of drugs in the 2 studies. Both Silber and colleagues (2005) and Downey and colleagues (2012) administered low doses of amphetamines under controlled conditions in a laboratory setting, whereas the current investigation was based on the results of DEC evaluations on suspected drug-impaired drivers who had self-administered drugs. The amount of drugs administered in the real world by drug users typically exceeds that ethically allowed in laboratory settings. Thus, higher doses of CNS stimulants were likely responsible for the differences in results between studies.

The findings observed in the current study provide support for the use of the SFST as a screening tool for law enforcement to identify impairment in persons who have used CNS stimulants, CNS depressants, cannabis, or narcotic analgesics. It should be noted, though, that the pattern of impairment is not necessarily the same as that displayed by persons who are impaired by alcohol. Foremost among the differences is the fact that CNS stimulants, cannabis, and narcotic analgesics do not produce HGN. The types of errors made on the various components of the SFST also appeared to differ by drug category. If replicated and validated by further research using larger samples with known blood drug concentrations, these

patterns of SFST signs would prove beneficial in identifying drug impairment and the identification of particular drug categories.

Although not officially scored in the DEC protocol, DREs frequently note the presence of eyelid tremors during their evaluations. Though we did not have sufficient data to include this variable in our statistical models, we noted a significant association between the type of drug category and the presence of eye tremors, χ^2 (4, N = 1119) = 87.72, P < .0001. Interestingly, nearly 57 percent of the 419 cannabis cases with data on this variable had demonstrated the presence of eyelid tremors, suggesting that the inclusion of this indicator in the SFST and DEC program may increase the detection of cannabis impairment. Although further research is needed to examine the predictive validity of eyelid tremors, Papafotiou et al. (2005a) have previously reported that the inclusion of head movements or jerks in their study increased the number of subjects deemed to be impaired by cannabis.

It should be noted that the findings from the present investigation are limited by the availability of cases for analysis. The cases included in the current work were 2142 cases from the entire set of 3861 evaluations conducted between 1995 and 2009 that were submitted to the national coordinator of the DEC Program in Canada. Although the data used in the current study were collected over a 14-year period, we do not have any reason to believe that variability in the DREs' reporting or laboratory protocols may have affected the current findings. The DEC Program is a systematic and standardized protocol used throughout North America and there have been no major changes to this protocol over the years. Though laboratories may have improved their test detection over this time period, the high levels of drugs typically used by drivers would not be expected to be affected by such an improvement in test detection. It is important to note, however, that the results derived from the SFST are dependent on the manner in which the test is administered by the officers. Although the SFST is a systematic and standardized protocol that officers are trained to conduct, it is possible that some may not be as thorough in the administration of the test and/or the interpretation of the results.

This study is also limited by the fact that the results have not been cross-validated using a sample of DEC cases from another jurisdiction or by splitting the current sample in half due to concerns related to statistical power. We will continue our efforts to gain access to another set of cases so that we can apply our multivariate model to a set of DEC evaluations from another jurisdiction to determine the generalizability of our model. There is also a need to pursue a large-scale study to collect normative data on individuals' performance on the SFST while not under the influence of drugs. This is particularly important given that such drug-negative cases rarely appear on the record. If an individual shows no indication of impairment due to drugs and/or alcohol on the SFST, there are no grounds to proceed with further testing (either via the DEC protocol in the case of drugs or a breath test in the case of alcohol). As such, false-negative cases typically go unrecorded.

Acknowledgments

The authors acknowledge the Royal Canadian Mounted Police for providing the data for this work.

References

Beirness DJ, Beasley E, LeCavalier J. The accuracy of evaluations by drug recognition experts in Canada. *Can Soc Forensic Sci J.* 2009;42:75–79.

Bosker WM, Theunissen EL, Conen S, et al. 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.* 2012;223:439–446.

Burns M, Anderson E. *A Colorado Validation Study of the Standardized Field Sobriety Test (SFST) Battery. Final Report.* Boulder, CO: Colorado Department of Transportation; 1995.

Burns M, Moskowitz H. *Psychophysical Tests for DWI Arrest.* Washington, DC: National Highway Traffic Safety Administration; 1977. Report No. DOT HS 802 424.

Downey LA, King R, Papafotiou K, Swann P, Ogden E, Stough C. Examining the effect of dl-3,4-methylenedioxymethamphetamine (MDMA) and methamphetamine on the standardized field sobriety tests. *Forensic Sci Int.* 2012;220:e33–e36.

Hart CL, van Gorp W, Haney M, Foltin RW, Fischman MW. Effects of acute smoked marijuana on complex cognitive performance. *Neuropsychopharmacology.* 2001;25:757–765.

International Association of Chiefs of Police. *The International Standards of the Drug Evaluation and Classification Program.* Arlington, VA: The DEC Standards Revision Subcommittee of the Technical Advisory Panel of the IACP Highway Safety Committee; 1999.

Jackson PG, Tunbridge RJ, Rowe DJ. Drug recognition and field impairment testing: evaluation of trials. In: Laurell H, Schlyter F, eds. *Alcohol, Drugs and Traffic Safety. Proceedings of the 15th International conference on Alcohol Drugs and Traffic Safety in Stockholm, Sweden, May 21–26, 2000.* Utrecht, The Netherlands: ICADTS; 2000.

McKnight AJ, Langston EA, Lange JE, McKnight AS. *Development of Standardized Field Sobriety Test for Lower BAC Limits (SFST-LL).* Washington, DC: National Highway Traffic Safety Administration; 1995.

Papafotiou K, Carter JD, Stough C. An evaluation of the sensitivity of the Standardised Field Sobriety Tests (SFSTs) to detect impairment due to marijuana intoxication. *Psychopharmacology.* 2005a;180:107–114.

Papafotiou K, Carter JD, Stough C. The relationship between performance on the standardised field sobriety tests, driving performance and the level of Δ9-tetrahydrocannabinol (THC) in blood. *Forensic Sci Int.* 2005b;155:172–178.

Ramaekers JG, Theunissen EL, de Brouwer M, Toennes SW, Moeller MR, Kauert G. Tolerance and cross-tolerance to neurocognitive effects of THC and alcohol in heavy cannabis users. *Psychopharmacology.* 2011;214:391–401.

Silber BY, Papafotiou K, Croft RJ, Stough CKK. An evaluation of the sensitivity of the standardised field sobriety tests to detect the presence of amphetamine. *Psychopharmacology.* 2005;182:153–159.

Stuster J. *The Detection of DWI at BACs Below 0.10. Final Report.* Washington, DC: National Highway Traffic Safety Administration; 1997.

Stuster J, Burns M. *Validation of the Standardized Field Sobriety Test Battery at BACs Below 0.10 Percent.* Washington, DC: National Highway Traffic Safety Administration; 1998.

Tabachnick BG, Fidell LS. *Using Multivariate Statistics.* 5th ed. Boston, MA: Allyn & Bacon; 2007.

Tharp V, Burns M, Moskowitz H. *Development and Field Test of Psychophysical Tests for DWI Arrest.* Washington, DC: National Highway Traffic Safety Administration; 1981. Report No. DOT HS 805 864.

to be the One Leg Stand (OLS), Walk and Turn (WAT), and Horizontal Gaze Nystagmus (HGN). This reduced battery of 3 tests could correctly identify more than 80 percent of people who had a BAC in excess of 80 mg/dL (Burns and Moskowitz 1977).

A series of subsequent studies served to further refine the battery of 3 tests into a standardized procedure that could be used at roadside to accurately and reliably identify impaired drivers (Burns and Anderson 1995; McKnight et al. 1995; Stuster 1997; Stuster and Burns 1998; Tharp et al. 1981). These studies have served to establish the validity of this battery of tests, which has come to be known as the Standardized Field Sobriety Test (SFST). The SFST has been widely implemented across Canada, the United States, and parts of Australia. Individual components of the battery have also been incorporated into the field impairment testing procedures used in many other countries. The 3 tests that comprise the SFST are also included as part of the 12-step protocol of the Drug Evaluation and Classification (DEC) program to detect impairment due to drugs (International Association of Chiefs of Police 1999). Although the SFST is sensitive to alcohol impairment, few studies have examined its sensitivity to the impairing effects of other psychoactive substances. The validity of using the SFST as part of the DEC program has to a large extent been inferred from studies of the overall accuracy of the DEC program to identify persons impaired by drugs other than alcohol. The problem with this approach is that the DEC program employs a much wider range of tests and measurements than the 3 tests of the SFST to identify drug impairment. Nevertheless, the SFST has come to be viewed as a general test of impairment, regardless of the substance responsible for the impairment.

A group of researchers (Downey et al. 2012; Papafotiou et al. 2005a, 2005b; Silber et al. 2005) in Australia have conducted a series of studies to determine the sensitivity of the SFST in detecting impairment due to substances other than alcohol. To examine the effect of amphetamines on SFST performance, Silber et al. (2005) administered 0.42 mg/kg of D,L-dexamphetamine, D,L-methamphetamine, or D-methamphetamine (plus a placebo condition) to volunteers and examined SFST performance at 120 and 170 minutes after ingesting the drug. None of the 3 amphetamines showed any evidence of impairment on the SFST. The authors concluded that the SFST was not an efficient means of identifying drivers who had used small doses of amphetamines.

Downey and colleagues (2012) recently explored the effects of DL-3,4-methylenedioxymethamphetamine (MDMA) and D-methamphetamine on SFST performance. In this double-blind, counterbalanced, and placebo-controlled study, the authors administered 100 mg of MDMA, 0.42 mg/kg D-methamphetamine, or placebo to participants and examined their performance on the SFST 4 and 25 hours following drug ingestion. The results showed that D-methamphetamine did not impair performance on the SFST, a result that is consistent with that previously obtained by Silber and colleagues (2005). However, MDMA was found to significantly impair overall performance of the SFST in comparison to the placebo condition, with 22 percent of the participants being judged impaired on 2 or more components of the SFST 4 hours postdrug consumption.

Papafotiou and colleagues (2005a, 2005b) conducted 2 placebo-controlled studies that assessed whether performance on the SFST provides a sensitive measure of impaired driving behavior following the administration of either a low (1.74%) or high dose (2.93%) of THC. In the first study (Papafotiou 2005b), the participants performed a driving simulation task and the 3 component tests of the SFST. The results showed that driving performance was significantly impaired 80 minutes after the consumption of THC. Performance on the SFST correctly identified up to 76 percent of participants as being either impaired or not impaired, indicating that the SFST was a good predictor of driving impairment and an appropriate screening tool to assess and identify drivers whose abilities are impaired by the use of cannabis.

The second placebo-controlled study conducted by Papafotiou et al. (2005a) involved a more thorough examination of the 3 components of the SFST after administration of the high or low dose of cannabis. The researchers also recorded head movement or jerks as a potential indicator of cannabis impairment. The findings revealed a positive relationship between the dose of THC administered and the number of participants classified as impaired. The inclusion of head movement or jerks increased the number of subjects deemed to be impaired. Interestingly, lack of smooth pursuit (the first stage of HGN) was significantly related to cannabis use 55 and 105 minutes following administration of the drug but not 5 minutes after cannabis smoking. This result is inconsistent with the DEC protocol; the only drug categories known to produce HGN are depressants, inhalants, and dissociative anaesthetics. The authors noted that blood samples in their study were only tested for THC and, as such, it is possible that the lack of smooth pursuit displayed by the participants may have occurred as the result of their consumption of drugs other than cannabis. Papafotiou and colleagues (2005a) also reported that subjects' performance on the WAT test was significantly related to THC condition, with 2 signs of this test being observed at all times: no balance and using arms to balance. Three signs of the WAT test were found to be unrelated to the level of THC during all administrations of this test, including misses heel to toe, improper turn, and incorrect number of steps. The authors also suggested that the OLS test provided the best indicator of impairment associated with the administration of THC.

More recently, Bosker and colleagues (2012) assessed the effects of smoking cannabis with and without alcohol on SFST performance in a double-blind, placebo-controlled study of heavy cannabis users. The results from this investigation showed that cannabis use (dose of 400 μg/kg body weight THC) was significantly related to impairment on the OLS test, whereas impairment on the HGN test only approached statistical significance. When cannabis was combined with alcohol (BACs of 50 and 70 mg/dL), participants' performance on the HGN was significantly impaired. Performance on the WAT test was not found to be impaired by cannabis either alone or in combination with alcohol.

As the use of the SFST and the components of the SFST that are embedded in the DEC protocol become more widespread, it is important that the tests be shown to be valid indicators of impairment due to drugs other than alcohol.