

* * *

A Validation Study of the DRI-II in a Large Sample of DUI Offenders

Frederick A. Marsteller, Ph.D.
Emory University School of Medicine
Atlanta, Georgia

Donald D. Davignon, Ph.D.
Behavior Data Systems, Ltd.
Phoenix, Arizona

October 14, 1997

Introduction

The Driver Risk Inventory (DRI) is an automated computer scored DUI/DWI offender assessment instrument. The DRI is a test uniquely suited for identifying problem drinkers, substance (alcohol and/or other drugs) abusers and high-risk drivers. The DRI was originally released in 1985 and contained five empirically based measures or scales which included the Truthfulness Scale, Alcohol Scale, Drug Scale, Driver Risk Scale and the Stress Coping Abilities Scale. A sixth measurement (or classification) was recently added to the DRI, called the Substance Dependency Scale, and this new release is called the DRI-II. The purpose of the present study was to validate the new DRI-II.

The new DRI-II Substance Dependency Scale incorporates the seven DSM-IV criteria for substance dependence classification. Also, equivalent items were added to the Alcohol and Drug scales. When a person admits to three or more of the seven DSM-IV criteria for substance dependence they are classified as dependent. The present study also investigated the validity of this new Substance Dependency Scale along with the predictive accuracy of the Alcohol and Drug scales in identifying offenders classified as dependent.

The original Driver Risk Inventory (DRI) was validated in a series of studies (Behavior Data Systems, unpublished) that began in the 1980's which involved primarily test item selection. A 1987 validation study that included 563 DUI offenders demonstrated the relationship between ratings of experienced DUI evaluators and DRI scale scores. DUI evaluators employed their "normal screening procedures" which included test results and an interview before they rated DUI/DWI offenders on the same five scales that are represented in the DRI. Raters had no knowledge of DRI scores. The relationships between staff ratings and DRI scale scores were as follows: Alcohol Scale ($r=.63$, $p<.001$); Drug Scale ($r=.54$, $p<.001$); Driver Risk Scale ($r=.44$, $p<.001$); Truthfulness Scale ($r=.09$, $p<.02$); and Stress Coping Abilities Scale ($r=.02$, n.s.).

Agreements between DRI scales (Alcohol, Drug and Driver Risk) and experienced evaluator ratings were highly significant. The less significant relationship between the Truthfulness Scale and evaluator ratings was not surprising. Without a Truthfulness Scale the evaluator is largely at the mercy of what the DUI/DWI offender says and the evaluator's training and experience. Keistner and Speight (1975) pointed out that drinking drivers tend to minimize alcohol-related problems if test outcomes play a major factor in sentencing. The nonsignificant coefficient between stress coping abilities and evaluator ratings is in marked contrast to earlier studies that showed highly significant relationships between the Stress Coping Abilities Scale and MMPI scales and the Social Readjustment Rating Scale ($r=.40$, $p<.001$). It was hypothesized that the DUI/DWI evaluators were not trained or experienced in evaluating DUI/DWI offenders' "stress coping abilities."

The validity of the DRI was again demonstrated in this 1987 study which showed that the DRI Alcohol Scale was highly correlated with the MAST ($r=.68$, $p<.001$). The study also presented highly significant Cronbach Alpha reliability coefficients for each DRI scale. The reliability coefficients were as follows: Truthfulness Scale=.81, Alcohol Scale=.89, Drug Scale=.74, Driver Risk Scale=.75, and Stress Coping Abilities Scale=.89.

Another DRI validation study (Behavior Data Systems, 1988, $N=1,299$) compared the Alcohol Scale and Drug Scale with the Mortimer-Filkins, MAST, and the MacAndrew. The correlation coefficients between the Mortimer-Filkins and DRI Alcohol Scale ($r=.45$, $p<.001$) and Drug Scale ($r=.24$, $p<.001$) were significant. Similar significant correlation coefficients were found between the MAST and DRI Alcohol Scale ($r=.38$, $p<.001$) and Drug Scale ($r=.20$, $p<.001$). The correlation coefficients between DRI Alcohol Scale ($r=.17$, $p<.02$) and Drug Scale ($r=.17$, $p<.02$) and MacAndrew were lower than those with the MAST and Mortimer-Filkins, yet were significant. Reliability coefficients for all of the DRI scales were again high and nearly identical or higher than those in the 1987 study. These results showed that the DRI is a valid and reliable DUI assessment instrument.

A report by the National Highway Traffic Safety Administration (NHTSA) published in 1988 (DOT HS 807 475) rated the Driver Risk Inventory the best of all the DUI assessment instruments evaluated. The following instruments were reviewed and evaluated: Addiction Severity Index (ASI), Alcohol Use Inventory (AUI), CAGE (Cut Down, Annoyed, Guilty, Eye-opener), Craig Analysis of the Substance Abuse Syndrome (CASAS), Driver Risk Inventory (DRI), Hopkins 20 Question Test, Life Activities Inventory (LAI), MacAndrew MMPI Scale (MAC), Minnesota Assessment of Chemical Health (MACH), Michigan Alcoholism Screening Test (MAST), Modified Criteria -National Council on Alcoholism Diagnosis (MOD-CRIT), Mortimer-Filkins Test (Court procedures for Identifying Problem Drinkers), and Substance Abuse Life Circumstances Evaluation/Automated Drinking Evaluation (SALCE/ADE). The DRI is used in over 37 states and two foreign countries. Over 420,000 DUI assessments are represented in the DRI database (Behavior Data Systems, personal communication, September 26, 1997).

A study by Leshowitz and Meyers (1996) applied decision theory to determine the accuracy of the DRI and a proposed DUI interview instrument (Clayton, et al., 1994). The DRI was found to be far more accurate than the new instrument by a wide margin. Using the data presented by Clayton et al., Leshowitz and Meyers showed that the interview instrument performed at near chance, whereas, the DRI categorized DUI offenders as either "high risk" or "low risk" at an

overall accuracy rate of about 70 percent. This is an interesting finding because the data used in the Leshowitz and Meyers analysis was presented by the authors (Clayton, et al.) of their new interview instrument and those authors were openly hostile to the DRI. Rather than discredit the DRI, Clayton, et al. provided validation of the DRI. Indeed, using their (Clayton, et al.) own data, Leshowitz and Meyers showed that the DRI was the better DUI assessment instrument.

Validation of the DRI-II

In general terms, a test is valid if it measures what it is supposed to measure. The process of confirming this statement is called validating a test. A common practice when validating a test is to compute a correlation between it and another (criterion) test that purports to measure the same thing and that has been previously validated. For the present study, the DRI-II Truthfulness, Alcohol, Drug and Dependency scales were validated with the following respective measures, the Minnesota Multiphasic Personality Inventory (MMPI-2) L-Scale, MMPI-2 MacAndrew Scale (Greene, 1991), Drug Abuse Screening Test (DAST, Skinner, 1982), and a DSM-IV substance use dependency scale devised for this study. The copyrighted material in the MMPI scales was used with permission of the University of Minnesota.

The DRI-II Driver Risk Scale and Stress Coping Abilities Scale were not included in this study because of time constraints involved in testing. These criterion items (or tests) would have increased testing time dramatically. The Driver Risk Scale was changed very little from its DRI scale and the Stress Coping Abilities Scale is identical to its DRI scale. These scales were not the focus of this study and each of these scales has been studied extensively. These practical matters contributed to the decision to limit the focus of this study to the DRI-II Truthfulness Scale, Alcohol Scale, Drug Scale and Dependency Scale.

Methods

For concurrent validity comparisons the following tests were incorporated into a 159 item "criterion test." MMPI-2 L-Scale, MacAndrew, Drug Abuse Screening Test (DAST), MMPI F-Scale, and the DSM-IV substance dependency items. All criterion test items were written in a True/False format. The MMPI-2 F-Scale was included in the criterion test because it indicates a haphazard approach to testing or a wish to put self in a bad light. Florida DUI evaluation agencies deal with the courts and therefore it would be expected that no offender would want to fake bad. In contrast, the MMPI-2 L-Scale detects clients attempting to present an unusually good front (fake good). DUI/DWI research literature consistently demonstrates DUI/DWI offenders attempt to minimize their problems and fake good--particularly in court-related settings. These findings help explain the MMPI-2 F-Scale and L-Scale differences.

Four established Florida certified DUI screening agencies participated and provided a representative sample of Florida DUI offenders. All participating staff were experienced in providing DUI screening services, including administration of the Driver Risk Inventory (DRI), and making DUI program recommendations to their courts. The DRI-II and the criterion test were administered in counterbalanced order to all participants as part of their normal DUI screening procedure. DUI examiners could score DRI-II tests, yet they had no knowledge regarding criterion test results. Criterion test answer sheets were returned by mail, matched with

DRI-II data by name and scored only after all tests were administered. Both DRI-II and criterion tests were scored by the researchers when the data analysis was undertaken--after data gathering was completed.

Population

There were 1,014 DUI offenders included in the present study. There were 811 males (80%) and 203 females (20%). The offenders are broadly defined as Caucasian (83.3%), between the ages of 21 and 40 (65.7%), High School graduate or better (75.2%) and single (49.4%).

Table 1. Blood Alcohol Concentration (BAC) at Time of Arrest

BAC	Males		Females		Total	
	N	Percent	N	Percent	N	Percent
0 - .01	3	0.4	2	1.0	5	0.5
.02 - .07	20	2.5	7	3.4	27	2.7
.08 - .14	198	24.4	53	26.1	251	24.8
.15 - .17	89	11.0	18	8.9	107	10.6
.18 - .19	31	3.8	16	7.9	47	4.6
.20 - .25	80	9.9	14	6.9	94	9.3
Over .25	29	3.6	2	1.0	31	3.1
Not Available	231	28.5	61	30.0	292	28.8
Refused	130	16.0	30	14.8	160	15.8

Table 2. Average Blood Alcohol Concentration (BAC) level as reported by the DUI offender.

Offender Classification	N	Mean BAC
All Offenders	562	0.157
Males	450	0.159
Females	112	0.148
First Offenders	401	0.152
Multiple Offenders	161	0.170
Not Available	292	
Refused	160	

Note: The “Not Available” classification category refers to DUI offenders that either did not remember their BAC or chose not to report it at the time of their DUI evaluation. The “Refused” category includes DUI offenders that refused the BAC test at the time of their arrest.

Table 3. Percent of First and Multiple Offenders by Gender

	Male		Female		Total	
	N	Percent	N	Percent	N	Percent
First Offenders	511	63.0	149	73.4	660	65.1
Multiple Offenders	300	37.0	54	26.6	354	34.9

Note: A Multiple Offender is an offender who reported two or more lifetime DUI’s.

Results

Scale Scores

Two measures were used to assess the agreement among the continuous-score scales used in this study. *Pearson product-moment correlations* measure the extent to which two scores tend to differ from their means by the same *relative* amount. Product moment correlations range from – 1 for exact agreement in opposite directions, to 0 for no agreement, to 1 for exact agreement in the same direction. The *intraclass correlation* measures the proportion of the combined variance of the scores of the two scales which is due to differences among individuals, rather than to differences between the scores on the two scales within individuals. The intraclass correlation ranges from, 0 when all of the variation is between scales, to 1 when the scales are identical. To compute intraclass correlations, all scales were standardized to have a mean of 0 and a standard deviation of 1. This procedure prevents differences in the range of scores for different scales from diluting the intraclass correlation. The intraclass correlations were identical to the product-moment correlations to the second decimal place when the product-moment correlations were positive and the intraclass correlation is undefined when the product-moment correlations were negative. For simplicity, only the product moment correlations are shown in Table 4.

Table 4. Product-moment correlations. All product-moment correlations shown are significant at $p<.001$.										
	DRI-II					Criterion				
	2	3	4	5	6	7	8	9	10	11
1 - DRI-II - Alcohol	.421	-.244	-.301	.356	.599	.291	.508	-.154	.241	.625
2 - DRI-II - Drug	--	.120	-.134	.275	.256	.152	.618	n.s.	.359	.276
3 - DRI-II - Truthfulness		--	.371	-.209	-.324	-.371	-.289	.668	n.s.	-.324
4 - DRI-II -Stress Coping			--	-.240	-.313	-.215	-.220	.323	-.363	-.315
5 - DRI-II - Driver Risk				--	.232	.213	.252	n.s.	.246	.234
6 - DRI-II - Dependency					--	.352	.371	-.251	.229	.964
7 - MacAndrew						--	.383	-.379	.135	.339
8 - DAST							--	-.273	.234	.380
9 - MMPI-L								--	.093	-.255
10 - MMPI-F									--	.232
11 - DSM-IV										--

Categorical Ratings

It is often desirable to simplify the use of the assessment scales by providing cutoff scores, above which a problem is deemed to be present. Each of the alcohol and drug scales in this study have such cutoffs defined based on previous research. The strength of association between the categorical outcomes can be assessed in several ways. The simplest is just the percent

agreement. This measure actually overestimates the extent of agreement because it includes the agreement which would occur if one measure were categorizing the outcome at random. A widely used measure of categorical agreement is the *kappa* coefficient (see e.g. Dunn, 1989). *Kappa* estimates the strength of agreement excluding that expected due to chance. There are several recommended “benchmarks” for assessing the strength of agreement using *kappa*. Those of Landis and Koch (1977) are as follows:

<u>Kappa</u>	<u>Strength of Agreement</u>
0.00	Poor
0.01-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost perfect

The following cutoff scores, defined as the score above which a problem is present were used in this analysis and the percentage of respondents scoring above the cutoff are as follows:

Scale	Cutoff score	Reference	% Positive
DRI-II Alcohol Scale - Problem	12	Behavior Data Systems, Ltd.	27.7
DRI-II Alcohol Scale - Severe	27	Behavior Data Systems, Ltd.	10.9
DRI-II Drug Scale - Problem	4	Behavior Data Systems, Ltd.	33.1
DRI-II Drug Scale – Severe	8	Behavior Data Systems, Ltd.	10.4
DRI-II Substance Dependency Scale	Categorical	Behavior Data Systems, Ltd.	25.9
MacAndrew Alcoholism Scale	23	Greene, 1991	25.0
DAST	7	Staley and El-Guebaly, 1990	11.4
DSM-IV Dependence	Categorical	APA, 1996	8.5
DSM-III-R Dependence	Categorical	APA, 1989	14.6

Truthfulness Scale

The MMPI-2 L-Scale assesses whether respondents are attempting to present an unusually good appearance by denying even the most minor personal flaws. In a similar vein, the DRI-II Truthfulness Scale is designed to detect denial, minimization of problems and reveal “faking good.” In a DUI test setting the constructs of minimization and denial are important factors that if not measured often hinder accurate assessment. The correlation between the DRI-II Truthfulness Scale and the MMPI-2 L-Scale is highly significant ($r=.668$, $p<.001$) and in the expected positive direction. It is rare to find correlation coefficients in validation testing above

.60. Usually they are much lower. These results support the validity of the DRI-II Truthfulness Scale. They also indicate that the DRI-II Truthfulness Scale and the MMPI-2 L-Scale measure essentially the same attitudes and behaviors. In other words, the DRI-II Truthfulness Scale measures what it is designed to measure, i.e., problem minimization and “faking good.”

Alcohol Scale

The MacAndrew Alcoholism Scale (MacAndrew, 1965) was derived from the MMPI as a measure of alcoholism. The MacAndrew Scale used in this study is the revised version applicable to the current version of the MMPI, the MMPI-2. MacAndrew Scale items were selected because, as a group, they successfully discriminated alcoholics from non-alcoholics in validation samples. The MacAndrew scale items have little face validity with respect to alcohol use, with only one item referring directly to alcohol. The opinion of researchers using the MacAndrew scale is that it reflects both a) personal attitudes which represent a risk of alcohol and drug problems, and b) behaviors and symptoms which are common among alcoholics. The DRI-II Alcohol Scale measures alcohol use and identifies alcohol-related problems. DRI-II Alcohol Scale items specifically refer to alcohol use and alcohol-related symptoms. The DRI-II Alcohol Scale correlates significantly with the MacAndrew Scale ($r=.291$, $p<.001$), in the predicted direction.

The *kappa* coefficients of the Alcohol Scale with the MacAndrew Scale at both problem and severe cutoffs are rather small (.248 and .166, respectively). The two dependence scales used in this study also support the validity of the Alcohol Scale. The Alcohol Scale had a correlation of .599 with DRI-II Substance Dependency Scale and *kappas* of .699 and .478 for the problem and severe cutoffs, respectively. The correlation with the sum of the DSM criterion items was .625 and the *kappas* were .320 and .414 for the problem and severe cutoffs for DSM-IV “dependence” and .450 and .460 for DSM-III-R “dependence”.

These results support the concurrent validity of the DRI-II Alcohol Scale. In other words, the Alcohol Scale demonstrates a statistically significant association with other recognized measures of alcohol problems.

The relatively small correlation coefficient with the MacAndrew Scale may reflect several differences between the scales. The MacAndrew Scale was developed to detect alcoholism per se. Its items are generally not directly related to alcohol use and alcohol-related problems, but refer instead to secondary symptoms and characteristics which have successfully discriminated alcoholics from non-alcoholics in clinical validation samples. The MacAndrew Scale was also devised to identify alcoholism among White males (Greene, 1991) and females and ethnic minorities have been shown to respond differently from White males.

The items in the DRI-II and criterion dependency scales represent major physical and social problems associated with alcohol and drug use and are designed to identify individuals in need of clinical attention. They are, therefore, designed to assess substance related problems at the upper end of severity. Because of its designed application, the DRI-II Alcohol Scale is designed to assess alcohol use across a full spectrum from minimal risk through severe dependence. It is not surprising that the Alcohol Scale has stronger *kappa* coefficients at the higher cutoff score.

The DRI-II Alcohol Scale, on the other hand, is very direct in asking about alcohol use and alcohol-use related symptoms. It is also designed to assess alcohol-related problems across a

broad range of severity, not just differentiate alcoholics from non-alcoholics. Furthermore, the DRI-II Alcohol Scale incorporates truth-correction, whereas non-DRI-II scales do not.

Drug Scale

The DAST is a drug abuse questionnaire that directly refers to drug use and abuse. It was designed to screen clinical populations for significant drug abuse problems. The DRI-II Drug Scale measures drug (marijuana, crack, cocaine, barbiturates, amphetamines, heroin) use and abuse problems. The DRI-II Drug Scale correlates significantly with the DAST ($r=.618$, $p<.001$), in the predicted direction. The substantial *kappa* for the association between the Drug Scale and the DAST at the higher Drug Scale cutoff (.681) compared to the small *kappa* (.286) at the lower cutoff again may reflect a difference in orientation between the scales, with the DRI-II Drug Scale providing assessment across the full spectrum, while the DAST focuses on major problems or extreme cases.

These results support the validity of the DRI-II Drug Scale. The DRI-II Drug Scale accurately measures illicit drug use and abuse. Again, the truth-corrected scores of the DRI-II Drug Scale may reduce the correlation with the DAST which is not truth-corrected.

Dependency Scale

DSM Classification of Substance Use Disorders

Substance dependence as operationalized in both the DRI-II Substance Dependency Scale and the DSM-IV Criterion scale is based on the DSM-IV criteria, which are presented in Table 5 below. Offenders who answer positively to items reflecting 3 or more DSM-IV criteria (test questions 146 through 155) are classified as having “*substance dependence disorder*”. Those who are not classified as dependent are classified as having a “*substance abuse disorder*” if they answer positively to any of the four substance abuse items (156-159). Although the clinical terms “abuse” and “dependence” are used in this discussion actual diagnosis can only be made by a qualified clinician based on a face-to-face interview. What are presented here are *screening classifications* and they are presented in quotes to avoid confusion.

The DSM-IV symptom most commonly endorsed is Symptom 4: a persistent desire to stop or reduce use or repeated attempts to stop or reduce use. This was endorsed by 36.1% of respondents. Reporting of Symptom 2 -- withdrawal symptoms or use to relieve or prevent withdrawal symptoms -- and Symptom 6 -- reducing other important activities to use -- were each reported by fewer than 5% of respondents. Using these criteria, 8.5% of respondents were classified as “dependent” using the DSM-IV criterion test.

Prior to the recent introduction of the DSM-IV, the standard for diagnosing substance use disorders was the DSM-III-R. Although the criteria under the two systems are similar, there is a very substantial difference in the definitions of dependence and abuse between the two. In DSM-III-R, a classification of dependence required that any three of nine symptoms be present during the past year. Abuse was a residual diagnosis based on continued use despite a wide range of problems which might be exacerbated by use or repeated use when hazardous. The DSM-IV dependence criteria depend more strongly on either physical dependence or loss of control over use than did the DSM-III-R criteria. In theory, using DSM-IV criteria will reduce the estimated prevalence of *substance dependence* and increase the estimated prevalence of

substance abuse diagnoses, due to the more stringent criteria for dependence. Almost all available estimates of the prevalence of alcohol dependence both in the general population and DWI populations are based on the DSM-III-R criteria. Accordingly, it may be useful to examine a recoding of the DSM-IV criterion items to DSM-III-R standards (Table 6).

Using DSM-III-R criteria increases the number of dependence “classifications” at the expense of abuse “classifications” and classifies slightly fewer individuals as having neither “classification.” The *kappa* coefficient for the two scorings (.703) indicates that although the agreement between the two is very substantial, they are not assessing exactly the same thing.

Table 5. DSM-IV Symptom definitions, diagnostic criteria and positive response rates.

DSM-IV Symptom			% Positive
Dependence Symptoms	1	Tolerance (146 or 147)	5.8
	2	Withdrawal or use to prevent withdrawal (148 or 149)	4.1
	3	Inability to control use (150)	11.5
	4	Desire or attempts to quit or reduce use (151 or 152)	36.1
	5	A great deal of time spent getting, using, recovering (153)	7.9
	6	Reduced other important activities to use (154)	2.8
	7	Continued use despite serious problems (155)	11.2
Dependence "Classification" (three or more of the above)			8.5
Dependence "Classification" w/ physical symptoms (I or II)			5.6
Abuse Symptoms	1	Continued use despite social problems (156)	9.4
	2	Continued use despite role impairment (157)	6.0
	3	Repeated use when dangerous (158)	25.1
	4	Continued use despite legal problems (159)	15.7
Abuse "Classification" (one of the above and not dependent)			27.5

Table 6. DSM-III-R symptom definitions, classification criteria and positive response rates.

DSM-III-R Symptoms			% Positive
Dependence Symptoms	1	Inability to control use (150)	11.5
	2	Desire or attempts to quit or reduce use (151 or 152)	36.1
	3	A great deal of time spent getting, using, recovering (153)	7.9
	4	Impaired role functioning or use when dangerous. (157 or 158)	27.1
	5	Reduced other important activities to use (154)	2.8
	6	Use despite recurrent psych., social or phys. problems (155, 156)	16.8
	7	Tolerance (146 or 147)	5.8
	8	Characteristic withdrawal symptoms (148)	3.4
	9	Use to relieve or avoid withdrawal (149)	2.5
Dependence "Classification" (three or more symptoms)			14.6
Abuse "Classification" (not dependent and item 4 or item 6)			18.6

It should also be noted that the estimated proportion of this DUI offender population which is classified as DSM-IV “dependent” (8.5%) or DSM-III-R “dependent” (14.6%) is very low with respect to other estimates, which tend to be in the range of 9-12% for the general population (e.g. Kessler, et al, 1995) and 40-60% in DUI populations. All published estimates of the prevalence

of alcohol dependence among DWI offenders to date were based on DSM-III-R criteria, rather than the recently implemented DSM-IV criteria. The prevalence of DSM-III-R alcohol dependence has been reported to be as low as 23.9% (Veneziano and Veneziano, 1992) and as high as 74% (Wieczorek et al, 1990).

In this testing environment, when respondents perceive that they will gain (not be referred for treatment) by underreporting problems, the validity of item responses may be limited. For example, 5.8% of respondents respond positively to at least one of the two items addressing tolerance in the DSM-IV criterion questions. However, 14% of respondents reported BACs of .18 or greater and 25.3% of those who reported non-missing BAC values reported BAC of .18 or greater. This apparent several-fold underestimation of this symptom is a cause for concern. Use of the BAC to correct (i.e. including any reported BAC of .18 or greater as positive for tolerance) the tolerance symptom score increases the proportion positive to 21.2% and the proportion with a dependence “classification” to 11.7% (a 37.7% increase). This result indicates that a Substance Dependence Scale used in this testing environment should 1) incorporate all available information, not just responses to direct questions, and 2) would likely benefit from correction based on a truthfulness scale such as that in the DRI-II.

The seven DSM-IV items were reworded (along with equivalent alcohol and drug items) to create the DRI-II Substance Dependency Scale. Consequently, DSM-IV substance dependency items were compared with DRI-II Dependency Scale items. There was a high positive correlation between the DRI-II Substance Dependency Scale and the DSM-IV Criterion items ($r=.964$, $p<.001$). This high correlation reflects their very strong overlap. The DRI-II Substance Dependency Scale found a larger fraction of “dependent” subjects than either the criterion scale did with either the DSM-III-R or DSM-IV coding. Its *kappas* with the other alcohol and drug scales were higher than those of the alcohol and drug scales with the criterion “classifications” using either coding. The relatively small *kappas* for the association between the DRI-II Substance Dependency Scale dependency classification and the two DSM criterion-based ratings may be due to the *better* detection rate of the DRI-II Substance Dependency Scale. This finding supports the validity of the DRI-II Substance Dependency Scale. In other words, clients answer DSM-IV substance dependency criteria items in the same way they answer DRI-II Dependency Scale items (and their equivalents).

In addition, the DRI-II Alcohol Scale is significantly correlated ($r=.625$, $p<.001$) with the DSM-IV substance dependency criterion (7 classification items). This correlation was predicted because the DSM-IV substance dependency items refer to alcohol symptoms. “The essential feature of Substance Dependence is a cluster of cognitive, behavioral, and physiological symptoms indicating that the individual continues use of the substance despite significant substance-related problems” (p. 176, DSM-IV). What this means is that the DRI-II Alcohol Scale and DSM-IV substance dependency criteria are to a large extent measuring the same thing. Specifically, the DRI-II Alcohol Scale is a measure of the severity of alcohol abuse. This is true for the DRI-II Drug Scale as well. The DRI-II Alcohol and Drug scales measure the severity of substance abuse. The DSM-IV Substance Dependency criteria refer to both alcohol and/or drugs. Consequently, the DRI-II Drug Scale also correlates significantly ($r=.276$, $p<.001$), in the expected direction, with the DSM-IV Substance Dependency criteria. However, a less significant correlation coefficient is demonstrated for the DRI-II Drug Scale ($r=.276$, $p<.001$) than the DRI-II Alcohol Scale ($r=.625$, $p<.001$).

The very high correlation between the DRI-II Alcohol Scale and DSM-IV criteria is in contrast to the much lower correlation between MacAndrew and DSM-IV ($r=.339$, $p<.001$). It is important to note that the term “substance” is more generic in application than the terms “alcohol” or “drugs.” The MacAndrew is referred to as an alcohol scale. As noted earlier, of the 46 MacAndrew items only one of them refers to alcohol, and none refer to drugs. In 1988 the MacAndrew scale was shown to correlate with both the DRI-II Alcohol and Drug Scales at the $p<.02$ level of significance. In the present study the MacAndrew Scale also significantly correlates ($r=.152$, $p<.001$) with the DRI-II Drug Scale. In retrospect the MacAndrew scale may be more of a generic substance (alcohol and drugs) use scale. A similar logic may apply to the DAST which significantly correlates ($r=.508$, $p<.001$) with the DRI-II Alcohol Scale.

Discriminant Validity

To assess the ability of the different scales used in this study to distinguish among subjects rated as “no classification”, “substance abuse” or “substance dependent” based on the criterion scale, ANOVAs comparing the mean scores for each scale among the “classification” groups were computed. Post hoc comparisons among the groups were made using Tukey’s Least-Significant Difference test.

The question addressed here is whether the different scales used in this study can discriminate among the “classification” groups. Keep in mind that the groups are established on the basis of self-report responses to DSM-IV criteria on the criterion test.

Table 7. Mean scale scores for subjects identified as having “no classification” (None) or “classifications” of substance abuse or dependence based on DSM-IV criterion and responses to the criterion scale items.				
	None	Abuse	Depend.	Significant Differences
DRI-II Alcohol	9.4	12.5	28.7	None<Abuse<Dependence
MacAndrew	20.2	21.7	24.0	None<Abuse<Dependence
DRI-II Drug	4.1	3.8	8.5	None=Abuse<Dependence
DAST	3.4	4.1	7.2	None<Abuse<Dependence
DRI-II Truthfulness	12.7	9.1	8.1	None>Abuse=Dependence
MMPI L-Scale	7.3	5.7	5.0	None>Abuse=Dependence
MMPI F-Scale	3.5	3.7	6.7	None=Abuse<Dependence
DRI-II Driver Risk	7.4	8.9	10.8	None<Abuse<Dependence
DRI-II Stress Coping	152.7	135.0	107.8	None>Abuse>Dependence

An ANOVA comparison among the “No Classification”, “Abuse” and “Dependence” groups found that for each scale, the “classification” groups were very significantly different (all p ’s $<.0001$). It is noteworthy that for the DRI-II Alcohol Scale, the differences among the “classification” groups are larger than those for the MacAndrew scale. This is reflected by the much larger R^2 for the Alcohol Scale (23.5%) than for the MacAndrew (8.3%). The R^2 is a measure of the proportion of the total variance in scale scores that are attributable to group differences. This finding supports the conclusion that the DRI-II Alcohol Scale accurately discriminates between “classification” categories and does so better than the MacAndrew. A similar comparison of the DRI-II Drug Scale and the DAST shows more comparable differences, with the $R^2=5.2\%$ for the Drug Scale and 8.5% for the DAST. The relatively small R^2 values probably reflect the fact that this sample was predominantly referred from problem use of

alcohol (i.e. getting a DUI). The DRI-II Truthfulness, Driver Risk and Stress Coping Abilities Scales are also all significantly associated with “classification,” supporting the validity of their contribution to the screening efficacy of the DRI-II.

A t-test comparison between First Offenders and Multiple Offenders (2 or more DUI's) with the DRI-II Alcohol Scale reveals that the Alcohol Scale scores of Multiple Offenders are significantly higher than First Offenders scores ($t=9.51$, $p<.001$). Similar findings are consistently found when First Offender's Blood Alcohol Concentrations (BAC) are compared to Multiple Offenders BAC. First Offenders BAC are consistently lower, on average, than Multiple Offenders BAC. The t-test comparison between these offender groups on the MacAndrew again showed a less significant difference ($t=2.77$, $p<.006$) than the DRI-II Alcohol Scale.

Summary of Validating the DRI-II

Each of the DRI-II scales (Truthfulness, Alcohol, Drugs and Dependency) correlate highly significantly with their respective criterion tests. These large correlation coefficients support the validity of the DRI-II. Indeed, these significant correlations provide strong support for the validity of the DRI-II scales.

Those scales which are most like the DRI-II associated scales (L-Scale, DAST and DSM-IV) have very high correlation coefficients and provide very strong support for the validity of the DRI-II scales. The MacAndrew Scale is not a direct measure of alcohol use and problems; it is instead a heterogeneous assemblage of items associated with either risk of alcoholism or with heavy alcohol use. In contrast, the DRI-II Alcohol Scale is specific to alcohol use and alcohol-related problems. The DRI-II Alcohol Scale measures alcohol-related problems and proneness to alcohol abuse. The DRI-II Alcohol Scale includes severity of abuse or alcoholism. The severity of abuse DRI-II Alcohol Scale items appear to be related to some MacAndrew items. Therefore, it would be expected that the MacAndrew and the DRI-II Alcohol Scale would not show a highly significant relationship. The correlation is, nonetheless, significant.

DRI-II Reliability

Reliability in testing refers to a test's stability or consistency. Test reliability refers to the consistency of scores obtained by the same person when retested with the same or equivalent test. In most testing environments a reliability coefficient of .80 or higher is accepted as satisfying reliability standards. All of the DRI-II scales exceed this standard. The weakest reliability is demonstrated in the DSM-IV Substance Dependency classification scale that consists of the seven DSM-IV dependence criterion items in reformatted or reworded format. These items were reworded for many reasons, including their reading levels and ease of understanding. As explained earlier, the Substance Dependency Scale is a classification (not a measurement) scale as presented in the DSM-IV.

In the present study, Cronbach's Coefficient Alpha (an important index of reliability and internal consistency) was computed on the sample of DUI offenders' responses to scale items. The following table summarizes the results of this analysis. It should be noted that all six DRI-II scales are included in this reliability (internal consistency) analysis. These six DRI-II scales include: Truthfulness Scale, Alcohol Scale, Driver Risk Scale, Drug Scale, Substance Dependency Scale and the Stress Coping Abilities Scale.

DRI-II Scales Internal Consistency (N=1,014, 1997)

<u>DRI-II Scale</u>	<u>Cronbach's</u> <u>Alpha</u>	<u>Significance</u> <u>Level</u>
Truthfulness Scale	.87	p<.001
Alcohol Scale	.93	p<.001
Driver Risk Scale	.83	p<.001
Drug Scale	.87	p<.001
Substance Dependency Scale*	.81*	p<.001
Stress Coping Abilities Scale	.92	p<.001

* This is a DSM-IV classification (not a measurement) scale.

All other DRI-II scales demonstrate very large Cronbach Alpha reliability coefficients. These results strongly support the internal consistency of the DRI-II scales. The Alcohol Scale demonstrates very high internal consistency, which again raises questions regarding the statistical properties of the MacAndrew Scale. It is reasonable to conclude that the DRI-II (and the scales contained therein) is a reliable assessment instrument or test.

For comparison purposes, Cronbach Alpha reliability coefficients were also obtained for the criterion test scales. These results were as follows: L-Scale=.72, MacAndrew=.56, DAST=.85, DSM-IV items=.81. Only the DAST and DSM-IV exceeds accepted reliability standards. Those criterion test scales with reliability coefficients of .7 or above show high correlation coefficients with their respective DRI-II scales. The MacAndrew Scale has a low reliability coefficient and its correlation with the DRI-II Alcohol Scale is relatively low (yet significant). Again, the MacAndrew Scale is shown to lack many acceptable statistical properties for assessment of DUI offenders. A test that has weak reliability usually has weak validity. Criterion test reliability coefficients provide additional insight into this study's findings. For example, criterion tests having good reliability (.80 or higher) coefficients could show even more substantial DRI-II scale correlation coefficients. Unfortunately, the MacAndrew Scale has a low reliability coefficient.

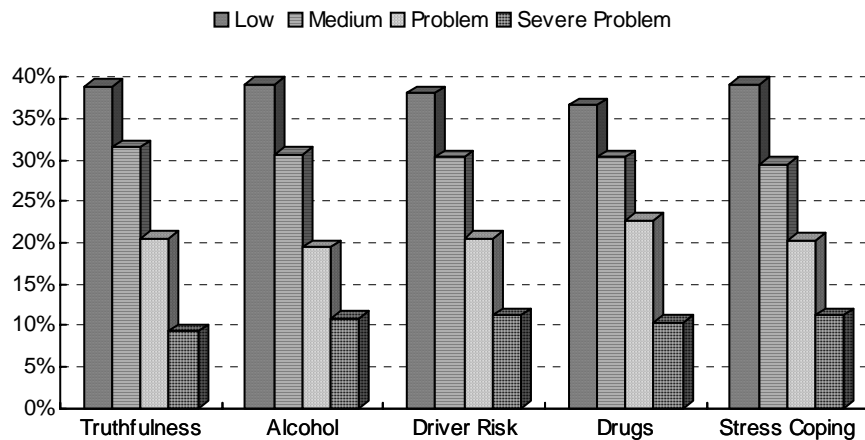
DRI-II Accuracy

DRI-II accuracy is determined by the closeness of obtained scale risk range percentages to predicted percentages. DRI-II predicted risk range percentages are presented in the table below. The actual or "obtained" percentage of offenders that scored in each scale's risk range are presented in the graph and table below. Obtained risk range percentages are based on offenders' scale scores which are comprised of test item totals for the scale with truth-correction factored in, then converted to a percentile score.

PREDICTED RISK RANGE PERCENTAGES FOR EACH DRI-II SCALE		
RISK CATEGORY	RISK RANGE	PREDICTED PERCENTAGE
Low Risk	zero to 39th percentile	39%
Medium Risk	40 to 69th percentile	30%
Problem Risk	70 to 89th percentile	20%
Severe Problem	90 to 100th percentile	11%

The results show that all of the DRI-II obtained risk range percentages were within 2.7 percent of the predicted percentages. There are only two instances where the differences between obtained and predicted scores are more than two percentage points. These results show that the DRI-II accurately classified severity or risk in this sample of DUI offenders.

DRI-II Scales Risk Ranges



	Truthfulness	Alcohol	Driver Risk	Drugs	Stress Coping	Predicted
Risk Range	%	%	%	%	%	%
Low	38.8	39.1	38.0	36.6	39.1	39%
Medium	31.5	30.7	30.3	30.3	29.5	30%
Problem	20.5	19.4	20.5	22.7	20.2	20%
Severe Problem	9.2	10.8	11.2	10.4	11.2	11%

The percentage of offenders falling into each risk range for each of the five scored Driver Risk Inventory-II scales is presented for the DUI offenders included in this study. It was noted earlier that the Substance Dependency Scale is a classification and not a measurement scale. Results demonstrate the accuracy of the DRI-II.

Summary

DRI-II validity was examined in this study of 1,014 DUI offenders presenting for mandatory alcohol and drug evaluation at four Florida DUI agencies. The DRI-II Alcohol, Drug and Truthfulness scales were compared to the MacAndrew Alcoholism scale of the MMPI-2 (MAC-R), the Drug Abuse Screening Test (DAST) and the L-Scale of the MMPI-2. The DRI-II

Truthfulness Scale significantly correlated with the MMPI L-Scale ($r=.668, p<.001$). The DRI-II Alcohol Scale significantly correlated with the MAC-R ($r=.291, p<.001$). The DRI-II Drug Scale significantly correlated with the DAST ($r=.618, p<.001$). The new DRI-II Substance Dependency Scale very significantly correlated with DSM-IV substance dependency criterion items which were developed for this study ($r=.964, p<.001$). Criterion validity is a measure based upon a test's correspondence with another established measure (or test) of the same thing (criterion or variable). These results support the validity of the DRI-II.

Very large Cronbach Alpha reliability coefficients were found for the DRI-II scales: Truthfulness Scale ($.87, p<.001$), Alcohol Scale ($.93, p<.001$), Drug Scale ($.87, p<.001$), Driver Risk Scale ($.83, p<.001$), Substance Dependency Scale ($.81, p<.001$), and Stress Coping Abilities Scale ($.92, p<.001$). This study strongly supports the internal consistency or reliability of the DRI-II.

Comparison between offenders DSM-IV "classifications" based on responses to the criterion items developed for this study shows that the DRI-II scales have very high discriminant validity. Analyses of variance comparing offenders classified as "no classification" or no diagnosis, "substance abuse" or "substance dependence" were highly significant. Mean scores of all five DRI-II scales differed significantly among the "classification" groups and showed patterns of differences that not only support their individual discriminant validity, but also demonstrated their strength as a group.

The original intent with the new Dependency Scale was to classify DUI offenders as dependent or nondependent. The DSM-IV also contains criteria for classifying substance abuse. Admission to one of the four substance abuse criteria classifies an individual as substance abuse. Because an offender may not meet the criteria for dependence but may still meet the criteria for substance abuse, it was decided to include the DSM-IV criteria for substance abuse in the DRI-II. Again, the DSM-IV criteria were reformatted and included in the DRI-II along with additional equivalent alcohol and drug items. The DSM-IV criteria for both substance dependence and substance abuse are now represented in this scale that has been renamed the "Substance Dependency/Abuse Scale". Consequently, in the future this DSM-IV based scale is called the Substance Dependency/Abuse Scale. DSM-IV symptoms were substantially underreported. DUI offender court history and other information show the offenders' minimized their problems on the DSM-IV items of the criterion test. And these findings support future modifications of the scale. Such modifications will include 1) using all available information (e.g. BAC) in assessing each symptom, and 2) adjusting classifications with the Truthfulness Scale to avoid minimization of problems or underreporting.

The DRI-II measures DUI offender alcohol and drug severity levels, risk or proneness toward problems. Severity level (risk) is assigned on the basis of scale scores and is divided into four risk ranges (low, medium, problem and severe problem). Risk range percentages are as follows: low risk, 0 to 39 percent; medium risk, 40 to 69 percent; problem risk, 70 to 89 percent; severe problem risk, 90 percent and above. The accuracy or closeness of obtained scores with predicted scores was demonstrated for each DRI-II scale. These risk range percentages are predicted and may be adjusted to comply with each states DUI/DWI program. Scale score-related recommendations may also be adjusted for each state's statutes and DUI/DWI program.

The results of this study show that the DRI-II accurately classifies alcohol/drug severity or offender (alcohol and/or drugs) risk. Using the above risk range percentages, the DRI-II

accurately predicts offender risk to within two percentage points for nearly all scales and all risk ranges. The DRI-II has very high predictive validity and accuracy.

Very good DSM-IV dependency classification was demonstrated. Yet utilization of DSM-III-R classification criteria resulted in even more significant results. Differences between DSM-III-R and DSM-IV dependency criteria were discussed. Similarly, some criterion measures were found to be lacking. Substitution of more reliable and valid alcohol criterion measures would likely result in even more substantial results. In conclusion, the DRI-II is a reliable, valid and accurate DUI/DWI offender assessment instrument or test.

References

- Behavior Data Systems. DRI: An Overview of Scientific Findings. Unpublished manuscript available upon request.
- Clayton, R.R., Huffman, M., Wester, S., & Caywood, M. (1994). Final Report Assessment of DUI Offenders in Fayette County: An evaluation using random assignment. *University of Kentucky, Center for Prevention Research*.
- Dunn, G. (1989). *Design and Analysis of Reliability Studies*. London: Edward Arnold.
- Greene, R.L. (1991). *The MMPI-2/MMPI: An Interpretive Manual*. Boston: Allyn and Bacon.
- Keistner, and Speight, L. (1975). Oregon Multivariate Prediction of Drinking Drivers. *Oregon Department of Transportation, Motor Vehicles Division*.
- Kessler R.C., McGonagle K.A., Zhao S., Nelson C.B., Hughes M., Eshleman S., Wittchen H.U., Kendler K.S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry* **51**(1):8-19.
- Landis, J.R., Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics* **33**, 159-174.
- Leshowitz, B. & Meyers, J.M. (1996). Application of Decision Theory to DUI Assessment. *Alcoholism: Clinical & Experimental Research* **20**:1148-1152.
- MacAndrew, C. (1965). The differentiation of male alcoholic outpatients from nonalcoholic psychiatric outpatients by means of the MMPI. *Quarterly Journal of Studies on Alcohol*, **42**, 604-625.
- National Highway Traffic Safety Administration (NHTSA), Washington, D.C., (1988). Assessment of Classification Instruments Designed to Detect Alcohol Abuse, DOT HS 807 475. Authors: C.L. Popkin, C.H. Kannenberg, J.H. Lacey, and P.F. Waller. Sponsoring agency: U.S. Department of Transportation.
- Staley, D. and El-Guebaly, N. (1990). Psychometric properties of the Drug Abuse Screening Test in a psychiatric patient population. *Addictive Behaviors*, **15**, 257-264.
- Veneziano, C., Veneziano, L. (1992) Psychosocial characteristics of persons convicted of driving while intoxicated. *Psychological Reports*. **70**: 1123-1130.
- Wieczorek, W.F., Miller, B.A., Nochajski, T.H. (1990). Alcohol diagnoses among DWI offenders. *The Problem-Drinker Driver Project*. Research Note 90-6. Albany: New York State Division of Alcoholism and Alcohol Abuse.

* * *