The LSI-R and the Compas: Validation Data on Two Risk-Needs Tools

Tracy L. Fass, Kirk Heilbrun, David DeMatteo and Ralph Fretz

Criminal Justice and Behavior 2008; 35; 1095 originally published online Jul 8, 2008;
DOI: 10.1177/0093854808320497

THE LSI-R AND THE COMPAS

Validation Data on Two Risk-Needs Tools

TRACY L. FASS
Drexel University and Villanova Law School

KIRK HEILBRUN
DAVID DEMATTEO
Drexel University

RALPH FRETZ
Community Education Centers, Inc.

Over the past two decades, the role of risk-needs assessment in the criminal justice system has increased substantially. This study provides validation data on the Level of Service Inventory–Revised (LSI-R) and the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) using a large male cohort ($N = 975$) with a substantial proportion of ethnic minority offenders. In comparing the predictive validity of these tools, the authors employed a retrospective, archival, known-groups design to study outcomes of offenders released into the community from New Jersey prisons between 1999 and 2002, with a postrelease outcome period of 12 months. The results indicate that both the LSI-R composite score and the COMPAS recidivism score have inconsistent validity when tested on different ethnic/racial populations. Furthermore, the results suggest that different ethnic/racial groups have varying risk and needs factors that predict recidivism.

Keywords: risk; recidivism; risk-needs tools; LSI-R; COMPAS

Actuarial risk-needs assessment tools have become increasingly important in the field of criminal justice (Andrews & Bonta, 2003; Holsinger, Lowenkamp, & Latessa, 2006). Previously, the risk assessment of offenders incorporated professional judgment to a large extent (Andrews & Bonta, 2003; Holsinger et al., 2006). Over the past two decades, however, researchers have developed a number of empirically based risk assessment and classification tools (Andrews & Bonta, 2003). These tools are used with increasing frequency and have become an integral part of many correctional interventions (Holsinger et al., 2006; Whiteacre, 2006).

Bonta (1996) described three generations of risk assessment instruments. The first generation involved assessment conducted through an unstructured or semistructured interview, based largely on the assessor's experience and qualitative observations, which may involve gathering information with regard to relevant criminogenic variables (Holsinger et al., 2006). Second-generation assessments are empirically based risk instruments that use risk factors empirically related to future antisocial behavior. However, these assessments use primarily static risk factors and yield little information on rehabilitation needs. Third-generation assessments are also empirically based but incorporate both static and dynamic

AUTHORS' NOTE: *Please address correspondence to Kirk Heilbrun, Department of Psychology, Drexel University, MS 626, 245 N. 15th Street, Philadelphia, PA 19102; e-mail: kirk.heilbrun@drexel.edu.*

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 35 No. 9, September 2008 1095-1108
DOI: 10.1177/0093854808320497

risk factors related to future antisocial behavior. Through ongoing assessment, third-generation tools can reflect risk-relevant change in the person and his or her situation.

More recently, researchers have described fourth-generation risk assessment tools (Andrews & Bonta, 2003). These tools are designed to follow the offender from intake through case closure (Andrews & Bonta, 2003; Andrews, Bonta, & Wormith, 2006). Such fourth-generation tools aim to further the principles of effective treatment and help design supervision that can protect society from recidivistic crime (Andrews et al., 2006).

There is a large and growing literature with regard to risk factors that are related to antisocial behavior. A key distinction found in the literature is between "static" and "dynamic" risk factors. Static risk factors are typically historical, unlikely to change, and not amenable to intervention efforts; dynamic factors, by contrast, may change over time (Andrews & Bonta, 2003). In recent years, researchers have focused a great deal of attention on identifying static and dynamic factors that are predictive of violence and other forms of antisocial behavior (see Conroy & Murrie, 2007, for a review). A review of the literature reveals several examples of empirically established static risk factors, including a history of violent and antisocial behavior, presence of psychopathy, age (with younger individuals being more at risk), and previous substance use. Although there is considerably less research on dynamic risk factors, recent research has identified several promising dynamic risk factors. In a recent comprehensive review of the relevant literature, Douglas and Skeem (2005) identified the following seven robust dynamic risk factors: impulsiveness, negative affect, psychosis, antisocial attitudes, current substance use, interpersonal relationship problems, and poor treatment compliance.

The MacArthur Violence Risk Assessment Study (Monahan et al., 2001) also provided valuable information with regard to risk factors that are most predictive of violent behavior. Monahan and colleagues studied more than 1,100 admissions to acute inpatient psychiatric facilities in three different cities and then conducted follow-up interviews every 10 weeks for 1 year after discharge. Out of the 134 risk factors examined in the study, psychopathy was more strongly associated with violence in the community after discharge than any other risk factor. Other robust risk factors found in the study included prior violence, a co-occurring diagnosis of substance abuse or dependence, and anger.

Other researchers have also examined the relationship between risk factors and antisocial behavior. Andrews and Bonta (2003) identified the most important of these factors and labeled them "the Big Four": antisocial attitudes, antisocial associates, antisocial behavioral history, and antisocial personality. They further described "the Big Eight" risk factors that are relevant to criminal behavior. The Big Eight include the Big Four plus problems at home, problems at school or work, problems in leisure circumstances, and problems with substance abuse (Andrews & Bonta, 2003), all of which are risk factors that have been empirically related to criminal offending (e.g., Gendreau, Little, & Goggin, 1996).

LEVEL OF SERVICE INVENTORY–REVISED (LSI-R)

The LSI-R (Andrews & Bonta, 2001) is a third-generation assessment tool that measures offenders' characteristics and situations and is used to inform decisions concerning level of service necessary for a given offender. It measures the Big Eight and other relevant criminogenic factors through 54 items that are grouped into 10 subscales: Criminal History, Education/Employment, Finances, Family/ Marital, Accommodations, Leisure/Recreation,

Companions, Alcohol/Drug, Emotional/ Personal, and Attitude Orientation. Unlike second-generation assessment tools, the LSI-R measures both static and dynamic risk factors (Andrews & Bonta, 2001). It was developed to assist correctional professionals in making decisions concerning necessary levels of supervision and can also aid in decisions concerning sentencing, program or institutional classification, release from institutional custody, bail, and security ratings (Kroner & Mills, 2001; Lowenkamp & Latessa, 2002). In addition, it is designed to provide an overall estimate of the risk of reoffending (Andrews & Bonta, 2001; Gendreau, Goggin, & Smith, 2002).

In 2002, Lowenkamp and Latessa conducted a validation study on the LSI-R in community-based correctional facilities in Ohio. They found that an increase in LSI-R composite score was positively related to the likelihood of an offender having a new arrest or a technical violation reported to the court, being sentenced to prison, or unsuccessfully terminating a community-based correctional facility program. In addition, they found that a larger percentage of higher risk individuals (as measured by LSI-R composite score) had negative outcomes (e.g., rearrest, technical violation reported to the court, incarcerated) than lower risk individuals, although some of the increases were slight. Overall, they reported inconsistent results in terms of predictive validity across the programs they evaluated, with LSI-R scores from some programs predicting certain outcomes very well and LSI-R data from other programs performing less accurately. They attributed the variation in results to inconsistent data collection at some sites or the possibility that the outcome measure did not measure aberrant behavior as expected. They recommended that program sites develop specific cutoff scores based on their own populations (Lowenkamp & Latessa, 2002).

In another study of the validity of the LSI-R, Holsinger et al. (2006) reported that their data supported the predictive validity of the LSI-R for the whole sample but provided mixed results when the sample was divided by race. The results suggested that the LSI-R predicted rearrest most accurately for White offenders and male offenders overall and White men and White women compared with non-White (Native American) men and women (Holsinger et al., 2006).

Another recent study considered whether the LSI-R accurately classified African American, Caucasian, and Hispanic residents at a federal community corrections center, with the investigator seeking cutoff scores for each racial/ethnic group to minimize over- and underclassification (Whiteacre, 2006). He reported that the types and rates of classification errors depended on both cutoff score and outcome. When the predicted outcome was programmatic success, African Americans were more likely to be false positives than Caucasians or Hispanics. However, African Americans were more likely to be false negatives when predicting disciplinary incidents. Overall, the results reflected a consistent pattern of more classification errors for African Americans than for Caucasians or Hispanics. Whiteacre (2006) recommended that institutions evaluating risk assessment tools consider the appropriate cutoff score in light of their population, predicted outcome, and purpose for which the tool was used.

CORRECTIONAL OFFENDER MANAGEMENT PROFILING FOR ALTERNATIVE SANCTIONS (COMPAS)

The COMPAS (Brennan & Oliver, 2000) is one of the best known fourth-generation assessment instruments (Andrews et al., 2006). It is a risk assessment tool that was created to measure key risk and needs factors in adult correctional populations in order to provide

information to aid in decision making with regard to placement of offenders in the community. Unlike other risk assessment instruments, which provide a single risk score, the COMPAS provides separate risk estimates for violence, recidivism, failure to appear, and community failure. In addition to the Overall Risk Potential, as represented by those four scales, the COMPAS provides a Criminogenic and Needs Profile for the offender. This profile provides information about the offender with respect to criminal history, needs assessment, criminal attitudes, social environment, and additional factors such as socialization failure, criminal opportunity, criminal personality, and social support (Brennan & Oliver, 2000).

There are apparently no peer-reviewed published data with regard to the predictive validity of the COMPAS. The initial validation study, described in the COMPAS manual, followed a sample of 241 offenders from a New York probation sample released into the community and collected data on whether each offender was rearrested within 1 year of the date of COMPAS administration (Brennan & Oliver, 2000). The researchers reported that their sample offended at a rate of 24.7% during the outcome period. Using receiver operating characteristics (ROC) analysis, they described very good predictive validity (area under the curve [AUC] = .79) for the COMPAS recidivism scale. However, they noted that the AUC value may have been inflated because the same data set was used to develop and validate the data, and the data set was relatively small. This initial study did not include analyses using race/ethnicity or gender, however, so it is difficult to gauge its generalizability.

PURPOSES OF THIS STUDY

This study had three main goals: first, to provide what was apparently the first empirical study on the COMPAS that was independent of its development; second, to describe the criminogenic variables most strongly related to rearrest within 1 year following release from prison in a relatively large (N = 975) sample with a substantial proportion of racial/ethnic minority offenders; and third, to compare the predictive validities of the COMPAS recidivism score, the LSI-R composite score, and the relevant criminogenic variables, respectively, with regard to the performance of each in predicting rearrest within 1 year. Previous studies have provided data on the LSI-R in midwestern states in the United States (Lowenkamp & Latessa, 2002), in Canada (Andrews & Bonta, 2001, 2003), and with a primarily Native American sample (Holsinger, Lowenkamp, & Latessa, 2003; Holsinger et al., 2006). This study provides data using a U.S. sample from a northeastern, primarily urban population. The comparison of the LSI-R and the COMPAS is made to help gauge the accuracy of these respective tools. Using standard criminogenic variables compared with these tools gives some indication of how much the particular tools add to such accuracy, as contrasted with the consideration of separate variables that are well-recognized risk factors for criminal offending.

METHOD

PARTICIPANTS

The sample was made up of 975 male offenders (COMPAS N = 276; LSI-R N = 696; criminogenic variables N = 975) released from two assessment and treatment centers in

New Jersey between 1999 and 2002. Approximately half of the offenders with COMPAS and half of the offenders with LSI-R scores came from each facility. The criminogenic variables were total number of previous adult arrests, total number of previous juvenile arrests, total number of previous adult convictions, and total number of prior parole violations. All participants were subsequently placed in community halfway houses in New Jersey.

Participants were randomly selected from a list of all male offenders released from the two facilities during the specified period. Researchers randomly chose a starting position on the list and then selected every other file. A power analysis revealed that for a logistic regression with four independent variables, with an alpha of .05, and a medium effect size (w = .3), 297 participants yielded power of .99. This indicates that with 975 participants, it is highly likely that a result was detected if it existed. Potential participants were eliminated if they did not have scores for the LSI-R or COMPAS (one of which, but not both, was administered to all participants, so the only cases eliminated on this basis were the few for whom LSI-R or COMPAS results were missing from the file). Participants were also eliminated if their files were missing information concerning the relevant criminogenic variables or if they had been released less than 12 months prior to data collection. Women were not included in this study because we intend to study these cohorts separately. Finally, participants who were not African American, Caucasian, or Hispanic were excluded because there was not a sufficient number of individuals of any other racial/ethnic group to perform statistical analyses.

Ages of the participants ranged from 18 to 63 years, with a mean of 32.5 years (*SD* = 7.58). The racial/ethnic breakdown of the participants was as follows: 71.4% African American (*n* = 696), 15.0% Hispanic or Latino (*n* = 146), and 13.6% Caucasian (*n* = 133). (Racial/ethnic status was recorded according to the individual's self-identified category.) The mean level of education completed was 11.21 years (*SD* = 1.96), with 46.3% of the participants having completed either 12th grade or equivalent and 47.7% having completed less than 12th grade.

MATERIALS

Participants' institutional files and New Jersey Department of Corrections (NJDOC) records were used for data collection purposes. Information in those files was collected as part of the routine assessment and classification procedure at both assessment and treatment facilities. These data were entered into a database created in SPSS 11.

DESIGN AND PROCEDURE

A list of male offenders released from two assessment and treatment centers between 1999 and 2002 was compiled. Participants' files were randomly selected from that list for inclusion in the study. Upon receiving the files, a coding manual was created to describe each variable to be collected. Selected information from the files was then entered into a database created in SPSS 11 according to the coding manual. A list of the selected participants was then sent to the NJDOC so that the outcome data could be obtained. Once the outcome data were entered into the database, all identifying information was removed to protect participant confidentiality. After completing the data entry, the amount of time between each participant's date of release and date of rearrest was calculated to determine whether participants were rearrested within 12 months of release into the community. Rearrest was recorded based on official records, obtained through the Department of Corrections.

RESULTS

To determine whether participants differed based on racial group, a series of independent samples *t* tests was conducted. The dependent variables were the background variables (number of adult arrests, number of juvenile arrests, number of adult convictions), assessment data (LSI-R composite score and COMPAS recidivism score), and the outcome variable (rearrest within 12 months of release into the community). The results showed that on average, Caucasian participants had fewer previous parole violations ($M = 0.43$, $SD = 0.71$) than African American participants ($M = 0.69$, $SD = 0.98$; $t = –3.50$, $df = 234.61$, $p = .001$, equal variances not assumed). This remained significant after alpha was reduced to .002 with a Bonferroni correction. African American participants also differed from Hispanic participants ($M = 0.52$, $SD = 0.70$) on number of previous parole violations ($t = 2.51$, $df = 284.523$, $p = .013$, equal variances not assumed). This did not remain significant after the Bonferroni correction. The results also showed that Caucasian participants obtained lower LSI-R composite scores ($M = 26.01$, $SD = 5.80$) than African American participants ($M = 27.64$, $SD = 5.22$; $t = –2.73$, $df = 603$, $p = .007$). This was no longer significant after alpha was adjusted with a Bonferroni correction. In addition, Caucasian participants had more previous adult arrests ($M = 12.29$, $SD = 7.74$) than Hispanic participants ($M = 10.15$, $SD = 7.21$; $t = 2.40$, $df = 273.77$, $p = .017$, equal variances not assumed). African American participants also had more previous adult arrests ($M = 11.77$, $SD = 8.43$) than Hispanic participants ($t = 2.18$, $df = 851$, $p = .029$). These, too, were no longer significant after the Bonferroni correction. Tables 1 and 2 present the results of these analyses.

In terms of rearrest within 12 months of release into the community, Caucasian participants differed significantly from African American participants ($\chi^2 = 6.72$, $df = 1$, $p = .009$), and African American participants differed significantly from Hispanic participants ($\chi^2 = 5.23$, $df = 1$, $p = .022$). However, neither of these results remained significant after alpha was adjusted to .002 with a Bonferroni correction. Table 3 shows the descriptive information with regard to rearrest by race. Table 4 depicts the results of these significance tests.

Before conducting the main analyses, we calculated the base rate of offending for the entire sample. That base rate (0.21 or 21%) was used as the predicted probability cutoff for the logistic regression analyses. Table 5 depicts a contingency table comparing predicted rearrest based on LSI-R composite score with actual rearrest within 1 year of community release. With a predicted probability of 0.21, the LSI-R composite score correctly predicted outcomes for 48.4% of the sample as a whole, 80.4% of Caucasians, 43.4% of African Americans, and 82.4% of Hispanics. African Americans were more likely to be false positives than Caucasians or Hispanics (51.8% vs. 7.6% and 0%, respectively). In contrast, Hispanics and Caucasians were more likely to be false negatives than African Americans (17.7% and 12% vs. 4.78%, respectively). ROC analysis revealed an AUC of 0.6 with regard to the sample as a whole, with AUC values of 0.55 for Caucasians, 0.61 for African Americans, and 0.54 for Hispanics.

Data on the COMPAS may be seen in Table 6, comparing predicted rearrest based on COMPAS recidivism score with actual rearrest within 1 year of community release. With a predicted probability of 0.21, the COMPAS recidivism score correctly predicted outcomes for 85% of the sample as a whole, 97.6% of Caucasians, 76.4% of African Americans, and 90.9% of Hispanics. African Americans were more likely to be false positives than Caucasians or Hispanics (7.32% vs. 0% and 0%, respectively). Likewise, African Americans were also more

TABLE 1: Offense History, Parole History, and Risk-Needs Tool Scores by Race

Variable	M	SD
Number of previous adult arrests		
Overall	11.60	8.18
Caucasian	12.29	7.74
African American	11.77	8.43
Hispanic	10.15	7.21
Number of previous juvenile arrests		
Overall	1.47	2.93
Caucasian	1.20	2.22
African American	1.53	3.12
Hispanic	1.45	2.54
Number of previous adult convictions		
Overall	7.01	5.52
Caucasian	7.50	5.31
African American	6.95	5.24
Hispanic	6.84	6.86
Number of previous parole violations		
Overall	0.63	0.92
Caucasian	0.43	0.74
African American	0.69	0.98
Hispanic	0.52	0.70
LSI-R composite score		
Overall	27.28	5.28
Caucasian	26.01	5.80
African American	27.64	5.22
Hispanic	26.68	4.90
COMPAS recidivism score		
Overall	62.42	27.02
Caucasian	57.56	27.28
African American	63.77	26.29
Hispanic	61.11	29.79

Note. LSI-R = Level of Service Inventory–Revised; COMPAS = Correctional Offender Management Profiling for Alternative Sanctions.

likely to be false negatives than Caucasians or Hispanics (16.2% vs. 2.4% and 9.1%, respectively). ROC analysis yielded an AUC of 0.53 with regard to the sample as a whole, with AUC values of 0.81 for Caucasians, 0.48 for African Americans, and 0.67 for Hispanics.

Table 7 offers predicted rearrest based on the criminogenic variables (previous juvenile arrests, previous adult arrests, previous adult convictions, and previous parole violations) with actual rearrest within 1 year of community release. With a predicted probability of 0.21, the criminogenic variables correctly predicted outcomes for 59.6% of the sample as a whole, 78.2% of Caucasians, 37.2% of African Americans, and 78% of Hispanics. African Americans were more likely to be overclassified (false positives) than Caucasians or Hispanics (57.9% vs. 9.8% and 13%, respectively). In contrast, Caucasians were more likely to be underclassified (false negatives) than African Americans or Hispanics (12% vs. 4.9% and 8.9%, respectively). The ROC AUC value was 0.57 for the sample as a whole, 0.55 for Caucasians, 0.56 for African Americans, and 0.64 for Hispanics.

Logistic regression analysis was performed using rearrest within 1 year of community release as the dependent variable and number of previous adult arrests, number of previous juvenile arrests, number of previous adult convictions, and number of previous parole

TABLE 2: Racial Differences in Offense History, Parole History, and Risk-Needs Tool Scores

Source	df	t	p
Previous adult arrests			
Caucasian vs. African American	838	0.65	.514
Caucasian vs. Hispanic	273.77[a]	2.40	.017[b]
African American vs. Hispanic	851	2.18	.029[b]
Previous juvenile arrests			
Caucasian vs. African American	838	–1.18	.240
Caucasian vs. Hispanic	281	–.87	.386
African American vs. Hispanic	851	0.31	.760
Previous adult convictions			
Caucasian vs. African American	838	1.12	.265
Caucasian vs. Hispanic	281	0.89	.375
African American vs. Hispanic	851	0.20	.839
Previous parole violations			
Caucasian vs. African American	234.61[a]	–3.50	.001[c]
Caucasian vs. Hispanic	281	–1.01	.313
African American vs. Hispanic	284.52[a]	2.51	.013[b]
LSI-R composite score			
Caucasian vs. African American	603	–2.73	.007[b]
Caucasian vs. Hispanic	195	–0.88	.381
African American vs. Hispanic	612	1.72	.087
COMPAS recidivism score			
Caucasian vs. African American	84	–0.58	.567
Caucasian vs. Hispanic	281	–0.87	.386
African American vs. Hispanic	61.15[a]	0.55	.584

Note. LSI-R = Level of Service Inventory–Revised; COMPAS = Correctional Offender Management Profiling for Alternative Sanctions.
a. Equal variances not assumed.
b. Indicates significant difference between groups, but not significant after Bonferroni correction of alpha to .002.
c. Indicates significant difference between groups that remains significant after Bonferroni correction.

TABLE 3: Participant Rearrest Within 12 Months of Release Into the Community

Race	*# Rearrested*	*# Not Rearrested*	*Total*
Caucasian	18	115	133
African American	165	531	696
Hispanic	22	124	146

TABLE 4: Racial Differences in Rearrest Within 12 Months Following Release Into the Community

Source	df	χ^2	p
Caucasian vs. African American	1	6.72	.010[a]
Caucasian vs. Hispanic	1	1.33	.715
African American vs. Hispanic	1	5.21	.022[a]

a. Indicates significant difference between groups, but not significant after Bonferroni correction of alpha to .002.

violations as predictors. As can be seen in Table 8, a total of 975 cases was analyzed and the full model was significantly reliable ($\chi^2 = 11.13$, $df = 4$, $p = .025$). This model accounted for between 1.1% (Cox & Snell R^2) and 1.8% (Nagelkerke R^2) of the variance in rearrest status. Of those who were not rearrested, 63.6% were correctly predicted to

TABLE 5: Level of Service Inventory–Revised (LSI-R) Predicted Recidivism and Actual Rearrest

	Predicted Rearrest			
	Rearrested		*Not Rearrested*	
	%	n	%	n
Actual rearrest				
Total	17.8	124	5.6	39
Caucasian	6.5	6	12.0	11
African American	20.7	104	4.8	24
Hispanic	0.0	0	17.7	18
No actual rearrest				
Total	46.0	320	30.6	213
Caucasian	7.6	7	73.9	68
African American	51.8	260	22.7	114
Hispanic	0.0	0	82.4	84

TABLE 6: Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) Predicted Recidivism and Actual Rearrest

	Predicted Rearrest			
	Rearrested		*Not Rearrested*	
	%	n	%	n
Actual rearrest				
Total	0.0	0	14.9	41
Caucasian	0.0	0	2.4	1
African American	2.6	5	16.2	31
Hispanic	0.0	0	9.1	4
No actual rearrest				
Total	0.0	0	85.1	235
Caucasian	0.0	0	97.6	40
African American	7.3	14	73.8	141
Hispanic	0.0	0	90.9	40

TABLE 7: Criminogenic Variables Predicted Recidivism and Actual Rearrest

	Predicted Rearrest			
	Rearrested		*Not Rearrested*	
	%	n	%	n
Actual rearrest				
Total	9.3	91	11.7	114
Caucasian	1.5	2	12.0	16
African American	18.8	131	4.9	34
Hispanic	6.2	9	8.9	13
No actual rearrest				
Total	28.7	280	50.3	490
Caucasian	9.8	13	76.7	102
African American	57.9	403	18.4	128
Hispanic	13.0	19	71.9	105

TABLE 8: Logistic Regression Analyses for Criminogenic Variables Predicting Rearrest Within 12 Months Following Release Into the Community

Variable	N	B	SE B	β
Overall[a]				
Previous adult arrests	975	.006	.014	1.006
Previous juvenile arrests	975	.038	.024	1.039
Previous adult convictions	975	.020	.021	1.020
Previous parole violations	975	.147	.082	1.158
Caucasian[b]				
Previous adult arrests	133	.059	.048	1.061
Previous juvenile arrests	133	–.248	.240	0.781
Previous adult convictions	133	–.091	.084	0.931
Previous parole violations	133	.231	.304	1.260
African American[c]				
Previous adult arrests	696	.009	.018	1.009
Previous juvenile arrests	696	.048	.026	1.050
Previous adult convictions	696	.011	.029	1.011
Previous parole violations	696	.087	.088	1.091
Hispanic[d]				
Previous adult arrests	146	–.076	.081	0.926
Previous juvenile arrests	146	–.055	.121	0.946
Previous adult convictions	146	.153	.109	1.165
Previous parole violations	146	.589	.308	1.803

a. Full model is significantly reliable ($\chi^2 = 11.13$, $df = 4$, $p = .025$; Cox & Snell $R^2 = .011$, Nagelkerke $R^2 = .018$).
b. Model is not significantly reliable ($\chi^2 = 4.24$, $df = 4$, $p = .374$; Cox & Snell $R^2 = .031$, Nagelkerke $R^2 = .057$).
c. Model is not significantly reliable ($\chi^2 = 6.86$, $df = 4$, $p = .143$; Cox & Snell $R^2 = .010$, Nagelkerke $R^2 = .015$).
d. Model is significantly reliable ($\chi^2 = 10.65$, $df = 4$, $p = .031$; Cox & Snell $R^2 = .070$, Nagelkerke $R^2 = .123$).

not be rearrested. However, only 44.4% of those who were rearrested were accurately predicted as such. Overall, 59.6% of cases were correctly predicted to either be rearrested or not rearrested, respectively.

Logistic regression analysis was also performed using the same dependent and predictor variables with the sample divided by race. Table 8 shows that a total of 146 Hispanic cases was included and the model was significantly reliable ($\chi^2 = 10.65$, $df = 4$, $p = .031$). The model for the Hispanic participants accounted for between 7.0% (Cox & Snell R^2) and 12.3% (Nagelkerke R^2) of the variance in rearrest status. Of those who were not rearrested, 84.7% were accurately predicted to not be rearrested. However, only 40.9% of those who were rearrested were accurately predicted to be rearrested. Overall, 78.1% of cases were correctly predicted to be either rearrested or not rearrested, respectively. A total of 133 Caucasian cases was included and the model was not significantly reliable ($\chi^2 = 4.24$, $df = 4$, $p = .374$). A total of 696 African American cases was included; again, the model was not significantly reliable ($\chi^2 = 6.86$, $df = 4$, $p = .143$).

To determine the relationship between LSI-R composite score and rearrest, logistic regression was run using rearrest as the dependent variable and LSI-R composite score as the predictor. As shown in Table 9, some 696 cases were included, and the full model was significantly reliable ($\chi^2 = 12.40$, $df = 1$, $p < .001$). This model accounted for between 1.8% (Cox & Snell R^2) and 2.7% (Nagelkerke R^2) of the variance in rearrest status. Forty percent of the nonrearrested offenders were accurately predicted to not be rearrested, and 76.1% of the rearrested offenders were successfully predicted to be rearrested. Overall, 48.4% of the

TABLE 9: Logistic Regression Analyses for Level of Service Inventory–Revised (LSI-R) Composite Score Predicting Rearrest Within 12 Months Following Release Into the Community

Variable	N	B	SE B	β
Whole sample[a]	696	.062	.018	1.064
Caucasian[b]	92	.026	.048	1.026
African American[c]	502	.071	.021	1.073
Hispanic[d]	102	.019	.054	1.019

a. Full model is significantly reliable ($\chi^2 = 12.40$, $df = 1$, $p < .001$; Cox & Snell $R^2 = .018$, Nagelkerke $R^2 = .027$).
b. Model is not significantly reliable ($\chi^2 = 0.29$, $df = 1$, $p = .590$; Cox & Snell $R^2 = .003$, Nagelkerke $R^2 = .005$).
c. Model is not significantly reliable ($\chi^2 = 12.02$, $df = 1$, $p = .001$; Cox & Snell $R^2 = .024$, Nagelkerke $R^2 = .035$).
d. Model is significantly reliable ($\chi^2 = 0.122$, $df = 1$, $p = .727$; Cox & Snell $R^2 = .001$, Nagelkerke $R^2 = .002$).

TABLE 10: Logistic Regression Analyses for Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) Recidivism Score Predicting Rearrest Within 12 Months Following Release Into the Community

Variable	N	B	SE B	β
Whole sample[a]	276	.003	.006	1.003
Caucasian[b]	41	.058	.059	1.060
African American[c]	191	–.003	.007	0.997
Hispanic[d]	44	.023	.020	1.023

a. Full model is not significantly reliable ($\chi^2 = 2.13$, $df = 1$, $p = .645$; Cox & Snell $R^2 = .001$, Nagelkerke $R^2 = .001$).
b. Model is not significantly reliable ($\chi^2 = 1.613$, $df = 1$, $p = .204$; Cox & Snell $R^2 = .039$, Nagelkerke $R^2 = .188$).
c. Model is not significantly reliable ($\chi^2 = 0.21$, $df = 1$, $p = .644$; Cox & Snell $R^2 = .001$, Nagelkerke $R^2 = .002$).
d. Model is not significantly reliable ($\chi^2 = 1.48$, $df = 1$, $p = .224$; Cox & Snell $R^2 = .033$, Nagelkerke $R^2 = .072$).

cases were accurately predicted to either be rearrested or not, respectively ($B = .062$, $SE\ B = .018$, $p = .001$).

When the LSI-R sample was divided by race, logistic regression with 92 Caucasian cases yielded a model that was not significantly reliable ($\chi^2 = 0.29$, $df = 1$, $p = .59$). As shown in Table 9, a logistic regression with 502 African American cases yielded a model that was significantly reliable ($\chi^2 = 12.02$, $df = 1$, $p = .001$). This model accounted for between 2.4% (Cox & Snell R^2) and 3.5% (Nagelkerke R^2) of the variability in rearrest status. Of those who were not rearrested, 30.5% were correctly predicted to be not rearrested. In addition, 81.3% of the rearrested offenders were accurately predicted to be rearrested. Overall, 43.4% of cases were accurately predicted to either be rearrested or not, respectively ($B = .071$, $SE\ B = .021$, $p = .001$). A logistic regression with 102 Hispanic cases resulted in a model that was not significantly reliable ($\chi^2 = 0.122$, $df = 1$, $p = .727$).

Finally, logistic regression analysis was conducted to assess the relationship between the COMPAS recidivism score and rearrest. Table 10 reveals that 276 cases were included, and the full model was not significantly reliable ($\chi^2 = 0.21$, $df = 1$, $p = .645$).

When the sample was divided by race, logistic regression including 41 Caucasian cases yielded a model that was not significantly reliable ($\chi^2 = 1.613$, $df = 1$, $p = .204$). A logistic regression including 191 African American cases resulted in a model that was also not significantly reliable ($\chi^2 = 0.213$, $df = 1$, $p = .644$). Finally, a logistic regression including 44 Hispanic cases yielded a model that was not significantly reliable ($\chi^2 = 1.48$, $df = 1$, $p = .224$).

DISCUSSION

Previous research has investigated the extent to which the LSI-R can predict future criminal behavior (Andrews & Bonta, 2003). However, there are limited data on the LSI-R using a northeastern U.S., primarily minority and urban cohort. To date, there are no published data on the predictive validity of the COMPAS. This study investigates the predictive abilities of the LSI-R, the COMPAS, and selected criminogenic variables for rearrest within 1 year of community release. This study used data drawn from a sample from a northeastern state that was substantially made up of minorities from urban areas.

The results indicate that the criminogenic variables predicted rearrest for the whole sample at a rate slightly better than chance. These criminogenic variables performed best for Hispanic individuals, correctly predicting approximately 78% of the outcomes.

These results suggest that there is predictive inaccuracy driven by racial/ethnic status. African Americans were more likely to be overclassified (predicted to be rearrested when they actually were not) than Caucasians or Hispanics across all three predictive approaches. By contrast, underclassification errors (predicting no rearrest when the participant was actually rearrested) varied by measure. For the LSI-R, Hispanics and Caucasians were more likely to be underclassified than the other two groups. For the COMPAS, African Americans were more likely to be underclassified. For the criminogenic variables, Caucasians were more likely to be underclassified.

The results provide further support for previous empirical evidence that criminal history is strongly related to future offending (Andrews & Bonta, 2003). In this case, the results demonstrate that prior juvenile and adult arrests, adult convictions, and parole violations are predictive of rearrest within 12 months of community release from prison.

These results have potential implications for jurisdictions that include substantial proportions of minority offenders. Researchers have expressed concern with regard to the validity of risk-needs assessment tools as applied to certain specific populations, especially nonmajority populations (Holsinger et al., 2003, 2006; Whiteacre, 2006). These results provide further support for previous findings that the LSI-R yields mixed predictive validity when used with different ethnic and racial populations (Holsinger et al., 2006; Whiteacre, 2006). In addition, this study suggests that the validity of the COMPAS and the criminogenic variables also varies with racial/ethnic group. It is possible that different groups may be influenced by different risk and needs factors that lead to recidivism. As part of the analysis, this study used contingency tables to test for racial/ethnic differences in prediction errors for the LSI-R composite score, COMPAS recidivism score, and four criminogenic variables. The results indicated an overall tendency to either over- or underclassify participants depending on ethnicity/race. This gives rise to concern about the use of different risk assessment instruments with certain specific populations.

In studies such as this, there is the possibility that the results of tools such as the LSI-R and the COMPAS would be available to parole officers in the community. This would be valuable for supervision and risk management purposes. However, it would also potentially contaminate the relationship between the risk-needs tool results and the risk of subsequent arrest, if parole officers were to focus more intensively on higher risk individuals and thus lower their reoffending risk. Because such results were not available to parole officers supervising this sample, however, this was not a problem in this study.

This study does have several limitations, however. First, it employed a relatively short outcome period of 12 months postrelease. Further research using these tools should expand that to include follow-up periods of varying length. Second, the outcome variable only represents rearrest records in a single jurisdiction (New Jersey). It is possible that some of the participants were rearrested in another state within the 1-year follow-up period. If that were so, then the base rate of offending would have been higher than was calculated. Third, the outcome in this study was rearrest. Although rearrest is commonly used and is often the only measure of recidivism available, it has been considered the least stringent of the possible recidivism measures (Holsinger et al., 2006). Perhaps, the major limitation to officially recorded rearrest is that some criminal offending that actually occurred will not be detected for the purposes of the study. Moreover, to the extent that there is racial bias in whether offenders are rearrested, this diminishes the sensitivity of rearrest as a measure of criminal behavior. Further research should consider other measures of recidivism that may more accurately reflect actual antisocial behavior, including self-report and collateral report. Some of the subsamples for the COMPAS predictions were small, so replication is important before deciding whether and how to use this tool in practice. Fourth, the timing of the LSI-R and COMPAS administration may have affected the predictive accuracy of both tools. They were administered after participants had been released from prison and had entered a community-based assessment center, where they typically stayed for 6 to 9 months. To the extent that some dynamic risk factors changed while participants were treated in the center prior to full release into the community, the profiles yielded by the LSI-R and COMPAS would have been somewhat different. Finally, the results of this study may not be generalizable beyond the immediate cohort. This study used data on an all-male, mostly urban sample made up solely of Caucasians, African Americans, and Hispanics. Therefore, the results may not be generalizable to women or offenders of other racial/ethnic backgrounds. Further research should investigate how well the LSI-R and the COMPAS can predict recidivism in other populations. Likewise, future research should further examine the risk and needs factors that are relevant to different specific populations.

REFERENCES

Andrews, D. A., & Bonta, J. (2001). *Level of Service Inventory–Revised (LSI-R): User's manual*. North Tonawanda, NY: Multi-Health Systems.

Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency*, *52*, 7-27.

Bonta, J. (1996). Risk-needs assessment and treatment. In A. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks, CA: Sage.

Brennan, T., & Oliver, W. L. (2000). *Evaluation of reliability and validity of COMPAS scales: National aggregate sample*. Traverse City, MI: Northpointe Institute for Public Management.

Conroy, M. A., & Murrie, D. C. (2007). *Forensic assessment of violence risk: A guide for risk assessment and risk management*. Hoboken, NJ: John Wiley.

Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law*, *11*, 347-383.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the "unparalleled" measure of offender risk? *Criminal Justice and Behavior*, *29*, 397-426.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology*, *34*, 575-607.

Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2003). Ethnicity, gender, and the Level of Service Inventory–Revised. *Journal of Criminal Justice*, *31*, 309-320.

Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2006). Exploring the validity of the Level of Service Inventory–Revised with Native American offenders. *Journal of Criminal Justice*, *34*, 331-337.

Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*, *28*, 471-489.

Lowenkamp, C. T., & Latessa, E. J. (2002). *Validating the level of service inventory revised in Ohio's community based correctional facilities*. Retrieved March 3, 2007, from http://www.uc.edu/criminaljustice/ProjectReports/OHIOCBCFLSI-R.pdf

Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P., Mulvey, E., et al. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York: Oxford University Press.

Whiteacre, K. W. (2006). Testing the Level of Service Inventory–Revised (LSI-R) for racial/ethnic bias. *Criminal Justice Policy Review*, *17*, 330-342.

Tracy L. Fass, JD, PhD, received her law degree from Villanova School of Law and her doctorate in psychology from Drexel University, completing her training in the Villanova-Drexel Law-Psychology Program in 2008. Her research interests include criminal and juvenile offending.

Kirk Heilbrun, PhD, is a professor and head of the Department of Psychology at Drexel University. His research interests include forensic mental health assessment and risk assessment.

David DeMatteo, JD, PhD, is an assistant professor in the Department of Psychology at Drexel University and a codirector of the Law-Psychology Program at Drexel. His research interests include psychopathy, specialty courts, and forensic mental health assessment.

Ralph Fretz, PhD, is the corporate director of assessment and research at Community Education Centers, Inc. His research interests include risk assessment, classification, and effective intervention.