

Review of Screening Instruments and Procedures for Evaluating **DWI** [Driving While Intoxicated/Impaired] Offenders



Prepared by

Iyiin Chang, MS
Cindy Gregory, PhD
Sandra C. Lapham, MD, MPH

Prepared for

AAA Foundation for Traffic Safety
1440 New York Avenue, NW, Suite 201
Washington, DC 20005
Tel: 202-638-5944
Fax: 202-638-5943
www.aaafoundation.org

November, 2002

Table of Contents

List of Tables	3	Evaluation Summary of	
List of Figures	3	Recommended Assessments	28
Executive Summary	5	<i>ASI</i>	28
Overview	5	<i>AUI</i>	29
Results Summary	6	<i>CAGE</i>	29
Recommendations Summary	7	<i>DRI</i>	30
Introduction	9	<i>LAI</i>	30
The Societal Cost of Drunk Driving	9	<i>MAC</i>	31
The Definition and Purpose of DWI		<i>MACH</i>	31
Screening	9	<i>MAST</i>	32
Purpose of this Report	10	<i>MF</i>	32
		<i>RIASI</i>	33
		<i>SALCE/NEEDS</i>	33
		<i>SASSI</i>	33
Methodology	11	Discussion and Recommendations	35
Rules for inclusion and exclusion		A Multidimensional Approach	
of instruments	11	to Screening	35
Compilations used to search for		Screening as Intervention	36
relevant instruments	11	Evaluation of Existing Studies on	
Reviewed instruments	12	Recidivism and AUD Prediction	37
Essential Principles and Issues	15	<i>Recidivism prediction</i>	37
<i>Outcome measures</i>	15	<i>AUD prediction studies</i>	38
<i>Limitations of recidivism as an</i>		Considerations for Choosing Instruments .	39
<i>outcome measure</i>	16	Acceptable Standards for Instrument	
<i>Limitations of AUD as an outcome measure.</i>	16	Performance	39
<i>The scope of screening procedures and</i>		Expanding Data to Include Broader	
<i>instruments</i>	17	Populations	40
Instrument Evaluation Characteristics	17	Further Research on Widely Used	
<i>Ensuring comparability to enable direct</i>		Assessments	41
<i>comparisons among instruments</i>	18	Addressing Client Defensiveness	41
<i>Sensitivity and specificity statistics</i>	18	Conclusion	42
<i>The scoring system for ranking instruments</i> ..	19		
Results	21	Acknowledgements	43
Inventory of Current Practices	21	Appendix A:	45
<i>Procedures used in US courts</i>	21	Terms used for Database Searches	
<i>Screening instruments used in US courts</i>	22	Appendix B:	47
<i>Screening as a Potential Intervention</i>	22	ROC Curve And Sensitivity/Specificity	
Results of the Comparative Study of		Analysis-Risk Score	
DWI Screening Instruments	23	Tables	49
<i>Validity for predicting subsequent DWI</i>	23	Figures	61
<i>Validity for predicting Alcohol Use</i>		Reference List	65
<i>Disorders (AUDs)</i>	24		
<i>Secondary evaluating criteria</i>	25		

This study was funded by the AAA Foundation for Traffic Safety in Washington, DC. Founded in 1947, the AAA Foundation is a not-for-profit, publicly supported charitable research and education organization dedicated to saving lives and reducing injuries by preventing traffic crashes. Funding for this study was provided by voluntary contributions from the American Automobile Association and its affiliated motor clubs, from individual AAA members, and from AAA-affiliated insurance companies.

This publication is distributed by the AAA Foundation for Traffic Safety at no charge, as a public service. It may not be resold or used for commercial purposes without the explicit permission of the Foundation. It may, however, be copied in whole or in part and distributed for free via any medium, provided the AAA Foundation is given appropriate credit as the source of the material. The opinions, findings, and conclusions expressed in this publication are those of the authors and are not necessarily those of the AAA Foundation for Traffic Safety or of any individual who peer-reviewed this report. The AAA Foundation for Traffic Safety assumes no liability for the use or misuse of any information, opinions, findings, or conclusions contained in this report.

If trade or manufacturers' names or products are mentioned, it is only because they are considered essential to the object of this report and their mention should not be construed as an endorsement. The AAA Foundation for Traffic Safety does not endorse products or manufacturers.

©2002, AAA Foundation for Traffic Safety

List of Tables

Table 1	Screening Instruments for DWI Offenders	49
Table 2	DWI Offender Assessment Fees, Treatment Referral Rates, and Sanctions by State	50
Table 3	DWI Offender Program Process Characteristics by State	51
Table 4	DWI Screening Instruments for DWI Offenders by State	52
Table 5	Validity of Predicting Subsequent DWI	53
Table 6	Validity of Predicting Alcohol Dependence/Abuse Problems	54
Table 7	Instrument Administration Characteristics	55
Table 8	Primary Testing Domains	56
Table 9	Profile Reporting and Recommendations for Treatment	57
Table 10	Reliability/Validity Statistics and Research Activities	58
Table 11	Evaluation Summary	59

List of Figures

Figure 1	Sensitivity and Specificity/ Predictive Value Positive and Negative	61
Figure 2	Prevalence of Screening Instruments Reported by State Court Systems	62
Figure 3	Validity of Predicting Subsequent DWI	63
Figure 4	Validity of Predicting Alcohol Use Disorder	64

Executive Summary

Overview

This report reviews the existing literature on DWI screening procedures and instruments for evaluating drunk-driving offenders and provides an inventory of procedures and instruments that state court systems report they are currently using. This review attempts to identify studies that examine whether screening procedures in and of themselves are effective interventions in deterring future drunk-driving behavior. The report's audience includes DWI researchers and practitioners as well as readers from administrative, legal, and other fields.

This report also provides an in-depth evaluation of DWI screening instruments and defines the extent of the field's knowledge about their efficacy and practical considerations for their use. The current empirical literature on self-report DWI screening questionnaires is reviewed, including psychometric research supporting the questionnaires' validity and reliability and predictive validity for correctly identifying offenders at risk for adverse outcomes. In addition, this report calls for a critical discussion about the DWI screening process. It also explains why research is urgently needed to assess the predictive validity of certain screening instruments in widespread use.

A review of the published literature reveals there is too little rigorous research on screening for drunk-driving offenders. This is particularly problematic considering that large numbers of people are arrested for this crime each year. Extensive literature searches found no published studies that specifically address the issue of the screening process as an intervention in itself.

The authors reviewed 15 articles that investigate the efficacy of the following self-report instruments for DWI screening: Alcohol Severity Index, Alcohol Use Inventory, CAGE, Drivers Risk Inventory, Mortimer Filkins, Michigan Alcoholism Screening Test, Minnesota Assessment of Chemical Health, RIA Self-Inventory Screening Instrument, Substance Abuse Subtle Screening Inventory, and the MacAndrew scale of the Minnesota Multiphasic Personality Inventory. The most commonly used assessments include the Mortimer Filkins, Michigan Alcoholism Screening Test, and Drivers Risk Inventory.

The primary criterion used was how well the instrument predicts DWI recidivism or determines alcohol-use disorders. The authors also reviewed and summarized evidence concerning four secondary criteria: 1) ease of administration and cost; 2) testing domains; 3) appropriateness for DWI screening; and 4) recommendations for treatment and profile reporting.

Results Summary

Based on available evidence, the best-rated instruments for DWI screening are the MacAndrew scale of the Minnesota Multiphasic Personality Inventory and the Alcohol Use Inventory, which were demonstrated to have the best predictive values for DWI recidivism. The MacAndrew scale was also demonstrated to determine concurrent alcohol-use disorder. However, none of these instruments evaluates drug-use disorders or other domains considered important in screening offenders. As these assessment tools have been evaluated in offender samples from only one state, further research is recommended to determine their usefulness in geographically and ethnically diverse populations. Two widely used screening instruments in the judicial system—the DRI and SALCE/NEEDS—have not been sufficiently validated; therefore there is an urgent need to evaluate them. The Mortimer Filkins and the Michigan Alcoholism Screening Test, together with the DRI, are the tests used most widely in the court system, despite the lack of published evidence that they are useful with the DWI population.

Questions remain about the accuracy of even the best-rated screening instruments. Predictive validity varies across instruments and receiver operator characteristic curves demonstrate that none of these instruments meets the stringent criteria for predictive validity that are an accepted standard in medical practice. The screening methods developed to date cannot accurately predict who will recidivate and who will not. Even the best assessments accurately detected only approximately 70% of recidivists and identified approximately 50% of offenders as problem drinkers. No evaluations have been shown to be valid for accurately determining drug-use disorders. Since drugs other than alcohol may impair a substantial proportion of drivers, it is critical that methods for determining drug-use disorders in this population be developed and evaluated.

Several studies have pointed out that drunk-driving offenders who have been mandated by the court to participate in screening programs, underreport their alcohol- or drug-related problems or offense histories. These and other factors reduce the accuracy of predictive tests. In addition, evidence is mounting that the vast majority of convicted drunk-driving offenders have serious alcohol-related problems and a substantial percentage have drug-related problems.

All of the conditions necessary for a brief intervention may be met when a DWI offender meets with the court assessor. Yet the research found no study that evaluates the effectiveness of face-to-face interviews in this population. Considerable evidence suggests that brief interventions with alcohol users help a substantial proportion to reduce their alcohol intake. More information is needed on the efficacy of different interviewing approaches for this population.

The demonstrated poor performance of screening instruments in predicting outcomes of interest raise fundamental questions about screening itself. Is an instrument with a sensitivity of 70% and a specificity of 40% acceptable? Such standards certainly would not meet the stringent criteria for a good medical test. But perhaps the more

relevant question is: Do these criteria meet the needs of DWI program administrators and interviewers to help triage offenders? These crucial questions need consideration.

Findings in this study have implications for the current operation of DWI screening programs and highlight the need for targeted research efforts. Moreover, these findings provide an impetus for suggesting a re-examination of the entire process of screening, monitoring, and treating drunk driving offenders.

Recommendations Summary

The authors of this report suggest the following recommendations:

- Determine the efficacy of screening procedures in deterring recidivism. To accomplish this, an experimental design could include a group of offenders who are not screened but who receive a single treatment option, without regard to risk of recidivism or need for rehabilitation. Similarly, different screening protocols could be paired with a single fixed treatment, so that instruments and processes could be compared.
- Include in studies of efficacy those subjects who were arrested but not convicted because of plea bargaining, court leniency, or simply not being referred for mandatory evaluation.
- Place greater emphasis on understanding how different groups of offenders, identified through typologies, fare within screening and treatment modalities. When possible, measure comparison groups in parallel that do not follow the same treatment path.
- Researchers are urged to address the problem of criterion validity. DWI researchers must determine what constitutes effectiveness in a screening program in order to guide the creation and evaluation of this and the next generation of DWI screening instruments.
- When making treatment disposition decisions, evaluators should augment data from self-report instruments and face-to-face interviews with externally valid sources of information. These should include offenders' court criminal/traffic records, arrest and pre-interview breath alcohol testing results, and possibly other biochemical tests in order to help identify those who are more likely to be highly defensive and under-report their involvement with alcohol or other drugs.
- DWI assessments should include disorders beyond AUD, and in particular drug use, using biochemical tests.
- Researchers should consider the very modest impact of treatment on mitigating drunk driving behavior. Practitioners should work collaboratively with criminal justice professionals and researchers to design and evaluate innovative programs that combine treatment strategies with sanctions and monitoring.

- DWI researchers and practitioners should initiate a dialogue regarding whether, because of their inherent limitations, self-report screening assessments provide sufficient information for appropriate triage to selective interventions. If so, the field needs to determine the appropriate balance between achieving maximal positive and negative predictive validity.

- DWI researchers and practitioners should also close the gap between research and practice. Practitioners who select screening instruments should understand that their decisions should be based on compelling scientific evidence of the tests' applicability to the DWI population. Researchers should rigorously investigate these existing instruments. Finally, journal editors should publish their findings, even when tests show low predictive validity.

- The authors recommend that DWI researchers, practitioners, and those who select screening instruments be informed regarding the normative data from which these instruments are derived. Most notably, the score interpretation will be incorrect if balanced against non-DWI respondents.

Introduction

The Societal Cost of Drunk Driving

The private and public cost of driving while impaired (DWI) in terms of lives lost, property damage, medical care dollars, lost work time, and increased insurance premiums is substantial.¹ The Department of Justice has estimated that almost 1.5 million (1,447,300) DWI arrests occurred in the U.S. in 1997.² Various strategies have been used to prevent first-time DWI offenses, including educational programs, server intervention, and warning labels.³ Once an offense has occurred, the focus changes to strategies for preventing recidivism.

Preventing drunk-driving recidivism is indeed an important goal. Repeat drunk-driving offenders pose a significant threat to public health. Many continue to drink and drive even after their driver's licenses have been revoked.⁴ These persistent drinking and driving behaviors often lead to traffic crashes and deaths.⁵ Fatally injured drivers with blood-alcohol levels (BAC) of 0.10 g/dl or greater were six times more likely to have a prior conviction for DWI compared with fatally injured sober drivers.⁶

Because repeat offenders pose a significant threat to the community and because it is essential that those offenders with alcohol abuse or dependence diagnosis (AUDs) be referred to appropriate treatment, many efforts have been directed toward identifying and intervening with arrested or convicted drunk drivers at high risk for re-offending.⁷⁻¹³ This report examines methods for determining which DWI offenders are at risk for re-offending.

The Definition and Purpose of DWI Screening

Screening is the use of easily and inexpensively administered tests and procedures in an attempt to establish the presence/absence of AUD, drug-use disorder, and recidivism risk.

Ultimately, the purpose of DWI screening is to protect society by preventing recurrence of drunk-driving *behavior*. Other benefits to society are apparent when persons with substance-use disorders are successfully treated, enabling them to lead full, productive lives. Prompt and appropriate intervention may allow offenders a unique opportunity to enter treatment without having to initiate the process themselves. Although some offenders eventually recognize the need for treatment and take advantage of the opportunity, others feel coerced and resist the process. This resistance makes it difficult for evaluators to ascertain the nature and severity of substance abuse in this population.

Another purpose of screening in some jurisdictions is to recommend appropriate sanctions that may reduce recidivism of individual offenders.

Purpose of this Report

This report reviews the literature on DWI screening procedures and instruments and provides an inventory of those currently used by state court systems. A particular focus of this review is to determine whether a screening procedure itself can act as an effective intervention in deterring future drunk-driving behavior.

This report also provides an in-depth evaluation of DWI screening instruments and defines the state of knowledge about the instruments' efficacy and practical considerations for their use. It reviews current empirical literature on self-report questionnaires used in DWI screening, including psychometric research supporting the instruments' validity and reliability and predictive validity for correctly identifying offenders at risk for adverse outcomes. For DWI screening instruments in general, and individually, the report

- categorizes and describes the instruments,
- describes and analyzes existing research, and
- defines strengths, weaknesses, and areas where more research is urgently needed.

Methodology

Bibliographic databases were systematically searched from 1965 to 2000 (see Appendix A). Text terms were first searched partnered with publication type and subsequently without any publication type qualifier. Electronic databases accessed include MEDLINE, HealthSTAR, PsychINFO, DIRLINE, Sociological Abstracts, WORLD CAT, Social Science Abstracts, ETOH, and HRSPoj. A search for reviews from the Cochrane Collaboration *Drug and Alcohol Group* found that none were published by the Drug and Alcohol Group during this time. Searches were limited to articles published in English or for which English translations of abstracts were present. Other reports were sought by citation tracking of relevant articles. In addition to traditional searching, informal email requests for other published and non-published studies were made to the *Research Society on Alcoholism* membership list (total 549 members with electronic addresses) and Internet searching of relevant government, law enforcement, traffic safety, and public interest archives.

Combinations of the search criteria were used to identify research reports on the topics under investigation for this report. We also searched for studies that included the terms *predictive validity*, *predictive value*, or *prediction*. Adding these terms did not identify any additional usable studies, primarily because the reports found did not include screening information.

Rules for inclusion and exclusion of instruments

The review is restricted to instruments evaluating alcohol use in the context of screening rather than diagnosis. Additional criteria restrict this review to the following:

- instruments applicable to the broader DWI setting and not designed for special populations, such as those for medical examination, psychological assessment, or unique populations such as college students or adolescents
- instruments with a publication record beyond the original release by the authors
- instruments currently in use by at least one state
- instruments in their original form. Hybrid instruments, which are created from subsets of items extracted from standardized instruments, are not included because they do not allow for psychometric evaluation.

Compilations used to search for relevant instruments

In addition to the literature search described above, three compilations were especially useful for identifying relevant instruments.

First, The National Institute on Alcohol Abuse and Alcoholism (NIAAA)¹⁴ periodically publishes an anthology of assessment instruments, the most recent enumerating 81 questionnaires helpful to clinicians and researchers in the area of alcohol abuse. An updated list is available via their web site (<http://www.niaaa.nih.gov/publications/instable.htm>). Eight instruments have been added since 1995. From this listing a total of 14 instruments were identified specifically for screening alcohol problems in adults.

Second, Popkin, in 1988, reviewed eight additional instruments targeted specifically to DWI populations.¹⁵ At that time, any instrument that appeared promising was included, although many were not well validated or psychometrically proven.

Third, Mayhew and colleagues¹⁶ more broadly reviewed scales and instruments used to assess factors identified as being associated with risky or problem driving. They critiqued 13 personality measures, six driver attitude scales, 16 alcohol abuse scales, and six comprehensive instruments.

In addition to instruments described in these published collections, those that appeared to be in current use were identified by court practices (see “Inventory of Current Practices” section). Several instruments listed by states as being used for screening were primarily diagnostic and were, therefore, excluded. These included the Minnesota Multiphasic Personality Inventory (MMPI), Substance Use Disorders Diagnostic Schedule (SUDDS), and criteria for alcohol and drug abuse and dependence from the Diagnostic and Statistical Manual, Version IV (DSM IV).¹⁷ In addition, customized instruments that appeared to be used only locally and without publication records were excluded.

Reviewed instruments

Twelve screening instruments that met our criteria for inclusion are enumerated in Table 1, along with highlights that may be of interest to assessment managers. A summary of the empirical data surrounding the use of these instruments in DWI populations follows.

ASI—Addiction Severity Index

The Addiction Severity Index, developed by McLellan, Luborsky, Woody and Obrien in 1980, is administered as a structured interview. It yields scores in chemical abuse, medical, family/social, employment/support, psychological, and legal problems.¹⁸ Its primary application has been in guiding treatment planning and outcome evaluations, because it provides an overview of problems related to substance abuse. It has been normed on alcohol and drug abusers in a treatment setting. Each domain is composed of subjective ratings derived by the interviewer about the severity of the problem and of composite scores based on individual item responses from the patient. The interviewer notes whether the patient appears to be misrepresenting information. Four states are currently using ASI.

AUI—Alcohol Use Inventory

The AUI, developed by Wanberg, Horn and Foster in 1977,^{19,20} presents 24 scales and 218 questions in four domains: benefits, styles, consequences, and concerns regarding alcohol use. It is especially useful for describing parameters in multidimensional models of abuse. The AUI was normed on clients diagnosed and hospitalized for alcohol abuse or dependence. Studies were conducted primarily with clients in residential care or outpatient settings.²¹ Chang and colleagues²² reported normative data of AUI on DWI offenders. One state is currently using AUI.

CAGE—Cut Down, Annoyed, Guilt, and Eye-Opener

The CAGE was developed in 1974 by Mayfield, McLeod and Hall. Convenient for its brevity, non-threatening nature, and ease of scoring, the CAGE²³ has been advocated primarily for screening of AUD in primary care. It was developed based on data from patients (39% alcoholic) admitted to a psychiatric service. The name of the instrument serves as an acronym for the content of the four items that comprise it: (1) Have you ever felt the need to cut down on your drinking? (2) Have people annoyed you by criticizing your drinking? (3) Have you ever felt bad or guilty about your drinking? and (4) Have you ever had a drink first thing in the morning (eye opener)? Five states are currently using CAGE.

DRI—Driver Risk Inventory

The DRI²⁴ was developed in 1987 by Behavior Data Systems for DWI screening and normed specifically for convicted drunk drivers. It consists of 5 independent scales: truthfulness, alcohol, drugs, driver risk, and stress coping, each with a risk assignment to one of four levels, and a percentile score. The recently released version II added a substance dependency scale built on DSM-IV criteria. Twelve states are currently using DRI.

LAI—Life Activities Inventory

The LAI, developed by Holden and Reis in 1981, was designed to obtain information on DWI offenders' life activities and personality characteristics prior to and during treatment.²⁵ It consists of nine life-situation scales: alcohol quantity frequency, alcohol problems, physical health, financial/employment, social interaction/involvement, family status/living situation, marriage, treatment receptivity, and residential stability, as well as six personality scales. They are: extroversion/introversion, sanguine, self-confident vs. anxious, depressed, moralistic, conservative vs. nontraditional, unconstrained, paranoid suspicious vs. naive trust, conforming compliance vs. acting-out aggressiveness, and improbability. It was normed on DWI offenders and was designed to assess treatment-induced changes in life circumstances over time. It includes a self-report and an interview. Texas is the only state that is currently using the LAI.

MAC—MacAndrew Alcoholism Scale

Developed by Hathaway and McKinley in 1943 and revised in 1989, the MAC scale of the Minnesota Multiphasic Personality Inventory (MMPI)^{26,27} screens for personality characteristics related to alcoholism without explicitly mentioning alcohol. It is possible, therefore, for a subject to score as high risk yet not have any history of drinking, since it does not directly query alcohol use. The items are extracted from the longer MMPI, which measures objective personality inventory for abnormal behaviors and was originally normed on members of the public who were friends or relatives of patients in the University Hospitals in Minneapolis. MMPI has been well validated in clinical populations. The MAC is a subscale and has been used for predicting DWI recidivism. Three states are currently using MAC.

MACH—Minnesota Assessment of Chemical Health

The MACH²⁸ was developed by Kincannon in 1984. It consists of interactive questioning that branches, depending on the subject's answers. MACH is fully computerized and must be interpreted by a counselor and explored with the client to arrive at a plan. It includes items from the MAST, MF, and DSM-IV criteria, and yields measures of pathological use, consequences, risk factors, stressors, and social functioning related to alcohol and drug use. It is less a quantified questionnaire than a means of objectively standardizing an interview. Two states currently use the MACH.

MAST—Michigan Alcoholism Screening Test

The MAST,²⁹ developed in 1971 by Selzer, is a popular 24-item questionnaire devised to provide a consistent, quantifiable, structured interview instrument to detect alcoholism. Shorter versions of MAST are the Brief Mast (10 items), Malmo Modification (nine items), and Short Mast (SMAST) (13 items). The MAST was originally tested on five groups including a control group, hospitalized alcoholics, convicted DWI offenders, persons convicted of drunk and disorderly behavior, and drivers whose licenses were under review. Fourteen states are currently using MAST.

MF—Mortimer-Filkins

The MF was explicitly designed for assessing DWI offenders, and contains both a self-report questionnaire and structured interview components,³⁰ although the interview is sometimes omitted. It was developed in 1971 by Mortimer, Filkins, and Lowery. Questions cover marital and family problems, recent stress, employment and finances, depression, nervousness, drinking, feelings, and ability to cope. It was developed using a sample of alcoholics (known problem drinkers) and a sample of controls (known non-problem drinkers) and field tested on DWI offenders during screening. Scores place a respondent into one of three risk categories—social drinker, presumptive problem drinker, or problem drinker. The questionnaire does not contain a correction or assessment for truthfulness. Twenty-one states are currently using MF.

RIASI—RIA Self-Inventory

Nochajski and Miller developed the RIASI in 1997. Empirically derived from three large samples of DWI offenders,³¹ it provides a problem-drinking score. Authors are currently validating a recidivism score based on follow-up and predictive ability of certain items.³² Administered by pencil and paper, it can be hand-scored and yields a single continuous score that can be interpreted at different cutoffs for different populations. One state is currently using RIASI.

SALCE/NEEDS—Substance Abuse Life Circumstances Evaluation

Also developed for DWI offenders, the SALCE,³³ developed in 1986 by ADE, Inc., assesses attitudes, life stress issues, alcohol and drug use, and driving records, and has been in use since 1986. NEEDS is an expanded version of SALCE that has been in use since 1990. It assesses attitudes, emotional stability, substance abuse, employment, relationships, health, education, and criminality. It includes a truthfulness estimation. The substance abuse scale and recommendations for both instruments are based on DSM-IV criteria. Currently nine states use the SALCE or the NEEDS instruments.

SASSI—Substance Abuse Subtle Screening Inventory

SASSI's first version was developed by Miller in 1985.³⁴ The third version, the SASSI-3,³⁴ presents 10 subscales assessing substance abuse, symptoms, legal difficulties, as well as a client's falsity on answers, concealing evidence, and undesired answering patterns. It is a screening instrument for substance-abuse disorders that is used for a variety of clinical and criminal populations. The developers described their clinical dataset as consisting of offenders and non-offenders from treatment centers, hospitals, and rehabilitation programs. Currently five states use SASSI.

Essential Principles and Issues

OUTCOME MEASURES

Researchers measure an instrument's validity against two outcomes: recidivism and the presence of AUD. To evaluate a screening procedure's validity in predicting recidivism, the investigator assigns DWI offenders into low-risk and high-risk groups according to the results of the instrument under evaluation. Efficacy in predicting recidivism is examined by searching traffic records and determining whether the high-risk offenders are more likely (and low-risk offenders less likely) to re-offend. To evaluate an instrument or procedure for its ability to establish AUD, the investigator determines if it correctly classifies offenders as having alcohol-use problems compared with a "gold standard" criterion measure.

Many investigators^{9,22,35-39} have evaluated how well screening procedures predict subsequent DWIs; others^{30,40-45} have examined how well they predict alcohol-misuse

problems. Subsequent DWI is a direct measure because the goal of DWI screening is to reduce drunk driving. Determination of AUD is an important endpoint of the screening process, as those who screen positive are most likely in need of treatment services. However, AUD is at best an indirect measure of recidivism risk, and not all recidivists have AUD. Offenders with AUD must not only be identified but also be successfully treated to lower their recidivism risk.

LIMITATIONS OF RECIDIVISM AS AN OUTCOME MEASURE

Studies using recidivism and AUD as outcomes, however, have significant limitations. Recidivism, measured by arrests or convictions for drunk driving, is not a good measure of actual drunk-driving behavior because only a small proportion of drunk drivers are stopped by law enforcement officers. Other problems include changes in law enforcement over time, timing and frequency of DWI countermeasures, and regional and temporal changes in the sociocultural climate surrounding the acceptability of DWI. Also, offenders are often required by the court or screening counselors to go through educational or therapeutic intervention in the follow-up period, which may affect their drinking-driving behaviors.

Because recidivism is customarily assessed from court records, the quality of record keeping is outside researchers' control. For example, recidivism data may be inaccurate or incomplete, records may be purged periodically, and access may be limited because of confidentiality issues. Moreover, records do not routinely capture out-of-state offenses. In a recent report, Lapham and colleagues⁴⁷ eliminated this potential source of bias by including only potential recidivists whose official residence was within the state. Recidivism based on self report rather than on court records has been shown to be similarly error prone, with subjects underreporting their past offenses.⁴⁶

To allow cross-study comparisons researchers must also agree on a standardized definition of recidivism. Little and Robinson⁴⁸ and Neff³⁷ counted a rearrest for any reason, including both traffic and non-traffic violations, in their recidivism studies of incarcerated offenders. Researchers often classify an offender as a recidivist only if he or she is arrested, and some only if the offender is convicted for drunk driving.

LIMITATIONS OF AUD AS AN OUTCOME MEASURE

The ability of evaluating screening instruments to determine concurrent AUD also has serious methodological problems. First, many studies are based on small samples. The second problem is test-taking defensiveness—offenders often under-report their problems in order to avoid further treatment or sanctions. This underreporting should be taken into account in the evaluation of an instrument. Chang and Lapham⁴⁶ found that 50% of clients underreport their criminal history and traffic violations during screening.

Third, the outcome measure may not be entirely valid or objective. There are no

agreed-upon standards for determining the severity of alcohol- and drug-related problems in this setting. Studies that rely on diagnosis of AUD by a self-report assessment lack an external criterion measure, as they depend solely on violator responses. Furthermore, different studies often use different instruments for determining what constitutes an AUD. Given all these limitations, using subsequent DWI (determined by traffic-record searches) as an outcome, while limited, probably represents a more accurate and objective measure than AUD.

The final methodological issue concerns the setting in which testing for AUD takes place. Because DWI offenders may report differently in different settings, instruments should be tested under the same conditions and in the same settings as those offenders experienced in the DWI screening process. In this way, an instrument could take defensiveness into account, thus making predictions more accurate. Because several of the comparative studies administered the instruments in a non-DWI setting (e.g., during treatment or in an education program), their results could potentially be less valid than when the setting is a DWI screening program.

THE SCOPE OF SCREENING PROCEDURES AND INSTRUMENTS

Another issue involves the scope of the domains covered in screening; many screening procedures look only at alcohol use, ignoring or discounting other factors such as drug use and personality. Current guidelines for sentencing drunk-driving offenders recommend that all DWI offenders should be screened for alcohol and drug-use problems and for recidivism risk.⁴⁹ Research has shown that in order to predict recidivists and nonrecidivists, dimensions beyond alcohol use (e.g., other drug use, demographics, personality, emotional and motivational factors, driving style, lifestyle, and cognitive factors) must be considered.⁵⁰ Even so, many instruments do not consider these other dimensions. State-of-the-art alcohol research recognizes that drunk driving is multidimensional in nature and screening instruments are available that take these additional dimensions into account.

Although most DWI offenders have AUD, a substantial proportion has other drug-use disorders, either in isolation or combined with AUD.⁵¹ Although it is recommended that all offenders should be screened for other drug-use disorders, no articles in the published literature have evaluated the efficacy of screening instruments for determining drug-use disorders in DWI offender populations. Therefore this measure will not be discussed further in this report.

Instrument Evaluation Characteristics

The comprehensive literature search identified a number of investigations examining the efficacy of screening instruments in predicting recidivism and AUD. A study was then performed to compare these instruments. The main focus of this analysis is predictive validity, i.e., how well the instruments categorized offenders as high or low risk. While developers of these instruments may claim that their instruments can iden-

tify DWI offenders likely to re-offend, objective evaluation is needed to confirm this. For this report, data from each investigation were converted into congruent statistics, enabling direct comparison among the instruments.

The instruments were critically reviewed using two criteria. First, and principally, researchers examined the instruments' validity in terms of predicting DWI recidivism and AUD. Second, researchers examined the instrument development process and the following four administration characteristics:

- **Administration:** Is the instrument easy to use, does it have a short processing time, and is it economical?
- **Testing:** Are the testing domains of the instruments relevant to DWI screening?
- **Reliability and Validity:** Does the instrument have acceptable reliability and validity and an adequate record of subsequent DWI-related research?
- **Relevance:** How helpful and relevant are the profile summary and treatment recommendations?

We conclude with a final recommended list of instruments, which we believe should help administrators, interviewers, and researchers become better informed of the state of the art in DWI screening instruments.

ENSURING COMPARABILITY TO ENABLE DIRECT COMPARISONS AMONG INSTRUMENTS

Researchers in the original studies typically present their findings in terms of the clients' placement in high- and low-risk groups and as percentages that were or were not re-arrested for DWI or were or were not positive for AUD. In other words, 2x2 tables were usually enumerated.

If the original researchers did not generate such a table it was generated at this time, using data presented or raw data provided by the investigators. In some cases, the original researchers did not report a cutoff point between high and low risk. Instead they used statistical modeling for their investigations; their best predictive model was reported as a continuous mathematical equation. In these cases, we obtained the raw data and created a cutoff point by dichotomizing at many points on the equation and choosing a cutoff where the sensitivity and specificity statistics were reasonably optimized.

SENSITIVITY AND SPECIFICITY STATISTICS

Two sets of statistics were calculated: sensitivity and specificity and positive and negative predictive values (Figure 1).

Sensitivity measures how well the instrument predicts who will recidivate (or be diagnosed with AUD), that is, how often the instrument placed those who would eventually recidivate into the high-risk group. *Specificity* measures how well the instrument predicts who will *not* recidivate (or be diagnosed with AUD), that is, how often the instrument placed those who would not recidivate into the low-risk group. Either sensitivity or specificity alone does not measure how well the instrument predicts. It is obligatory that both sensitivity and specificity are high to ensure a good instrument.

Predictive value positive (PP+) measures the percentage of those identified correctly as high risk and who actually do recidivate or test positive for AUD. The *False Positive* equals one minus PP+ (i.e., those who were falsely identified as high risk but did not recidivate or test positive for AUD). *Predictive value negative (PP-)* is the percentage of offenders who are identified correctly as low risk and actually do not recidivate or test positive for AUD. The *False Negative* equals one minus PP- (those who were falsely identified as low risk but in fact did recidivate or test positive for AUD). A good screening instrument should have low false positives and low false negatives.

THE SCORING SYSTEM FOR RANKING INSTRUMENTS

A scoring system was defined as follows:

- Dark dots are given to instruments with acceptable predictive validity for subsequent DWIs
- Gray dots are given to instruments with acceptable predictive validity for AUD
- Small dark dots are given to instruments evaluated favorably for any of the four secondary evaluating categories.

To be recommended, an instrument had to receive either a dark or gray dot (i.e., have predictive validity for subsequent DWI or for AUD). Secondary criteria are discussed in the individual recommendations in the next section.

Recommended list for instruments predicting subsequent DWIs—dark dots

Only one of the instruments achieved sensitivity as high as 70%. For an instrument to be placed on the recommended list for evaluating subsequent DWIs, it had to have a sensitivity of at least 65% and a specificity of at least 40%. These standards are low simply because most of the instruments had low sensitivity and specificity.

Recommended list for instruments predicting AUD—gray dots

For evaluating AUD, the same minimum requirements were used: sensitivity of at least 65% and specificity above 40%. Two gray dots were given to instruments with the sensitivity above 75% and specificity above 65%.

Recommended list for administrative criteria—small dark dots

As for the secondary evaluating criteria, a small dark dot is given to instruments which (1) are easy to use, are less costly, and take less processing time, (2) have at least four of the relevant testing domains identified from the literature, (3) have reasonable reliability and validity statistics and research activities, and (4) have at least three of the five features for profile summary and treatment recommendations.

Results

Inventory of Current Practices

PROCEDURES USED IN US COURTS

This section summarizes the literature that evaluates DWI screening procedures. Screening usually entails three activities: 1) testing, 2) interviewing/assessment, and 3) referral/monitoring. The multiple steps of this process may involve a variety of agencies and personnel from private and public institutions.

Testing refers to the use of self-report instruments to evaluate the offender's use of alcohol and drugs. Interviewing uses specially trained personnel who meet with offenders to further elucidate the circumstances of the arrest and to learn about family, medical, personal, or legal problems that may indicate a need for treatment. (Most screening instruments include just a self-report, though some incorporate interview information into the instrument's outcomes.) The screener assesses the available information from the self-report instrument or interview and decides whether the offender would benefit from treatment or further evaluation. The screener then matches the client's needs, based on the severity of identified alcohol and drug-related problems, with corresponding treatment options available in the community and arranges for follow-up of their compliance with court-initiated treatment.

To provide a contextual basis for the review, the authors cataloged current court processes for DWI screening (Table 2). Information on screening was primarily obtained from the National DWI Compendium survey, conducted by the National Organization of State Impaired-Driving Programs. The effort was coordinated and compiled by the Bureau of Driver Education and DWI programs, Florida Department of Highway Safety and Motor Vehicles. A report from the National Commission Against Drunk Driving⁵² was also used.

In the chronology of adjudication, one of the first opportunities for reducing recidivism occurs during the screening process, which in large part determines the rehabilitative and correctional response. Most jurisdictions in the US use alcohol-screening programs to evaluate the substance-abuse problems of DWI offenders and to determine offenders' needs for further assessment and treatment^{15,50} in order to prevent recidivism by identifying and treating those at high risk. Most states mandate screening for drunk-driving offenders and 90% report procedures that include both in-person interviews and self-report assessments. In addition, most states have specific criteria for determining which offenders will receive treatment referral or be monitored. Thirty-two states screen both pre- and post-trial; 15 screen at post-trial only. Many jurisdictions employ personal interviews in addition to self-report instruments.

Table 3 reveals that most states in the U.S. currently have laws in place for mandatory DWI screening. All states reported using results from self-report assessments in making recommendations to the court, reporting to probation and safety officers, and determining need for further treatment or education. Nearly all states (45 of 50, or 90%) use some form of personal (face-to-face) interview with screening assessments, a figure that has risen over the last ten years. One state, Mississippi, reported using the interview only for repeat offenders; Georgia and South Dakota were the only states that explicitly reported not using interviews as part of the screening process; and Nevada and Utah did not respond. Most states have set criteria for treatment referral and some form of monitoring function to ensure offenders complete all requirements in the process.

Thirty-one states also use pre-trial screening in addition to post-trial screening (Table 3). Sixteen states use only post-trial. Clients pay 100% of the fees and support most of the DWI programs. Four states report that their programs are supported by the state (Alaska, Arkansas, Louisiana) and other funding sources (Hawaii). Most states do not report the referral rates for treatment. Of those that do report, the rates range widely, from 20% to 100%. All states report driver license suspension/revocation for various lengths. Fifteen states report using one or more sanctions, including vehicle impoundment, monitoring devices, or interlock devices.

Regarding the referral/monitoring or case management phase, not all states require screening programs to monitor client compliance with their recommendations. Some state program personnel reported that monitoring is considered part of treatment. Others reported that it was considered part of probation monitoring, where treatment providers or probation officers are required to report non-compliance. No studies were found that evaluate this phase of the screening process.

SCREENING INSTRUMENTS USED IN US COURTS

Table 4 summarizes data on standardized instruments used by courts. It shows that many state jurisdictions allow assessors to choose from among several instruments. Of the standardized screening instruments, the Mortimer Filkins (MF),³⁶ Michigan Alcoholism Screening Test (MAST),⁵³ and Drivers Risk Inventory (DRI)²⁴ are most commonly used for DWI screening (41%, 27%, and 24% respectively, Figure 2). These percentages were calculated by dividing the number of states currently using the specific instruments by the total number (n=50) of states that responded to the survey. States that limit their assessments to a single instrument are more likely to use the MF or DRI than the MAST, which is often partnered with an additional instrument since it only measures alcohol dependence. These figures contrast with those reported by an earlier survey¹⁵ showing 70% of the states using the MF and 62% the MAST.

SCREENING AS A POTENTIAL INTERVENTION

Most states use face-to-face interviews with offenders, which raises the possibility

that screening itself might be an effective brief intervention. Extensive literature searches found no published studies that specifically address this issue. Unfortunately researchers and practitioners have not yet critically evaluated screening itself; there has been no research to evaluate the efficacy, for instance, of combining self-report instruments with personal interviews. Moreover, little information is available on the percentage of offenders who are triaged to receive treatment services.

Results of the Comparative Study of DWI Screening Instruments

Investigators representing 15 studies evaluated 12 screening instruments. Six out of the 12 instruments were examined for predicting subsequent DWIs. These results are summarized below in “Validity for Predicting Subsequent DWI.” Seven instruments were examined for predicting AUD. These results are reported in “Validity for Predicting Alcohol Use Disorders.” Four are evaluated for both, and three could not be evaluated.

VALIDITY FOR PREDICTING SUBSEQUENT DWI

A total of eight studies examining six instruments were found in the literature. These generally classified DWI offenders into two groups according to their cutoff scores: low risk and high risk. Most investigators³⁵⁻³⁸ reported percentages of recidivists and non-recidivists in the low- and high-risk groups. Lapham⁹ used a log-rank test, which takes into consideration time since the first arrest. She also reported percentages rearrested according to risk group. Lapham,⁹ Chang,⁵⁴ and C’de Baca³⁹ used statistical modeling approaches and evaluated the AUI, MAC, and MAST in conjunction with offender demographics and other variables in an attempt to better predict outcomes. Wells-Parker^{55,56} used a statistical modeling approach to evaluate the LAI. The study design by Wells-Parker and colleagues was different in that the offenders were randomized into the LAI intervention and non-LAI groups. Little & Robinson⁴⁸ evaluated the MAC, and reported a positive correlation between the risk groups and recidivism, but they defined recidivism as all subsequent offenses not restricted to DWI. For this reason, the authors excluded this study.

Sensitivity (Table 5) ranged from 14% to 71%. Specificity ranged from 46% to 89%; most were above 40%. AUI and MAC were the only two instruments that met both criteria, sensitivity greater than 65% and specificity greater than 40%. Results are displayed graphically in Figure 3. An instrument with predictive validity would be placed in the upper right quadrant of the graph. However, many of the instruments fall in the lower right quadrant, indicating they had better specificity than sensitivity, that is, they were better at predicting who would *not* recidivate than who *would* recidivate, an undesirable characteristic. Only AUI and MAC hovered on the border of the upper right quadrant. MAC has a lower specificity, indicating that a higher percentage of non-recidivists would be referred for treatment.

The effects of these low prediction rates can be elucidated with an example. The

best instruments had a sensitivity of 70% and a specificity of 50%. In real-world situations, the estimated recidivism rate after five years is about 30%. Therefore, in a population of 1,000 offenders, about 300 will re-offend. A test with 70% sensitivity will accurately predict 210 of these 300 recidivists and miss 90. In this same population, a test with 50% specificity will accurately predict 350 and miss 350 of the 700 non-recidivists. In this real-world sample of 1,000 offenders, the recidivism rate among those with a positive test is only 38% (8% above the rate for the entire population).

For more information, Appendix B presents a graphical example²² that includes cutoffs, sensitivity, and specificity statistics. A ROC (receiver operator characteristic) curve is also included that shows the full range of cutoffs. In all but one study of the MF, PP+ statistics ranged 24% to 37%. This means only 24-37% of persons tested positive were truly recidivists—the other 63-76% were false positives. With one exception (also a study of MF), PP- ranged 62% to 86%. To illustrate, out of every three to six people who tested negative and were truly non-recidivists, one to two additional offenders tested negative but were re-arrested.

VALIDITY FOR PREDICTING ALCOHOL USE DISORDERS (AUDS)

A total of seven studies examining seven instruments were found in the literature. They classify DWI offenders into problem drinkers and non-problem drinkers by using two instruments, the screening instrument and the criterion instrument.

The studies evaluate the screening instrument according to how well its results corresponded to the results of a criterion instrument. This methodology assumes that the criterion instrument represents a “gold standard” by which a screening instrument can be judged (issues surrounding this methodology are discussed in “Discussion and Recommendations”).

Criterion instruments vary widely and include the DSM-III or other diagnostic criteria, an expert survey, counselor’s diagnosis, and other local criteria using blood alcohol concentration (BAC), prior arrests, and other test results. Investigators^{40,41,44,30,42,38,43} reported percentages of concordance and discordance between the results of the two instruments to indicate the validity of the instrument being evaluated. One study, using a small sample size of 54 offenders,⁵⁷ presented phi coefficients instead to assess the concordance rates among four instruments: CAGE, MAC, MAST, and SASSI-2. All of the phi coefficients were between .43 and .59 and were statistically significant. We were unable to contact the authors of this study. Therefore, computing sensitivity and specificity was not possible. This study was not included.

Sensitivity statistics reported in Table 6 ranged from 59% to 92% with most above 70%. Specificity statistics ranged from 26% to 85% with most falling above 60%. Results of all studies are presented graphically in Figure 4. An instrument with predictive validity would have both high sensitivity and high specificity, i.e., falling in the upper right quadrant of the graph. Almost all of the instruments fell in the right upper

quadrant, indicating they agreed with the criterion measures.

PP+ statistics ranged 58% to 94%. This means 58-94% of persons who tested positive concurred with the results of the criterion instrument, indicating that false positive rates were relatively low. In general, the false positive rates were lower than those calculated in the recidivism study (6%-42%). Except for one study by Lucker and Gold,⁴¹ PP- ranged from 67% to 85%. This means 67-85% of those who tested negative were truly non-problem drinkers by the criterion measure.

SECONDARY EVALUATING CRITERIA

As for the secondary evaluating criteria, a small dark dot was given to instruments that met any of the four criteria: (1) are easy to use, are economical, and require relatively brief processing time, (2) have at least four of the five relevant testing domains identified from the literature, (3) have reasonable reliability and validity statistics as shown in subsequent independent research, and (4) have at least three of the five features for profile summary and treatment recommendations.

Administration

Dots were given to instruments that are easy to use and economical, and that require minimum processing time and resources. CAGE, MAC, MF, MAST, and RIASI met these criteria (Table 7). They have no minimum charge, require less time for testing and scoring, and have no minimum training requirement.

Table 7 is supplied to give DWI program administrators/interviewers a quick review of the administrative ease of these instruments. Because this is a secondary consideration, it should not by itself be used to conclude that certain instruments are necessarily better.

Testing domains

A number of investigators have discussed content domains that are critical to screening for DWI recidivism or AUD. Nochajski and Miller³¹ conducted an extensive literature search and collapsed relevant domains into five areas: psychological factors (alcohol expectancies, aggression, impulsivity, and psychological symptoms), personal factors (physical health and interpersonal competence), classic symptoms (heavy consumption, tolerance/withdrawal, physical problems, and social problems), predisposing factors (familial risk factors and childhood risk factors), and physical markers (biochemical markers and neurological markers). Similarly, Beirness and colleagues⁵⁰ reported four critical domains for all problem drivers including personality factors (aggression, sensation seeking, impulsivity, and depression), behavioral factors (high risk problem behaviors, alcohol consumption patterns, extent of problems due to drinking and stressful life events), attitudinal factors (driving-related attitudes), and sociodemographic factors (age, sex, marital status, etc.). Mortimer and colleagues³⁰ reported three domains,

including personal history, personality, and drinking patterns. An instrument's ability to detect the test taker's dishonesty and defensiveness is also critical. Almost all authors of the instruments discussed this issue in the development process.

Based on the results of these investigations, this review established five critical testing domains. They include substance abuse, defensiveness, driving attitudes and risk, coping skills, and personality and psychological factors. Table 8 lays out the scales/problem areas used in the development of the various screening instruments. Four instruments (ASI, DRI, RIASI, and SALCE) covered at least four of five critical testing domains and received a dot for doing so.

Profile reporting and treatment recommendations

Developers of recent screening instruments concur that any instrument used for DWI offenders should include normative statistics based on DWI offenders, an interpretive profile, a recidivism predictive estimate, treatment recommendations, and critical responses requiring special attention.

Normative statistics are computed from a reference sample of DWI offenders to establish the distributions of individual scores. *An interpretive profile* indicates the severity of the individual's substance abuse as compared with the reference population. A *recidivism predictive estimate* reports the probability that the individual will recidivate; it mathematically derives this probability by comparing the individual's scores with those of a similar population of subjects. *Treatment recommendations* are based on severity measures regarding what kinds of treatment would be appropriate for the individual. *Critical items* report miscellaneous urgent or otherwise important issues, such as behavioral or mental disruption (such as danger for suicide attempts or violent behavior). Developing these features requires research, which indicates an ongoing commitment by the developers.

Instruments with summary reports showing at least three of the five items include DRI, MACH, and SALCE. These instruments are indicated with a triangle in Table 9. An *x* indicates instruments that provide demographic information.

Of those with an interpretive profile, DRI and SALCE/NEEDS were developed based on a DWI population; AUI, SASSI, and other instruments were not. If the underlying profiles were computed based on alcoholics who seek treatment or are in an educational setting (see normative samples in Table 9), the profiles may be different from those of DWI offenders.²² Therefore, the interpretive profile may not be as informative or as reliable.

Reliability/Validity statistics and research activities

This section focuses on the original author/creator's reliability and validity statistics.

Reliability statistics pertinent to the development of a screening instrument include:

- Test-retest reliability: the stability of results when the same subject completes the instrument at two different times.
- Inter-rater reliability: the stability of the test results when different interviewers administer it.
- Cronbach alpha coefficients: Also called internal consistency reliabilities, these refer to how well the items/questions correlate with each other within the same concept. The Cronbach alphas are expected to be between .70 and .90 because the items are assumed to pertain to a single concept even though the concepts are expressed differently.²¹

Not all authors reported all three reliability statistics. Therefore, focus in this section is on Cronbach alphas, presented in column 3, Table 10.

Overall, the Cronbach alphas were within acceptable limits. Authors of the ASI did not report Cronbach alphas, probably because some questions in the ASI are descriptive rather than multiple choice, which makes such computation impossible. Some AUI primary scales have lower reliability, though the secondary and third level scales all have reliability statistics between .74 and .93. However, DWI administrators should note that reliability statistics are lower for DWI offenders.²² The MAC has a relatively low reliability, which has been reported by a number of investigators.^{9,58,57} Butcher et al.⁵⁹ indicate that the MAC is not uni-dimensional, implying it may contain more than one concept. Authors of the MF reported split-half reliability. Authors of SALCE/NEEDS have not reported their reliability statistics; thus, the developmental stage of the instrument is unclear. Some of the SASSI reliability statistics were below .70.

Validity statistics pertinent to the development of a screening instrument include construct validity, discriminate validity, and concurrent validity. Construct validity refers to the degree of association between a new test and known measures of basic theoretical concepts. Discriminate validity is the extent to which a test discriminates between known groups of those who exhibit the risk factor and those who do not. Concurrent validity refers to the accuracy of a test as judged by comparison with some external source of information. Validity studies by original assessment authors are included in column 4, Table 10. All authors demonstrate reasonable validity statistics for their instruments but different authors present different comparisons. Some provide sensitivity and/ specificity statistics, some show the percent agreement with another criterion measure, and most show correlation coefficients that are statistically significant.

Some instruments have been well researched for DWI screening since their development. Others, however, have not. Independent quantitative studies are reported in Table 10. To receive a dot for this criterion, a study must present reasonable reliability/ validity statistics and must have been evaluated by independent investigators. AUI, CAGE, MAC, MAST, and MF met this requirement.

Evaluation Summary of Recommended Assessments

The MacAndrew scale and the AUI showed positive validity in predicting DWI recidivism (Table 11). The MacAndrew scale, CAGE, DRI, MAST, MF, RIASI, and SALCE/NEEDS showed positive validity in predicting AUD (Table 6), which was considered as an indirect measure of DWI recidivism risk.

MAC is the only instrument that showed positive validity in predicting both outcomes. Overall, the MAC and the AUI were rated favorably because of their positive validity for predicting recidivism. DRI, MAST, MF, and RIASI met the criteria using the indirect measure to be in the medium recommended category.

Administratively, CAGE, MAC, MAST, MF, and RIASI are more economical, shorter, and easier to use. ASI, DRI, RIASI and SALCE/NEEDS have the most relevant testing domains for DWI screening. Reliability and validity statistics reported by the original authors for all instruments are acceptable. Research on DRI and SALCE/NEEDS is especially lacking. Of particular concern is that none of the available instruments has been evaluated with respect to scales for detection of non-alcohol drug-use problems and whether offenders with both alcohol and drug-use problems have higher recidivism rates than those with alcohol problems alone. DRI, MACH, and SALCE/NEEDS have the capability of printing a summary report.

The results of the literature review and comparative analysis for each of the twelve screening instruments is provided below. These evaluations are based on our best analysis of the available literature. The reader is urged to consult Table 11 for a summary of how each instrument fared in regard to the scoring system described under “Instrument Evaluation Characteristics.”

ASI

Overview: Since 1977, ASI has been used for a variety of clinical populations, including opioid users, crack and cocaine abusers, those with mental illness, the homeless, pathological gamblers, and those in rehabilitation, detoxification, penitentiaries, and various drug treatment programs.⁶⁰⁻⁶⁷ Numerous publications were found reporting the reliability and validity statistics of the instrument for these samples.^{62,68-71}

Current Use: West Virginia, Delaware, North Dakota, Oklahoma.

Limitations and Further Research: Administered as part of a structured interview, ASI can only be used during the face-to-face interview stage. It cannot be used when group testing or fast screening is necessary. The ASI has not been tested or researched for use with a DWI sample; therefore, the normative statistics reported by these investigators may not be appropriate in DWI settings. Interviewers should be aware that the composite scores derived with ASI may not be as informative or as reliable as those derived with instruments developed specifically for use in DWI-offender populations.

Furthermore, because the instrument was designed as an assessment tool and not a screening tool, it may be too lengthy to be practical in some DWI offender programs. Validity of the instrument in DWI screening and DWI normative statistics should be established before the ASI is used in DWI-offender populations.

AUI

Overview: Developed in 1977, the AUI was designed as an assessment tool for treatment planning. It has 218 questions and requires training to interpret, though it can be used as a self-report. The AUI has been well researched in various clinical settings.⁷²⁻⁷⁵ Further validity statistics have also been reported.^{76,77} It is one of the only two instruments in this report to show predictive validity for screening DWI. It detects 71% of recidivists and 49% of potential problem drinkers; therefore, it is in the top recommended category. Its predictive validity is enhanced when the test is combined with demographic and arrest BAC information.⁵⁴

Current Use: West Virginia.

Limitations and Further Research: Designed as an assessment tool rather than a screening tool, AUI may be too lengthy for practical use in some DWI offender programs. It is commercially available and is the most expensive of all the instruments evaluated. Unfortunately, there is also a gap between current research results and what is available through the distributors of AUI. Specifically, its underlying profiles were computed based on hospitalized alcoholics. New findings⁵⁴ that include the DWI normative statistics and prediction equation based on DWI offenders are not included in the information/user packets provided with the purchase of the AUI. Because the AUI was normed on a hospitalized population rather than a DWI population (who may score lower than inpatient alcoholics due to defensiveness), lower cutoffs for determining alcohol abuse and dependent drinking patterns should be used when using the AUI with DWI populations. Interviewers may also consider requesting that the developer incorporate DWI normative statistics into the AUI. Another limitation is that the AUI does not evaluate other drug use. However, since it has demonstrated predictive validity for DWI recidivism, its use is recommended.

CAGE

Overview: Developed in 1974, many studies report on use of the CAGE⁷⁸⁻⁸² with four authors reporting on its use for DWI screening.^{40,43,57,83}

Current Use: New Hampshire, Vermont, West Virginia.

Limitations and Further Research: The CAGE has limited usefulness for DWI screening. Its scope is limited to AUD. Research has shown that the CAGE has problems in the areas of reliability and validity. Mischke and colleagues reported that CAGE findings suggest little support for its use as a single index of problem drinking. Hays

and coworkers investigated the response time between CAGE and three other tests, finding that although CAGE has the shortest response time (31-32 seconds), its estimated reliability was the lowest and its standard error the highest. Myerholtz and coworkers reported only moderate correlations for problem drinker status between CAGE and three other tests (MAC, MAST, and SASSI). Lacey et al.⁴³ compared CAGE+6 and CAGE+C with four other tests (MAST, DRI, SALCE, and MF) and found them inferior to the other instruments (Table 6). (CAGE+6 includes the four CAGE questions plus the six questions relating to daily consumption of alcohol and CAGE+C includes the four CAGE questions plus five of the six questions relating to daily consumption of alcohol.) Overall, CAGE has been evaluated less favorably than other instruments.

DRI

Overview: Designed specifically for DWI screening, the DRI is computerized and commercialized. Many of the relevant DWI testing domains have been included and it generates profile summaries with treatment recommendations.

Current Use: Alabama, Arizona, Delaware, Florida, Hawaii, Illinois, Kentucky, Missouri, New Hampshire, North Carolina, Oklahoma, West Virginia.

Limitations and Further Research: Very little research has validated the DRI. While Popkin et al.¹⁵ opined just after its development in 1987 that the DRI was “the most psychometrically well-constructed instrument available,” she also warned that “there remains the need for careful follow-up validation research to determine whether it truly achieves what its developers purport.” Unfortunately, these validation studies have not been done. During the last 13 years, only two independent/quantitative studies have been found that compare DRI results to prior arrest status⁴² and a criterion measure.⁴³ Leshowitz et al.⁴² showed positive discriminate validity of the instrument, distinguishing between first-time and multiple offenders (Table 6). Lacey et al.⁴³ found good validity statistics between the test and the expert survey. However, the methodological issues of the study were also discussed in the same section. The developmental stage of the instrument is not clear. A report by Marsteller et al.⁵⁸ summarized reasonable reliability and validity statistics. Although the DRI is popular because it meets the need of the DWI screening administrators for reasonable processing time, there has not been enough research to demonstrate its validity. Further research is urgently needed. It is placed in the medium recommended category.

LAI

Overview: The LAI has self-report and interview portions. Designed to assess how treatment of DWI offenders has altered their life circumstances over time, the LAI has been investigated as an intervention itself on DWI recidivism after two years of recidivism tracking.⁸⁴ After a long-term six to nine years of tracking,^{56,85,86} investigators found that at both short- and long- term follow-up the LAI has a marginal effect, about a 6% reduction in recidivism for low-risk drinkers, but had no effect on high-risk drinkers.

The LAI individual scale scores on a composite measure of recidivism showed some predictive validity; Arstein-Kerslake et al.⁸⁶ found only two out of the 15 scores to be significant predictors of recidivism and concluded that these psychometric measures do not play a major role in the prediction of recidivism.

Current Use: Tennessee.

Limitations and Further Research: Wells-Parker and colleagues⁸⁵ reported that repeated administration of the LAI was detrimental for screened high problem-risk women. The LAI has a relatively lengthy administration time (115 questions, 60 minutes to administer, 20 minutes to score). With its lengthy administration time and poor predictive validity, the LAI is not recommended.

MAC

Overview: MAC is the only instrument that has been evaluated positively for DWI recidivism and AUD. It detects about 67% of the recidivists and identifies an additional 48% as problem drinkers (Table 5). However, research has been carried out on only a single DWI-offender population. (For a discussion on the acceptability of these percentages, refer to “Results and Discussions—Validity for Predicting Subsequent DWI.”)

Current Use: Arizona, North Carolina, North Dakota.

Limitations and Further Research: Designed as a subscale of MMPI to detect alcoholics among members of the general population, its psychometric applicability to DWI screening is not clear. It assesses only one domain, alcohol use, and its usefulness has not been confirmed in other DWI-offender populations. Normative data for DWI offenders should be compiled and included in the MMPI packet. Interviewers should request that the developers provide DWI normative data so that standardized scores for DWI-offender populations can be computed. Three states use the MAC, although predictive validities have been demonstrated for only a single DWI-offender population. We encourage researchers to validate this instrument in other DWI offender populations. Since it has been demonstrated to have predictive validity, it is in the most recommended category.

MACH

Overview: The MACH instrument is different from other instruments in that it is not really an independent instrument but a computerized program that integrates the MAST, the MF, and the DSM-IV diagnosis. It does, however, have its own measure of drug abuse, MDIS. It uses all but two of the MAST questions, 38 of 46 of the MF interview portion, and the DSM-IV criteria for alcohol abuse and dependence. It is based on a social model⁸⁷ and describes the context and pattern of alcohol and other drug use in three ways: MACH referral grid, MACH assessment matrix, and MACH summary of an individualized action plan.

Current Use: Oklahoma, North Carolina.

Limitations and Further Research: There have been a few reports and presentations by the developer of the MACH, but we have found no independent, quantitative studies of its validity. Its validity and reliability statistics are probably similar to those of the MAST and MF. Independent validity studies on MACH are needed for DWI-offender populations.

MAST

Overview: One of the earliest instruments developed for DWI screening, the MAST has been used extensively in various clinical applications.⁸⁸ The MAST is a unidimensional instrument that evaluates AUD. Hedlund and coworkers did a comprehensive review of the MAST, covering over fifty studies on its administration, scoring, alternate forms, reliability, validity, face structure, and utility issues. Additional studies were also conducted by Zung and colleagues⁸⁹⁻⁹¹ and Tulevski.⁹²

Current Use: Arizona, California, District of Columbia, Kansas, Louisiana, Maine, Maryland, New Hampshire, North Dakota, Oklahoma, Tennessee, Vermont, Virginia, West Virginia.

Limitations and Further Research: While the MAST has been in widespread use, the design of the questionnaire has been criticized for the ease with which clients can deliberately and undetectably falsify responses.^{57,93} In the comparative study of this report, investigators^{9,38} examined its predictive validity for DWI recidivism and reported less than ideal results (Table 5). Other investigators^{40,41,43-44} compared its predictive validity for AUD with a criterion measure and reported inconsistent results (Table 6). Despite the favorable results reported by Lacey et al.⁴³ the study suffered from methodological problems. Inconsistent results have been reported in the literature. After decades of research, the MAST has not been shown to be in the top-rated group and its validity as a DWI screening instrument has not been established. Considering the widespread use of the MAST, rigorous studies of these issues are urgently needed.

MF

Overview: Developed in 1971, MF is one of the oldest instruments. It was designed specifically for DWI screening. Independent investigators^{35,36} evaluated its validity for predicting DWI recidivism and reported minimally acceptable results (Table 5). Lacey et al.⁴³ compared its validity for predicting AUD with a criterion measure and reported good agreement (Table 6). Inconsistent results concerning its validity make the use of MF problematic.

Current Use: Alabama, Alaska, Arizona, California, Colorado, Connecticut, Delaware, Illinois, Kansas, Louisiana, Massachusetts, Mississippi, New Hampshire, North

Carolina, North Dakota, Oklahoma, Oregon, Rhode Island, Tennessee, Texas, Virginia.

Limitations and Further Research: The inconsistency of findings may be due to an inadequacy in one or both of the studies, or due to an inadequacy in the MF itself. Whatever the reason for the inconsistency, it suggests the need for further research. Inconsistent results have been reported in the literature. Considering the widespread use of the MF, rigorous studies of these issues are urgently needed.

RIASI

Overview: One of the newer instruments, developed in 1997, RIASI is easily administered for DWI screening as a paper/pencil assessment. The 52 questions can be scored within 20 minutes.

Current Use: New York.

Limitations and Further Research: The authors have been engaged in research efforts^{38,44,94} that have demonstrated some degree of validity, but more independent research is still needed. While it can be administered and scored efficiently, it does not have the computer automation and summary printout with treatment recommendations.

SALCE/NEEDS

Overview: SALCE/NEEDS was designed specifically for DWI screening. It has a reasonable processing time, is computerized, and includes many of the relevant testing domains, DWI normative data, and profile summary with treatment recommendations. It is becoming more used for these reasons.

Current Use: Alabama, Georgia, Maine, Michigan, New Hampshire, New Mexico, North Carolina, Texas, West Virginia.

Limitations and Further Research: The instrument lacks sufficient rigorous research that would demonstrate its validity. Although Lacey et al.⁴³ reported reasonable validity statistics for SALCE/NEEDS, it should be noted that *all* the instruments evaluated by Lacey et al.⁴³ had reasonable statistics (the methodological issues of the study were discussed previously). Furthermore, the developmental stage of the instrument is unclear, and the developers do not appear to have engaged in any research since the instrument's development in 1987. Because it is commercially available but lacks solid data yet is in widespread use, more research is needed to examine its predictive validity for DWI recidivism.

SASSI

Overview: The authors of SASSI have been productive⁹⁵⁻¹⁰⁰ in research and re-

ported normative samples of SASSI on 2,015 respondents from various clinical settings including addiction treatment centers, general psychiatric hospitals, a dual diagnosis hospital, a vocational rehabilitation program, and a sex offender treatment program.

Current Use: Kansas, North Carolina, Oklahoma, Texas, West Virginia.

Limitations and Further Research: More research is needed to examine the instrument's norms and predictive validity for DWI screening, since it has not been validated for a DWI offender population. Only one study⁵⁷ reported moderate correlations with the CAGE, the MAC, and the MAST. Furthermore, although it has a reputation for being resistant to dissimulation, Myerholtz et al.⁵⁷ showed, to the contrary, that it was vulnerable to intentional faking. Interviewers need to be aware that the profile reported by SASSI is based on clinical samples and thus SASSI's profiles may not be as reliable or informative as instruments normed for DWI populations. Interviewers should consider requesting that the developers incorporate DWI normative statistics before they continue using the SASSI. Validity and normative statistics of the instrument need to be established in DWI screening.

Discussion and Recommendations

This report describes the first study in almost a decade—since Popkin et al.,¹⁵ Beirness et al.,⁵⁰ and Mayhew et al.¹⁶ collectively reviewed about 20 instruments to compile and critically review the literature and to evaluate the predictive validity of DWI screening instruments.

The results of the national survey by the National Organization of State Impaired-Driving Programs (National DWI Compendium), presented in this report, indicated that most states mandate screening for drunk-driving offenders. Of these, 90% report using procedures that include both in-person interviews and self-report assessments, and the two most widely used screening instruments continue to be the MF and MAST. Computerized and commercialized instruments, e.g., DRI and others, are gaining use that is more widespread.

A Multidimensional Approach to Screening

Too little rigorous research exists on screening for drunk-driving offenders. This is particularly problematic considering the large number of persons arrested for this crime in the United States each year.

The screening process is poorly described and evaluated. The paucity of evidence that screening processes and instruments can impact outcomes strongly suggests that screening programs should be evaluated more rigorously wherever possible. In a meta-analysis of DWI control research, Wagenaar and colleagues were able to summarize only the crudest elements of past studies (positive or negative apparent effect).¹⁰¹ They were not able to apply traditional meta-analytic methods for a number of reasons: single DWI strategies were embedded in broader packages for evaluation, outcome measures were too heterogeneous to combine, and reports were incomplete or ambiguous. These difficulties were also encountered in the present literature review.

The National DWI Compendium survey shows that current assessment techniques suffer from “tunnel vision,” and that assessment should move beyond alcohol-specific measures to include the offender’s alcoholic-related problems. For instance, instruments like the AUI, MAC, MAST, and CAGE have been shown to directly address problem drinking, but fail to address other drug misuse or the psychosocial correlates of drunk-driving behavior. In fact, the published literature is void of studies that evaluate screening instruments for their ability to detect concurrent drug-misuse disorders.

As our understanding of the complexity of problem drinking grows, screening instruments will need to measure and report on a greater number of dimensions, such as personality, emotional and motivational factors, lifestyles, cognitive factors, and psychiatric problems that often coexist with alcohol abuse and dependence.

Screening as Intervention

After extensive literature searches, the authors found no published articles that address the issue of the value of the entire screening process as intervention in itself. Two studies, however,¹⁰² did evaluate counselors' use of scores on standardized self-report instruments in making treatment referral decisions. They showed that some evaluators used these scores only informally, some attached importance to cutoff scores, and others used only selected item responses qualitatively to guide their referral decisions.¹⁰²

On the other hand, the use of brief interventions has been extensively researched and shown to be effective in motivating alcohol-dependent persons to reduce their alcohol consumption.¹⁰³ A brief intervention typically 1) involves four or fewer sessions; 2) is conducted in a non-treatment setting; 3) is performed by individuals who have not specialized in addiction treatment; and 4) is used with individuals who are at risk for dependence.¹⁰⁴ Something as simple as a handwritten note or a telephone call by a counselor expressing concern can act as a brief intervention.¹⁰³ Investigators have reported that simply asking about drinking and its consequences can have a positive effect on drinking patterns.¹⁰⁴ Despite this evidence, neither the value of the screening interview component nor the value of the monitoring phase has been evaluated.

It is, however, a challenge to design and execute an experiment that evaluates the potential therapeutic effects of screening. Because the purpose of screening is to direct subsequent procedures—continuing evaluation or triaging offenders to different treatment tracks—the effects of screening alone often cannot be isolated. A true test of an instrument's predictive validity would require that some subjects receive screening without treatment. Such a study might be difficult to justify or execute in a real-world setting. Still, there is a precedent for conducting randomized studies in partnership with court systems, and opportunities for creative research abound. For example, numerous convicted DWI offenders in various states are screened well before a judge mandates any treatment.

An ideal experimental design would include a group of offenders who are not screened and who would receive a single treatment option, without regard to risk of recidivism or need for rehabilitation. Similarly, different screening protocols could be paired with a treatment constant so that instruments and processes could be compared. These designs are difficult but not impossible to implement within existing court systems.

The authors recommend that a greater emphasis be placed on researching and understanding how different groups of offenders, identified through typologies, fare within screening and treatment modalities. When possible, comparison groups should be measured in parallel that do not follow the same treatment path.

Studies are also needed that include those who were arrested but not convicted because of plea bargaining, court leniency, or who were simply not referred for mandatory evaluation.

Evaluation of Existing Studies on Recidivism and AUD Prediction

RECIDIVISM PREDICTION

Screening personnel who use these current instruments should realize that, for the most part, the instruments are more valid at predicting who has an alcohol disorder than at predicting who is at risk for a repeat DWI. Subsequent recidivism was correctly identified in only one-quarter to one-third of respondents who were thought to be at risk because of their screening score.

Based on the evaluation results (Table 11), two instruments, the MacAndrew scale of the MMPI and the AUI, have predictive validity for DWI recidivism. The MacAndrew scale also rated well for predicting AUDs (alcohol abuse or dependence diagnosis). Based on these results, the MAC and the AUI were placed in the highest recommended category. Both instruments were able to predict about 70% of the recidivists and misclassified 30% of them. They were also able to predict about 50% of the nonrecidivists and misclassified 50% of them as recidivists.

New Mexico investigators reported a potential problem of data quality with traffic record matching.⁵⁴ Seven percent of the individuals screened for DWI did not have corresponding records in the NM DWI Citation and Tracking File, which is part of the driver's license master file system operated by the state's Motor Vehicle Division. Nonmatching was primarily due to names, birthdays, and Social Security numbers being presumably falsely reported by the offenders at screening. The nonmatch rate was estimated to be less than 11% in the studies conducted by the Mississippi investigators.^{10,56,84}

These studies were also potentially limited in that offenders are often required by the court or screening counselors to go through additional educational programs or therapeutic interventions, which may affect their drinking-driving behavior and, in turn, DWI recidivism. These programs included elements such as victim impact panels (VIP), treatment, and DWI schools. Recent research suggests that a high percentage of DWI offenders have AUDs and therefore should be referred to treatment. In a review of 22 studies on drinking-driving offenders and alcoholism, Vingilis found that the percentage of offenders considered to be "alcoholic" ranged from as low as 4% to as high as 87%.¹⁰⁵ Furthermore, the definitions of alcoholism varied in these studies. When AUD was broadened to include "alcohol abusers," "heavy drinkers," and "problem drinkers," 50% or more of offenders fell into this extended category.¹⁰⁵ In a 1975 study, Fine and colleagues¹⁰⁶ classified only 8% as having serious alcohol-related problems. In a 1991 study, however, clinical evaluations of first offenders showed that 82% were alcoholics or problem drinkers.¹⁰⁷

The authors' review found only three studies that used structural diagnostic interviews to verify related psychiatric diagnosis in DWI offender populations.¹⁰⁸⁻¹¹⁰ The most recent, by Lapham et al.,¹¹⁰ reported that 85% of female and 91% of male offend-

ers had a lifetime AUD. The study also reported that 32% of female and 38% of male offenders had a lifetime drug-use disorder and that a substantial proportion of offenders suffer from depression, post-traumatic stress disorder, and other psychiatric problems common among those with alcohol and drug problems.

This points out a need to evaluate DWI assessments for their ability to detect nonalcohol-use disorders. Since DWI offenders are notorious for underreporting their alcohol use, we suggest screening programs should possibly test offenders using biochemical measures that also detect drug use. Certainly, this is a promising area for future study.

Unfortunately, investigations of treatment programs show they have demonstrated very limited success. The basic assumption underlying treatment screening asserts that the treatment will have an impact. C'de Baca and colleagues¹¹¹ examined the effect of VIP referral on recidivism and concluded that VIP was not a strong predictor of recidivism. Similarly, Wells-Parker and colleagues¹¹² reported that the average effect of remediation on recidivism was minimal (around an 8-9% reduction over no remediation). These remediation programs included educational programs, contract probation, treatment, and Alcoholics Anonymous. Earlier, Foon¹¹³ also concluded there is no evidence in the literature that any treatment program for convicted drinking drivers is effective in reducing subsequent recidivism. Referral to treatment, then, is unlikely to moderate outcomes significantly.

We strongly urge researchers to consider the very modest impact of treatment on mitigating drunk driving behavior. We recommend that practitioners work collaboratively with criminal justice professionals and researchers to design and evaluate innovative programs that combine treatment strategies with sanctions and monitoring. We encourage the development and evaluation of treatment programs that include pharmacological agents useful in treating substance use disorders and other psychiatric problems.

AUD PREDICTION STUDIES

Most of the instruments evaluated had relatively higher sensitivity and specificity for AUD prediction than for recidivism prediction.

However, as discussed in the Introduction, these studies have significant limitations. For one, DWI researchers and practitioners do not adhere to a single “gold standard” of AUD, let alone agree on which criterion instrument would best detect it. In DWI screening, the gold standard for AUD has been ambiguous, often consisting of the imprecisely defined “clinical judgment” of a counselor or another screening or diagnostic instrument. Lacey and coworkers⁴³ addressed the lack of a universally recognized standard by creating a new instrument using an expert panel, which was then used to judge the validity of four existing instruments. While the concept of a better instrument is appealing, researchers must establish the instrument’s reliability and validity. As a result, screening instruments are judged according to a criterion that may not itself be valid.

The quest for a good criterion by which to judge screening instruments should lead theoreticians beyond merely verifying counselor decisions to treat or not to treat. As Wells-Parker and Popkin¹¹⁴ point out, the underlying purpose of screening needs to be clearly identified, —whether to predict risk of repeat offense or for triage to appropriate treatment.

Because AUD screening is crucial for referring DWI offenders to appropriate treatment, researchers are urged to address the problem of criterion validity. DWI researchers and practitioners are urged to begin a dialogue designed to address these fundamental issues and, at the very least, to draw tentative conclusions. Indeed, the authors believe researchers and practitioners must determine what constitutes effectiveness in a screening program in order to guide the creation and evaluation of this and the next generation of DWI screening instruments.

Considerations for Choosing Instruments

The authors recommend that DWI researchers, practitioners, and those who select screening instruments be informed regarding the normative data on which these instruments are derived.

Profile rankings and severity scores, when based on those of a clinical population, may not be applicable to the DWI offender population. End users of the instruments also need to be informed regarding the setting where the instrument was used and evaluated, because DWI offenders are often more defensive than populations who take an instrument in other circumstances such as a clinical setting, for instance. Some instruments—DRI, RIASI, and SALCE/NEEDS—were developed specifically for screening DWI offenders. Other instruments—ASI, AUI, MAC, SASSI—were borrowed from other applications.

Results of self-report instruments for DWI offenders who complete screening instruments in a criminal justice setting may differ significantly from the population on which a test was normed. Table 7 summarizes the populations on which the 12 instruments were normed.

The authors urge evaluators to augment data from self-report instruments and face-to-face interviews with externally valid sources of information when making treatment disposition decisions. These include offenders' court criminal/traffic records, arrest and pre-interview breath alcohol testing results, and possibly other biochemical tests in order to help identify those who are more likely to be highly defensive and under-report their involvement with alcohol or other drugs.

Acceptable Standards for Instrument Performance

The cutoff points of 70% for sensitivity and 40% for specificity chosen for this analysis are arbitrary, based more on the limited performance of the instruments than any scientific rationale.

Indeed, the poor results of the instruments reported in this investigation raise fundamental questions about screening. Is an instrument with a sensitivity of 70% and a specificity of 40% acceptable? Such standards certainly would not meet the stringent criteria for a good medical test. Perhaps an even more relevant question is: Do these criteria meet the needs of DWI program administrators and interviewers to help triage offenders?

These crucial questions must be carefully considered, discussed, and addressed. Such discussions are essential for guiding the creation and evaluation of the next generation of DWI screening instruments. Indeed, before effective instruments can be designed it is important to select appropriate standards for determining an instrument's effectiveness.

Expanding Data to Include Broader Populations

Although rigorous research on some instruments has shown predictive validity, this research was conducted on limited samples from isolated DWI populations. Before these promising instruments can be recommended, they need to be proven effective with populations from other regions of the United States.

It must also be determined whether promising screening instruments and procedures are appropriate for diverse populations. Subgroups may exist for which screening processes might interact and show differential effects. Authors and editors are urged to report descriptive screening statistics of homogenous subgroups, by gender and age and, if possible, by educational attainment, socioeconomic status, and ethnicity.

Even though the subsample may be too small for statistical power in primary analysis, the figures can contribute to secondary analysis by providing aggregate numbers that allow some generalization. For example, 18 studies report quantitative information for the MAST (the most studied instrument in the literature), but the very high percentage of males in the study populations preclude analysis of female offenders in the primary studies. Yet these studies account for an estimated 719 women from at least eight states. It can be argued that this under-representation of women in DWI samples will be an ongoing problem for researchers, even though early evidence indicates these offenders have similar profiles and risk factors compared with male offenders.^{47,115}

Using meta-analytic aggregation of these small samples, researchers may begin to detect differences among offender subsets with some precision and make generalizations about them. This will only be possible to the extent that researchers and publishers report these figures, however. Moreover, researchers whose samples are by necessity limited in size, either by economic or geographic constraints, should be encouraged to publish.

Further Research on Widely Used Assessments

Some of the most widely used screening instruments have little or no research to demonstrate their predictive validity for predicting DWI re-arrest. Because these tests are in such widespread use, further validation research is needed immediately.

Research abounds for some DWI screening instruments, but the most widely used instruments have received little research attention. The AUI, CAGE, MAC, MF, and MAST, for example, have been well studied for diverse groups of offenders/patients. A literature search on the MAC turned up more than 100 published articles. Other newer instruments, however, such as the DRI (developed in 1987) and SALCE/NEEDS (developed in 1986), are commercialized and are increasing in popularity but published research on them is lacking.

Twelve states use the DRI, yet only two independent quantitative studies of it exist.^{42,43} that have evaluated its validity, and those studies suffer from troubling methodological issues.

No published articles exist on SALCE/NEEDS. Although a few unpublished reports were obtained from the developers, no apparent effort has been made to publish these for scientific scrutiny.

SASSI, which is commercialized, is increasing in popularity as well. A literature search found several publications by the developers, but these studies¹¹⁶ focus on clinical populations in various service settings.

As these new screening instruments gain prominence the driving force propelling sales and use may be marketing rather than solid data. As the national survey shows, computerized instruments are becoming more prevalent, and with time others may be better studied and utilized.

DWI researchers and practitioners are urged to close the gap between research and practice. Those practitioners who select screening instruments should base their decisions on compelling scientific evidence for the tests' applicability to the DWI population. Researchers are urged to rigorously investigate these existing instruments. Journal editors need to publish their findings, even when low predictive validities are found.

Addressing Client Defensiveness

Most screening instruments are self-reporting, although some also incorporate face-to-face interviews. Self-report methods are seriously limited, since DWI offenders in the criminal justice system are likely to be motivated to under-report or deliberately falsify their alcohol use.²² Although some self-report instruments have developed scales to measure and compensate for under-reporting, many of the most widely used instruments do not. Other screening approaches use biochemical testing to determine heavy

drinking and drug use and employ additional information for external validation of offender self-reports. In any case, defensiveness is a serious issue that has a significant impact on the predictive power of screening instruments.

In order to compensate for defensiveness, researchers should follow the recommendations of Lapham and colleagues.^{51,54} They suggest several techniques that may be useful adjuncts in DWI screening. These include: examination of court records for previous DWI and other offenses or use of collateral interviews in conjunction with a self-report measure; requiring the offender to submit to biochemical tests to detect alcohol or drug abuse; use of routine alcohol breath testing at screening appointments; and training of interviewers in motivational interviewing techniques. More studies are needed to determine whether these adjuncts are useful in better identifying at-risk offenders.

Conclusion

Between a quarter and a third of first offenders are later re-arrested for drunk driving. A recent report also suggests that many more, in fact the vast majority of drunk driving offenders, have diagnosable alcohol disorders and a large proportion use drugs or have psychiatric problems.⁵¹ Perhaps more intense interventions with all offenders at an earlier stage may reduce recidivism and crash rates. Given the state of the art in screening and the inability of any currently available instrument to predict which offenders will recidivate, perhaps what is called for is a re-examination of the entire system of adjudicating offenders, including screening, sanctions, and treatment.

Acknowledgements

The authors would like to thank Ms. Barbara Lauer, Bureau of Driver Education and DWI Programs, Florida Department of Highway Safety and Motor Vehicles for releasing the National DWI Compendium Survey, Dr. Chuck Paine for technical writing and editing, Dr. Catherine Taylor for proofreading, Ms. Joyce Welt and Ms. Rachael Vidales for reference management, graphic and word processing support, and Mr. Don Peyton for literature/author searches and tracking. Without them, this report would not have been possible.

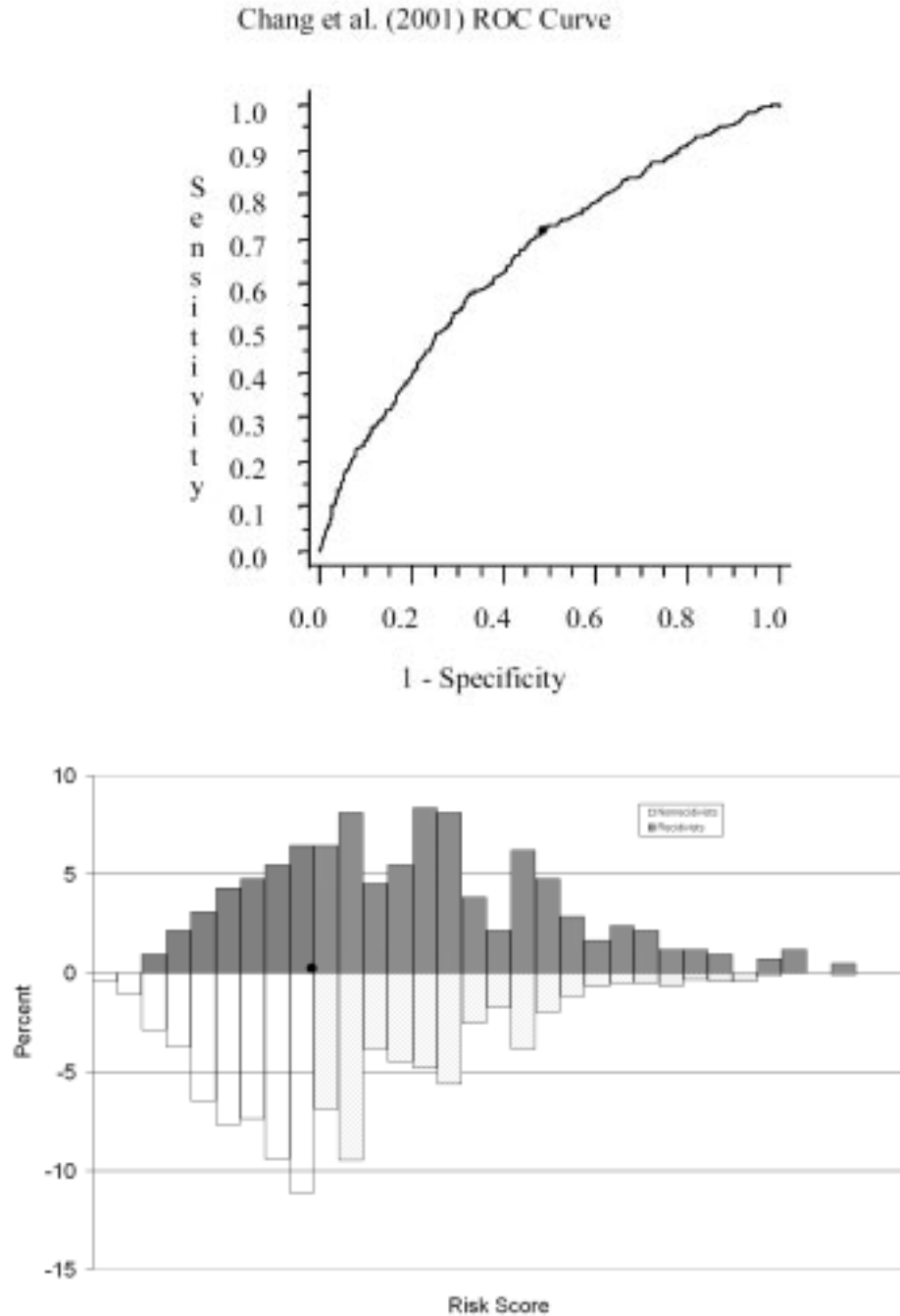
Appendix A:

Terms used for Database Searches

TERM	VARIATIONS	
1 quantitative	quant# methodologic# systematic best evidence	text word
2 review	review# overview#	publication type
3 meta-analysis	meta-analy# metaanaly# meta analy#	publication type
4 clinical trial	clinical trial# randomized clinical # RCT# random# control#	text word
5 DWI/DUI	DUI# DWI# driving wh# intox# driving wh# impair# dr#nk driving driving drunk automo# (AND) alcohol# (OR) substance# (OR) ETOH driv# (AND) alcohol# (OR) substance# (OR) ETOH	text word
6 screening	screen# instrument# diagno# assess# evaluat# predict#	text word
7 instrument	(each of the instruments indicated in Section I, both by acronym and by exact phrase)	text word
8 compliance	complian# complet#	text word
9 school	educat# school	text word
10 counseling	couns# treatment therap#	text word
11 Alcoholics Anonymous	AA Alc# Anon#	text word

Appendix B:

ROC Curve And Sensitivity/Specificity Analysis-Risk Score



The receiver operating characteristics (ROC) curve (top figure) shows the predictive accuracy for recidivists (y-axis) and misclassification of nonrecidivists (x-axis) based on a logistic regression model by Chang et al.⁵⁴ for AUI. For a reasonably optimal screening threshold (risk score) indicated by the dot on the figure, 71% of the recidi-

vists were identified and 51% of the nonrecidivists were incorrectly identified as recidivists. Therefore, if the ROC curve rises quickly, more recidivists will be identified and fewer nonrecidivists will be identified incorrectly. For a model with high predictive accuracy, the ROC curve rises quickly so that the area under the curve is large. Conversely, a slowly-rise ROC curve has a smaller area under the curve with low predictive accuracy. The area under the curve is 0.657 in the top figure and is not considered to be a fast rising one. This test is by no means meeting the stringent criteria of an accepted standard in medical practice.

The second figure above shows the population distribution by risk score. The upper portion of the figure shows the distribution of recidivists and the bottom portion is for nonrecidivists. The dot indicates the screening threshold used to derive the 71% and 51% values above. Although the 51% value indicates a misidentification for recidivism, it may have some significance in identifying problem drinkers because, according to a review by Vingilis et al.,¹⁰⁵ 50% of the DWI offenders are likely to be problem drinkers.

The analysis discussed above raised two fundamental issues: (1) Is detecting 71% of the recidivists good enough? (2) Is misclassification of 51% of the nonrecidivists acceptable? As discussed above, a screening instrument should optimize its accuracy by having a fast-rising ROC curve. Given a curve, a reasonable screening threshold may be defined to allow ~ 50% misclassification to optimize the identification of recidivists and problem drinkers. Whether the threshold should be lowered further to detect more problem drinkers to keep potential drunk drivers off the street should be a decision based on a realistic cost-benefit analysis. It is also recognized that other interpretations and optimization approach may be used for different applications.

Tables

Table 1: Screening Instruments for DWI Offenders

Instrument	Fee	Length ^a	Primary Domain	Considerations
Alcohol Severity Index (ASI)		●●●● during interview	Problem areas in substance abusing patients	More for treatment planning than screening
Alcohol Use Inventory (AUI)	\$	●●●●●	Perceptions, benefits, styles of drinking	More for treatment planning than screening
CAGE		●	Alcoholism	Simple and brief
Driver Risk Inventory (DRI) (Newer version DRI-II)	\$	●●● ●●●	Alcohol, drugs, driver risk, stress/coping	Truthfulness correction, Designed for DWI screening
Life Activities Inventory (LAI)		●●●+ interview	Life situation and Personality scales	Designed for DWI offenders to assess treatment induced changes in life circumstances over time
MacAndrew Alcoholism Scale (MAC)	\$	●	Personality characteristics	Subscale of MMPI
Minnesota Assessment of Chemical Health (MACH)	\$	●●●(varies by branching)	Severity, stressors, obstacles, referral	MAST, MF, DSM-III-R questions
Mortimer-Filkins Questionnaire (MF)		●+ interview	Problem drinking	Widely used since 1971, Designed for DWI screening
Michigan Alcoholism Screening Test (MAST) ^b		●	Alcoholism	Widely used since 1971
RIA Self-Inventory Screening Instrument (RIASI)		●	Alcohol use, problem areas, personality and behavior questions	Non-obvious indicators, Designed for DWI screening
Substance Abuse Life Circumstances Evaluation (SALCE)/NEEDS	\$	●● ●●●	Alcohol dependency, life stress issues, attitudes	Truthfulness scale, Designed for DWI screening
Substance Abuse Subtle Screening Inventory (SASSI)	\$	●●	Chemical dependence, related psychosocial domains	Designed for screening of a variety of clinical populations

\$ Fee per test

^a 50 items = ●

^b Does not include briefer and specialized variants of the MAST

Table 2: DWI Offender Assessment Fees, Treatment Referral Rates, and Sanctions by State

State	Assessment Fees ^a	Treatment Referral Rate ^b (First Offenders)	License Suspension/Revocation	Vehicle Impoundment, Monitoring Device and/or Interlock Device	Estimated Number of DUI Arrests 1997	Rate of DUI Arrests per 100,000 drivers 1997
Alabama (AL)	C	22%	▲		21,600	638
Alaska (AK)	S	65%	▲	▲	5,200*	1,196*
Arizona (AZ)	nr	nr	▲	▲	33,100	1,062
Arkansas (AR)	S	nr	▲	▲	19,400	1,032
California (CA)	C	100%	▲	▲	193,400	949
Colorado (CO)	C	95%	▲	Voluntary	23,500	827
Connecticut (CT)	C	nr	▲	nr	12,500	551
Delaware (DE)	C	nr	▲	Voluntary	nr	nr
District of Columbia (DC)	nr	nr	▲	nr	3,400*	927*
Florida (FL)	C	61%	▲	▲	52,000*	473*
Georgia (GA)	C	nr	▲	nr	58,600*	1,216*
Hawaii (HI)	O	nr	▲		5,300	721
Idaho (ID)	C	nr	nr	nr	10,400	1,227
Illinois (IL)	C	nr	▲	▲	48,600**	660**
Indiana (IN)	C	nr	▲	nr	26,700	681
Iowa (IA)	C	nr	▲	nr	14,200	725
Kansas (KA)	C	nr	▲	nr	25,400**	1,429**
Kentucky (KY)	C	nr	▲	nr	17,200*	685*
Louisiana (LA)	S	nr	▲	nr	17,300	645
Maine (ME)	C	55%	▲	▲	7,600	839
Maryland (MD)	nr	nr	▲	▲	23,700	709
Massachusetts (MA)	C	nr	▲	▲	18,400	419
Michigan (MI)	C	nr	▲	▲	55,600	823
Minnesota (MN)	C	nr	▲	▲	39,300	1,385
Mississippi (MS)	C	0%	▲	nr	nr	nr
Missouri (MO)	C	nr	▲	▲	30,800	823
Montana (MT)	C	42%	▲	nr	5,800**	995**
Nebraska (NE)	nr	nr	▲	▲	13,600	1,156
Nevada (NV)	C	nr	▲	▲	6,600*	654*
New Hampshire (NH)	C	66%	▲		6,900*	786*
New Jersey (NJ)	nr	38%	▲		25,700	461
New Mexico (NM)	C3	nr	▲	nr	3,800	1,156
New York (NY)	C	36%	▲	nr	55,500*	535*
North Carolina (NC)	C	60%	▲	▲	67,600	1,252
North Dakota (ND)	C	35%	▲		3,700	829
Ohio (OH)	C	40%	nr	nr	44,500	544
Oklahoma (OK)	C	nr	▲	▲	20,400	895
Oregon (OR)	C	83%	▲	▲	16,400	719
Pennsylvania (PA)	C	85%	▲	nr	39,400*	485*
Rhode Island (RI)	C	nr	▲	▲	2,100	306
South Carolina (SC)	C	20%	▲	nr	14,500	557
South Dakota (SD)	C	nr	nr	nr	7,200*	1,418*
Tennessee (TN)	C	nr	▲	nr	nr	nr
Texas (TX)	C	nr	▲	nr	84,300	657
Utah (UT)	C	nr	nr	nr	7,800	577
Vermont (VT)	C	65%	▲	nr	2,900*	640*
Virginia (VA)	C	50%	nr	nr	31,500	643
Washington (WA)	nr	nr	nr	nr	50,600	1,261
West Virginia (WV)	C	77%	▲		9,800	760
Wisconsin (WI)	C	40%	▲	nr	33,200	904
Wyoming (WY)	C	nr	▲	nr	4,300	1,214

^a C: Primarily paid by client; S: Primarily supported by State funding; O: Supported by other funding sources.

^b Some states reported rates for all offenders

nr - no response; ▲ - present; * - 1994 DOJ Report; ** - 1991 DOJ Report

Table 3: DWI Offender Program Process Characteristics by State

State	Mandated Screening ^a	Pre-Trial	Testing	Face to Face Interview	Compliance Monitoring	Criteria for Treatment Referral
Alabama (AL)	▲	▲	▲	▲	▲	▲
Alaska (AK)	▲	▲	▲	▲	▲	▲
Arizona (AZ)	▲		▲	▲	▲	▲
Arkansas (AR)	▲		▲	▲	▲	▲
California (CA)	▲	▲	▲	▲	▲	▲
Colorado (CO)	▲	▲	▲	▲	▲	▲
Connecticut (CT)		nr	▲	▲	▲	▲
Delaware (DE)	▲	▲	▲	▲	▲	▲
District of Columbia (DC)		▲	▲	▲	▲	▲
Florida (FL)	▲	▲	▲	▲	▲	▲
Georgia (GA)			▲		▲	
Hawaii (HI)	▲	▲	▲	▲	▲	▲
Idaho (ID)	▲	▲	▲	▲	nr	▲
Illinois (IL)	▲	▲	▲	▲	▲	Nr
Indiana (IN)		▲	▲	▲	▲	▲
Iowa (IA)	▲		▲	▲	nr	Nr
Kansas (KA)	▲	▲	▲	▲	▲	▲
Kentucky (KY)	2 nd offense		▲	▲	▲	▲
Louisiana (LA)		▲	▲	▲	▲	▲
Maine (ME)	▲		▲	▲	▲	▲
Maryland (MD)	▲	▲	▲	▲	▲	▲
Massachusetts (MA)	▲	▲	▲	▲	▲	▲
Michigan (MI)	▲	▲	▲	▲	▲	▲
Minnesota (MN)	▲	▲	▲	▲	▲	▲
Mississippi (MS)		▲	▲	N-1 st Y-multiple	▲	N-1 st Y-multiple
Missouri (MO)	▲	▲	▲	▲	▲	▲
Montana (MT)	▲	▲	▲	▲	▲	▲
Nebraska (NE)	▲	▲	▲	▲	▲	
Nevada (NV)	▲	▲	▲	Nr	▲	Nr
New Hampshire (NH)	▲		▲	▲	▲	▲
New Jersey (NJ)	▲		▲	▲	▲	▲
New Mexico (NM)	▲	▲	▲	▲	▲	▲
New York (NY)		▲	▲	▲	▲	▲
North Carolina (NC)	▲	▲	▲	▲	▲	▲
North Dakota (ND)	▲		▲	▲	▲	▲
Ohio (OH)	▲		▲	▲	▲	▲
Oklahoma (OK)	▲	▲	▲	▲	▲	▲
Oregon (OR)	▲		▲	▲	▲	▲
Pennsylvania (PA)	▲		▲	▲	▲	▲
Rhode Island (RI)		▲	▲	▲	▲	▲
South Carolina (SC)			▲	▲	▲	▲
South Dakota (SD)	BAC >.17	▲	▲		nr	Nr
Tennessee (TN)	2 nd offense	▲	▲	▲	▲	▲
Texas (TX)	▲	▲	▲	▲	▲	▲
Utah (UT)		▲	▲	nr	nr	Nr
Vermont (VT)	▲		▲	▲	▲	▲
Virginia (VA)	▲		▲	▲	▲	▲
Washington (WA)	▲	nr	nr	nr	nr	Nr
West Virginia (WV)	▲	nr	▲	▲	▲	▲
Wisconsin (WI)	▲	▲	▲	▲	▲	▲
Wyoming (WY)		nr	▲	▲	nr	▲

^a National Commission Against Drunk Driving (2001)

Table 4: DWI Screening Instruments for DWI Offenders by State

State	Instrument(s)
Alabama (AL)	DRI, SALCE, MF
Alaska (AK)	MF
Arizona (AZ)	DRI, MAST MAC, MF
Arkansas (AR)	
California (CA)	MF, MAST
Colorado (CO)	MF
Connecticut (CT)	MF
Delaware (DE)	MF, DRI, ASI
District of Columbia (DC)	SMAST
Florida (FL)	DRI
Georgia (GA)	SALCE
Hawaii (HI)	DRI
Idaho (ID)	
Illinois (IL)	DRI, MF
Indiana (IN)	
Iowa (IA)	
Kansas (KA)	SASSI, MAST, MF
Kentucky (KY)	DRI
Louisiana (LA)	MF, SMAST
Maine (ME)	MAST, SALCE
Maryland (MD)	MAST
Massachusetts (MA)	MF
Michigan (MI)	NEEDS ^a
Minnesota (MN)	
Mississippi (MS)	MF
Missouri (MO)	DRI
Montana (MT)	
Nebraska (NE)	
Nevada (NV)	
New Hampshire (NH)	MAST, MF, CAGE, DRI, SALCE,
New Jersey (NJ)	
New Mexico (NM)	NEEDS ^a
New York (NY)	RIASI
North Carolina (NC)	MF, DRI, MAC, MACH, SALCE, SASSI
North Dakota (ND)	MAST, MAC, MF, ASI
Ohio (OH)	
Oklahoma (OK)	MF, ASI, SASSI, MACH, MAST, DRI
Oregon (OR)	MF
Pennsylvania (PA)	
Rhode Island (RI)	MF
South Carolina (SC)	
South Dakota (SD)	
Tennessee (TN)	MF, MAST, LAI
Texas (TX)	SALCE, SASSI, MF
Utah (UT)	
Vermont (VT)	MAST, CAGE
Virginia (VA)	MF, MAST
Washington (WA)	nr
West Virginia (WV)	ASI, AUI, MAST, SASSI, SALCE, CAGE, DRI
Wisconsin (WI)	
Wyoming (WY)	

a: NEEDS is a later version of SALCE

Table 5: Validity of Predicting Subsequent DWI

Instrument	Recidivism (months)	Setting	Sample	Quantitative Measure	Sensitivity	Specificity	PP+	PP-	Authors	Score
AUI	48	DWI screening	1,496 New Mexico	Logistic Regression (using AUI D2 & MAC)	66%	51%	31%	82%	C' de Baca et al. ²⁹	●
	60		1,644 New Mexico	Logistic Regression	71%	49%	32%	83%	Chang et al. ²²	
	48		1,384 New Mexico	Life Table Analysis (using AUI D2 & MAC)	69%	42%	21%	86%	Lapham et al. ⁹	
					ALCINVOL score	57%	53%	24%		82%
			DISRUPT I score	52%	59%	25%	82%			
			DISRUPT II score	18%	89%	30%	80%			
			ENHANCED score	54%	57%	25%	82%			
LAI	24	DWI screening	3,050 Mississippi	LAI vs. no LAI	Low risk drinkers: 6% reduction in recidivism High risk drinkers: no effect				Neff ²⁷	
	6-9 years		3,035 Mississippi	Logistic Regression	Low risk drinkers: Estimated 6% reduction in recidivism High risk drinkers: no effect				Wells-Parker ²⁵	
MAC	48	DWI screening	1,496 New Mexico	Logistic Regression (using MAC & AUI)	66%	51%	31%	82%	C' de Baca et al. ²⁹	●
	48			11-22 vs. 23+	67%	48%	25%	69%		
					Life Table Analysis (using MAC & AUI D2)	69%	42%	21%	86%	Lapham et al. ⁹
MAST	48-53	Educational program	1,592 New York	Life Table Analysis (using MAC & MAST)	72%	35%	20%	72%	Noehjnski and Wietczorek ¹⁸	
				high risk	22-46%	-	-	-		
				low risk	47%	58%	25%	81%		
MF	36	DWI screening	1,384 New Mexico	Life Table Analysis (using MAST & MAC)	33%	73%	25%	81%	Lapham et al. ⁹	
				Non- + Possible Problem Drinkers	72%	35%	20%	72%		
				Problem Drinkers	22%	86%	73%	39%		Struckman-Johnson and Mashill ¹⁵
RIASI	48-53	Educational program	9,920 ^b Texas, South Dakota	Non- or Possible Problem Drinkers	30%	74%	16%	86%	Struckman-Johnson and Mashill ¹⁵	
				Problem Drinkers	29%	81%	36%	75%		
				Social + Excessive drinkers	57-78%	<<50%	-	-		Wendling & Kolody ²⁶
			1,709 ^b California	Problem drinkers				Noehjnski and Wietczorek ¹⁸		
			1,592 New York	high risk						
				low risk						

^btotal score; Questionnaire score

Table 6: Validity of Predicting Alcohol Dependence/Abuse Problems

Instruments	Setting	Sample	Quantitative Measure	Sensitivity	Specificity	PP+	PP-	Authors	Score
CAGE CAGE-6	DWI screening	375 Washington	Counselor	60%	86%	67%	83%	Mischke & Vennart ³⁰	●
	Treatment or Educational Programs	609 Boston and Pittsburgh	Problem Drinking, Expert Survey	91%	26%	58%	71%	Lacey et al. ³¹	
CAGE-IC	Treatment or Educational Programs	609 Boston and Pittsburgh	Problem Drinking, Expert Survey	67%	76%	78%	67%	Lacey et al. ³¹	● ●
	DWI Screening	333 Kentucky	Alcohol Status	63%	70%	47%	82%	Lembarwitz & Meyers ³²	
MAC	Inpatient and treatment program	1,283 Military, Texas	DSM III	73%	42%	94%	18%	Lasker & Gold ³³	●
MAST + Repeat DWI	Educational Program	97 New York	Problem Drinking, Local Criteria	73%	73%	-	-	Nochajski & Wierczok ³⁴	●
	Treatment or Educational Programs	609 Boston and Pittsburgh	Problem Drinking, Expert Survey	76%	85%	84%	85%	Lacey et al. ³¹	● ●
MAST	Inpatient and treatment program	1,283 Military, Texas	DSM III	60%	72%	94%	20%	Lasker & Gold ³³	● ●
	Educational Program	97 New York	Problem Drinking, Local Criteria	59%	75%	-	-	Nochajski & Wierczok ³⁴	
MF	DWI screening	90 Washington	Counselor	85%	68%	53%	91%	Mischke & Vennart ³⁰	● ●
	Treatment or Educational Programs	609 ^a Boston and Pittsburgh	Problem Drinking, Expert Survey	78%	66%	73%	72%	Lacey et al. ³¹	
REASE REASE + Repeat DWI	DWI screening	41 ^b Michigan	Counselor	82%	57%	98%	40%	Mortimer et al. ³⁵	● ●
	DWI screening	253 ^b Washington	Counselor	65%	77%	88%	74%	Mischke & Vennart ³⁰	
SALCE	Educational Program	113 New York	Problem Drinking, Local Criteria	82%	81%	-	-	Nochajski & Wierczok ³⁴	● ●
	Treatment or Educational Programs	609 Boston and Pittsburgh	Problem Drinking, Expert Survey	92%	83%	-	-	Nochajski & Wierczok ³⁴	
				92%	57%	71%	57%	Lacey et al. ³¹	●

^aquestionnaire score; ^btotal score

Table 7: Instrument Administration Characteristics

Instrument	Target Population	Fees	Length (# of Questions)	Testing Time (min.)	Computerized or Pencil/Paper Administration	Computerized or Hand Scoring	Scoring Time (min)	Training Required	Score
ASI	Adults	0	200	50-60	Both	Both	5	Yes, self training possible	
AUI	Adults & Adolescents 16+	\$1-15*	228	35-60	Both	Both	3-5	Yes	
CAGE	Adults & Adolescents 16+	0	4	1	Both	Hand	<1	No	●
DRI	Adults	\$7	140	30-35	Both	Computerized	2	No	
LAI	Adults	0	115	60	Pencil/Paper	Hand	5-20	nr	
MAC	Adults	\$1	49	10	Both	Hand	2-3	No	●
MACH	Adults & Adolescents 16+	\$5	11->200	30	Computerized	Computerized	<1	No	
MF	Adults	0	58	45-90	Pencil/Paper	Hand	5	No	●
MAST	Adults	0	25	10	Pencil/Paper	Hand	5	No	●
RIASI	Adults	0	52	14	Pencil/Paper	Hand	3	No	●
SALCE/NEEDS	Adults	\$4.5	98	20	Pencil/Paper	Computerized	3	No	
SASSI	Adults & Adolescents 12+	\$1-7 ^a	78	10-15	Both	Both	10	No	

^a The estimated price range represents a minimum charge if administered using paper and pencil and scored manually and a higher charge if computer administered, scored, and reported.

Table 8: Primary Testing Domains

Instrument	Normative Sample	Summary of Scales/Problem Areas	Comments	# of Areas Covered	Score
ASI	Alcoholics	Substance Abuse, Medical, Employment/Support, Legal, Family/Social, Psychological	Confidence rating	4	●
AUI	Alcoholics	Sociability, Mental Functioning, Mood Management, Marital Problems, Gregarious Drinking, Compulsive Obsession, Sustained/Periodic Drinking, Loss of Control, Social Role Maladaptation, Perceptual Withdrawal Symptoms, Somatic Withdrawal Symptoms, Quantity, Guilt, Attempts to Deal with Drinking, Readiness, Awareness, Life Disruption, and Broad Involvement	Indirect Measure of Life Disruption	3	
DRI	DWI Offenders	Truthfulness, Alcohol, Drug, Driver Risk, Stress Coping Abilities, Substance Dep/Abuse		5	●
CAGE	Psychiatric Patients (39% Alcoholics)	Alcoholism		1	
LAI	DWI Offenders	Alcohol Quantity-frequency, Alcohol Problems, Physical Health, Financial/employment, Social interaction/involvement, Family status/living Situation, Marriage, Treatment Receptivity, Residential Stability, Extroversion/introversion, Sanguine, Self-Confident vs. Anxious, Depressed, Moralistic, Conservative vs. Nontraditional Unconstrained, Paranoid Suspicious vs. Naive Trust, Conforming Compliance vs. Acting-out Aggressiveness, and Improbability	Lie Scale Item	3	
MAC	Members of the Public	Personality and Attitudinal Characteristics related to Alcoholism		2	
MACH	DWI Offenders and Others	Drug Involvement Scale, Alcoholism, Problem Drinking		2	
MAST	5 Groups including DWI Offenders	Alcoholism		1	
MF	Alcoholics/DWI Offenders	Problem Drinking	Questions on Family/Social Problems	2	
RIASI	DWI Offenders	Drinking habits, Number of Drinking Locations, Number of Drinks, Family History, Alcohol Beliefs, Preoccupation with Alcohol, Health Issues, Personal Competence, Aggression, Depression, Anxiety, Deviant Behavior, Sensation Seeking	Non-obvious Indicators	5	●
SALCE/NEEDS	DWI Offenders	Test Testing Attitude, Problem Solving, Emotional Health, Substance Abuse, Employment, Relationship, Physical Health, Education, Criminal Record, Overall Summary		4	●
SASSI	Alcoholics	Face Valid Alcohol (FVA), Face Valid Drug (FVOD), Symptoms (SYM), Obvious Attributes (OAT), Subtle Attributes (SAT), Defensiveness (DEF), Supplemental Addiction Measure (SAM), Family vs. Control Subjects (FAM), Correctional (COR), Random Answering Pattern (RAP)		3	

Table 9: Profile Reporting and Recommendations for Treatment

Instrument	Normative Sample	Participant Demographics	Interpretive Profile	Recidivism Predictive Estimate	Treatment Recommendations	Critical Items	Comments	Score
ASI	Alcoholics	X	▲			▲		
AUI	Alcoholics	X	▲			▲		
CAGE	Psychiatric Patients (39% Alcoholics)						Cutoff	
DRI	DWI Offenders	X	▲		▲	▲		●
LAI	DWI offenders	X	▲					
MAC	Members of the Public						Cutoff	
MACH	Alcoholics, DWI Offenders	X	▲		▲	▲		●
MAST	Alcoholics, DWI Offenders						Cutoff	
MF	5 Groups Including DWI Offenders						Cutoff	
RIASI	DWI Offenders						Cutoff	
SALCE/NEEDS	DWI Offenders	X	▲		▲	▲		●
SASSI	Alcoholics	X	▲					

Table 10: Reliability/Validity Statistics and Research Activities

Instrument	Normative Sample	Internal ^a Reliability/ Consistency	Construct/ Criterion/ Discriminate Validity	Source	# of Independent Studies in DWI	Score
ASI	Alcoholics		.29-.61	McLellan et al. ¹¹⁷	0	
AUI	Alcoholics	.54-.93 / .29-.90	.72	Horn et al. ²¹ ; Chang et al. ²²	5	●
CAGE	Psychiatric Patients (39% Alcoholics)	NA / .71	.60-.89 ^b	Mayfield et al. ²³ ; Mischke & Venneri ⁴⁰	5	●
DRI	DWI Offenders	.81-.93 / NA	.29-.96	Marsteller & Davignon ⁵⁸	2	
LAI	DWI Offenders	.55-.87 / NA	NA	Holden & Reis ²⁵	3	
MAC	Members of the Public	.45-.56 / .56	.82 ^b	Graham ¹¹⁸ ; Lapham et al. ⁹	5	●
MACH	DWI Offenders and Others	NA/NA	.70-.71 ^b	Kincannon ⁸⁷	0	
MAST	5 Groups including DWI Offenders	.83-.95 / .84	.98 (sensitivity) .95 (specificity)	Selzer et al. ⁵³ ; Selzer et al. ¹¹⁹ ; Mischke & Venneri ⁴⁰	9	●
MF	Alcoholics and DWI Offenders	.82-.95 / .73-.83	.85-.92 ^b	Filkins et al. ¹²⁰ ; Webb et al. ¹²¹	5	●
RIASI	DWI Offenders	.82 / NA	.42-.69 ^b	Nochajski & Miller ³¹	0	
SALCE/ NEEDS	DWI Offenders	.93 / NA	.57-.77 (exact agreement) .96-.97 (general agreement)	Correspondence with developer	1	
SASSI	Alcoholics	.33-.93 / .03-.93	.94 (sensitivity) .94 (specificity)	Miller ¹¹⁶ ; Myerholtz & Rosenberg ⁵⁷	1	

^aStatistics reported by developers / statistics reported by independent researchers

^bCorrelation coefficients if not labeled

Table 11: Evaluation Summary

Instrument	Primary Evaluating Criteria		Secondary Evaluating Criteria					Final Scores	Final Rating Category
	Predictive Validity DWI Recidivism	Predictive Validity Alcohol Use Disorder	Administration	Testing Domains	Reliability and Validity Statistics and Research Activities	Profile Reporting / Treatment Recommendations			
ASI				●			●		
AUI	●				●		●	Top	
CAGE		●	●		●		●		
DRI		●		●		●	●	Medium	
LAI									
MACH					●		●		
MAC	●	●	●		●		●	Top	
MAST		●	●				●	Medium	
MF		●	●		●		●	Medium	
RIASI		●	●	●			●	Medium	
SALCE/NEEDS		●		●			●		
SASSI									

Figures

Figure 1: Sensitivity and Specificity/Predictive Value Positive and Negative

	Rearrested for DWI*	Not Rearrested	
Place in High Risk Group by instrument under consideration	a=Number placed in High Risk Group who were rearrested	b=Number placed in High Risk Group who were not rearrested	a + b
Placed in Low Risk Group by instrument under consideration	c=Number placed in Low Risk Group who were rearrested	d=Number placed in Low Risk Group who were not rearrested	c + d

a + c

b + c

a+ b + c + d

Sensitivity = $a/a+c * 100$ or the percentage of all offenders who are rearrested for DWI who were accurately placed in the high risk group by the test

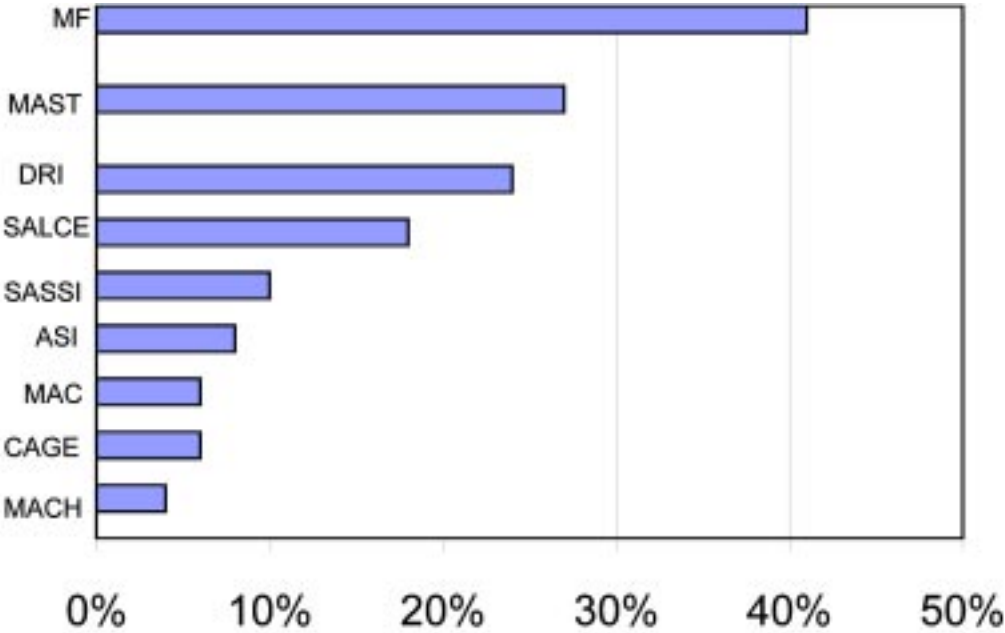
Specificity = $d/b+c * 100$ or the percentage of all offenders who are not rearrested who were accurately placed in the low risk group by the test

Predictive Value Positive (PP+) = $a/a+b * 100$ or the percentage of all those who were placed in the high risk group by the test who subsequently were arrested for DWI

Predictive Value Negative (PP-) = $d/c+d * 100$ or the percentage of all those who were placed in the low risk group by the test who subsequently were not rearrested for DWI

*Calculations are the same for AUD

Figure 2: Prevalence of Screening Instruments Reported by State Court Systems



*Percentages were calculated by dividing the number of states currently using the instrument by the total number of states (n=50) that responded to the survey

Figure 3: Validity of Predicting Subsequent DWI

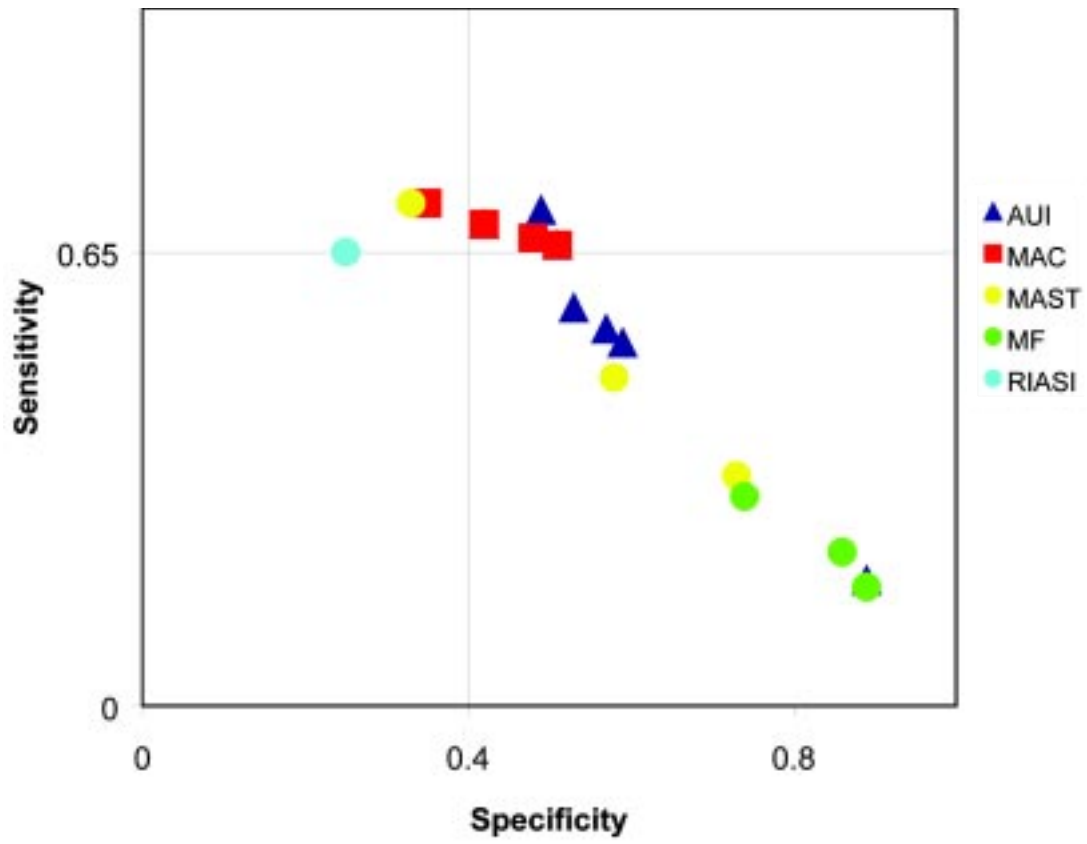
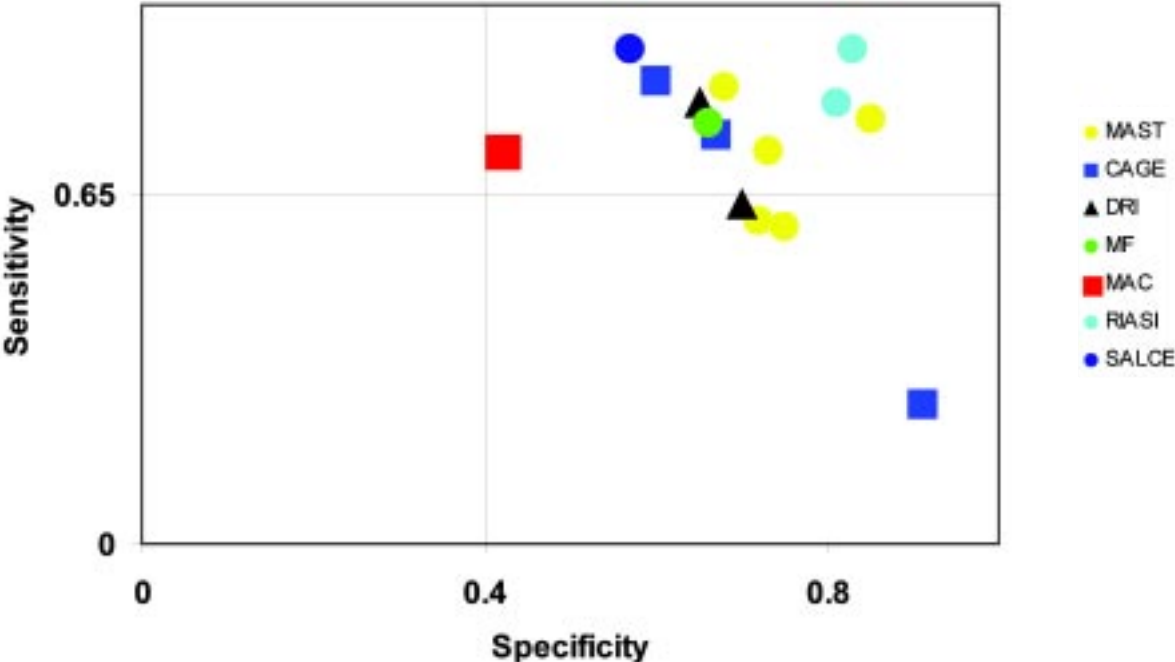


Figure 4: Validity of Predicting Alcohol Use Disorder



Reference List

- 1 National Highway Traffic Safety Administration (1997). Traffic safety facts 1996: A compilation of motor vehicle crash data from the fatal accident reporting system and the general estimates system. (DOT HS 808 649), US Department of Transportation, Washington, D.C.
- 2 US Department of Justice (4-10-2001). Bureau of Justice Statistics. www.ojp.usdoj.gov/bjs/pub/ascii/dwiocs.txt
- 3 National Institute on Alcohol Abuse and Alcoholism (1993). *Eighth Special Report to the US Congress on Alcohol and Health*. US Government Printing Office.
- 4 Waller, P.F. (1985). Licensing and other controls of the drinking driver. *J Stud Alcohol*. (Supplement No.10) 150-160.
- 5 Simpson, H.M. and Mayhew, D.R. (1991, Dec.). The hard core drinking driver, Traffic Injury Research Foundation of Canada.
- 6 National Highway Traffic Safety Administration (2000). Traffic Safety Facts 1999-Alcohol. (DOT HS 809 086), US Department of Transportation, Washington, D.C.
- 7 Gijsbers, A.J., Raymond, A. and Whelan, G. (1991). Does a blood alcohol level of 0.15 or more identify accurately problem drinkers in a drink-driver population? *Med J Aust*. 154 (7) 448-452.
- 8 Lapham, S.C., Chang, I., Berger, L.R. and Skipper, B.J. (2000). Blood alcohol concentrations at arrest and the subsequent diagnosis of alcohol dependence, The International Council on Alcohol, Drugs & Traffic Safety. Stockholm, Sweden.
- 9 Lapham, S.C., Skipper, B.J. and Simpson, G.L. (1997). A prospective study of the utility of standardized instruments in predicting recidivism among first DWI offenders. *Journal of Studies on Alcohol*. 58 (5) 524-530.
- 10 McMillen, D.L., Pang, M.G., Wells-Parker, E. and Anderson, B.J. (1991). Behavior and personality traits among DUI arrestees, nonarrested impaired drivers, and nonimpaired drivers. *Int J Addict*. 26 (2) 227-235.
- 11 Peck, R.C. (1994). Evaluating DUI program impact in quasi experiments. *Alcohol, Drugs and Driving*. 10 (3-4) 207-215.
- 12 Perrine, M.W.; Peck, R.C. and Fell, J.C. (1989). Epidemiologic perspectives on drunk driving. 35-76. Department of Health and Human Services, Public Health Service, Office of the Surgeon General.
- 13 Wieczorek, W.F., Miller, B.A. and Nochajski, T.H. (1992). The limited utility of BAC for identifying alcohol-related problems among DWI offenders. *J Stud Alcohol*. 53 (5) 415-419.
- 14 National Institute on Alcohol Abuse and Alcoholism (1995). *Assessing alcohol problems: A guide for clinicians and researchers*, NIH Publication No. 95-3745, US Department of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD.

- 15 Popkin, C.L., Kannenberg, C.H., Lacey, J.H. and Waller, P.F. (1988, Dec.). Assessment of classification instruments designed to detect alcohol abuse. (DOT HS 807 475), US Department of Transportation, NHTSA, Washington, DC.
- 16 Mayhew, D.R., Beirness, D.J., Simpson, H.M. and Traffic Injury Research Foundation of Canada (1992, June). Diagnostic assessment of problem drivers: State of assessment and treatment techniques., Ontario Ministry of Transportation, Ontario.
- 17 American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders: Fourth Edition*. Williams, J.B.W. (ed.), American Psychiatric Association.
- 18 McLellan, T.A., Luborsky, L., Woody, G.E. and O'Brien, C.P. (1980). An improved diagnostic evaluation instrument for substance abuse patients: The Addiction Severity Index. *J Nerv Ment Dis*. 168 26-33.
- 19 Wanberg, K.W., Horn, J.L. and Foster, F.M. (1977). A differential assessment model for alcoholism: The scales of the Alcohol Use Inventory. *Journal of Studies on Alcohol*. 38 (3) 512-543.
- 20 Horn, J.L., Wanberg, K.W. and Foster, F.M. (1986). Alcohol Use Inventory. P.O. Box 1416, Minneapolis, MN 55440, National Computer Systems, Inc.
- 21 Horn, J.L., Wanberg, K.W. and Foster, F.M. (1990). Guide to the Alcohol Use Inventory (AUI). Minneapolis, MN, National Computer Systems, Inc.
- 22 Chang, I., Lapham, S.C. and Wanberg, K.W. (2001). Alcohol Use Inventory: Screening and assessment of first-time driving-while-impaired (DWI) offenders, I. Reliability and profiles. *Alcohol & Alcoholism*. 36 (2) 112-121.
- 23 Mayfield, D., McLeod, G. and Hall, P. (1974). The CAGE questionnaire: Validation of a new alcoholism screening instrument. *Am J Psychiatry*. 131 (10).
- 24 Behavior Data Systems (1987). Driver Risk Inventory (DRI). P. O. Box 32308, Phoenix, AZ 85064.
- 25 Holden, C.A. and Reis, R.E. (1981). A description of the Life Activities Inventory and scoring procedures 1980 annual report: Volume VI. (DOT HS-806 556), US Department of Transportation, Washington, D.C.
- 26 Hathaway, S.R.; McKinley, J.C. (1943). Minnesota Multiphasic Personality Inventory, MMPI. Minneapolis, MN: University of Minnesota Press, National Computer Systems, Inc.
- 27 Hathaway, S.R.; McKinley, J.C. (1989). Minnesota Multiphasic Personality Inventory 2 (MMPI2). Minneapolis, MN: University of Minnesota Press, National Computer Systems, Inc.
- 28 Kincannon, J. (1984). Minnesota Assessment of Chemical Health (MACH). Chaska, Minnesota, International Professional Services, Inc.
- 29 Selzer, M.L. (1971). The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *American Journal of Psychiatry*. 127 (12) 1653-1658.
- 30 Mortimer, R.G., Filkins, L. and Lowery, F. (1971). Court procedures for identifying

problem drinkers. Final report. DOT Contract FH-11-7615., Highway Safety Research Institute, University of Michigan.

- 31 Nochajski, T.H. and Miller, B.A. (1997). Training manual for the Research Institute on Addictions Self Inventory (RIASI)., Research Institute on Addictions, Buffalo, New York.
- 32 Nochajski, T.H., Miller, B.A. and Wieczorek, W.F. (1997). The Research Institute on Addictions Self-Inventory: A Screening Instrument for Convicted Drinking-Drivers. 1-31. Buffalo, New York, Research Institute on Addictions.
- 33 ADE Inc (1986). Substance Abuse Life Circumstances Evaluation (SALCE). 20 West Washington Street, Suite 12B, Clarkston, MI 48016.
- 34 Miller, G.A. (1985). The Substance Abuse Subtle Screening Inventory (SASSI) Manual. Springville, IN, The SASSI Institute.
- 35 Struckman-Johnson, D.L. and Mushill, E.F. (1976). Program level evaluation of ASAP diagnosis, referral and rehabilitation efforts. Vol. 2. Analysis of ASAP diagnosis and. (DOT HS 802 043), National Highway Traffic Safety Administration.
- 36 Wendling, A., Kolody, B. (1982). An evaluation of the Mortimer-Filkins test as a predictor of alcohol-impaired driving recidivism. *J Stud Alcohol.* 43 (7) 751-766.
- 37 Neff, R.L. (1983). The Life Activities Inventory as a countermeasure for driving while intoxicated. *J Stud Alcohol.* 44 (5) 755-769.
- 38 Nochajski, T.H., Wieczorek, W.F. (2000). Identifying potential drinking-driving recidivists: Do non-obvious indicators help? *Sage Urban Studies Abstracts.* 28 (3).
- 39 C'de Baca, J., Miller, W.R. and Lapham, S.C. (2001). A multiple risk factor approach for predicting DWI recidivism. *Submitted.*
- 40 Mischke, H.D., Venneri, R.L. (1987). Reliability and validity of the MAST, Mortimer-Filkins Questionnaire and CAGE in DWI assessment. *J Stud Alcohol.* 48 (5) 492-501.
- 41 Lucker, G.W., Gold, J.D. (1995). The diagnostic value of BAC for identifying alcohol-related problems among DWI offenders: Another look. *Am J Drug Alcohol Abuse.* 21 (4) 573-580.
- 42 Leshowitz, B., Meyers, J.M. (1996). Application of decision theory to DUI assessment. *Alcohol Clin Exp Res.* 20 (7) 1148-1152.
- 43 Lacey, J.H., Jones, R.K. and Wiliszowski, C.H. (1999). Validation of problem drinking screening instruments for DWI offenders. (Technical Report No. DTNH22-90-C-07287), US Department of Transportation, Washington, D.C.
- 44 Nochajski, T.H., Miller, B.A., Augustino, D.K. and Kramer, R.J. (1995). Use of non-obvious indicators for screening DWI offenders., Buffalo, New York.
- 45 Nochajski, T.H., Walter, J.M. (1998). Association of licensing agency and clinical evaluation outcomes. *Alcoholism: Clinical & Experimental Research.* 22 (3) 77A.
- 46 Chang, I., Lapham, S.C. (1996). Validity of self-reported criminal offences and traffic violations in screening of driving-while-intoxicated offenders. *Alcohol Alcohol.* 31 (6) 583-590.


- 47 Lapham, S.C., Skipper, B.J., Hunt, W.C. and Chang, I. (2000). Do risk factors for rearrest differ for female and male drunk-driving offenders? *Alcoholism: Clinical and Experimental Research*. 24 (11) 1647-1655.
- 48 Little, G.L., Robinson, K.D. (1989). Relationship of DUI recidivism to moral reasoning, sensation seeking, and MacAndrew alcoholism scores. *Psychol Rep*. 65 (3 Pt 2) 1171-1174.
- 49 NHTSA (1996). National Highway Traffic Safety Administration Compendium. National Highway Traffic Safety Administration.
- 50 Beirness, D.J., Simpson, H.M. and Mayhew, D.R. (1991). Diagnostic Assessment of Problem Drivers: Review of Factors Associated With Risky and Problem Driving. Report to Transport Canada TP11549E. Ottawa: Transport Canada, Road Safety and Motor Vehicle Regulation.
- 51 Lapham, S.C., C'de Baca, J., Chang, I., Hunt, W.C. and Berger, L.R. (2001). Are drunk-driving offenders referred for screening accurately reporting their drug use? *Submitted*.
- 52 National Commission Against Drunk Driving (4-23-2001). National Commission Against Drunk Driving Survey. www.ncadd.com/implementation/index3.html
- 53 Selzer, L.M., Vanosdall, F.E. and Chapman, M. (1971). Alcoholism in a problem driver group: A field trial of the Michigan Alcoholism Screening Test (MAST). *Journal of Safety Research*. 3 (4) 176-181.
- 54 Chang, I., Lapham, S.C., C'de Baca, J. and Davis, J.W. (2001). Alcohol Use Inventory: Screening and assessment of first-time DWI offenders. II. Typology and predictive validity. *Alcohol & Alcoholism*. 36 (2) 122-130.
- 55 Wells-Parker, E. (1988). Behavioral and employment consequences of driver's license suspension for drinking driving offenders. *Journal of Safety Research*. 19 (1) 5-20.
- 56 Wells-Parker, E., Anderson, B.J., Landrum, J.W. and Snow, R.W. (1988). Long-term effectiveness of probation, short-term intervention and LAI administration for reducing DUI recidivism. *Br J Addict*. 83 (4) 415-421.
- 57 Myerholtz, L.E., Rosenberg, H. (1997). Screening DUI offenders for alcohol problems: Psychometric assessment of the Substance Abuse Subtle Screening Inventory. *Psychol Addict Behav*. 11 (3) 155-165.
- 58 Marsteller, F.A. and Davignon, D.D. (1997). A validation study of the DRI-II in a large sample of DUI offenders., Emory University School of Medicine, Atlanta, Georgia.
- 59 Butcher, J.N., Dahlstrom, W.G., Graham, J.R., Tellegen, A. and Kaemmer, B. (1989). Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring, University of Minnesota Press, Minneapolis.
- 60 Brooner, R.K., King, V.L., Kidorf, M., Schmidt, C.W. and Bigelow, G.E. (1997). Psychiatric and substance use comorbidity among treatment-seeking opioid abusers. *Archives General Psychiatry*. 54 71-80.
- 61 Kleinman, P.H.; Woody, G.E.; Todd, T.C.; Millman, R.B.; Kang, S.; Kemp, J. and

- Lipton, D.S. (1990). Crack and cocaine abusers in outpatient psychotherapy. *Psychotherapy and Counseling in the Treatment of Drug Abuse*. Psychotherapy and counseling in the treatment of drug abuse, (ed.). 24-35. US Government Printing Office, Washington, D.C.
- 62 Drake, R.E., Mchugo, G.J., Biesanz, J.C. and Validity and Reliability Committee of the NIAAA (1995). The test-retest reliability of standardized instruments among homeless persons with substance use disorders. *Journal of Studies on Alcohol*. 56 (2) 161-167.
- 63 Lesieur, H.R., Blum, S.B. (1992). Modifying the Addiction Severity Index for use with pathological gamblers. *American Journal on Addictions*. 1 (3) 240-247.
- 64 McLellan, A.T., Erdlen, F.R., Erdlen, D.L. and O'Brien, C.P. (1981). Psychological severity and response to alcoholism rehabilitation. *Drug Alcohol Depend.* 8 23-35.
- 65 Hayashida, M., Alterman, A.I., McLellan, A.T., O'Brien, C.P., Purtill, J. and Volpicelli, J. (1989). Comparative effectiveness and costs of inpatient and outpatient medical alcohol detoxification. *New England Journal of Medicine*. 329 (6) 358-365.
- 66 Amoureux, M.P.S.R., van den Hurk, A.A., Schippers, G.M. and Breteler, M.H.M. (1994). The Addiction Severity Index in penitentiaries. *International Journal of Offender Therapy & Comparative Criminology*. 38 (4) 309-318.
- 67 Corse, S.J., Zanis, D. and Hirschinger, N.B. (1995). The use of the addiction severity index with people with severe mental illness. *Psychiatric Rehabilitation Journal*. 19 (1) 9-18.
- 68 Fureman, I., McLellan, A.T. and Alterman, A.I. (1994). Training for and maintaining interviewer consistency with the ASI. *Journal of Substance Abuse Treatment*. 11 (3) 233-237.
- 69 Hendricks, V.M., Kaplan, C.D., Van-Limbeek, J. and Geerlings, P. (1989). The Addiction Severity Index reliability and validity in a Dutch addict population. *Journal of Substance Abuse Treatment*. 6 (2) 133-141.
- 70 Hodgins, D.C., El-Guebaly, N. (1992). More data on the Addiction Severity Index: reliability and validity with the mentally ill substance abuser. *Journal of Nervous and Mental Disease*. 180 (3) 197-201.
- 71 Joyner, L.M., Wright, J.D. and Devine, J.A. (1996). Reliability and validity of the Addiction Severity Index among homeless substance misusers. *Substance Use and Misuse*. 31 (6) 729-751.
- 72 Donat, D.C. (1995). Empirical groupings of perceptions of alcohol use among alcohol dependent persons: a cluster analysis of the Alcohol Use Inventory (AUI) scales. *Assessment*. 1 (1) 103-110.
- 73 Rychtarik, R.G., Koutsky, J.R. and Miller, W.R. (1998). Profiles of the Alcohol Use Inventory: a large sample cluster analysis conducted with split-sample replication rules. *Psychological Assessment*. 10 (2) 107-119.
- 74 Stueland, D., Hart, L. (1979). A canonical correlation analysis of the alcohol-use inventory and the human service scale. *Journal of Behavioral Medicine*. 2 (3) 275-283.

- 75 Lee, G.P., DiClimente, C.C. (1985). Age of onset versus duration of problem drinking on the Alcohol Use Inventory. *Journal of Studies on Alcohol*. 46 (5) 398-402.
- 76 Isenhart, C.E. (1992). Further support for the criterion validity of the Alcohol Use Inventory. *Psychology of Addictive Behaviors*. 4 (2) 77-81.
- 77 Jenkins, D.E. (1993). Construct validations of the Alcohol Use Inventory using an adult outpatient sample: an exploratory factor analysis. *Dissertation Abstracts Int*. 54 (6-B) 3386-3386.
- 78 Wetterling, T., Kanitz, R., Rumpf, H., Hapke, U. and Fischer, D. (1998). Comparison of CAGE and MAST with the alcohol markers (gamma)-GT,ALAT,ASAT and MCV. *Alcohol and Alcoholism*. 33 424-430.
- 79 Girela, E., Villanueva, E., Hernandez-Cueto, C. and Luna, J.D. (1994). Comparison of the CAGE questionnaire versus some biochemical markers in the diagnosis of alcoholism. *Alcohol and Alcoholism*. 29 337-343.
- 80 Murray, M., McMillan, C. (1993). Problem drinking in northern Ireland: results of a community survey using the CAGE questionnaire. *Alcohol and Alcoholism*. 28 477-483.
- 81 Ford, D.E., Klag, M.J., Whelton, P.K., Goldsmith, M. and Levine, D. (1994). Physician knowledge of the CAGE alcohol screening questions and its impact on practice. *Alcohol and Alcoholism*. 29 329-336.
- 82 MacKenzie, D., Langa, A. and Brown, T.M. (1996). Identifying hazardous or harmful alcohol use in medical admissions: a comparison of AUDIT, CAGE, and Brief MAST. *Alcohol and Alcoholism*. 31 591-599.
- 83 Hays, R.D., Hill, L., Gillogly, J.J., Lewis, M.W. and Bell, R.M. (1993). Response times for the CAGE, Short-MAST, AUDIT, and Jellinek alcohol scales. *Behavioral Research Methods, Instruments Computers*. 25 (2) 304-307.
- 84 Neff, R.L., Landrum, J.W., Windham, G.O. and Miles, S.M. (1983). Traffic safety evaluation research review. (Vol. 2, No. 3), U.S. Department of Transportation, NHTSA.
- 85 Wells-Parker, E., Pang, M.G., Anderson, B.J., McMillen, D.L. and Miller, D.I. (1991). Female DUI offenders: A comparison to male counterparts and an examination of the effects of intervention on women's recidivism rates. *J Stud Alcohol*. 52 (2) 142-147.
- 86 Arstein-kerslake, G.W. and Peck, R.C. (1985, Oct.). A typological analysis of California DUI offenders and DUI recidivism correlates. (Research and Development Office CAL-DMV-RSS-85-100), State of California, Department of Motor Vehicles.
- 87 Kincannon, J.C. (1999). MACH: Minnesota Assessment of Chemical Health., International Professional Services, Inc., Chaska, Minnesota.
- 88 Hedlund, J.L., Vieweg, B.W. (1984). The Michigan Alcoholism Screening Test (MAST): A comprehensive review. *J Operational Psychiatry*. 15 55-64.
- 89 Zung, B.J., Charalampous, K.D. (1975). Item analysis of the Michigan Alcoholism Screening Test. *J Stud Alcohol*. 36 (1) 127-132.
- 90 Zung, B.J. (1978). Factor structure of the Michigan Alcoholism Screening Test. *J Stud Alcohol*. 39 (1) 56-67.

- 91 Zung, B.J. (1984). Correlates of the Michigan Alcoholism Screening Test (MAST) among DWI offenders. *J Clin Psychol.* 40 (2) 607-612.
- 92 Tulevski, G.I. Michigan Alcoholism Screening Test (MAST)—its possibilities and shortcomings as a screening device in a pre-selected non-clinical population. *Drug Alcohol Depend.* 24 255-260.
- 93 Otto, K.R., Hall, E.J. (1988). The utility of the Michigan Alcoholism Screening Test in the detection of alcoholics and problem drinkers. *J Pers Assess.* 52 499-505.
- 94 Nochajski, T.H., Miller, B.A., Wiczorek, W.F. and Parks, K.A. (1993). The Utility of Non-Obvious Indicators for Screening of DWI Offenders. Research Society on Alcoholism Annual Meeting . 1-1.
- 95 DiNitto, D.M.; Schwab, A.J. (1991). Substance Abuse Factors Which Interfere With Entry into Employment. University of Texas, Austin.
- 96 Fox, M.R. (1993). MMPI and SASSI scores as predictors of sexual abuse and substance abuse in adolescent psychiatric inpatients. (Unpublished Manuscript) Doctoral Dissertation; University of Toledo.
- 97 Hale, D. (1994). The impact of a substance abuse treatment group on defensiveness in adolescent psychiatric inpatients. (Unpublished Manuscript) Master's Thesis; Southern Connecticut State University, New Haven, CT.
- 98 Lazowski, L.E., Miller, F.G., Boye, M.W. and Miller, G.A. (1998). Efficacy of the Substance Abuse Subtle Screening Inventory-3 (SASSI-3) in identifying substance dependence disorders in clinical settings. *Journal of Personality Assessment.* 71 (1) 114-128.
- 99 Levine, E.B.; Wish, E.D. and Lorion, R.P. (1999). Validation of the Substance Abuse Subtle Screening Inventory (SASSI) with female juvenile detainees. *Drug Testing Technology: Assessment of Field Applications.* Drug testing technology: assessment of field applications, (ed.). CRC Press, Boca Raton, FL.
- 100 Piazza, N.J., Martin, N. and Dildine, R.J. (2000). Screening instruments for alcohol and other drug problems. *Journal of Mental Health Counseling.* 22 (3) 218-227.
- 101 Wagenaar, A.C., Zobeck, T.S., Williams, G.D. and Hingson, R. (1995). Methods used in studies of drink-drive control efforts: A meta-analysis of the literature from 1960 to 1991. *Accid Anal Prev.* 27 (3) 307-316.
- 102 Spiller, B., Rosenberg, H. (1987). Diagnostic practices of evaluators of drunken drivers. *J Subst Abuse Treat.* 4 (2) 79-83.
- 103 Bien, T.H., Miller, W.R. and Tonigan, J.S. (1993). Brief interventions for alcohol problem: A review. *Addiction.* (88) 315-336.
- 104 US Department of Health and Human Services (1997). Ninth special report to the US Congress on alcohol and health. (NIH Pub. No. 97-4017), US Government Printing Office.
- 105 Vingilis, E. (1983). Drinking drivers and alcoholics: Are they from the same population? *Addiction Research Foundation.* 299-342.

- 106 Fine, E.W., Scoles, P. and Mulligan, M. (1975). Under the influence: Characteristics and drinking practices of persons arrested the first time for drunk driving, with treatment implications. *Public Health Rep.* 90 424-429.
- 107 Kramer, A.L. (1991). Judge calls the last shot. *Traffic Safety.* 10-13.
- 108 Wiczorek, W.F., Callahan, C.P. and Miller, B.A. (1994). Victimization experiences and problem drinking among DWI offenders. *Alcohol Clin Exp Res.* 18 (2) 475.
- 109 Pristach, E.A., Nochajski, T.H., Wiczorek, W.F., Miller, B.A. and Greene, B. (1991). Psychiatric symptoms and DWI offenders. *Alcohol Alcohol Suppl.* 1 493-496.
- 110 Lapham, S.C., Smith, E., C'de Baca, J., Chang, I., Skipper, B.J., Baum, G. and Hunt, W.C. (2001). Prevalence of psychiatric disorders among persons convicted of driving while impaired (DWI). *Archives of General Psychiatry-in Press.*
- 111 C'de Baca, J., Lapham, S.C., Paine, S. and Skipper, B.J. (2000). Victim impact panels: Who is sentenced to attend? Does attendance affect recidivism of first-time DWI offenders? *Alcoholism: Clinical and Experimental Research.* 24 (9) 1420-1426.
- 112 Wells-Parker, E., Bangert-Drowns, R., McMillen, R. and Williams, M. (1995). Final results from a meta-analysis of remedial interventions with drink/drive offenders. *Addiction.* 90 907-926.
- 113 Foon, A.E. (1988). The effectiveness of drinking-driving treatment programs: A critical review. *International Journal of Addictions.* 23 (2) 151-174.
- 114 Wells-Parker, E., Popkin, C. (1994). Deterrence and rehabilitation: rehabilitation and screening- research needs for the next decade. *J Traffic Med.* 23 (2) 71-78.
- 115 Wells-Parker, E.; Popkin, C.L. and Ashley, M. (1996). Drinking and driving among women: Gender trends, gender differences. 215-238. National Institutes of Health.
- 116 Miller, G.A. (1999). The SASSI Manual second edition substance abuse measures. The SASSI Institute, Springville, IN.
- 117 McLellan, A.T., Luborsky, L., Cacciola, J., Griffith, J., Evans, F., Barr, H.L. and O'Brien, C.P. (1985). New data from the Addiction Severity Index: reliability and validity in three centers. *J Nerv Ment Dis.* 173 (7) 412-423.
- 118 Graham, R.J. (1990). *MMPI-2 Assessing Personality and Psychopathology.* Oxford University Press.
- 119 Selzer, M.L., Vinokur, A. and van Rooijen, L. (1975). A self-administered short Michigan Alcoholism Screening Test (SMAST). *J Stud Alcohol.* 36 (1) 117-126.
- 120 Filkins, L.; Mortimer, R.; Post, D. and Chapman, M. (1973). *Field Evaluation of Court Procedures for Identifying Problem Drinkers.* National Highway Traffic Safety Administration.
- 121 Webb, G.R., Redman, S., Hennrikus, D. and Sanson-Fisher, R.W. (1992). The reliability and stability of the Mortimer-Filkins test. *J Stud Alcohol.* 53 (6) 561-567.

Funded through the generosity of  *and its
members*

AAA FOUNDATION FOR TRAFFIC SAFETY

1440 New York Avenue, NW, Suite 201, Washington, DC 20005
202-638-5944 • 800-993-7222 • www.aaafoundation.org

©2002
Printed in USA