Exploring the validity of the Level of Service Inventory-Revised with Native American offenders

Alexander M. Holsinger a,⁎, Christopher T. Lowenkamp b, Edward J. Latessa c

a 208 Haag Hall, Department of Sociology/Criminal Justice and Criminology, University of Missouri - Kansas City, Kansas City, MO 64110-2499, United States
b Corrections Institute, University of Cincinnati, P. O. Box 210389, Cincinnati, OH 45221-0389, United States
c Division of Criminal Justice, University of Cincinnati, P. O. Box 210389, Cincinnati, OH 45221-0389, United States

Abstract

The use of actuarial risk/need assessment tools is an increasingly important part of the correctional landscape. Actuarial tools ideally will provide a valid, dynamic assessment of an offender’s overall risk/need level, and will identify their most prevalent criminogenic needs. What results is typically a number or score that can be used to assign an offender to a risk level that is associated with an assumed likelihood of recidivism. Testing the predictive validity of actuarial risk/need assessment tools is of paramount concern, particularly when they are utilized with new (and under-researched) populations. The current study assessed the predictive validity of the Level of Service Inventory-Revised using a sample of Native American and White offenders in a northern midwestern state. Results showed the instrument to have modest predictive validity utilizing the entire sample of offenders, with varying results for subsequent subgroups.

© 2006 Elsevier Ltd. All rights reserved.

Introduction

Dynamic risk/need assessment processes have become an increasingly important part of virtually any correctional intervention. In previous decades, the most common method by which a correctional professional would get to know an offender would be through an unstructured (or semi-structured) interview, that may or may not involve information pertaining to relevant criminogenic factors. This ‘first generation’ of offender assessment is largely based on the prior experience and qualitative observations of the correctional professional. This method also possesses a high potential for bias to enter into decision-making (Bonta, 1996; Clear & Gallagher, 1985).

The practice of offender classification has undergone many changes via advances in research identifying the best predictors of antisocial behavior. Specifically, since the ‘first generation’ method of classification referenced above, correctional practice has involved the use of many different actuarial-based instruments involving both static (historic/unchanging) and dynamic (current/changeable) factors. Most of these tools provide a summary score that dictates a certain level of supervision or intervention (see for example the salient factors score, or the Wisconsin risk/need-CMC) (Baird, Heinz, & Bemus, 1979; Hoffman, 1994).

Perhaps of most recent note, the Level of Service Inventory (LSI-R) has been used with increasing frequency over the last ten years (Jones, Johnson, Latessa, & Travis, 1999). Developed by Andrews and...
Bonta (1995a), the LSI-R is a fifty-four-item assessment tool that results in a composite score indicating an offender’s overall level of risk, and criminogenic need. The composite score is rendered via the assessment of ten ‘domains,’ most of which are assessed in a dynamic way—meaning, as the domain exists in the offender’s life (e.g., over the last twelve-month period the offender was in the community) (Andrews & Bonta, 1995b; Bonta & Motiuk, 1985). The ten domains covered via the LSI-R include: criminal history, education/employment, financial, family/marital circumstances, accommodations, leisure/recreation activities, companions, alcohol/drug use, emotional/personal, and attitudes/orientations (criminogenic thinking). In theory, assessing an offender using the LSI-R allows for a valid classification (e.g., high, medium, or low-risk), as well as the identification of the most prevalent criminogenic needs that may facilitate case planning and treatment intervention (Andrews & Bonta, 2001).

The predictive validity of the LSI-R had been demonstrated within several different correctional settings (Andrews, 1982; Andrews & Robinson, 1984; Bonta & Andrews, 1993; Bonta & Motiuk, 1985; Gendreau, Goggin, & Smith, 2002). In addition, there was evidence that the LSI-R had predictive validity for various subgroups of the offender population such as female offenders, and African-American offenders (Coulson, Ilacqua, Nutbrown, Giulekas, & Cudjoe, 1996; Lowenkamp, Holsinger, & Latessa, 2001; Lowenkamp & Latessa, 2002). In theory, assessing an offender using the LSI-R allows for a valid classification (e.g., high, medium, or low-risk), as well as the identification of the most prevalent criminogenic needs that may facilitate case planning and treatment intervention (Andrews & Bonta, 2001).

The following section defines the participants included in the sample, the procedures for data collection and sample selection, and a review of the measures utilized.

Participants

The participants in this study were drawn from a larger sample of assessments (N = 1,243) included in the norming and profiling study conducted in 2003 (Holsinger et al., 2003). For the current study, in order to ensure representation of female offenders and Native American offenders, these two groups were over-sampled. This was done to ensure that separate analyses could be conducted on these two groups specifically.

The final N for the sample was 403 (35 percent of the offenders were female, 65 percent were male). The total sample was 65 percent White (263) and 35 percent Native American (140). Of the 140 Native American offenders included in these analyses, 10 were female and 100 were male. Of the 263 White offenders, 62 percent were male (162) and 38 percent female (101). Finally, all of the offenders in the sample (N = 403) were under supervision in the community when they were assessed using the LSI-R.

Procedures

As indicated above, certain groups of offenders were over-sampled for the current research in order to ensure adequate representation and the provision of group size large enough to conduct separate statistical analyses. As such, the sampling was divided into three sampling frames: one containing all Native American offenders, one containing White male offenders, and one containing White female offenders. All (100 percent) Native American offenders were selected for analyses, whereas 60 percent and 25 percent of the White female and White male offenders were randomly selected, respectively.

The LSI-R scores obtained were the result of interviews and collateral reviews of documentation for the sample used in Holsinger et al. (2003). As such, the predictive validity of the LSI-R was tested for the entire sample, as well as several subgroups contained within the sample (specifically Native Americans, non-Native Americans, male offenders, and female offenders).
conducted by staff members trained on the use of the LSI-R. Data on the outcome measure were collected by the State Department of Corrections using standardized internal databases.

Measures

Data on the LSI-R included the scoring of each of the fifty-four items on the instrument. A subcomponent score for each of the ten sections listed above along with an overall composite score was based on these data.

Race was coded as 0 for White and 1 for Native American. Sex was coded as 0 for male and 1 for female. Finally, the outcome variable was developed using data from the State Bureau of Criminal Investigation database system. The outcome variable was coded 0 if the event did not occur, and 1 if the event did occur (the event being any new arrest while at risk in the community).

Results

Table 1 presents the results for correlation analyses between the LSI-R composite score and any new arrest while at risk in the community. Overall, for the entire sample a correlation of .18 (N = 403) was observed. While the relationship was not of a great magnitude regarding strength, it was in the expected (positive) direction, and was statistically significant. The 95 percent confidence intervals are reported for each correlation as well. For all the cases in the analysis, the 95 percent confidence intervals ranged from .11 to .30.

When White offenders were isolated in the analyses, a correlation of .23 (N = 263) was revealed. This indicated a stronger relationship for White offenders when compared to the total sample, and was also in the expected (positive) direction. This relationship was statistically significant, and the 95 percent confidence intervals ranged from .12 to .34. A somewhat weaker relationship (r = .11) was revealed when the Native American offenders were isolated in the analyses (N = 140). In addition to demonstrating a weaker nonsignificant relationship overall, the 95 percent confidence intervals included .00, and ranged from −.06 to .27. A z-test was conducted in order to test the statistical difference between the two correlation coefficients, comparing the respective relationships for White offenders and Native American offenders. The z-test did not reveal a significant difference between the correlation coefficients for these two groups (p = .06).

Sex of the offender was also used as a comparative variable when investigating the relationship between LSI-R score and recidivism. For all male offenders, a correlation of .20 was revealed (N = 262). The relationship was statistically significant, of approximately moderate strength, and in the expected direction. The 95 percent confidence intervals for this correlation revealed a range of .08 to .31. Female offenders rendered a weaker relationship between LSI-R score and recidivism (r = .15). The relationship was not statistically significant (p = .075), and the 95 percent confidence intervals included .00 (−.01 to .31). Similar to the analysis comparing White offenders to Native American offenders, the z-test for the difference between these two correlation coefficients (males compared to females) was not statistically significant.

The analyses were further disaggregated by both sex and race. When utilizing only White males in the analysis (N = 162), a correlation coefficient of .22 was revealed. The coefficient was statistically significant, and the 95 percent confidence intervals ranged from .06 to .36. White females rendered a stronger relationship than White males, with a correlation of .26 (N = 101). Like White male offenders, the coefficient was statistically significant, and the relationship was positive. The 95 percent confidence intervals ranged from .07 to .43. The z-test for testing the difference between correlation coefficients was not significant indicating a lack of difference between the coefficients for White males and White females.

Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>r</th>
<th>p</th>
<th>N</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>.18</td>
<td>.001</td>
<td>403</td>
<td>.11</td>
<td>.30</td>
</tr>
<tr>
<td>White</td>
<td>.23</td>
<td>.001</td>
<td>263</td>
<td>.12</td>
<td>.34</td>
</tr>
<tr>
<td>Native American</td>
<td>.11</td>
<td>.197</td>
<td>140</td>
<td>−.06</td>
<td>.27</td>
</tr>
<tr>
<td>Male</td>
<td>.20</td>
<td>.002</td>
<td>262</td>
<td>.08</td>
<td>.31</td>
</tr>
<tr>
<td>Female</td>
<td>.15</td>
<td>.075</td>
<td>141</td>
<td>−.01</td>
<td>.31</td>
</tr>
<tr>
<td>White males</td>
<td>.22</td>
<td>.006</td>
<td>162</td>
<td>.06</td>
<td>.36</td>
</tr>
<tr>
<td>White females</td>
<td>.26</td>
<td>.009</td>
<td>101</td>
<td>.07</td>
<td>.43</td>
</tr>
<tr>
<td>Native American males</td>
<td>.19</td>
<td>.060</td>
<td>100</td>
<td>−.01</td>
<td>.37</td>
</tr>
<tr>
<td>Native American females</td>
<td>−.13</td>
<td>.428</td>
<td>40</td>
<td>−.42</td>
<td>.19</td>
</tr>
</tbody>
</table>

a z-test for difference between correlation coefficients was not statistically significant (White versus Native American; male versus female; White males versus White females).

b z-test for difference between correlation coefficients was statistically significant (Native American males versus Native American females).
When Native American males were isolated in the analysis, a correlation coefficient of .19 (N = 100) was rendered when assessing the relationship between LSI-R score and recidivism. The coefficient was not statistically significant (p = .06), and the 95 percent confidence intervals included .00 (95 percent C.I. = -.01 to .37). While the correlation for Native American males was of less than moderate strength and in the expected direction, the opposite was found when isolating Native American female offenders in the analyses. Specifically, a correlation of −.13 was revealed (N = 40). The relationship was opposite from the expected direction, however, the coefficient was not statistically significant (p = .428). Although these results were contrary to what one might expect, it is necessary to point out that the size of this subsample was small, particularly when compared to the other groups in the analyses. The 95 percent confidence intervals for this correlation coefficient ranged from −.42 to .19. In addition, the z-test for differences between correlation coefficients revealed a significant difference between Native American males, and Native American females.

Table 2 presents the recidivism rates by risk category according to the LSI-R. Recidivism rates are presented by race and sex. Although the proprietors of the LSI-R assessment and its supporting materials put forth recommended cut-off scores for the purpose of offender classification, the current analyses utilized different cut-off scores than those previously published. Attempts were made to create cut-off scores from the composite scale that adequately fit all four subgroups of offenders in the sample (White males, White females, Native American males, and Native American females). Ultimately, categorizations were utilized that rendered four classifications of risk. The four classifications were low-risk offenders (LSI-R scores ranging from 0 to 20), low/moderate (LSI-R scores ranging from 21 to 30), moderate (LSI-R scores ranging from 31 to 39), and high (40 and above).

Table 2 includes the recidivism rates for each subgroup of offenders, for each LSI-R categorization. Also present on Table 2 are the chi-square analyses testing the independence of the four-category classification and the dichotomous measures of recidivism, for each subgroup of the offender sample.

For White males, the chi-square analysis revealed a significant difference in recidivism rates across the four categories of risk. Low-risk White males had a 20 percent recidivism rate, low/moderate-risk White males had a 32 percent recidivism rate, moderate-risk White males had a 52 percent recidivism rate, and high-risk White males had a 60 percent recidivism rate. Similarly, for White females, the chi-square analysis revealed a significant difference in recidivism rates across the LSI-R risk categories. It should be noted that the risk categorizations that were used did not place any White females in the high-risk category. Fisher’s exact test was observed due to low cell frequency in this particular analysis. For White females, 16 percent of the low-risk offenders recidivated, 42 percent of the low/moderate-risk offenders recidivated, and 50 percent of the moderate-risk offenders recidivated.

The chi-square analysis for Native American males did not reveal a significant difference in rates of recidivism across categories of risk. Despite the lack of statistical significance, the recidivism rates were notably different across categories (low-risk offenders = 11 percent recidivism rate, low/moderate-risk offenders = 31 percent recidivism rate, moderate-risk offenders = 36 percent recidivism rate, and high-risk offenders = 57 percent recidivism rate). For Native American females, a significant difference in recidivism rates across risk categorizations was revealed. The results were not intuitive, which was not surprising in light of the negative correlation coefficient noted above. For example, 50 percent of the low-risk offenders recidivated, while a substantially lower proportion (16 percent) recidivated from the low/moderate risk categorization. Moderate-risk Native American female offenders had a 20 percent recidivism rate, while 100 percent (1 out of 1) of the high-risk Native American female offenders recidivated. Although the Native American female sample was the smallest group utilized in the current analyses, these results might reveal a necessity.

<table>
<thead>
<tr>
<th>Category</th>
<th>White males</th>
<th>White females</th>
<th>Native Am. males</th>
<th>Native Am. females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>20% (15/76)</td>
<td>16% (9/56)</td>
<td>11% (2/18)</td>
<td>50% (5/10)</td>
</tr>
<tr>
<td>Low/moderate</td>
<td>32% (16/50)</td>
<td>42% (13/31)</td>
<td>31% (11/36)</td>
<td>16% (3/19)</td>
</tr>
<tr>
<td>Moderate</td>
<td>52% (16/31)</td>
<td>50% (7/14)</td>
<td>36% (14/39)</td>
<td>20% (2/10)</td>
</tr>
<tr>
<td>High</td>
<td>60% (3/5)</td>
<td></td>
<td>57% (4/7)</td>
<td>100% (1/1)</td>
</tr>
<tr>
<td>X²</td>
<td>12.68*</td>
<td>10.11*</td>
<td>6.00</td>
<td>6.7</td>
</tr>
</tbody>
</table>

* = p less than or equal to .05.
for further investigation into the nature of the LSI-R, or risk/need assessment in general, and the Native American female population.

Based on the correlation analyses initially presented, and the chi-square analyses that followed, it appeared that the LSI-R score had some predictive validity for two of the four subgroups within the sample of offenders (White males and White females). In addition, White offenders as an entire group rendered predictive validity for the composite LSI-R score, and valid categorizations for classification purposes. According to the current data, the results appeared less promising when examining Native American offenders as a whole, and particularly when the Native American results were disaggregated by sex. Of particular note were the nonsignificant but somewhat intuitive categorizations of risk for Native American males, and the significant but contradictory risk-categorization results for Native American female offenders (as well as the negative correlation coefficient that was observed previously).

Fig. 1 presents the survival rates over a period of several months, for each of the four risk categories that were developed for the current research sample. The survival rates were calculated via Cox regression analysis, where offender sex, race, and the LSI-R score (used as the four-tiered categorization) were entered as predictors. The primary purpose of the survival analysis was to provide a more rigorous test of the predictive validity of the risk classifications using the LSI-R, for the entire sample while controlling for sex and race.

For the first six months (of the seventeen-month follow-up period), the four groups appeared to be very close together regarding the proportion that had not recidivated at that time (low = .92, low/moderate = .90, moderate = .92, and high = .92). After six months, however, substantive differences between the four categories began to emerge. At one year, there were noticeable differences between the survival rates of the four risk categorizations: .88 for low-risk offenders, .84 for low/moderate-risk offenders, .78 for moderate-risk offenders, and .76 for high-risk offenders. By the end of the follow-up period (seventeen months), the differences in proportions surviving become even more stark: low = .63, low/moderate = .35, moderate = .35, and high = .19. These results lent some support to the possibility that in reality, the data might reveal three meaningful gradations of risk (low, moderate, and high). As a whole, for the entire sample, the LSI-R composite score appeared to offer the function of risk classification supported by appropriate survival rates, even when controlling for race and sex.

Discussion

The analyses above offered support for the LSI-R as a composite tool on the entire sample of offenders, but mixed support when the results were disaggregated by certain subgroups of the sample. In terms of predictive validity, the instrument appeared to show the most promise for White offenders overall, males offenders overall, and White males and White females in particular.

Using any new arrest as the only measure of recidivism may provide one explanation for the non-intuitive results discovered for Native American females in particular, and the nonsignificant results overall for the Native American sample. Of all potential outcome variables for a study of recidivism, ‘any new arrest,’ while a potentially adequate indicator of antisocial/criminal behavior, is perhaps the least rigorous of recidivism measures (Latessa & Holsinger, 1998). Unfortunately, this recidivism measure is commonly used, and often the only available outcome criteria, as was the case in the current study. Future studies involving the LSI-R (or any risk/need assessment tool) should take steps to include additional measures of recidivism.
The available data did not allow for an in-depth exploration as to why the relationship between LSI-R score and recidivism appeared to break down for Native American offenders (both male and female). The factors measured by the LSI-R did have strong support in the correctional literature base (see for example Gendreau, Little, & Goggin, 1996; Motiuk, 1993; Shields & Simourd, 1991; Simourd & Andrews, 1994). It was unlikely, however, that Native American offenders were included in much of the data utilized to identify those factors, let alone isolated as a subgroup in and of themselves for analyses. The results presented above might serve as a call for additional research within the Native population of the United States, regarding the most relevant factors that are predictive of antisocial behavior.

Even if there is some currently unknown subset of criminogenic risk/need factors that are specifically relevant to the Native American population, many would argue that there has to be at least some overlap with the factors that are assessed by the LSI-R. This may be particularly relevant when considering the LSI-R assesses previous antisocial behavior, functionality at work/school, financial situation, significant familial support relationships, leisure activities, living situation, friendship network, substance abuse issues, emotional/mental health, and antisocial values and beliefs.

Some existing research comparing Native to non-Native offender populations had found similarities in the factors predicting recidivism (Bonta, LaPrairie, & Wallace-Capretta, 1997). Further, the LSI-R had been previously validated using a Native (Canadian) offender sample, although results did vary between non-Native and Native offenders (Bonta, 1989). Despite this existing validation, the question still remains regarding whether or not Native offenders differ between Canada and the United States. Regardless, additional research using Native offender populations is necessary to further examine general predictors of recidivism, and the predictive validity of specific risk/need assessment tools.

Aside from low sample size (at least in the case of Native American females), the answer to these confounding results may involve the responsivity principle of offender classification and intervention. Briefly, the responsivity principle involves the consideration that different offenders (even those that possess the same set of criminogenic needs) may need to be dealt with in different ways. Motivation, cognitive development, interpersonal maturity, mental illness, personality type, race, age, and ethnicity are well known responsivity factors that can influence the delivery of services, and interaction with the treatment/supervision environment. Correctional strategies are most effective when they take responsivity factors into account while processing offenders, gathering information, or delivering therapeutic interventions. Within the correctional environment, responsivity may be of most use when considering the matching of offenders with therapists, assignment of offenders to certain groups, and assignment of offenders to work with certain staff for supervision and assessments.

The LSI-R relies heavily on a one-on-one interview between the correctional professional and the offender. As such, there may be a need for extra care to be taken when assessing Native American offenders, regarding relational expectations, communication styles, cultural heritage and customs, and perhaps even jargon or vernacular. The professionals who conducted the LSI-R assessments utilized in the current study were trained in the use and technical aspects of the instrument. The extent to which they had received potentially necessary training regarding the Native American population however, was unknown. Also unknown was the extent to which Native Americans were represented within the correctional staff conducting the assessments. Further research using LSI-R data and Native American offender samples may shed some more light on this issue. In light of the results presented above, future research where specific subgroups of the offender population are isolated should also include data on, and statistical control for, the race/ethnicity of the professionals conducting the assessments, as well as data regarding the extent to which the staff had received training on specific cultures represented within the offender population.

Notes

1. The context of previous research examining correlations between risk assessments and measures of recidivism is typically used to determine whether or not a correlation is “strong,” “moderate,” or “weak.” Hanson (2000) for example noted that moderate levels of predictive accuracy corresponded with correlations ranging between .25 and .30.

2. An item-by-item analysis was conducted examining the relationship between each LSI-R item and the outcome criteria for Native American female offenders only. These analyses were strictly exploratory in nature, but nonetheless revealed that only two of the items were significantly associated with recidivism. Of the fifty-four items, two items from the criminal history section revealed a significant zero-order correlation with recidivism. The remaining correlations between each individual item and recidivism were nonsignificant and many were negative.

References


