

The Predictive Validity of the LSI-R on a Sample of Native American Females

Research Brief

An Understudied Population

The use of Risk-Need Assessments (RNAs) in criminal justice has proliferated throughout the last three decades. One of the largest advantages that RNAs provide is the ability to objectively predict the risk that an individual will recidivate. In theory, this should reduce recidivism by identifying high risk individuals to target for treatment and reduce the bias that is inherent in human decision making (Taxman & Dezember, 2016). Although research shows that using RNAs can provide more accurate and equitable predictions of risk than professional judgement (Lin et al., 2020), the degree of accuracy and fairness varies across racial/ethnic groups (Butler et al., 2022) and between genders (Miller et al., 2021). Specifically, research suggests that RNAs tend to overclassify non-White groups and women (Hamilton et al., 2016; Miller et al., 2021). This issue is especially salient for Native Americans (Holsinger et al., 2003). To date, Native Americans are one of the most under-studied justice populations (Fremon, 2021). The lack of criminal justice research on the risk and protective factors of Native Americans obscures our ability to accurately assess their risk.

Summary

As risk assessments seek to predict the likelihood of recidivism across demographic populations, it is essential that their measurements are racially and gender invariant. Specifically, it is important to validate assessments using different populations and demographic groups to ensure that their predictive accuracy is commensurate. The current project used an all-female sample from North Dakota to assess the predictive accuracy of the LSI-R for Native American females.

Findings revealed that recidivism rates for Native Americans did not differ significantly between risk level classifications. Further, it was found that the tool's predictive accuracy was greatly diminished for the Native American sample. Native Americans scored an average of two points higher on the LSI-R, despite them possessing a 12% reduced probability of recidivating, although the latter finding did not reach statistical significance. Finally, it was found that two demographic variables may explain the overclassification in LSI-R scores for Native Americans. Being homeless and not having a college education significantly increased LSI-R scores and were more strongly correlated with being Native American. Findings highlight the need for RNA developers and practitioners to identify risk and protective factors that are more representative for the female Native American population.

Inaccurate assessments of risk can exacerbate existing racial/ethnic disparities in corrections. Further, systematically overclassifying subpopulations of individuals can strain already limited correctional resources. Finally, imposing intensive correctional interventions on those who are not truly high risk has been shown to have adverse effects, precluding proper reintegration into the community, and increasing the potential for recidivism (Sperber et al., 2013). This brief seeks to evaluate the predictive validity of the Level of Service Inventory – Revised (LSI-R) on Native American populations. Specifically, predictive validity will be assessed using a female-only sample from a community corrections center in North Dakota.

Background

The LSI-R is billed as a “demographically blind” assessment, meaning that it does not measure demographic information like race or gender (Taxman et al., 2013, pg. 83). The creators of the LSI-R posit that the LSI-R demonstrates, 'gender-neutral' and equitable predictive validity between genders and across races/ethnicities (Bonta & Andrews, 2016). More recently, risk assessment research has cast doubt on this assumption of invariance among subgroups (Ejk et al., 2017; Van Voorhis, 2009). For example, in a meta-analysis, Wormith and colleagues (2015) found that the LSI-R demonstrated diminished predictive validity for Aboriginal populations in Canada. This finding and others like it (Gutierrez et al., 2013; Olver et al., 2014) demonstrate that testing the predictive validity of assessments, especially how they perform on populations that differ from their development samples, is important for assessment developers and agencies seeking to adopt such tools (Bucklen et al., 2021; Holsinger et al., 2006).

Unequal predictive accuracy leads to overclassification

i.e., being scored as higher risk than is appropriate) through individuals being scored as similar risk, despite possessing different probabilities of reoffending. This is referred to as intercept *bias* (American Educational Research Association, 2018). Similarly, the term *slope bias* is used to describe a relationship where a one unit increase in risk score is associated with a commensurate increase in the probability of recidivism for one group, but not another. In other words, slope bias can occur when recidivism rates between two groups are similar at lower levels of risk, but an increase in risk score is associated with a greater increase in the probability of recidivism for one group over another.

Current study

Differences across races/ethnicities lead to inequitable predictions of risk, which result in overclassification. This issue has shown to be especially salient for Canadian Aboriginals (Wormith et al., 2015) and Native Americans (Holsinger et al., 2006). As such, this research brief tests the predictive validity of the LSI-R on groups of Native American and non-Native American women. First, overclassification is examined by comparing the probability of recidivism across risk level categories. Next, the relationship between LSI-R scores and the probability of reoffending is compared between race/ethnicities. Finally, implications and recommendations for future research are discussed.

Sample

Data for this study were collected from a female community correctional facility in a North Dakota (N=375). Individuals were referred to this facility by the North Dakota Department of Corrections, or by the Bureau of Prisons. The facility acts as a reentry halfway house, where residents can leave the facility for pro-social activities, such as work.

Most of the individuals in the sample were near the end of their sentences and were transferred to the community facility for transitional services. The other participants were parolees or probationers who were referred due to technical violations. The majority of those admitted have many rehabilitative needs and are either homeless or have domestic situations that are inhibitive to their recovery or are high-risk for substance use relapses (North Dakota Department of Corrections, 2018). **Table 1** below provides sample descriptive statistics for the three described measures.¹

Recidivism was measured dichotomously and defined as any new criminal conviction² that resulted in a new sentence. Recidivism was assessed from the date that individuals were admitted to the community corrections center, with a 2-year follow-up. Recidivism data was gathered from North Dakota and Minnesota³ public open court records databases. The main variables of interest were race/ethnicity and LSI-R score. Participants were assessed prior to transfer to the community facility. Race/Ethnicity was coded as Native American (N=156) and non-Native American⁴ (N=219).

Analytic Strategy

First, cross-tabulations were conducted for the Native American sample to assess if recidivism rates were different between the Low (≤ 30), Moderate (31-

39) and High (≥ 40) risk-level categories.⁵

Next, a binary logistic regression model was computed to assess intercept bias. The binary race/ethnicity variable and continuous LSI-R scores were used to predict recidivism. Area Under the Curve (AUC) statistics⁶ were calculated to assess the LSI-R's predictive accuracy across the two racial/ethnic groupings. Further, an interaction was created for both racial categories and LSI-R score to assess the combined impact of LSI-R score and the racial/ethnic grouping on probability of recidivism to assess slope bias. As a means to explore how specific assessment items can lead to the overclassification of Native American females, linear regression models were constructed to examine demographic variables that were significantly associated with LSI-R scores, and bivariate correlations were computed to assess whether these predictors were associated with one subgroup stronger than the other.

Table 1. Descriptive Statistics of the Population (N=375)

	Native American (n=156)	Non-Native American (n=219)	Total
Variables	%/Mean	%/Mean	%/Mean
Recidivated	47%	52%	50%
Risk Level	39	37	38
Low (≤ 30)	17%	29%	46%
Moderate (31-39)	44%	36%	39%
High (≥ 40)	40%	35%	37%

¹ See Appendix 1 for the full table.

² Only convictions more severe than a petty misdemeanor were recorded (e.g., more severe than a traffic citation).

³ Records were collected from Minnesota due to the close proximity of the correctional center to the Minnesota boarder.

⁴ The majority of the non-Native American group are White individuals. A small number of Black (n=10) and Hispanic (n=13) are also included in this category.

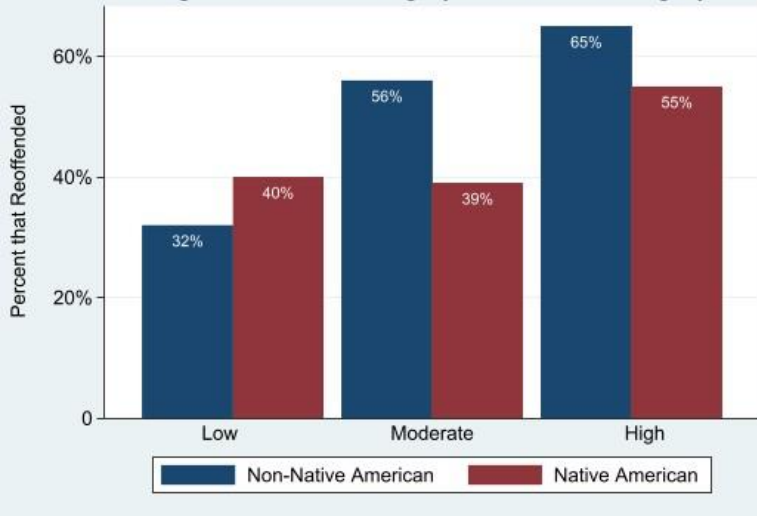
⁵ While the LSI-R provides four, only three risk level categories were assessed in this study due to the small sample size of Native American individuals having LSI-R scores less than 20. The Low (≤ 20) and Low/Moderate (21-30) categories were collapsed into a single risk level. Due to the small sample size for the Low category, a Fisher's exact test was used.

⁶ AUCs are used to determine if assessments demonstrate (.556), moderate (.639), or strong (.714) predictive validity (Rice & Harries, 2005).

Results

Figure 1 displays a bar chart of the cross-tabulation findings indicating non-significant differences in recidivism rates between risk level categories for Native Americans.⁷ Recidivism rates were higher for Native Americans categorized as low risk compared to moderate risk, indicating inaccurate measurements of risk. Additionally, Native Americans recidivated 17% and 10% less than non-Native Americans who were categorized as Moderate and High risk.⁸

Figure 1. Reoffending by Risk Level Category

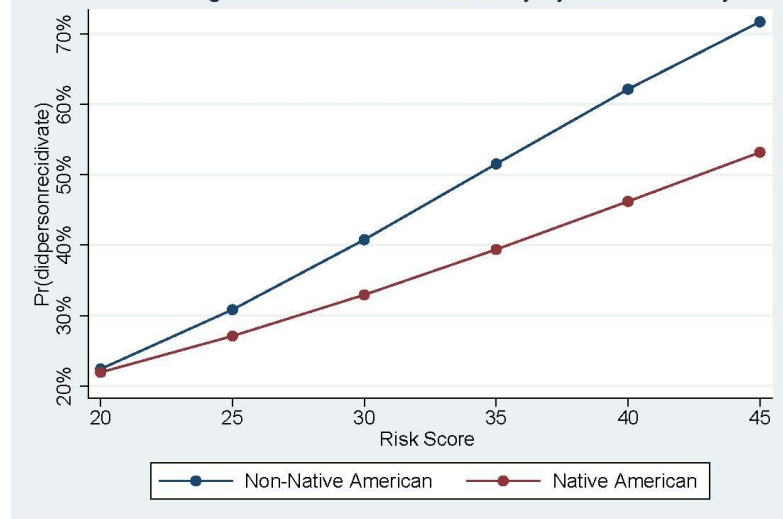


Predictive Validity

The LSI-R demonstrates moderate predictive validity for the non-Native American group (AUC=.66) and weak predictive validity for the Native American group (AUC=.60). The predictive validity of the LSI-R for diminished for Native Americans as risk scores increased. The assessment demonstrated moderate predictive strength for low-risk Native Americans (AUC=.66), and the predictive validity for moderate-

risk individuals was fairly equivalent between Native Americans and non-Native Americans (AUCs =.57 and .55 respectively). Unfortunately, the disparity in risk prediction increased between the high risk non-Native American (AUC=.58) and Native American groups (AUC=.49), revealing the presence of slope bias. Intercept bias was also observed, where Native American females have a lower probability of reoffending. **Figure 2** displays the probability of reoffending as LSI-R scores increase. While overclassification averages 10%, it becomes more severe as risk levels increase, which is indicative of slope bias. The degree of overclassification is roughly 1-5% for lower risk and 15-20% for higher risk females. Further, high risk Native American females recidivate at a rate that is similar to that of moderate risk non-native American females.

Figure 2. Recidivism Probability by Race/Ethnicity



⁷ Low and Moderate ($\chi^2 < .01$, $p = .93$), Moderate and High ($\chi^2 = 3.40$, $p = .065$). Appendix Table 2 displays the results.

⁸ This relationship was further examined using binary logistic regression, revealing non-significant differences in recidivism ($p = .69$). Appendix table 3 displays the results.

Influential Predictors in LSI-R Classification

Given the identified bias among the Native American sample, a linear regression was computed to explore the possibility that specific items measured by the LSI-R do not strongly predict recidivism for Native American females, which would lead to overclassification. The findings show that being Native American is significantly associated ($p=.01$) with over a two-point increase ($b=2.19$) in LSI-R risk score when controlling for age, education, marital status, homelessness, employment, and having a violent history⁹. This finding is notable considering no significant differences in recidivism were found. Other significant predictors of LSI-R score were being homeless ($p=.02$, $b=.18$) and not having a college education ($p<.01$, $b=-2.4$). A larger proportion of Native Americans were homeless (31% vs. 29%). Similarly, a slightly larger proportion of non-Native Americans had received at least some college education (44% vs 40%). Notably, level of education was unknown for 26% of the Native American sample, compared to just 19% for non-Native Americans. Addressing the higher level of missing data for Native Americans, as well as the disproportionality in homelessness and education between the two groups, may well serve as a starting point for future research regarding the accurate assessment and classification for this population.

Conclusion

Overall, these findings demonstrate the LSI-R's diminished predictive validity for Native American when compared to non-Native American females. Findings indicate that recidivism rates for Native American females were not significantly different between risk level categories, suggesting that the

assessment was unable to consistently distinguish risk of recidivism. Further, bias is indicated on two levels, by intercept and by slope, where Native Americans are overclassified as higher risk and this over classification increases as risk scores increase.

Since intensive supervision measures are typically reserved for higher risk individuals, overclassification leads correctional agencies to misallocate supervision resources. Further, imposing intensive interventions on individuals who are not high risk is known to increase their likelihood of reoffending (Bonta & Andrews, 2016).

Further, the LSI-R demonstrated weaker predictive validity for Native Americans. This study also found that Native Americans in the sample overall had a 12% reduction in the odds of recidivating despite scoring an average of two points greater on the LSI-R than non-Native Americans, although only the latter finding was statistically significant. Finally, research indicated that the other variables that were significantly associated with LSI-R score were inversely correlated between the Native American and non-Native American groups, with the direction of the association favoring Native Americans to score higher on the LSI-R.

Although this final finding is far from conclusive, it is a step in the direction of identifying demographic characteristics that may cause Native Americans to be less accurately assessed for risk. RNA developers should carefully consider the characteristics of local populations and subgroups to create fair and accurate assessments. Novel research in this area has found that considering the characteristics of the local jurisdiction, as well as racial/ethnic and gender

⁹ See appendices table 4 for the tabled results of the linear regression model.

differences, can reduce disparities and enhance predictive accuracy (Duwe, 2014; Duwe & Rocque 2019; Hamilton et al. 2016, 2019, 2020). Crafting jurisdiction-specific assessments also allows RNA developers to better identify and account for unique risk and protective factors for substantive populations, reducing bias (Hamilton et al., 2022). Additionally,

agencies seeking to adopt an RNA should consider working with researchers to create an assessment that is tailored to each agency's population and specific needs (Bucklen et al., 2021). By reducing bias and increasing accuracy, agencies will be better able to allocate resources appropriately, which, in turn, will decrease recidivism rates.



Appendices

Table 1. Descriptive Statistics of the Population (N=375)-

	Native American (n=156)	Non-Native American (n=219)	Total
Variables	%/Mean(SD)	%/Mean(SD)	%/Mean(SD)
Recidivated	47%	52%	50%
LSI-R Score	39_(5.5)	37_(6.8)	38_(7.3)
0-20	4%	5%	5%
21-30	13%	24%	19%
31-39	44%	36%	39%
>39	40%	35%	37%
Age	33_(6.9)	34_(8.8)	33_(7.7)
18-27	22%	25%	23%
28-37	56%	48%	51%
38-47	18%	22%	20%
48-58	5%	6%	5%
Violent History	26%	32%	29%
Homeless	31%	29%	30%
Employed	63%	66%	65%
Marital Status			
Single	51%	60%	56%
Divorced	11%	11%	11%
Married	8%	7%	7%
Separated	11%	4%	10%
Engaged	2%	2%	2%
Unknown	18%	11%	18%
Education			
No Highschool			
Diploma/GED	17%	7%	11%
Highschool/GED	40%	44%	42%
Some College/			
Degree	17%	30%	25%
Unknown	26%	19%	22%

Suggested citation: Ursino, J. (2023). The Predictive Validity of the LSI-R on a Sample of Native American Females (Doc. # 2023-002). Nebraska Center for Justice Research. Omaha, NE.

Table 2. Cross Tabulations of Risk Level Categories for Native Americans_(N=156).

Recidivated	Low_(n=10)	Moderate(n=70)	Total%	χ^2 (phi)
Yes	4	27	61%	<.01(.01)
No	6	43	39%	

p=.93

Recidivated	Moderate_(n=70)	High_(n=76)	Total%	χ^2 (phi)
Yes	27	42	47%	3.4(.15)
No	43	34	53%	

p=.065

Table 3. Binary Logistic Regression Model Predicting Recidivism (N=375)¹⁰.

Variables	b	SE	p-value	OR
Native American	0.12	0.31	0.69	0.88
LSIR	0.06	0.02	0.008*	1.07
Age	0.04	0.02	0.03*	0.97
Homeless	0.23	0.28	0.41	1.26
Marital Status				
Divorced	0.23	0.43	0.60	1.26
Married	0.25	0.60	0.67	0.78
Separated	0.73	0.60	0.22	2.08
Engaged	0.06	0.76	0.93	0.94
Violent History	0.13	0.28	0.65	0.88
Employed	0.11	0.30	0.71	0.89
Education				
Some College/Degree	0.198	0.326	0.54	1.22
Less than Highschool/No GED	0.49	0.44	0.27	0.61

* Indicates p-value <.05

¹⁰ Reference categories are non-Native American, not homeless, single, not having a violent history, unemployed, and not having a high school degree/GED.

Table 4. Linear Regression Model of Variables Related to LSI-R Scores (N=375)¹¹

Variables	b	SE	p-value	β
Native American	2.19	0.85	0.01*	0.15
Age	0.01	0.05	0.76	0.019
Homeless	0.18	0.74	0.02	0.14
Marital Status				
Divorced	1.77	1.16	0.13	0.096
Married	-0.33	1.48	0.82	-0.01
Separated	1.20	1.51	0.43	0.05
Engaged	-0.86	2.1	0.68	-0.02
Violent History	0.84	0.74	0.26	0.07
Employed	0.19	0.79	0.81	0.014
Education				
Some				
College/Degree	-2.4	0.83	<0.001*	-0.18
Less than				
Highschool/No GED	0.543	1.2	0.65	0.03

* Indicates p-value <.05

¹¹ Reference categories are non-Native American, not homeless, single, not having a violent history, unemployed, and not having a high school degree/GED.

 | NEBRASKA CENTER FOR JUSTICE RESEARCH

6001 Dodge Street, Omaha, NE 68182 | 402.554.3794 | justiceresearch.unomaha.edu
