VALIDATION AND APPLICATION OF THE LS/CMI IN NEBRASKA PROBATION

ALISHA CALDWELL JIMENEZ
R. HAZEL DELGADO
TRACE C. VARDSVEEN
RICHARD L. WIENER
University of Nebraska–Lincoln

This study examined the validity of the Level of Service/Case Management Inventory (LS/CMI), as probation officers in the state of Nebraska use the tool. Study 1 evaluated the predictive validity of the LS/CMI by examining 19,344 probationer records over a 5.5-year period (January 2007-July 2013), and found that the LS/CMI total risk score demonstrated moderate predictive validity. Consistent with past findings, logistic regression showed that the total risk score predicted recidivism (return to probation) differently for nonminorities than for minorities. Furthermore, minorities scored higher than nonminorities on seven of the eight criminogenic factors. Study 2, a true randomized experiment, explored probation officer bias as an explanation for these findings, and found that training increased officers' ratings of scores in some LS/CMI domains and decreased ratings in others. Most importantly, there was no evidence that officers demonstrated racial bias in administering the LS/CMI survey when scoring Black versus White target clients.

Keywords: LS/CMI; validation; race/ethnicity

CRIMINAL JUSTICE POLICY AND CORRECTION POPULATIONS

WHO IS IN AND WHO IS OUT?

At the end of 2015, there were 6.7 million offenders under the supervision of the correction system in the United States, which is one in 37 adults (Kaeble & Glaze, 2016). Of the 6.7 million offenders, 56.2% (3,789,800) were probationers, 12.9% were on parole, 22.6% were in prison, and 10.8% were in a local jail (Kaeble & Glaze, 2016). With regard to being locked up behind bars, in 2016 there were more than 2.3 million inmates in 3,283 local jails, 1,719 state prisons, 102 federal prisons, and 79 Indian County jails as well as in military prisons, immigration detention facilities, civil commitment centers, and in U.S. territory prisons (The Sentencing Project: Research and Advocacy for Reform, 2016). The high rates of incarceration affect racial minorities, especially men, disproportionately as compared

AUTHORS' NOTE: *This work was partially supported with funds from the Nebraska Administrative Office of Probation. The authors extend a special thanks to Deborah Minardi and Ellen Brokofsky of the Nebraska Administrative Office of Probation, whose support and interest made this work possible. We would also like to thank Krystia Reed and Evan McCracken who assisted with the earlier stages of this work. Correspondence concerning this article should be addressed to Richard L. Wiener, Department of Psychology, University of Nebraska–Lincoln, 238 Burnett Hall, Lincoln, NE 68588; e-mail: rwiener2@unl.edu.*

CRIMINAL JUSTICE AND BEHAVIOR, 201X, Vol. XX, No. X, Month 2018, 1–22.
DOI: 10.1177/0093854818763231

with Whites. In the state of Nebraska, the source of the current data, the ratio of incarceration rates of Blacks to Whites in 2014 was 8.4 to 1 (among the highest in the nation) and for Hispanics, it was 1.8 to 1 (The Sentencing Project: Research and Advocacy for Reform, 2016). Furthermore, national probation statistics show the opposite disparate effects, such that 54% of probationers were White, while only 30% are Black (Kaeble, Glaze, Tsoutis, & Minton, 2015), suggesting that the criminal justice system treats Blacks more severely (i.e., sending them to prison) than Whites (i.e., putting them on probation).

Why are so many adults currently in the corrections system, and why so many Blacks and Hispanics? The answer to that question begins with a look back at the last 50 years in the U.S. criminal justice system. Prior to the 1970s, our criminal justice system was largely rehabilitative in nature; however, when researchers, commentators, and public officials concluded that the rehabilitative efforts were unsuccessful, the system shifted to a punishment-focused approach in an effort to deter criminal activity (Andrews & Bonta, 2010). Reacting to an increased crime rate in the 1970s and the failure of a rehabilitation model, policy makers in the past several decades responded by increasingly adopting "get-tough on crime" laws along with the policies of the so called "War on Drugs" initiative, which increased the rates and duration of incarceration to deter crime (Andrews & Bonta, 2010; Walmsley, 2009). Thus, the policy approach to crime in the latter years of the 20th century favored deterrence as the best method of preventing recidivism. Ironically, instead of deterring a criminal lifestyle, the "get-tough on crime" approach accompanied by high rates of incarceration contributed to producing offenders who were often more "criminal" once they left prison (Andrews & Bonta, 2010; Burnett & Maruna, 2004; Freeman, 2003; Gendreau, Cullen, & Goggin, 1999; Green & Winik, 2010; Nagin, Cullen, & Jonson, 2009). In fact, there is little evidence that punitive restrictions reduce recidivism (Andrews & Bonta, 2010; Andrews, Bonta, & Wormith, 2011).

RISK–NEED–RESPONSIVITY (RNR)

In part as a reaction to the failure of the "tough-on-crime" approach, recent years have witnessed a renaissance of rehabilitation in corrections with the advance of Andrew and Bonta's RNR model (Andrews & Bonta, 2010; Andrews & Dowden, 2006; Andrews et al., 1990; Gendreau & Andrews, 1990; Luther, Reichert, Holloway, Roth, & Aalsma, 2011; Olver, Stockdale, & Wormith, 2014; Parent, Guay, & Knight, 2012; Taylor-Gooby & Zinn, 2006). According to RNR, the principles of assessing risk level through measuring criminogenic needs and intervening through cognitive behavioral learning techniques are the most effective ways to bring about desistance (Andrews & Bonta, 2010). The risk principle proposes that the level of treatment should match the level of risk, so that high-risk offenders should receive stronger doses of intervention, while low-risk offenders should receive minimal or no treatment. The need principle states that treatments should focus on criminogenic needs, which are the factors most predictive of decisions to engage in criminal activity. The responsivity principle also advises that correctional programs should match the characteristics of the offenders (e.g., learning style, motivation, intensity, etc.) with treatments that fit their individual strengths and weaknesses. Several studies have provided evidence to support the RNR model as a generally effective means of reducing recidivism (Andrews et al., 1990) with general and special populations, such as violent offenders (Dowden & Andrews, 2000), women (Dowden & Andrews, 1999), and juveniles (Dowden & Andrews, 1999b,

2003). The first component of the RNR model requires objective and unbiased assessment of offender risk, which is the focus of the current manuscript.

Assessing Risk in Probation

A general review of the risk assessment literature includes little published research addressing recidivism among probationers in the United States. Most studies that do exist are unpublished and limit their analyses to single wave snapshots of probation records in a single state (Clarke, Lin, & Wallace, 1988; McGaha, Fichter, & Hirschburg, 1987; Petersilia, Turner, Kahan, & Peterson, 1985). Making the literature even more difficult to interpret are the often incomparable measures of recidivism that researchers report. For example, Maltz (as cited in Benedict & Huff-Corzine, 1997) identified nine dimensions that researchers sometimes report as indicators of recidivism: arrests, reconvictions, incarcerations, parole or probation violations, parole or probation suspensions, parole or probation revocations, offenses, absconding, and return to probation. In part, because researchers rely on so many different indicators of failure, published studies of recidivism among probationers show disparate failure rates, ranging anywhere from 22% (Vito, 1987) to 65% (Petersilia et al., 1985). For example, while a study in Iowa in 2005 showed a rearrest rate of 43% for male and 27% for female probationers during the first fiscal quarter of 2001 (Stageberg & Wilson, 2005), Petersilia's (1985) study of recidivism rates in several counties in California revealed a 65% recidivism rate during a 3-year window. Supporting the California findings, Texas statistics similarly showed a 64.5% rearrest rate for individuals who were under adjudicated probation supervision from 2005 until 2008 (The State of Texas, Legislative Budget Board, 2013).

Complicating the findings even more, The Sentencing Project: Research and Advocacy Forum's (2010) summary of studies conducted from 1995 to 2009 shows that the unpublished rates of recidivism for probationers vary even more dramatically than those that appear in the published literature. An unpublished Idaho study reported that 74% of its probationers recidivated within 5 years of finishing probation. Recidivism started out at a higher rate for the first 30 months (82%), but then steadily leveled off until it reached a much lower rate of 12% of those remaining outside the system after 3 years (The Sentencing Project: Research and Advocacy Forum, 2010). At the same time, data from New York state (New York State, Division of Criminal Justice Services, 2009) indicated a recidivism rate for probationers of only 25.7% in 3 years, but that study defined recidivism as a felony rearrest and did not consider misdemeanors. The lessons learned from the probation recidivism literature are three: First, although there is little agreement on the best way to measure probationer risk, most jurisdictions settle on some measure of recidivism as a core indicator. Second, definitions of recidivism and time frames for measuring recidivism vary greatly from study to study and from jurisdiction to jurisdiction. Third, it is critical that comparisons within jurisdictions over time, or within jurisdictions between different types of programs, use a set of agreed-upon indicators defined similarly across data collection waves, and that these indicators are comparable across programs within a jurisdiction. In this report, we report on a broad indicator of recidivism, namely any return to probation or jail after discharge.

RISK ASSESSMENT TOOLS

There are a variety of published and unpublished risk assessment tools in the literature that attempt to predict general criminal behavior. Desmarais, Johnson, and Singh (2016)

reviewed and described all the actuarial approaches (i.e., mechanical models that interpret objective scores on risk-related factors), and structured professional judgment approaches (i.e., assessment guides for professionals to weigh factors that are risk related) that criminal justice practitioners currently use to assess risk in U.S. jurisdictions. After an exhaustive literature review, they found 53 multistudy reports in journals and government articles in which researchers and practitioners implemented and validated 19 risk assessment tools in the United States. There were surprisingly few independent studies of risk assessment tools with U.S. offender samples, and only four, the Level of Service Inventory (LSI; 27 studies), the Salient Factor Score (15 studies of three versions), the Wisconsin Risk and Needs (10 studies), and the Ohio Risk Assessment System (seven studies of its various components), registered more than two investigations. Some like the Static Risk and Offender Needs Guide (STRONG-R) showed only one study with limited credibility due to the lack of independent scholarship; that is, the one and only published study of the STRONG-R's utility was the one that the instrument developers had completed and published themselves (Hamilton et al., 2016).

Desmarais et al. (2016) reviewed the predictive validity of these four instruments, and concluded that they were all rather similar with respect to area under the curve (AUC), binary correlations, and odds ratio estimates of effect sizes. They found some differences in the strength of validity predictions depending upon the type of recidivism (e.g., new arrests, new convictions, probation violations, or parole violations) that the authors used as criteria. The overwhelming conclusion of the Desmarais et al. review was that there were relatively few investigations other than those done with the LSI scales, and that we need much more work studying the reliability and validity of current risk tools with U.S. samples. To that point, Monahan and Skeem (2016) argue that up to the current time, we have understudied our instruments, given the current surge of interest in using risk assessment in sentencing, and most importantly, to divert low-risk offenders from jail and prison into community corrections. They argue that we need some clarification on the difference between risk and blame, and the role each should play in assignment of punishment and rehabilitation, as well as additional clarification on the limits of using group-level data to make individual inferences. Perhaps most importantly, the criminal justice system carefully ought to consider whether existing risk instruments may give rise to disparities in imprisonment, probation, and parole because of the racial or economic limitation, or even bias in the instruments that we use. This is extremely important given the disparate rate of arrest, conviction, especially imprisonment in minority populations as discussed above. This article takes a detailed look at one instrument, the LS/CMI, as probation officers administer it to clients in Nebraska.

LSI Scales

The LSI is a measure of offender (i.e., inmates, probationers, and paroles) risk of recidivism, which includes an assessment of offenders' criminogenic needs resulting in service recommendations and level of required supervision recommendations. The inventory consists of a commonly used set of scales with more than 1 million administrations (internationally) in 2010 alone (Andrews et al., 2011). In the last 30 years, there have been six different published versions of the LSI scales, all of which the authors originally developed in Canada. Nonetheless, criminal justice practitioners have used these scales with youth and adults not only in Canada but also in the United States and in a number of countries outside

North America (Olver et al., 2014). Including the most recent study completed in 2014 (Olver et al., 2014), researchers have conducted eight meta-analytic studies of the LSI scales with youth and adults (Gendreau, Goggin, & Smith, 2002; Olver, Stockdale, & Wormith, 2009, 2014; Schwalbe, 2007, 2008; Singh, Grann, & Fazel, 2011; Smith, Cullen, & Latessa, 2009; Yang, Wong, & Coid, 2010).

In the most recent meta-analysis of the LSI scales, Olver et al. (2014) coded effect sizes from 128 studies with 151 independent samples ($N = 137{,}931$ offenders) that examined the power of the LSI to predict recidivism among offenders. The studies came from Canada, the United States, and outside North America (Australia, United Kingdom, Singapore, Germany, Japan, New Zealand, and Pakistan). Overall, the effect size using a fixed model (one that does not adjust for the differences in sample sizes) produced a $r = .30$ across countries for general recidivism (a moderate predictor) and $r = .21$ for violent recidivism. However, and most importantly for our purposes, the effect size varied significantly, and substantially by country and region. For general recidivism, the effect size in Canada, where the instrument originated, was .43 (moderate) while it was .29 outside North America, also moderate, but significantly lower ($p < .05$). In the United States, the effect size, $r = .22$, was moderate but significantly lower than that in either Canada ($p < .05$) or countries outside North America ($p < .05$; Olver et al., 2014). All effect sizes were statistically significant ($p < .05$) or statistically greater than 0.

Given the variation in results, researchers and practitioners need to be aware of the possibility that the risk assessment instruments are more valid in some populations than in others. The Olver et al. (2014) meta-analysis did examine differences in the predictive validity of the LSI scales in men and women as well as in minorities and nonminorities. The effect sizes for general recidivism in the Olver et al. (2014) meta-analysis did show slight gender differences across regions with $r = .32$ for women and $r = .30$ for men. The Olver et al. (2014) meta-analysis tested these differences and found them significant at the .05 level of significance, suggesting that the scales do a slightly better job of predicting risk for women than they do for men. Much more importantly, the effect sizes showed large and statistically significant differences ($p < .05$) for ethnicity worldwide with $r = .23$ for minorities and $r = .32$ for nonminorities. The effect size disparities for ethnicity showed that the instrument is more predictive of the nonminority than of the minority offenders. One of the major purposes of the current study was to see if this racial difference in predictive validity holds true for the use of the LS/CMI in our Nebraska sample.

The discrepancies in the LSI scales as a function of minority status took on a new importance recently in the criminal law when a Canadian court in *Ewert v. Canada* (2015) questioned the continued use of risk assessment instruments because some were unreliable predictors of recidivism for minority offenders; in this case, for Canadian Aboriginal inmates. The court challenged the use of the Hare Psychopathy Checklist–Revised (PCL-R), Violence Risk Appraisal Guide (V-RAG), Sex Offender Risk Appraisal Guide (SORAG), Static 99, Violence Risk Scale–Sex Offender (VRSSO) because of the lack of evidence that these instruments, normed on nonminority samples, were accurate predictors of recidivism for Aboriginal inmates. The court prohibited the Correctional Service Canada from using these five instruments to predict future risk of violence or future risk of sexual violence with First Nation offenders, until researchers have tested the instruments in cross-cultural samples and verified their validity for Canadian Aboriginals. Although a Federal court later overturned this decision based upon misplaced burden of proof (i.e., requiring states to

prove lack of discrimination, Canada v. Ewert, 2016), the fact that a trial court found assessment tools violated the law should serve as a warning in Canada as well as in other Western countries. Specifically, there is a need to verify in separate jurisdictions that risk instruments perform similarly for minority and nonminority populations. The current research examines this issue for the LS/CMI in Nebraska.

Supporting the need for validation at the jurisdiction level, Shepherd and Lewis-Fernandez (2016) argued that culture and subculture might influence the construction, application, and interpretation of results of risk assessment instruments in some very specific ways. The source of biases includes constructing and norming instruments based upon the behaviors and life experiences of White offenders, generating item content from the psychology of Western culture and the failure to address issues of cultural competency, so that the minority group norms are misaligned with the assumptions built into the instruments. Most importantly and closer to the purpose of the current study, Chenane, Brennan, Steiner, and Ellison (2015) found that the LSI-R predicted incidents of institutional misconduct better for White inmates than for Black inmates. We note that Chenane et al. (2015) may not generalize to differential validity in community corrections, where offenders spend most of their time acting freely outside the direct supervision of probation and parole officers, which is, of course, not the case in prison. The current study helps close that gap by examining the validity of the LS/CMI in predicting recidivism, defined here as the return to probation or jail after release from probation, for individuals completing a probation sentence in Nebraska. One of the main purposes of our analysis was to examine whether the LS/CMI predicts return to probation equally well for minority and nonminority populations.

STUDY 1

Based upon the prior literature, we predicted that the total score on the LS/CMI would show predictive validity in the low to moderate effect size range, similar to the effects that others have found in validity studies of the instrument in the United States (Olver et al., 2014). One of the major purposes of our study was to test whether minority status tainted the use of the LS/CMI in the state. Therefore, we hypothesized ethnic/racial differences and interactions with risk on return to probation to test possible ethnic/racial disparities in the risk assessment process: First, if there were minority disparities, minorities would score higher on most if not all of the eight criminogenic factors that the LS/CMI measures, and second the LS/CMI might be a better predictor of return to the criminal justice system for nonminority probationers than for minority probationers. Thus, we tested for a main effect of the LS/CMI risk level on probation recidivism, main effects of minority status on probation recidivism, and an interaction between the LS/CMI and minority status on probation recidivism.

METHOD

Description of Nebraska Probation and Use of the LS/CMI

The first study evaluates the validity of the Level of Service/Case Management Inventory (LS/CMI; Andrews, Bonta, & Wormith, 2004), which Nebraska probation officers administer to all offenders charged with a felony or a serious misdemeanor for two purposes: (a) to determine which evidence-based services to provide to offenders and (b) to form the basis

of a presentencing investigation—an interview style investigation where probation officers determine the offender's level of risk and supervision requirements. Officers collect answers to the LS/CMI questions during the interview process and submit their findings to the court for consideration at the time of sentencing (i.e., a presentence investigation). The court then decides the type of probation programming that is appropriate for each offender, after taking into consideration his or her LS/CMI risk level.

Participants

The Nebraska Administrative Office of Probation (NAOP) maintains a database of all LS/CMI administrations for all probationers across the state. The present study used data that we obtained from NAOP over a 5½-year window (January 2007-July 2013), which included 19,344 probationers for whom probation officers had administered the LS/CMI. Nebraska law considers juvenile offenders to be 18 years or younger; we did not include data from juvenile offenders in this report. The data consist of measures for both men (77%) and women (23%) with a mean age of 33 years. The sample included White European Americans (73%), Black African Americans (12.7%), Asian Americans (0.7%), Native Americans (2.6%), and 10% who identified as other. Considering the full sample of probationers across the 5½-year window, 10.7% of the probationers were of Hispanic descent as self-reported and 89.3% were not. For purposes of our analyses, we considered nonminority status to include those probationers who self-identified as White European Americans of non-Hispanic descent (77% or 14,895), and we considered those probationers to be of minority status if they self-identified as not being White Europeans or were White, but were of Hispanic descent (23% or 4,449). The index charges for the probationers in our sample ranged from lower level misdemeanors (Misdemeanor 5) to serious felonies (Felony 1) with the majority (68%) falling into the midrange categories of serious misdemeanors (Misdemeanor W, Misdemeanor 1) and lower level felonies (Felony 4).

Measures

LS/CMI. The LS/CMI provides a comprehensive assessment of overall risk level, criminogenic needs, and responsivity for adult offenders, those above the age of 16 (Andrews et al., 2004). The LS/CMI focuses on eight criminogenic needs (static and dynamic risk factors) for purposes of risk classification and management, as well as for planning interventions and services. The first criminogenic factor, Criminal History, consists of four questions that measure the offender's past criminal behavior (i.e., frequency, onset, settings, and severity of crimes). The second factor examines Education and Employment with six employment items that measure the offender's capability to enter or reenter the labor market, and six additional items that quantify academic achievements in both traditional and technical programs. The third component, Family/Marital Issues, includes four items that examine positive interpersonal relationships. Leisure and Recreation, the fourth component made up of two items, examines how offenders use their free time. The fifth subsection, Offender Companions, utilizes four items to examine the quality and influence of acquaintances and friends (i.e., both positive and negative) with whom the offender has chosen to interact and engage. Eight items measure the sixth subcomponent, Alcohol and Drug use of all intoxicating substances excluding nicotine and caffeine. The seventh subcomponent utilizes four questions to assess the offender's Procriminal Attitudes. Finally, Antisocial

Pattern uses four items to determine whether offender's personality and associated behavior are antisocial.

Together, these eight subcomponents inform sentencing, case management, and supervision. To administer this instrument, professionals (e.g., counselors, psychologists, probation officers) undergo training to learn how to score items using information obtained during a one-on-one interview. Nebraska probation functions under the supervision of the state judiciary and answers directly to the Nebraska Supreme Court. The judiciary operates its own Judicial Branch Education, which conducts all the training of new probation officers. Beginning officers attend a specialized 3-day training session on the administration of the LS/CMI during which time expert trainers, who have themselves undergone "training the trainer" sessions, instruct new officers on how to administer the instrument through both didactic lessons and hands-on experience (i.e., practice with written summaries and taped interviews with actual probationers). Officers must pass minimal criteria before they conduct their own interviews and write their own reports in the field. Once in the field, officers attend periodic booster training sessions, and submit samples of their investigative interviews and reports to quality assurance reviewers for corrective feedback.

Trained LS/CMI officers conduct client interviews and mark corresponding answers to LS/CMI questions for both dichotomous options (i.e., Yes or No) and rating scales (i.e., 0, 1, 2, 3). Scores for each of the eight factors and the overall risk are the summed responses. Total risk scores range from 0 to 42, which the LS/CMI manual categorizes as very low, low, medium, high, and very high (Andrews & Bonta, 1995; Rettinger, 1998; Rowe, 1999). Nebraska probation currently utilizes a similar five-level cutoff categorization of risk for administrative purposes, which consists of very low (0-4), low (5-10), medium (11-19), high (20-29), and very high (30-42) levels. We used this system when examining the predictive validity of risk level. The number of probationers in each of these five levels in the analyses are respectively: very low (n = 836), low (n = 2,909), medium (n = 7,641), high (n = 6,139), and very high (n = 1,775).

Recidivism as a measure of risk. We collected follow-up probation recidivism data (i.e., subsequent involvement with the probation system after the first offense resulting in additional probation sentences or time in jail) for all individuals with LS/CMI scores at the time of sentencing or shortly thereafter. Recidivism in this report refers to a failure due to return to Nebraska probation; that is, any offender with a subsequent probation or jail sentence following the index offense during the 5½-year window was a probation recidivist. Return to probation refers to any new contact with probation after the index offense, including new offenses committed after release and new offenses committed while the client was in probation serving out her or his time for the index offense, if it triggered a new conviction as long as the offense occurred within the study's 5.5-year time frame.

RESULTS

Instrument Validity

To test the validity of the LS/CMI total risk score, we calculated an outcome variable "failure," and coded it as "1" if a probationer returned to probation or went to jail in the 5½-year window following the index conviction. We coded failure as a "0" if the client did not return to probation or jail during that same period. We used the total risk score on

TABLE 1: Predicting Probation Failure With the LS/CMI Total Scores Controlling for Time Following Probation

Factor	*Beta*	*SE*	Wald	*df*	Prob.	OR
Time after LS/CMI admin	.001	0.001	35.773	1	.001	1.000
Total LS/CMI risk score	.069	0.002	835.747	1	.001	1.072
Constant	–.028	0.053	0.277	1	.599	0.972

Note. LS/CMI = Level of Service/Case Management Inventory; OR = odds ratio.

LS/CMI that probation officers administered immediately following the index offense to predict failure using logistic regression. However, because probationers may have started their index probation any time during the window and completed it at variable times, we calculated a time variable for each probationer that measured the number of days from the administration of the LS/CMI to the end of the 5½-year window to use as a control variable. After the time in probation variable, we entered into the regression equation the probationer's total score on the LS/CMI. Table 1 shows the results of the logistic regression.

The full regression model included two significant predictors, time since probation and total risk score, such that, as each increased so did the likelihood that a probationer would fail; that is, return to probation during the 5½ year window, $\chi^2(2) = 934.045$, $p < .001$, Nagelkerke $R^2 = .073$. Converting the χ^2 associated with the risk score to a r value showed the effect size to be .21. Thus, the validity coefficient of the LS/CMI total risk score for this sample of Nebraska probationers was similar to that observed in the Olver et al. (2014) meta-analysis for the United States ($r = .22$). To examine the shape and nature of the relationship in detail, we replaced the continuous LS/CMI risk score with the five risk-level classifications that Nebraska probation uses to measure risk of reoffense in its clients, using the lowest risk level as the comparison group in the logistic regression. Table 2 displays those results showing a significant model in which all risk levels are significantly different from very low risk, Model $\chi^2(5) = 934.191$, $p < .001$; Nagelkerke $R^2 = .073$. Table 2 demonstrates that a probationer selected at random with a very high LS/CMI risk level is 8.84 times as likely to fail (return to probation or jail) as one selected at random with a very low-risk score.

Finally, to depict this illustration graphically, Figure 1 displays the probability of failing for each level of the LS/CMI risk categorizations. Note that all columns are significantly different from each other at the .001 level of significance.

Minority Outcomes

We concluded that the LS/CMI predicted probation recidivism, as well as other studies reflected in the Olver et al. (2014) meta-analysis, but took this opportunity to determine if the prediction equations were different for minorities and nonminorities. We examined the subsequent probation or jail predicted by LS/CMI risk level and some basic demographic characteristics of the probationers, including age, gender, minority status, the interaction of LS/CMI level with gender, and the interaction of LS/CMI level with minority status. Table 3 shows the results. It shows an overall significant logistic regression model, Model $\chi^2(16) = 1{,}003.421$, $p < .001$, Nagelkerke $R^2 = .085$, with significant effects for risk level, age of probationer (with older probationers being more likely to return), sex of probationer (with

TABLE 2: Logistic Regression Analysis: Predicting Subsequent Jail or Probation by LS/CMI Risk Level (Reference Group = Very Low Risk)

Predictor	*Beta*	*SE*	Wald	*df*	OR
Time after LS/CMI admin	.000	.000	35.756***	1	1.00
LS/CMI overall			896.136***	4	
LS/CMI (low)	.687	.080	74.187***	1	1.987
LS/CMI (medium)	1.320	.075	310.067***	1	3.742
LS/CMI (high)	1.795	.078	523.318***	1	6.019
LS/CMI (very high)	2.180	.105	433.741***	1	8.843
Constant	–.207	.078	7.063*	1	.813

Note. LS/CMI = Level of Service/Case Management Inventory; OR = odds ratio.
*$p < .05$. ***$p < .001$.

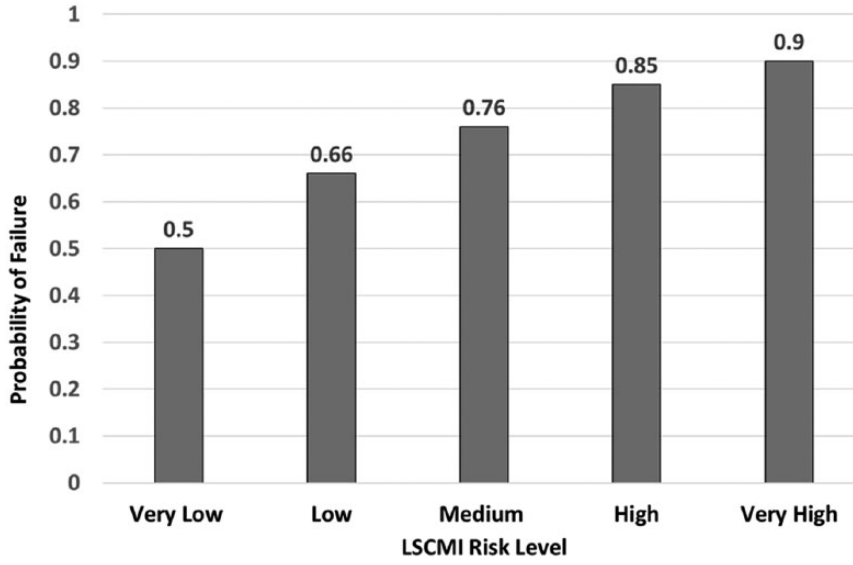

Figure 1: Probability of Subsequent Return to Probation or Jail as a Function of LS/CMI Risk Level
Note. All columns are significantly different from each other. LS/CMI = Level of Service/Case Management Inventory.

men more likely to return), and an interaction between risk level and minority status. The interaction between LS/CMI risk level and gender was not significant in our Nebraska sample, but the interaction between minority status and risk level was significant.

To test the difference in predictive validity of the risk levels for minorities (i.e., non-White European Americans or White of Hispanic descent) and nonminorities (i.e., White Europeans of non-Hispanic descent), we conducted separate follow-up logistic regressions for each. Risk level predicted failure slightly better for minorities, Wald(4) = 129.133, $p < .001$, Nagelkerke $R^2 = .094$, than for nonminorities, Wald(4) = 423.535, $p < .001$, Nagelkerke $R^2 = .085$. Using the LS/CMI score rather than level, risk again predicted slightly better for minorities, Wald(1) = 198.234, $p < .001$, $r = .212$, than for nonminorities, Wald(1) = 198.234, $p < .001$, $r = .208$. The difference in validity coefficient (effect sizes, r) in the latter analyses is trivial. Furthermore, the interaction between minority status and the risk using

TABLE 3: Logistic Regression Analysis: Predicting Subsequent Jail or Probation by LS/CMI Risk Level With Age, Gender, and Minority Status (Reference Group = Very Low Risk)

Predictor	*Beta*	*SE*	Wald	*df*	OR
Time after LS/CMI admin	.001	0.001	50.056***	1	1.000
LS/CMI overall			451.873***	4	
Age of probationer	.010	0.002	33.125***	1	1.010
Sex of probationer	.846	0.173	25.444***	1	2.331
Minority status	.315	0.173	3.321[a]	1	1.370
LS/CMI overall by sex			6.245[a]	4	
LS/CMI overall by minority status			13.067*	4	
Constant	–1.265	0.167	57.269***	1	0.282

Note. LS/CMI = Level of Service/Case Management Inventory; OR = odds ratio.
[a]Not significant.
*$p < .05$. ***$p < .001$.

TABLE 4: Effects of Minority Status on the Eight LS/CMI Criminogenic Factors

Source	Sum of squares (*df* = 1)	Mean square effect	Mean square error (*df* = 17,858)	*F* value	Eta square
Criminal history	62.341	62.341	0.806	77.376***	.004
Education	724.545	724.545	1.432	505.857***	.028
Family	44.791	44.791	1.466	30.553***	.002
Leisure	40.248	40.248	1.229	32.753***	.002
Companions	101.190	101.190	2.262	44.744***	.002
Substances	141.408	141.408	1.676	84.357***	.005
Procriminal	98.775	98.775	1.676	58.951***	.003
Antisocial	192.841	192.841	1.413	136.492***	.008

Note. LS/CMI = Level of Service/Case Management Inventory.
***$p < .001$.

the LS/CMI score is actually not statistically significant, Wald(1) = 2.062, $p = .151$, only the interaction between risk level and minority status is significant—largely because of the very large sample sizes.

Next, to drill down on LS/CMI indices as a function of minority status we examined the scores on each of the eight LS/CMI criminogenic factors for nonminority and minority probation with a MANOVA, in which minority status was the factor and each of the risk-level scores for the criminogenic factors was the criteria measure. Table 4 shows the results for univariate follow-ups after a significant multivariate model, Wilks' lambda = .958, $F(9, 17851) = 97.49$, $p < .001$, $\eta^2 = .042$, in which all the minority status differences for each of the criminogenic factors were significantly different between minorities and nonminorities ($p < .001$). As seen in Figure 2, minorities scored higher on all the risk factors except for alcohol and drug problems, on which minorities scored lower than nonminorities in our Nebraska sample.

STUDY 1 DISCUSSION

The results of Study 1 demonstrated that the use of the LS/CMI in Nebraska resulted in a similar validity coefficient ($r = .21$) predicting probation recidivism as it did in the meta-analysis results in the other jurisdictions in the United States, $r = .22$ (Olver et al., 2014).

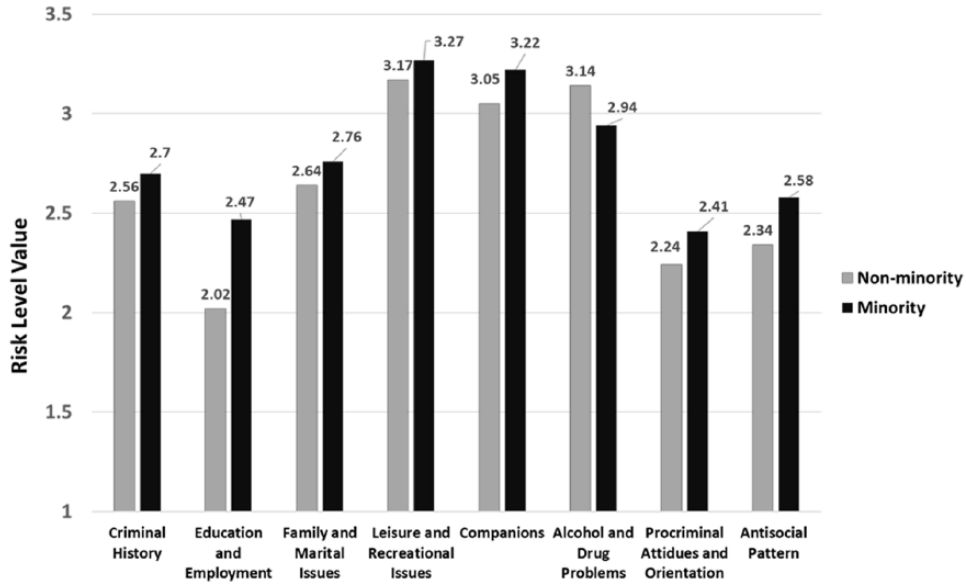

Figure 2: Risk-Level Scores for Minority and Nonminority Probationers on the Eight Criminogenic Factors of the LS/CMI

Note. All differences in each domain are significantly different. LS/CMI = Level of Service/Case Management Inventory.

Furthermore, in the Nebraska probation system the LS/CMI predicted risk levels equally well for women and men as evidenced by the absence of a significant interaction between the LS/CMI and sex on probation recidivism, despite the fact that other studies in other jurisdictions found LSI scales to predict recidivism differently for men and women. For example, Reisig, Holtfreter, and Morash (2006) reported that the LSI-R showed adequate predictive validity for men and for some women but not for those women whose criminal activity resulted from a gendered pathway into a criminal lifestyle. Furthermore, Hannah-Moffat (2006, 2009) persuasively opines that the prediction of risk for women offenders should take a different approach than the same prediction for men. Nonetheless, the recent Olver et al. (2014) meta-analysis showed similar LSI predictions of recidivism for men and women with some domains predicting slightly better for men and others slightly better for women. Altogether, these results point to the importance of testing the validity of risk assessment tools jurisdiction by jurisdiction because instruments with known psychometric properties may perform differently across jurisdictions. In that vein, our results show that LS/CMI is well suited to Nebraska probation as a validated indicator of risk of recidivism despite the fact that other studies in other jurisdictions found it to predict recidivism differently for men and women.

Equally important, we found some evidence, but not conclusive evidence, of an interaction between minority status and risk. The interaction was significant in the risk level analysis but not significant in the risk score analysis. Furthermore, to our surprise, our data suggest that the LS/CMI risk levels as configured in Nebraska actually predicted return to probation slightly better for minorities than for nonminorities. The slight differences in prediction are minimal, but not easily dismissed considering that further analyses showed

minorities scored higher on seven of eight risk domains on the LS/CMI. While the effect sizes for these differences are not large, when taken together they add up to a meaningful difference in risk estimates. Our analyses show that the largest effect size was for education needs on the LS/CMI, suggesting that the driver of minority differences in our sample was those education needs which were higher for minorities than for nonminorities. However, the reader should exercise caution not to overinterpret the impact of this one domain because the effect size for the education domain still reflects a relatively small difference between minorities and nonminorities.

Even if the LS/CMI predicts recidivism similarly for minorities and nonminorities, officers' risk predictions for the former may still be biased if officers perceive minorities at higher risk than is justified in their interview data (i.e., if officers are biased against minorities). Yet, there are other reasons for raised domain scores for minorities beside officer bias. For example, higher domain scores could indicate that the minorities are at higher risk than nonminorities, perhaps because of the sometimes low socioeconomic status (SES) positions that minorities occupy in society. Such an explanation would still favor a recommendation for continued use of the risk instrument but would invite further work to determine the specific social and economic conditions, and threats to minorities that increase their criminogenic risk.

Alternatively, probation officers who administer and score the LS/CMI might do so in a biased manner by misinterpreting interview data because of social stereotypes, which the officers hold that lead to higher risk judgments for minorities than for nonminorities with similar backgrounds. Finally, it is possible, as Shepherd and Lewis-Fernandez (2016) suggest, that test constructors like those who developed the LS/CMI inadvertently built in inherent biases because they based the instrument's items on an understanding of majority Western values and not on the realities of minority life styles in society. While the latter explanation is a possibility, it is inconsistent with the LS/CMI, predicting slightly more accurately for minorities than for nonminorities.

Given the fact that the LS/CMI, if anything, was a better predictor for minorities than for nonminorities, a full analysis of the item construction of the LS/CMI is unwarranted and, in any case, is beyond the scope and data in this research project. Instead of drilling down on the scale construction, we conducted an experiment to explore whether officers showed prejudice in their scoring of risk factors for Black African American as opposed to White European American probationers. Setting the stage for the experiment was the fact that the main effect of minority status on recidivism was not significant (see Table 3), Wald(1) = 3.321, but yet the score on the full LS/CMI was significantly greater for minorities ($M = 19.03$) than for nonminorities ($M = 17.47$), $t(17{,}930) = 11.28$, equal variances not assumed, $p < .001$. This leaves open the possibility that officers' ratings might result in higher risk predictions for minorities because they perceive elevated risk levels for minority clients, even though the minorities did not return to probation at a higher rate than did nonminorities.

To test the possibility of biased officer scoring on the LS/CMI, we proceeded to examine a specific race/ethnicity category to capture any potential officer bias due to stereotyping. The experimental task required manipulating a specific race and/or ethnic background because presenting a generic or nonspecific minority client to the officers to evaluate with the LS/CMI would not measure stereotyped bias in test administration against a specific social group. Therefore, we manipulated offender race in Study 2 presenting a Black African

American and a White European American offender, choosing race as the factor of interest because Blacks and Whites comprised the largest populations in the original Study 1 sample.

STUDY 2

TRAINING STUDY

As part of its effort to assure that officers are consistent and accurate in administering and scoring the LS/CMI, Nebraska probation conducted a training study in 2015 requiring all officers who administer the instrument to attend booster sessions. Before the training session began, the expert trainers conducted group meetings with officers who administer the LS/CMI to identify common disagreements in scoring, and to establish rules for resolving disagreements and coding errors. The product of these meetings was an addendum to the LS/CMI coding manual that offered additional rules for resolving coding problems that arose for officers in the field. At the conclusion of these group meetings, all officers administering the instrument attended daylong training sessions during which the training experts went over the coding manual, including the new addendum, and offered examples of how to apply each of the new addendum codes. The expert trainers reviewed the logic and theory underlying the LS/CMI, they discussed and reviewed all the administration and scoring rules and procedures, worked through scoring disagreements, and allowed time for officers to discuss and resolve disagreements among themselves.

The trainers administered a test to measure their efforts by asking participants to score one of two interview protocols that were identical except that the race of the client for some officers was African American and for others, White European American. The interview protocol needed to identify a specific racial/ethnic minority to present a meaningful client to the officers. Presenting a minority versus nonminority to test the effects in Study 1 would not have allowed the officers to activate their own race/ethnicity stereotypes, and therefore would not have tested officer bias. We chose to label the minority target as African American because Blacks represented the largest minority in Study 1.

The protocol was a transcription of an actual interview that an officer had recently conducted with a then, new client as part of a presentence investigation report. The client was a 24-year-old, unemployed high-school graduate who committed a Class I Misdemeanor—possession of a controlled substance with the intent to deliver. The offender had a lengthy criminal history going back to his juvenile years, including having been on probation on two separate occasions for prior charges of theft and driving under the influence of drugs. The transcript also provided additional information relevant to each of the criminogenic factors.

METHOD

Design and Test Stimuli

The research team revised the transcript of the 9-page protocol by leaving all the information about the offender and the offense record intact, but altering the names and demographics of the offender and his associates, so that they were either Black African American (Tyrone Ferguson) or White European American (Tanner Fitzgerald). In consultation with

the researchers, Nebraska probation trainers randomly assigned participants to one of two training conditions, either a pretest or a posttest condition. In the pretest condition, half of the participants received the stimulus materials and scored the target using the LS/CMI instrument before they received specialized training. In the posttest condition, the other half of the participants received specialized training before they received the stimulus materials and scored the same target using the LS/CMI. In summary, this was a 2 (Race of Target: White vs. Black) × 2 (Testing Status: Pretest—testing before training vs. Posttest—testing after the training) completely crossed between subjects experimental design. Thus, the trainers with the assistance of the authors randomly assigned each of the 213 participating officers to one of the four cells of the design.

Based upon the results of Study 1, we expected some biased scoring, but we anticipated that the booster training, designed to improve quality control and thereby constrain some of the officer discretion, would reduce any observed racial bias in the LS/CMI scores. Thus, we expected an interaction showing officer bias in the pretests with higher risk scores for Black than for White offenders but a reduction in this bias in the posttests.

Participants and design. Two hundred thirteen Nebraska probation officers (M_{age} = 42.19, *SD* = 11.03, range = 23-76) participated in Study 2.[1] As self-reported, there was a moderately uneven split between females (*n* = 113, 56.5%) and males (*n* = 87, 43.5%). The participants on average were all familiar with scoring and using the LS/CMI, with the average number of months administering the instrument for the 194 who reported this information equal to 63.64 months (or 5 years and 4 months), *SD* = 45.12 months, and range from 0 to 340 months (or 0-20 years). The participants largely identified their ethnicity as White European American (*n* = 183, 91.5%), supporting our concern that officers might be biased against the minority (largely Black offenders) because of their outgroup stereotypes. Much smaller numbers identified as Hispanic (*n* = 8, 4.0%), African American (*n* = 3, 1.5%), Native American (*n* = 1, 0.5%), and other (*n* = 5, 2.5%). The participants were well educated with a large majority having a bachelor's degree (*n* = 164, 82%) and the rest holding an advanced degree (*n* = 36, 18%). This high level of education reflects Nebraska's requirement that staff have at least a bachelor's degree to serve as a probation officer.

RESULTS AND DISCUSSION

Manipulation of Race

The probationer's name on the protocol was either one strongly associated with an African American (Tyrone Ferguson) or one associated with a White European American (Tanner Fitzgerald) as were the names of family members and friends (Shanice, Jamal, and Sylvia vs. Shirley, John, and Chris). In addition, the protocol contained a direct statement describing the client's race as either "Black" or "White" as is a standard practice on Nebraska presentence investigation materials. To check that the officers took notice of the client's race, the test forms asked them to recall the names of the offenders as they completed the LS/CMI scores for the client. Of the 213 officers completing the forms, 195 (92%) identified the correct Black or White name. None of the officers misidentified the name, but 18 (8%) left the name of the client blank. We took this as evidence that the manipulation of different racial names was successful.

LS/CMI Total Score and Criminogenic Analyses

We performed separate 2 (Race of Target: White vs. Black) × 2 (Testing Status: Pretest—testing before training vs. Posttest—testing after the training) between-groups ANOVAs on the scores of the LS/CMI's eight criminogenic factors, as well as on the total score. Table 5, which displays the results of these analyses, shows significant training effects for the following criminogenic factors: Leisure/Recreation, Offender Companions, Procriminal Attitudes, Antisocial Pattern, Education and Employment, and Family and Marital Issues. Table 6 displays the means and standard deviations for each of the four training status by race of the client cells: First, it is notable that the booster training reduced overzealous scores in some domains (i.e., Companions, Procriminal Attitudes, and Education/Employment) while increasing deflated scores in other domains (i.e., Leisure/Recreation and Family/Marital). It had no significant impact on Alcohol/Drug or Criminal History scores, and because training reduced some inflated scores and increased other deflated scores, it did not change the overall risk scores. Thus, it would appear while booster training was helpful to recalibrate officer scoring on the LS/CMI, the overall risk level remains relatively stable. Second, as seen in Table 5 there were no main effects or interactions involving client race. The means in Table 6 demonstrate that total risk scores and scores on all the LS/CMI domains were stable across race, showing no effects for race either before or after training. Thus, the Nebraska officers do not display any evidence of minority scoring bias, at least not for Black African American versus White European American clients.

OVERALL DISCUSSION

SUMMARY OF THE PROBLEM AND FINDINGS

We started this article pointing to the observation that high rates of incarceration affect racial minorities disproportionately as compared with Whites. We offered some statistics to support this conclusion, and a second conclusion that even in community corrections outside of prison there is evidence of disproportionate treatment of Blacks compared with Whites (The Sentencing Project: Research and Advocacy for Reform, 2016). We argued that disproportionate treatment and overcrowding were byproducts of the "tough on crime" approach of earlier decades, but current efforts using the RNR model, which emphasize valid risk assessment, evidence-based practice, and tailored services in the name of rehabilitation, have the potential to reduce overcrowding and disparate outcomes in the criminal justice system. However, the potential of the RNR approach to bring about a reduction in recidivism and in reducing disparate outcomes turns on the utilization of assessment tools that are themselves verified as valid and unbiased predictors of outcomes in the specific jurisdictions that apply them. Desmarais et al. (2016) as well as Monahan and Skeem (2016) agree that we need more work on the reliability and validity of current risk assessment tools with U.S. samples. The current manuscript studied the use of one such assessment instrument, the LS/CMI, as probation officers make use of it in Nebraska (Study 1). While our findings are not the final word on these issues with regard to the LS/CMI, let alone other risk assessment tools, they do offer some strong evidence that warrants the continued use of LS/CMI. Furthermore, they point the way to the need for additional studies in this state (Nebraska), as well as in other jurisdictions, to understand the sources of disproportionate justice and overcrowding in corrections.

Based upon prior findings in the literature, we expected that the LS/CMI would be a better predictor of return to the criminal justice system for nonminority probationers, and that

TABLE 5: Results of Analysis of Variance on the LS/CMI Total Score and Separate Factors by Officer Training Condition and Race of the Client (Study 2)

Effects	*df*	Mean square	*F*	*p*	η^2p
Total LS/CMI score					
Training	1,202	17.00	0.40	.526	.002
Race of client	1,202	12.24	0.29	.590	.001
Training by race of client	1,202	34.15	0.81	.369	.004
Leisure/recreation					
Training	1,202	26.28	34.53	.000	.146
Race of client	1,202	0.44	0.58	.448	.003
Training by race of client	1,202	1.95	2.56	.111	.013
Companions					
Training	1,202	8.43	7.60	.006	.036
Race of client	1,202	0.17	0.56	.693	.001
Training by race of client	1,202	1.28	1.15	.284	.006
Alcohol and drug problems					
Training	1,202	1.85	0.62	.432	.003
Race of client	1,202	1.59	0.54	.465	.003
Training by race of client	1,202	0.61	0.21	.651	.001
Procriminal attitude					
Training	1,202	5.14	5.91	.016	.029
Race of client	1,202	0.03	0.03	.860	.000
Training by race of client	1,202	0.18	0.21	.645	.001
Antisocial pattern					
Training	1,202	7.60	5.04	.026	.024
Race of client	1,202	4.78	3.17	.077	.015
Training by race of client	1,202	0.97	0.64	.424	.003
Criminal history					
Training	1,202	2.21	3.05	.082	.015
Race of client	1,202	0.22	0.30	.584	.001
Training by race of client	1,202	0.49	0.67	.414	.003
Education/employment					
Training	1,202	204.83	60.34	.000	.230
Race of client	1,202	1.81	0.53	.467	.003
Training by race of client	1,202	1.03	0.30	.583	.001
Family/marital					
Training	1,202	79.40	60.51	.000	.231
Race of client	1,202	0.24	0.18	.672	.001
Training by race of client	1,202	1.26	0.96	.327	.005

Note. LS/CMI = Level of Service/Case Management Inventory.

minority probationers would score higher on most, if not all, of the eight criminogenic factors that the instrument measures (Olver et al., 2014). The data only partially confirmed our hypotheses. The instrument performs about as well in Nebraska as elsewhere in the United States, such that a Nebraska probationer selected at random with a very high LS/CMI risk level is more than 8 times as likely to fail (return to probation or jail) than one selected at random with a very low-risk level.

We, however, did find evidence of an interaction between risk level and minority status, but to our surprise the effects showed that there was a slight advantage for the LS/CMI predictions of failure in the minority sample relative to the White European American sample. Although the effect was small, when we drilled down deeper, we found that the

TABLE 6: Criminogenic Factor Scores for Race by Officer Training Written Pre-Sentence Investigation Conditions

	Test before training		Test after training	
	Black	White	Black	White
Criminogenic factor	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Total LS/CMI score	23.43 (4.21)	24.73 (4.21)	23.67 (8.04)	23.34 (8.20)
Leisure/recreation**	0.80 (1.02)	1.08 (1.16)	1.71 (0.58)	1.60 (0.60)
Companions**	3.46 (1.00)	3.56 (0.90)	3.22 (1.01)	3.00 (1.26)
Alcohol/drug problem	5.63 (1.36)	5.56 (1.73)	5.55 (1.78)	5.26 (1.98)
Procriminal attitude*	0.96 (1.12)	1.00 (0.86)	0.71 (0.88)	0.62 (0.84)
Antisocial pattern*	1.54 (1.13)	1.98 (1.00)	2.06 (1.41)	2.23 (1.33)
Criminal history	5.46 (0.77)	5.63 (0.84)	5.35 (0.66)	5.32 (1.07)
Education/employment**	5.13 (1.30)	5.46 (1.30)	3.27 (2.34)	3.32 (2.16)
Family/marital**	0.46 (0.95)	0.69 (0.99)	1.86 (1.37)	1.77 (1.22)

*Significant training effect at $p < .05$. **Significant training effect at $p < .01$.

minorities scored higher on all risk domains except substance abuse when compared with the nonminority applicants. This raised the possibility that officers scored Black clients differently than White clients and if that were indeed the case, then the administration of the LS/CMI, even if the instrument itself is unbiased, could still lead to disproportional outcomes in community corrections and sentencing.

Study 2 ruled out this possibility, at least for race, using an experimental methodology that varied both booster training on the LS/CMI and the race of a test client. It is important to remember that the Black version and White version of the written transcripts of the test interview were identical except for race of the probationer. First, we found that the booster training had different effects on different LS/CMI factors reducing overzealous scores in some domains (i.e., Companions, Procriminal Attitudes, and Education/Employment) while increasing deflated scores in other domains (i.e., Leisure/Recreation and Family/Marital). It had no significant impact for Alcohol/Drug or Criminal History scores, and because training reduced some inflated scores and increased other deflated scores, it did not change the overall risk scores of the total instrument. Most importantly though for our purposes, the Nebraska officers scored the Black client the same as the White client both before and after the training. These results suggest that officer bias is unlikely to be the reason why minorities in Study 1 scored differently than nonminorities on the LS/CMI domains.

Lessons Learned

The current research offers several lessons for applied research on risk assessment instruments. To begin with, there was no evidence that LS/CMI risk level interacts with probationer gender to predict recidivism in our Nebraska sample even though others (Reisig et al., 2006) find differential predictive validity as a function of client gender. Future studies should examine how risk assessment tools perform for both female and male offenders. Furthermore, although the Olver et al. (2014) meta-analyses of the LSI scales, one of which is the LS/CMI, showed a significant effect for minority status—risk levels predicted recidivism more strongly for nonminorities than for minorities—our Nebraska data showed a

much smaller interaction effect, and it was in the opposite direction. That is, if anything we found that the LS/CMI predicted outcomes better for minorities than for nonminorities. Lesson 1 from this current investigation is that it is important to investigate the validity of risk assessment instruments in the jurisdictions in which professionals use them (Desmarais et al., 2016; Monahan & Skeem, 2016). Although minorities scored higher on seven of eight criminogenic factors in Study 1, Study 2 draws into serious question whether this was the result of officer bias. Instead, it is more likely that either instrument construction bias (Shepherd & Lewis-Fernandez, 2016) or differences in social and economic status of minority clients led to higher scores on these factors. The latter explanation is more likely due to the weak interaction between race and risk level in Study 1. This brings us to Lesson 2—research exploring the sources of possible bias in the administration of instruments in a variety of jurisdictions is necessary to help understand why minorities and nonminorities sometimes score differently on risk-level instruments.

Finally, we conducted two different types of studies: the first, a large archival investigation of ongoing assessments over time and the second, a smaller but much better controlled field experiment, to examine the role that minority status might play in the risk assessment process in community corrections. The results of the two studies together added to our knowledge of how the RNR model works in Nebraska, allowing us to find support for the validity of the LS/CMI as utilized in that state. The third lesson emerging from this research is the importance of a multimethodological strategy for studying applied criminal justice issues. No one methodology, by itself, can ever tell us as much as do the converging, or even nonconverging results from different types of studies.

LIMITATIONS AND FUTURE RESEARCH

As is the case with all applied research, there are limitations to these studies' findings and, as is also often the case, those limitations point toward the need for additional investigations: First, our work did not examine closely the item composition of the LS/CMI scale and in light of Shepherd and Lewis-Fernandez's (2016) important cultural insights, future work would do well to examine each of the items in the scale composition of the criminogenic domains, both from an empirical item analysis perspective and from a cultural sensitivity point of view. Second, while Study 1 examined overall minority differences, Study 2 in our work focused on differences between Black African American and White European American clients. Future work needs to consider the differences in evaluator judgments for other types of minorities, including Latinos, religious minorities (Muslims), and lesbian, gay, bisexual, and transgender (LGBT) individuals, to name just a few. In the end, our criminal justice system should perform the same regardless of the social dimensions of the clients and victims that the system serves. Third, although our data represent probationers across a 5-year window, the current study did not examine changes in instrument validity over time or changes in the effects of minority status overtime. Longitudinal research examining the stability of risk assessment overtime is important to the criminal justice system to allow us to ensure that the system does not drift in systematic ways to the disadvantage of one or more minority populations. Fourth, Study 2 presented to officers one client who had committed a substance offense, which limits its generalizability. Additional experimental studies that present different types of probationers who have committed different types of offenses are most certainly needed in this research area.

FINAL CONCLUSION

This research, combining archival analyses with a field experiment, finds support for the validation and application of one common risk instrument as it operates in one state. However, it also demonstrates that minority offenders in that same state show elevated risk factors on the validated instrument. Perhaps, the most important conclusion of this work is the need for continued work to understand how the contexts in which they live, the social environments that they inhabit, and the economic conditions in which they struggle elevate the criminological risk of minorities. Only when we understand more about this bigger issue regarding scores on criminogenic factors, will we be able to address directly the problems of overrepresentation of minorities in the criminal justice system and the disproportionate outcomes that they endure.

NOTE

1. For the Demographics Questionnaire, some participants did not provide information leaving this section blank (n = 13, 6.1%). Thus, all sample descriptions are based upon 200 responses.

REFERENCES

Andrews, D. A., & Bonta, J. (1995). *The Level of Supervision Inventory–Revised* (Measurement instrument). Toronto, Ontario, Canada: Multi-Health Systems.

Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, *16*, 39-55. doi:10.1037/a0018362

Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *The Level of Service/Case Management Inventory (LS/CMI)* [Measurement instrument]. Toronto, Ontario, Canada: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The Risk-Need-Responsivity (RNR) model: Does adding the good lives model contribute to effective crime prevention? *Criminal Justice and Behavior*, *38*, 735-755. doi:10.1177/0093854811406356

Andrews, D. A., & Dowden, C. (2006). Risk principle of case classification in correctional treatment: A meta-analytic investigation. *International Journal of Offender Therapy and Comparative Criminology*, *50*, 88-100. doi:10.1177/0306624X05282556

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, *28*, 369-404.

Benedict, W. R., & Huff-Corzine, L. (1997). Return to the scene of the punishment: Recidivism of adult male property offenders on felony probation, 1986–1989. *Journal of Research in Crime and Delinquency*, *34*, 237-252. doi:10.1177/0022427897034002004

Burnett, R., & Maruna, S. (2004). So "prison works," does it? The criminal careers of 130 men released from prison under home secretary, Michael Howard. *The Howard Journal of Crime and Justice*, *43*, 390-404. doi:10.1111/j.1468-2311.2004.00337.x

Canada v. Ewert, 2016 FCA 203 (2016).

Chenane, J. L., Brennan, P. K., Steiner, B., & Ellison, J. M. (2015). Racial and ethnic differences in the predictive validity of the Level of Service Inventory–Revised among prison inmates. *Criminal Justice and Behavior*, *42*, 286-303. doi:10.1177/0093854814548195

Clarke, S. H., Lin, Y. H. W., & Wallace, W. L. (1988). *Probationer recidivism in North Carolina: Measurement and classification of risk*. Chapel Hill: Institute of Government, University of North Carolina.

Desmarais, S. L., Johnson, K. L., & Singh, J. P. (2016). Performance of recidivism risk assessment instruments in U.S. correctional settings. *Psychological Services*, *13*, 206-222. doi:10.1037/ser0000075

Dowden, C., & Andrews, D. A. (1999a). What works for female offenders: A meta-analytic review. *Crime & Delinquency*, *45*, 438-452. doi:10.1177/0011128799045004002

Dowden, C., & Andrews, D. A. (1999b). What works in young offender treatment: A meta-analysis. *Forum on Corrections Research*, *11*, 21-24. Ottawa, Ontario: Correctional Service of Canada.

Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*, *42*, 449-467.

Dowden, C., & Andrews, D. A. (2003). Does family intervention work for delinquents? Results of a meta-analysis. *Canadian Journal of Criminology and Criminal Justice*, *45*, 327-342. doi:10.3138/cjccj.45.3.327

Ewert v. Canada, 2015 F.C. 1093 (2015).

Freeman, R. (2003). *Can we close the revolving door? Recidivism vs. employment of ex-offenders in the U.S* (Unpublished manuscript). New York University Law School, New York University, New York, New York.

Gendreau, P., & Andrews, D. A. (1990). Tertiary prevention: What the meta-analyses of the offender treatment literature tell us about what works. *Canadian Journal of Criminology, 32*, 173-184.

Gendreau, P., Cullen, F. T., & Goggin, C. (1999). *The effects of prison sentences on recidivism*. Ottawa, Ontario: Solicitor General Canada.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the "unparalleled" measure of offender risk? A lesson in knowledge cumulation. *Criminal Justice and Behavior*, *29*, 397-426. doi:10.1177/0093854802029004004

Green, D. P., & Winik, D. (2010). Using random judge assignments to estimate the effects of incarceration and probation on recidivism among drug offenders. *Criminology*, *48*, 357-387. doi:10.1111/j.1745-9125.2010.00189.x

Hamilton, Z., Kigerl, A., Campagna, M., Barnoski, R., Lee, S., Van Wormer, J., & Block, L. (2016). Designed to fit: The development and validation of the STRONG-R recidivism risk assessment. *Criminal Justice and Behavior*, *43*, 230-263. doi:10.1177/0093854815615633

Hannah-Moffat, K. (2006). Pandora's box: Risk/need and gender-responsive corrections. *Criminology & Public Policy*, *5*, 183-192. doi:10.1111/j.1745-9133.2006.00113.x

Hannah-Moffat, K. (2009). Gridlock or mutability: Reconstructing "gender" and risk assessment. *Criminology & Public Policy*, *8*, 209-219. doi:10.1111/j.1745-9133.2009.00549.x

Kaeble, D., & Glaze, L. (2016). *Correctional populations in the United States, 2015* (DOJ Publication No. NCJ 248029). Washington, DC: Bureau of Justice Statistics. Retrieved from https://www.bjs.gov/content/pub/pdf/cpus15.pdf

Kaeble, D., Glaze, L. E., Tsoutis, A., & Minton, T. D. (2015). *Correctional populations in the United States, 2014* (DOJ Publication No. NCJ 249513). Washington, DC: Bureau of Justice Statistics. Retrieved from http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5519

Luther, J. B., Reichert, E. S., Holloway, E. D., Roth, A. M., & Aalsma, M. C. (2011). An exploration of community reentry needs and services for prisoners: A focus on care to limit return to high-risk behavior. *AIDS Patient Care and STDs*, *25*, 475-481. doi:10.1089/apc.2010.0372

McGaha, J., Fichter, M., & Hirschburg, P. (1987). Felony probation: A re-examination of public risk. *American Journal of Criminal Justice*, *12*, 1-9. doi:10.1007/BF02887535

Monahan, J., & Skeem, J. L. (2016). Risk assessment in criminal sentencing. *Annual Review of Clinical Psychology*, *12*, 489-513. doi:10.1146/annurev-clinpsy-021815-092945

Nagin, D. S., Cullen, F. T., & Jonson, C. L. (2009). Imprisonment and reoffending. *Crime and Justice*, *38*, 115-200. doi:10.1086/599202

New York State, Division of Criminal Justice Services. (2009). *2009 CrimeStat report*. Albany: Author. Retrieved from http://www.criminaljustice.ny.gov/pio/annualreport/2009-crimestat-report.pdf

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior*, *36*, 329-353. doi:10.1177/0093854809331457

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service Scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, *26*, 156-176. doi:10.1037/a0035080

Parent, G., Guay, J. P., & Knight, R. A. (2012). Can we do better? The assessment of risk of recidivism by adult sex offenders. *Criminal Justice and Behavior*, *39*, 1647-1667. doi:10.1177/0093854812451680

Petersilia, J. (1985). Probation and felony offenders. *Federal Probation*, *49*, 4-12.

Petersilia, J., Turner, S., Kahan, J., & Peterson, J. (1985). Executive summary of Rand's study, "Granting felons probation: Public risks and alternatives." *Crime & Delinquency*, *31*, 379-392. doi:10.1177/0011128785031003004

Reisig, M. D., Holtfreter, K., & Morash, M. (2006). Assessing recidivism risk across female pathways to crime. *Justice Quarterly*, *23*, 384-405. doi:10.1080/07418820600869152

Rettinger, J. (1998). *A recidivism follow-up study investigating risk and need within a sample of provincially sentenced women* (Unpublished doctoral dissertation). Department of Psychology, Carleton University, Ottawa, Ontario, Canada. Retrieved from https://www.collectionscanada.gc.ca/obj/s4/f2/dsk2/tape15/PQDD_0002/NQ37054.pdf

Rowe, R. C. (1999). *The prediction of recidivism in a parole sample: An examination of two versions of the Level of Service Inventory* (Unpublished manuscript). Carleton University, Ottawa, Ontario, Canada.

Schwalbe, C. S. (2007). Risk assessment for juvenile justice: A meta-analysis. *Law and Human Behavior*, *31*, 449-462. doi:10.1007/s10979-006-9071-7

Schwalbe, C. S. (2008). A meta-analysis of juvenile justice risk assessment instruments: Predictive validity by gender. *Criminal Justice and Behavior*, *35*, 1367-1381. doi:10.1177/0093854808324377

The Sentencing Project: Research and Advocacy Forum. (2010). *State recidivism studies*. Washington, DC: Author. Retrieved from https://www.prisonlegalnews.org/media/publications/the_sentencing_project_state_recidivism_studies_compendium_2010.pdf

The Sentencing Project: Research and Advocacy Forum. (2016). *Reducing racial disparity in the criminal justice system: A manual for practitioners and policymakers*. Washington, DC: Author. Retrieved from http://www.sentencingproject.org/wp-content/uploads/2016/01/Reducing-Racial-Disparity-in-the-Criminal-Justice-System-A-Manual-for-Practitioners-and-Policymakers.pdf

Shepherd, S. M., & Lewis-Fernandez, R. (2016). Forensic risk assessment and cultural diversity: Contemporary challenges and future directions. *Psychology, Public Policy, and Law*, *22*, 427-438. doi:10.1037/law0000102

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, *31*, 499-513. doi:10.1016/j.cpr.2010.11.009

Smith, P., Cullen, F. T., & Latessa, E. J. (2009). Can 14,737 women be wrong? A meta-analysis of the LSI-R and recidivism for female offenders. *Criminology & Public Policy*, *8*, 183-208. doi:10.1111/j.1745-9133.2009.00551.x

Stageberg, P., & Wilson, B. (2005). *Recidivism among Iowa probationers*. Des Moines: Iowa Division of Criminal and Juvenile Justice Planning. Retrieved from https://humanrights.iowa.gov/sites/default/files/media/CJJP_Recidivism_Among_Iowa_Probationers.pdf

The State of Texas, Legislative Budget Board. (2013). *Statewide criminal justice recidivism and revocation rates*. Retrieved from http://www.lbb.state.tx.us/Documents/Publications/Policy_Report/Statewide%20Criminal%20Justice%20Recidivism%20and%20Revocation%20Rates2009.pdf

Taylor-Gooby, P., & Zinn, J. O. (2006). Current directions in risk research: New developments in psychology and sociology. *Risk Analysis*, *26*, 397-411. doi:10.1111/j.1539-6924.2006.00746.x

Vito, G. F. (1987). Felony probation and recidivism: Replication and response. *Federal Probation*, *50*, 17-25.

Walmsley, R. (2009). *World prison population list* (8th ed.). London, England: International Centre for Prison Studies. Retrieved from http://www.prisonstudies.org/sites/default/files/resources/downloads/world_prison_population_list_11th_edition_0.pdf

Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*, 740-767. doi:10.1037/a0020473

Alisha Caldwell Jimenez, JD, is a graduate student in the Law and Psychology Program at the University of Nebraska-Lincoln. She earned her Juris Doctorate from the University of Nebraska College of Law in 2016 and expects to complete her PhD in May 2019, with an interest in juvenile justice, program development and implementation, and program evaluation, and public policy.

R. Hazel Delgado, MA, is currently employed at the Nebraska Court Improvement Project as a Research and Evaluation Specialist. She received her master's degree from the University of Nebraska-Lincoln where she is currently pursuing a PhD in social psychology. She is expected to complete her PhD in 2018, and has received the National Science Foundation, Law and Social Sciences Doctoral Dissertation award.

Trace C. Vardsveen, JD, received his Juris Doctorate from the University of Nebraska College of Law in 2015. He is currently pursuing a PhD in social psychology with an interest in psycho-legal research on criminal justice, punishment, and discrimination.

Richard L. Wiener, PhD, MLS, holds a doctoral degree in psychology and a master's in legal studies. He is currently a faculty member in the Law and Psychology Program at the University of Nebraska Lincoln. He works closely with the Nebraska judiciary and the Nebraska Administrative Office of Probation. He has published numerous articles in the area of social psychology and law, and has received funding from the National Science Foundation and the National Institute of Justice.