

Criminal Justice and Behavior

<http://cjb.sagepub.com/>

Does Gender Moderate the Predictive Utility of the Level of Service Inventory— Revised (LSI-R) for Serious Violent Offenders?

Sarah M. Manchak, Jennifer L. Skeem, Kevin S. Douglas and Maro Siranosian
Criminal Justice and Behavior 2009 36: 425
DOI: 10.1177/0093854809333058

The online version of this article can be found at:
<http://cjb.sagepub.com/content/36/5/425>

Published by:



<http://www.sagepublications.com>

On behalf of:



[International Association for Correctional and Forensic Psychology](http://www.iacfp.org)

Additional services and information for *Criminal Justice and Behavior* can be found at:

Email Alerts: <http://cjb.sagepub.com/cgi/alerts>

Subscriptions: <http://cjb.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://cjb.sagepub.com/content/36/5/425.refs.html>

DOES GENDER MODERATE THE PREDICTIVE UTILITY OF THE LEVEL OF SERVICE INVENTORY–REVISED (LSI-R) FOR SERIOUS VIOLENT OFFENDERS?

SARAH M. MANCHAK

JENNIFER L. SKEEM

University of California, Irvine

KEVIN S. DOUGLAS

Simon Fraser University

MARO SIRANOSIAN

University of California, Irvine

A prison inmate's risk of recidivism is central to decisions about whether to release her to serve a portion of her sentence on parole. Although several tools have been developed to improve accuracy in estimating the risk of future reoffense, few have been validated with the ever-growing population of female offenders. This study compared the utility of the Level of Service Inventory–Revised (LSI-R) in predicting recidivism for 70 female and 1,035 male offenders who had been convicted of serious violent offenses. The LSI-R predicted 1-year general recidivism quite well for women. Although gender did not moderate the utility of the LSI-R in predicting recidivism, risk factors that best predicted recidivism differed for men and women. Implications for risk assessment and parole decision making with female inmates are discussed.

Keywords: LSI-R; risk assessment; recidivism; female offenders

The population of women under supervision of the criminal justice system is growing rapidly. Between 1983 and 2003, the number of women supervised by California's Department of Corrections and Rehabilitation increased by a staggering 497% (Little Hoover Commission, 2004). Nationally, between 1996 and 2005 the number of women arrested increased by 7% whereas the number of men *decreased* by 8% (Federal Bureau of Investigation [FBI], 2005). Nevertheless, correctional research has long focused on men, who still comprise the vast majority of inmates (93%; Sabol & Couture, 2008). In fact, most risk assessment tools have been developed and validated with predominantly (or exclusively) male samples (e.g., Historical-Clinical-Risk Management–20 [HCR-20]; Webster, Douglas, Eaves, & Hart, 1997; Violence Risk Appraisal Guide [VRAG]; Quinsey, Harris, Rice, & Cormier, 1998).

The extent to which these tools will generalize from men to women is unclear. First, women tend to be arrested for different crimes than men (e.g., prostitution and embezzlement; FBI, 2005, 2006). When they are involved in violence, women are more likely than men to

AUTHORS' NOTE: *The authors thank Robert Barnoski of the Washington Institute of Public Policy for his assistance with the data set. Correspondence may be addressed to Jennifer L. Skeem, PhD, Associate Professor, Psychology & Social Behavior, University of California, Irvine, 3311 Social Ecology II, Irvine, CA 92697-7085; e-mail: skeem@uci.edu.*

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 36 No. 5, May 2009 425-442

DOI: 10.1177/0093854809333058

© 2009 International Association for Correctional and Forensic Psychology

target family members (rather than acquaintances or strangers), more likely to engage in violence in the home (rather than in public places), and less likely to be intoxicated during the commission of an offense (e.g., Greenfeld & Snell, 1999; Robbins, Monahan, & Silver, 2003). In short, the criterion that the tools must predict may be somewhat different for women than men. Second, risk factors for men and women may differ or be differentially predictive for recidivism. For example, childhood abuse is more prevalent among female inmates than male inmates. The rate of exposure to sexual abuse is up to 5 times higher among female (26%) than male (5%) inmates (McClellan, Farabee, & Crouch, 1997; see also Cutler & Nolen-Hoeksema, 1991; James & Glaze, 2006). Abuse, in turn, is a risk factor for violence and other crime (e.g., McClellan et al., 1997; Monahan et al., 2001; Trickett & McBride-Chang, 1995; Widom, 1989; Widom & Maxfield, 1996; cf. Bonta, Pang, & Wallace-Capretta, 1995). Female inmates also are twice as likely as their male counterparts to have a serious mental disorder (Hodgins, Lapalme, & Toupin, 1999). Mental disorder per se is a relatively weak risk factor for violence (Monahan et al., 2001), arrest (Juninger, Claypoole, Laygo, & Cristiani, 2006), and recidivism (Bonta, Law, & Hanson, 1998). However, when mental disorder is present, women are as likely to become involved in violence as men (Hiday, Swartz, & Swanson, 1998; Lidz, Mulvey, & Gardner, 1993; Robbins et al., 2003). At a minimum, mental disorder is important to take into account as a “responsivity” factor (see Andrews & Bonta, 2006) when tailoring supervision and intervention to offenders’ needs. In short, it is possible that tools designed to distill risk factors for men may poorly capture those of women.

One tool that may be particularly informative for examining differences in risk between genders is the Level of Service Inventory–Revised (LSI-R; Andrews & Bonta, 1995). As is the case with many risk assessment tools, the LSI-R was developed and initially validated on a predominantly male sample (83% male; Andrews, 1982). Although the vast majority of the tool’s inmate U.S. normative sample is male (89%), the Canadian normative sample is gender balanced (40% male; Andrews, Bonta, & Wormith, in press). The latter gender-balanced sample distinguishes the LSI-R from other tools (e.g., Quinsey et al., 1998).

In addition, the LSI-R includes “dynamic” risk factors for recidivism, that is, changeable factors that can be monitored and targeted in supervision and intervention (Douglas & Skeem, 2005). In contrast, several measures that are widely used as risk assessment tools (e.g., Hare, 2003; Quinsey et al., 1998; see Tolman & Mullendore, 2003) focus on more static risk factors that cannot be changed, such as prior felonies and misdemeanors (Lowenkamp, Holsinger, & Latessa, 2001). Such tools are ill suited for tracking any changes in an inmate’s risk during incarceration or community supervision and for providing offenders with incentive to participate in correctional programming, engage in prosocial behavior, avoid infractions or violations, and generally make an effort to change. A parole board cannot advise an inmate to “undo” her history of assault. They can, however, advise her to work on developing more prosocial attitudes or greater emotional stability (see Hollin & Palmer, 2006).

The LSI-R consists of 54 items that are grouped into 10 scales, most of which are dynamic: (a) Criminal History, (b) Education/Employment, (c) Financial Situation (financial stability and problems), (d) Family/Marital Relationships, (e) Accommodation (level of domestic stability), (f) Leisure and Recreation, (g) Companions (composition and nature of core social network), (h) Alcohol and Drug Use, (i) Emotional/Personal, and (j) Attitudes/Orientations (assess offender’s attitudes, values, and antisocial or prosocial

beliefs). The 54 items are scored based on an interview with the offender (and collaterals, if possible) and a review of her or his records. Scores are weighted and summed (via an actuarial formula) to obtain a total score.

Research that compares men's and women's scores on the LSI-R has produced mixed results. For example, women have been shown to obtain total LSI-R scores that are similar to (Andrews & Bonta, 2003; Lowenkamp et al., 2001) or lower than (Mihailides, Jude, & Van den Bossche, 2005) those of men. Although differences in scale scores vary by study, the most consistent findings indicate that men tend to obtain higher scores on Criminal History (Holsinger, Lowenkamp, & Latessa, 2003; Mihailides et al., 2005; Raynor, 2007), whereas women tend to obtain higher scores on the Emotional/Personal (Holsinger et al., 2003; Mihailides et al., 2005; Palmer & Hollin, 2007) and Financial (Holsinger et al., 2003; Mihailides et al., 2005; Raynor, 2007) scales.

A large body of research supports the predictive utility of the LSI-R for male offenders' general recidivism (e.g., $r = .44$; Simourd, 2004) and convictions and revocations (e.g., $r = .34$ and $.45$, respectively; Kroner & Mills, 2001). In a meta-analysis of 33 effect sizes incorporating more than 7,000 male and female participants, Gendreau, Goggin, and Smith (2002) found that the LSI-R predicted general recidivism relatively well ($r = .37$), on par with the tool that forensic psychologists use most widely to assess risk (Tolman & Mullendore, 2003), the Revised Psychopathy Checklist (PCL-R; Hare, 2003). In this meta-analysis, the LSI-R was more strongly predictive of general recidivism for women ($\phi = .45$) than men ($\phi = .33$). However, few studies were included in this subanalysis, and confidence intervals for the two genders overlapped (Gendreau et al., 2002). In a more recent meta-analysis of 27 effect sizes associated with more than 14,000 female offenders, Lowenkamp, Smith, Latessa, & Cullen (in press) found that the LSI-R predicted women's recidivism well ($r = .35$) and, in the subsample of studies that provided relevant data, was similarly predictive for women's and men's recidivism.

These findings are in keeping with the handful of peer-reviewed, published studies that specifically have examined the predictive validity of the LSI-R with adult female offenders, which have produced largely positive findings. On one hand, four studies suggest that the LSI-R moderately to strongly predicts women's recidivism. Based on a sample of 526 female inmates in a Canadian medium security prison, Coulson, Ilacqua, Nutbrown, Guilekas, and Cudjoe (1996) found that across types of prison discharge, LSI-R scores were strongly related to 1-year rates of recidivism ($r_{pb} = .51$), parole failure ($r_{pb} = .53$), and half-way house failure ($r_{pb} = .45$). Similar results were obtained with a sample of (a) 125 women in a minimum security community corrections facility who were at risk on average for 611 days ($r = .37$ for reincarceration; Lowenkamp et al., 2001), (b) 169 women on intensive community supervision probation followed for 1 year ($r = .41$ for new arrests; Kirkpatrick, 1999), and (c) 100 released female offenders followed for 6 years after release ($r = .30$ for new convictions; Folsom & Atkinson, 2007). Additionally, a study of probationers in the United Kingdom found that there was a significant difference in LSI-R total score between women who had and had not recidivated within 1 year (Raynor, Kynch, Roberts, & Merrington, 2000). On the other hand, the LSI-R produced smaller effects in two studies, including those with a sample of (a) 141 female community supervisees followed for up to 17 months ($r = .15$ for new arrests; Holsinger, Lowenkamp, & Latessa, 2006) and (b) 134 female community supervisees followed for 6 months ($r = .16$ for self-reported new arrest; Holtfreter, Reisig, & Morash, 2004).

As a group, these studies indirectly suggest that the LSI-R predicts women's recidivism as well as men's. However, this hypothesis rarely has been tested directly. In the present study, we went beyond an examination of the LSI-R's predictive utility for women's general recidivism (Aim 1) to address the more crucial issue of whether gender moderates the measure's predictive utility (Aim 2). We tested directly whether the LSI-R works less well for women than for men and do so with a relatively rarely studied sample of offenders who have been convicted of serious violent offenses. We also explored whether the LSI-R scales that maximally predict men's and women's recidivism differ (Aim 3). As shown earlier, research reveals some gender-related differences in scores on the LSI-R scales. Even if the predictive utility of the LSI-R is similar for men and women, the bases of that predictive utility may differ by gender. Those differences, in turn, may have important implications for gender-specific risk assessment and risk management strategies. Such information is crucial, given the rapid growth of the female offender population in the United States over recent years.

METHOD

We addressed these aims using a deidentified data set provided by the Washington State Institute for Public Policy (WSIPP). The present sample is drawn from a larger database maintained by WSIPP. For example, 71% of our female sample ($n = 49/70$) were included in a larger female sample ($n = 4,822$) described in an unpublished report available from WSIPP (Barnoski & Aos, 2003). The present sample includes only the particularly policy-relevant group of released offenders who had been incarcerated for such serious violent offenses as murder, rape, kidnapping, and arson. The database contains sentence length, demographic, and recidivism information for all offenders released from Washington State Prisons from 1999 through 2005. The database also includes recidivism data for all offenders who had recidivated by the predesignated follow-up period of 12 months within the state of Washington.

We assessed the utility of the LSI-R in a state with determinate sentencing (WSIPP, 2003) and no parole supervision (Reentry Policy Council, 2005). This seems an ideal context in which to assess the predictive utility of any risk-assessment tool. First, all offenders, whether classified as high risk or low risk by the tool, are released at the end of their term. Thus, high-risk classification is not conflated with low opportunity to reoffend. Second, offenders are not supervised by parole agents, providing maximum opportunity for them to behave naturally in the community after release.

PARTICIPANTS

Participants were 1,105 inmates, including 70 (7%) women and 1,035 (93%) men. Table 1 provides descriptive statistics for the sample. Both male and female samples were predominantly White and in their late 30s upon release from prison. By definition, both groups had serious index offenses, with the majority of men having committed homicide (33%), violent property crimes (27%), and sex crimes (14%) and the majority of women having committed homicide (60%), violent property crimes (20%), and kidnapping (13%). The groups did not differ significantly in age at release from prison or time at risk. The

TABLE 1: Demographic and Offense Features of Female, Male, and Total Offender Samples

| Variable | Total Sample (N = 1,105) | | Men (N = 1,035) | | Women (N = 70) | |
|-----------------------|-----------------------------|------|-----------------|------|----------------|------|
| | M | SD | M | SD | M | SD |
| <i>Continuous</i> | | | | | | |
| Time at risk (months) | 28.2 | 20.0 | 30.2 | 19.7 | 31.0 | 18.9 |
| Age at release | 39.2 | 12.1 | 39.30 | 11.9 | 37.0 | 12.3 |
| Prison years | 8.9 | 5.6 | 8.8 | 5.7 | 7.8 | 5.2 |
| <i>Categorical</i> | N | % | N | % | N | % |
| Race | 1,105 | | 1,035 | | 70 | |
| White | 874 | 72.0 | 745 | 72.0 | 47 | 67.1 |
| Black | 240 | 19.8 | 209 | 20.2 | 13 | 18.6 |
| Asian | 40 | 3.3 | 34 | 3.3 | 3 | 4.3 |
| Hispanic | 136 | 11.2 | 111 | 10.7 | 7 | 10.0 |
| Native | 49 | 4.0 | 40 | 3.9 | 3 | 4.3 |
| Recidivism | 872 | | 816 | | 56 | |
| General | 160 | 8.3 | 151 | 18.5 | 9 | 16.0 |

Note. Recidivism data are based on the number of participants who had been at risk for 1 year or recidivated within the first year postrelease.

groups did not differ significantly in their LSI-R total score, but women scored significantly lower than men on the Criminal History, Leisure and Recreation, and Attitudes/Orientations scales and significantly higher than men on the Emotional/Personal scale (see Table 2).

General recidivism data for men and women are presented separately at the bottom of Table 1. To determine general recidivism base rates, a slightly smaller sample size was used, reflecting total, female, and male offenders who were at risk for a year or more or who had recidivated within a year ($n = 872$, 56, and 816, respectively). At 1 year postrelease, 18% of our total sample ($n = 160$), including 16% of the female sample and 19% of the male sample, had recidivated.

MEASURES

LSI-R. The LSI-R (Andrews & Bonta, 1995) is a 54-item instrument in which trained raters score each risk/need item as *present* (1) or *absent* (0) on the basis of a file review and an interview with the offender. In Washington State, the LSI-R was scored according to the manual's instructions, with the exception of 1 employment item and 7 substance abuse items that were scored based on the inmate's behavior just prior to incarceration, rather than the "past year" in the institution. In keeping with the manual's instructions, items that referenced community adjustment (e.g., from the Accommodation, Leisure and Recreation, Financial Situation, and other scales) were rated based on the inmate's performance during the year prior to prison entry. Nevertheless, scoring of most of the LSI-R's dynamic factors took into account inmates' recent or "current" behavior (e.g., Attitudes/Orientations, Emotional/Personal, and Family/Marital Relationships scales).

As shown in Table 2, levels of internal consistency for women's scores on most LSI-R scales were acceptable ($\alpha = .62-.87$), although scores on the Criminal History ($\alpha = .48$),

TABLE 2: Level of Service Inventory–Revised (LSI-R) Scores of Female, Male, and Total Offender Samples

| Variable | Total Sample (N = 1,105) | | Men (N = 1,035) | | Women (N = 70) | | | | |
|---------------------------------|-----------------------------|-----|-----------------|-------------|----------------|----------|----------|-----|----------|
| | M | SD | M | SD | M | SD | | | |
| LSI-R score | 26.1 | 9.4 | 25.8 | 9.3 | 23.8 | 9.7 | | | |
| LSI-R Factor Scores | (N = 1,102) | | | (N = 1,034) | | | (N = 67) | | |
| | M | SD | α | M | SD | α | M | SD | α |
| 1. Criminal History** | 5.8 | 2.1 | .5 | 5.9 | 2.1 | .5 | 4.8 | 1.9 | .5 |
| 2. Education/Employment | 5.0 | 2.8 | .7 | 5.0 | 2.8 | .7 | 4.5 | 2.7 | .7 |
| 3. Financial Situation | 1.0 | 0.7 | .3 | 0.9 | 0.7 | .3 | 1.0 | 0.7 | .2 |
| 4. Family/Marital Relationships | 2.1 | 1.3 | .6 | 2.1 | 1.3 | .6 | 2.2 | 1.2 | .4 |
| 5. Accommodation | 0.8 | 1.0 | .6 | 0.8 | 1.0 | .6 | 1.0 | 1.1 | .6 |
| 6. Leisure and Recreation** | 1.6 | 0.8 | .7 | 1.6 | 0.7 | .7 | 1.2 | 0.9 | .8 |
| 7. Companions | 2.3 | 1.4 | .7 | 2.3 | 1.4 | .7 | 2.5 | 1.4 | .7 |
| 8. Alcohol and Drug Use | 2.8 | 2.5 | .9 | 2.9 | 2.5 | .9 | 2.4 | 2.5 | .9 |
| 9. Emotional/Personal* | 1.9 | 1.4 | .6 | 1.9 | 1.4 | .6 | 2.4 | 1.6 | .7 |
| 10. Attitudes/Orientations** | 2.4 | 1.6 | .8 | 2.4 | 1.6 | .8 | 1.7 | 1.5 | .8 |

* $p < .01$. ** $p < .001$.

Family/Marital ($\alpha = .45$), and Financial ($\alpha = .19$) scales manifested poor internal consistency. Although poor, the latter results are related to the brevity of the Financial scale (see discussion below) and comparable to those seen in the U.S. norms supplement (Andrews & Bonta, 2003). The predictive utility of the LSI-R was reviewed earlier.

Although the LSI-R's authors use normative data to provide four risk groupings for female offenders (minimum, medium, high/medium, and maximum), they recommend that these scores be assessed and refined for individual jurisdictions and populations (Andrews & Bonta, 2003). In the present study, we did so with a three-category risk grouping derived by trichotomizing the female sample at the 33rd and 66th percentiles so that the risk groupings were low (total score 0-19), medium (20-26), and high (27-48). In addition to this sample trichotomy, three other forms of LSI-R scores were used: (a) scores on the 10 LSI-R risk/needs scales (listed earlier), (b) the four-category risk grouping recommended by Andrews and Bonta (1995) for use with male inmates, and (c) LSI-R total scores (ranging from 0 to 54).

Recidivism. The primary criterion variable was general recidivism, defined here as conviction of any new offense in the state of Washington during the follow-up period after the index prison release. Although arrest data were not available, our stringent definition of recidivism (as reconviction) provides some assurance that a serious offense actually occurred: In approximately 20% of prosecuted crimes, offenders are not convicted (DeFrances, 2001).

Recidivism was coded as 0 (*no*) or 1 (*yes*) with respect to whether inmates recidivated during the entire period that they were followed in the community. Because inmates were released at different periods of time, their length of follow-up varied. For participants convicted during their follow-up, the date of the first conviction was coded. For those who had no new conviction during their follow-up, the end date of that follow-up period was coded. These date-based outcome variables were used in survival analyses described below.

For the 872 cases with an eligible follow-up period (56 of whom were female), we coded whether inmates recidivated generally within the 1st year of release. We chose a 1-year follow-up period because this is the period of maximum risk for recidivism (e.g., Langan & Levin, 2002), and setting a longer follow-up would reduce the sample size. These 1-year outcome variables were used in analyses that supplement the primary survival analyses—which use the full sample—described below.

PROCEDURE

Although the present study involved secondary analysis of an existing data set, details on the process for completing the LSI-R and its reliability among staff in Washington State are available. Typically, the LSI-R was completed as part of routine prison procedure within a year of inmates' release date. To train on the LSI-R, selected staff completed a 3-day training that included both didactic and experiential exercises. After presenting didactic material and video clips that covered each subcomponent of the LSI-R, trainers had staff view a full-length videotaped interview (provided by James Bonta) and render scores based on the videotape. Staff members were defined as "in agreement" when their LSI-R total scores fell within 5 points of the criterion score. The average level of agreement was 80% to 90% (Kevin Mauss, personal communication, October 4, 2006). In the last phase of training, staff interviewed individual offenders conjointly with a facilitator, scored the case, and obtained feedback. Staff not attaining reliability were referred to supervisors for remedial measures. In an effort to increase reliability, staff used a scoring guide developed for the state that provided sample interview questions and prototypical responses for particular scores on each item.

Three forms of quality control were implemented to guard against raters' drift. First, the central training staff completed "tune-up" training approximately once per year that involved rating and discussing a videotaped case (reliability = 80%-85%, as defined above; Kevin Mauss, personal communication, October 4, 2006). Second, supervisors were mandated to audit 2 cases per month per unit. The audit involved independently scoring the LSI-R and meeting with the original scorers to discuss any discrepancies. Third, beginning in 2003, the central training staff independently scored 3,800 cases that were eligible for a 50% reduction in their sentences, based on their level of risk, as assessed by the LSI-R. Although this sample overlaps only partially with the present sample of released offenders, these data convey the level of agreement that these staff maintain on the LSI-R. For these 3,800 cases, administrators computed the number of items line staff scored correctly (i.e., in keeping with the independent scores of training staff). Across items, the average level of items scored correctly was 93.9% (Kevin Mauss, personal communication, October 4, 2006). Staffs' LSI-R scores were classified as "accurate and valid" when there were 3 or fewer (out of 54) scoring errors. The majority (64%) of cases were classified as accurate and valid at this stringent criterion level.

RESULTS

Recall that this study has three specific aims. First, to assess the predictive utility of the LSI-R with women, we completed (a) receiver operating characteristic (ROC; Wright,

2005) analyses to describe the basic strength of the relation between LSI-R scores and recidivism for women who had been at risk for at least 1 year ($n = 56$) and (b) a series of survival analyses ($n = 70$) to determine whether LSI-R predicted risk of time to recidivism. Second, to assess whether gender moderated the effectiveness of the LSI-R total scores in predicting recidivism within 1 year of release, we conducted a logistic regression analysis (Baron & Kenny, 1986), using data on eligible general inmates ($n = 872$). Third, we applied survival analyses separately with men ($n = 1,034$; see Manchak, Skeem, & Douglas, 2008) and women ($n = 70$) to examine which LSI-R scales contribute most to prediction of risk of time to recidivism for each gender.

AIM 1: WHAT IS THE PREDICTIVE UTILITY OF THE LSI-R AMONG FEMALE SERIOUS OFFENDERS?

To estimate the overall strength of the relationship between the LSI-R and recidivism among women, we conducted an ROC analysis, using only women who had been at risk for 1 year or had recidivated during that time ($n = 56$). Overall, the LSI-R performed well, with an area under the curve (AUC) value of .77 (CI = .62-.91, $p < .01$). Using correlations (which are more affected by base rates than AUCs), the LSI-R weakly predicted recidivism at 1 and 3 years ($\eta = .23$ and $.34$, *ns*, respectively).

To address Aim 1, three Cox proportional hazards survival analyses were completed using the full female sample ($n = 70$), predicting time to general recidivism as a function of the LSI-R total score, the manual risk groupings (see Andrews & Bonta, 1995), and the sample trichotomy. The full sample was used because survival analysis can account for differential follow-up periods, allowing us to include 20% of the female sample at risk for less than 1 year. The results indicated that the LSI-R total score significantly predicted risk of time to recidivism, $\chi^2(1, n = 70) = 7.31, p < .01$. At the categorical level, our sample trichotomy significantly predicted risk of time to recidivism, $\chi^2(2, n = 70) = 6.35, p < .05$, but the manual risk groupings did not, $\chi^2(3, n = 70) = 6.08, ns$.

Effect size estimates in survival analysis use the hazard ratio (HR), which can be interpreted similarly to odds ratios. That is, the HR is a ratio of one hazard (e.g., high-risk LSI-R group) to another (e.g., low-risk LSI-R group), where *hazard* indicates, approximately, the risk for recidivism at any given point in time (taking into account the fact that some offenders are no longer eligible to recidivate because they did so at an earlier point in time). So if the hazard of recidivism in the high-risk group is .20, and in the low-risk group it is .10, then the HR is 2.0, meaning that those in the high-risk group are twice as likely to recidivate as those in the low-risk group. HRs represent the increase or decrease in risk for time to an event (here, recidivism), or how quickly the event of interest occurs, as a function of either group membership or some specified predictor.

Our results indicated that the hazard for recidivism increases 7.6% for every 1-point increase in the LSI-R total score, indicating that those who score higher on the LSI-R recidivate more quickly and more often than those who score lower on the LSI-R. Compared to the low-risk group in our sample trichotomy, the medium-risk group was at about 5 times greater risk (HR = 5.2, Wald = 2.26, *ns*), and the high-risk group was at about 9 times greater risk (HR = 9.5, Wald = 4.43, $p < .05$), at any given time point in the follow-up. The survival function for this trichotomy grouping may be seen in Figure 1. The y -axis of the graph is a measure of cumulative survival, where 1.0 indicates 100% survival and no recidivism. The x -axis indicates months to recidivism.

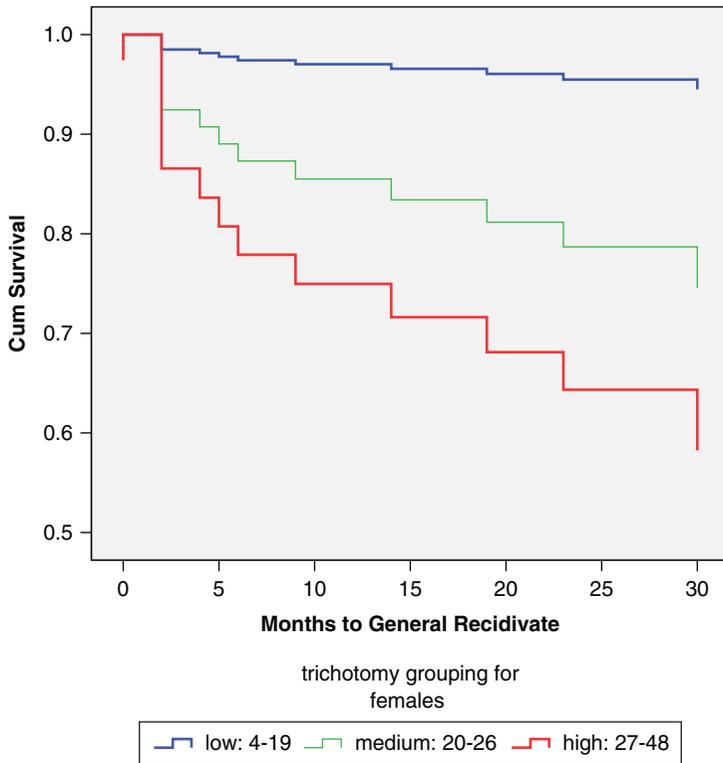


Figure 1: Survival Function for Women’s General Recidivism

AIM 2: DOES GENDER MODERATE THE PREDICTIVE UTILITY OF THE LSI-R?

Having shown that the LSI-R performs fairly well with women, we next assessed whether gender moderated the measure’s predictive utility. Does the LSI-R predict recidivism at 1 year less well for women than men? To address this aim, we conducted a logistic regression using all men and women in the sample who had been at risk for at least a year or more *or* who had recidivated within the 1st year ($n = 872$). Here, LSI-R total score and gender were entered in Block 1, and the interaction term for the two was entered in Block 2. The interaction term of gender and LSI-R in Block 2 was not significant ($B = .002$, $p = .89$). This suggests that the predictive utility of the LSI-R is not moderated by gender. In other words, the LSI-R appears to work as well for men as it does for women.

To rule out limited power ($n = 56$ women + 816 men) as a potential explanation for the lack of interaction, we (a) fit separate regression models for all men and women at risk for a year or more and (b) compared the effect sizes generated by the LSI-R in predicting recidivism. The fact that the effect sizes for the two genders were similar (men $\exp(B) = 1.05$, $p < .001$; women $\exp(B) = 1.12$, $p < .01$) provides some confidence that gender does not moderate the measure’s predictive utility. Additionally, there were no violations of the homogeneity of variances assumption between the male and female samples on LSI-R total score ($F = 0.00$, *ns*), adding to the credibility of reported results.

AIM 3: DO DIFFERENT SCALES DRIVE THE PREDICTIVE UTILITY OF THE LSI-R FOR MEN VERSUS WOMEN?

Even if the LSI-R is equally effective with women as it is with men, the scales that drive the predictive utility of the tool may differ by gender. To determine whether this was the case, we completed two survival analyses using all cases with available LSI-R scale scores, one for men ($n = 1,034$) and one for women ($n = 67$), in which we predicted time to general recidivism from the 10 LSI-R scale scores. The scales were allowed to enter stepwise, using the forward likelihood ratio criterion. (Notably, direct entry produced identical results.)

The overall model significantly predicted risk of (time to) general recidivism for both men, $\chi^2(10, n = 1,034) = 69.6, p < .001$, and women, $\chi^2(10, n = 67) = 20.7, p < .05$. However, we found that different scales predicted recidivism for men than for women. Specifically, the significant predictors for men were Criminal History (HR = 1.1, Wald = 10.4, $p < .01$), Financial (HR = 1.3, Wald = 8.7, $p < .01$), and Alcohol/Drug scale scores (HR = 1.1, Wald = 24.0, $p < .001$). In contrast, the only significant scale for women was the Financial scale (HR = 4.6, Wald = 9.8, $p \leq .05$).

DISCUSSION

Although the number of women entering the criminal justice system has been increasing dramatically over recent years (Glaze & Bonczar, 2006; Sabol, Minton, & Harrison, 2007), and the LSI-R is among the most widely used tools for assessing general offenders' risk of criminal recidivism, this study is among the first to assess directly whether gender moderates the tool's level of, or basis for, predictive utility. It is also relatively rare in its focus on serious violent offenders. The study produced three key findings. First, the LSI-R performs quite well in predicting women's general recidivism. In fact, when sample-derived (not manual-based) LSI-R risk groupings are applied, offenders are quite well differentiated in their rates of recidivism. Second, the predictive utility of the LSI-R does not depend on gender: The tool works as well for women as it does for men. Third, the scales of the LSI-R that maximally predict recidivism for men may differ from those that maximally predict recidivism for women. In this section, after discussing this study's limitations, we discuss each of these findings and their implications for research and practice.

LIMITATIONS

The study's findings come with four primary qualifications. First, precise estimates of interrater reliability on the LSI-R in this sample are not available, and there were minor deviations from the criteria for scoring employment and substance abuse. However, as noted above, staff members were carefully trained and subjected to rigorous measures of quality control. Moreover, the predictive utility of the LSI-R in this study (including that of the Alcohol/Drug Use scale) is comparable to that of other studies, and reliability is a prerequisite for predictive utility.

Second, our sample of women was relatively small ($n = 70$), given the focus on serious violent offenders. This sample size did not seem to unreasonably limit statistical power for analyses involving the total score of the LSI-R. In fact, effect sizes for total *and* scale scores were similar for the (small) female and (large) male sample. Nevertheless, it is likely that

a larger sample would have yielded different results with respect to scales that maximally predict women's recidivism. Specifically, with a large female sample, more LSI-R subscales would probably enter the multivariate equation to predict recidivism. Although we can be quite confident that the Financial subscale is an important predictor of recidivism for these women, we can be less confident that other LSI-R subscales are not important predictors of women's recidivism.

Third, the base rates of recidivism in this study were limited. First, we used official records of reconviction, which are known to underestimate the incidence of crime (Coleman & Moynihan, 1996). Nevertheless, our recidivism figures do not appear particularly questionable: The rate of general recidivism within 1 year of release in this study (16%) is comparable to that of a national sample of women (~17%; Deschenes, Owen, & Crow, 2007; secondary data analysis of Langan & Levin, 2002). Still, we recommend that the results be replicated using more liberal definitions of general and violent recidivism. Second, our records of reconviction were restricted to those that occurred within the state of Washington. Because some released prisoners recidivate in states other than their own (e.g., 7.6%; Langan & Levin, 2002), the results should be replicated in a study with access to FBI or other national records of recidivism.

Fourth, we were unable to assess the extent to which the predictive utility of the LSI-R may vary based on potentially gender-specific forms of recidivism. As mentioned earlier, women are more likely than men to be arrested for such offenses as prostitution. The extent to which the LSI-R predicts relatively "female" crimes is an important issue to address in future research.

BASIC PREDICTIVE UTILITY OF THE LSI-R WITH FEMALE SERIOUS OFFENDERS

The vast literature on female offenders suggests that they differ from male offenders in key ways, including their criminal histories, the antecedents to crime, and mechanisms of desistance from crime (Belknap & Holsinger, 2006; Daly, 1992; Laub & Sampson, 2003). Nevertheless, we found that the male-developed LSI-R performed quite well in predicting women's general recidivism (AUC = .77). Relative to women scoring in the lower third of our LSI-R distribution, the medium-scoring group was more than 5 times as likely to recidivate and the high-scoring group was more than 9 times as likely to recidivate. Typically, relative risks (or risk ratios) greater than 2 are regarded as substantial (Brignell, 2000; Cantor-Graae & Selton, 2005). The LSI-R's AUC for women's general recidivism is higher than those found for men's general recidivism (e.g., .69, Kroner & Mills, 2001; .66, Manchak et al., 2008) and comparable to that found for leading risk assessment tools in predicting (mostly) forensic patients' violence (e.g., VRAG AUC = .76, Quinsey et al., 1998; HCR-20, .65-.82, see Douglas, Guy, & Weir, 2007).

These results are notable in part because they are based on a particularly "deep-end" sample of women convicted of serious violent offenses. The majority of our women (60%) were convicted of homicide, compared to only 11% of incarcerated women nationally (Greenfeld & Snell, 1999). Still, the consistency of our results with those obtained in other LSI-R research (e.g., Lowenkamp et al., in press) suggests that the results may generalize to less serious female offenders.

Our findings also highlight the importance of local recalibration of actuarial formulas derived in other samples. Ideally, this recalibration would itself be cross-validated in new

samples from this same setting. In this study, risk classifications derived by trichotomizing the sample significantly predicted women's recidivism, whereas published risk groupings for the LSI-R did not. Our findings illustrate the danger associated with applying optimized risk classifications to a new sample, given that predictive accuracy tends to degrade in the process. It also demonstrates the utility of a customized approach to establishing risk classifications based on individual sample characteristics, as recommended by the LSI-R authors (Andrews & Bonta, 2003).

LACK OF GENDER MODERATION FOR THE LSI-R'S PREDICTIVE UTILITY

In this study, the LSI-R predicted women's recidivism as well as men's recidivism: Gender did not moderate the tool's predictive utility. The LSI-R may work as well for women as men because it captures eight broadband risk factors that might apply to both men and women. These risk factors have been termed the "central eight" predictors owing to their consistency in the prediction of crime across numerous studies (Andrews & Bonta, 2003). As reviewed by Poels (2007), there is some evidence that at least some of these risk factors (i.e., criminal history, substance use) are predictive of crime and violence among both men and women. However, as Poels also noted, it is far from clear (a) which risk factors predict crime and violence comparably across men and women and (b) how such risk factors might operate for men and women. Therefore, it remains possible—or even likely—that some risk factors are more or less important across gender (Holtfreter & Cupp, 2007).

DIFFERENCE IN THE LSI-R'S PREDICTIVE SCALES, AS A FUNCTION OF GENDER

In this study, financial problems were a salient risk factor for women. In contrast with prior research indicating that female inmates have more financial problems than male inmates (Holtfreter et al., 2004; see also Holsinger et al., 2003; Mihailides et al., 2005; Raynor, 2007), we found that men and women obtained similar scores on the LSI-R's Financial scale (i.e., .92 and .97, respectively). Although the Financial scale significantly predicted recidivism for both genders, the scale predicted women's time to recidivism (HR = 4.6) more strongly than men's time to recidivism (HR = 1.3).

Our finding that the Financial scale predicted recidivism for both genders is consistent with other research indicating a relationship between the Financial scale and rearrest ($r = .19$; Holtfreter et al., 2004) and past convictions ($r = .33$; Palmer & Hollin, 2007), and more generally with research showing that such factors as employment problems, poverty, and low socioeconomic status are important contributors to crime (Valdez, Kaplan, & Curtis, 2007). Despite such findings, the Financial scale was deleted from the LSI-R to create the Level of Service/Case Management Inventory (LS/CMI) on the grounds that it (a) exhibited low frequency, (b) had poor reliability (which by definition will limit predictive utility), and (c) was not considered one of the central eight risk factors (Andrews & Bonta, 2003; Andrews, Bonta, & Wormith, 2004).

This deletion may have been premature, particularly for women. Given that the typical female offender is impoverished, poorly educated, and has few vocational skills (see Covington & Bloom, 2006), she is likely to confront huge financial stressors when released from prison (see Holtfreter & Cupp, 2007). Female offenders may be even more vulnerable to the effects of poverty, unemployment, and homelessness than men (see Holtfreter &

Cupp, 2007). If the LSI-R Financial scale had been defined differently (perhaps as a multifaceted construct) and scored with better reliability, it may have (a) shown even greater predictive utility in the present study, (b) demonstrated more consistent predictive utility in other studies, and (c) been retained for the LS/CMI. Problems with internal consistency, for example, probably stem less from the financial construct than the short length of the corresponding LSI-R scale. Applying Nunnally and Bernstein's (1994) formula, α for the Financial scale would be considerably higher if it comprised 10 items (similar to other LSI-R scales) rather than just 2 items ($\alpha = .56$ vs. $.19$ for women, $\alpha = .68$ vs. $.30$ for men). The predictive utility of this scale for both men and women in this study, despite poor internal consistency, combined with more general support for this risk factor suggests that it should be more clearly defined and included in assessments of offenders' risk of recidivism.

The usefulness of the LSI-R and similar instruments lies not only in their predictive utility but also in their ability to point toward areas for intervention (see Holtfreter & Cupp, 2007). Given that correctional treatment is most effective when it targets dynamic risk factors for recidivism, effective programs for women should probably go beyond typical criminogenic needs targeted for men (e.g., criminogenic attitudes) to include the provision of skills and opportunities that may improve socioeconomic standing (Covington & Bloom, 2006). Interestingly, vocational skills and educational programming, which ostensibly would promote financial security postrelease, are more commonly found in correctional programs for men than for women (Gray, Mays, & Stoher, 1995).

Beyond gender-based differences in the most predictive risk factors captured by the LSI-R, there are also gender-based differences in scale scores that may have implications for effective supervision and intervention. First, relative to men, women obtained significantly lower scores than men on Attitudes/Orientations, Leisure and Recreation, and Criminal History (in keeping with Holsinger et al., 2003; Mihailides et al., 2005; Raynor, 2007). This suggests that contemporary correctional programs that focus on "criminogenic thinking" are better tailored to men than women—even when the focus is on offenders with serious violent offenses (Sorbelli, Eccleston, Ward, & Jones, 2002; see also Landenberger & Lipsey, 2005). Second, women obtain significantly higher scores on the LSI-R Emotional/Personal scale than men (in keeping with Holsinger et al., 2003; Mihailides et al., 2005; Palmer & Hollin, 2007). This content of this scale references psychological distress; symptoms of psychotic, affective, and anxiety disorders; and psychosocial functioning (Andrews & Bonta, 2003). Although the scale does not reference childhood abuse, childhood abuse is a risk factor for mental disorder (e.g., Messina & Grella, 2006; Messina, Grella, Burdon, & Prendergast, 2007) as well as violence and crime (e.g., Monahan et al., 2001; Widom, 1989). Indirectly, this suggests that the recent focus on providing "trauma-sensitive" treatment for female offenders—treatment that focuses on reducing the symptoms of sexual or physical victimization and improving mental health—*might* be well placed (Clark, 2002; Harris, 1998). Whether such programs actually reduce women's recidivism risk, however, is an open question.

The present results suggest that supervision and treatment programs that are gender responsive within a risk-needs-responsivity (Andrews & Bonta, 2006) framework may be helpful. There is an increasing awareness that effective programming for women should address financial and emotional vulnerabilities unique to women offenders. In their review of what works, Austin and colleagues found that the most effective programs included

elements that provide skills training to empower them and achieve independence and addressed women's psychological concerns such as trauma (Austin, Bloom, & Donahue, 1992). A number of programs across the United States are also geared toward and sensitive to trauma (for a review, see Clark, 2002).

CONCLUSIONS

Our finding that the utility of the LSI-R generalizes to serious female offenders does not mean that the LSI-R—particularly as applied in this study—is the *best* tool for women (see also Lowenkamp et al., in press). The tool may not capture all the dynamic risk factors that may be uniquely predictive of recidivism for women. Future research should examine some of the more promising dynamic risk factors found in the literature on women (e.g., abuse) and test whether they add significant predictive power above and beyond the utility of the LSI-R on recidivism. Research is needed to determine whether a risk assessment tool developed explicitly for women would be more applicable and useful than the LSI-R is or if additional items added to the LSI-R that would better capture traditionally women-specific risk/needs factors would be the best approach.

Gender-specific development and performance of contemporary risk assessment tools have been identified as a significant gap in the literature (Heilbrun, Douglas, & Yasuhara, in press). Key goals for future research include the identification of risk factors that (a) differ in the strength with which they predict men's and women's recidivism, (b) predict recidivism for one gender but not for the other, and (c) operate differently for men and women in *how* they drive criminal behavior (e.g., are financial problems manifested differently, or linked differently to crime, as a function of gender?). The last point is crucial for understanding men's and women's reasons for recidivism, as opposed to merely predicting its occurrence.

The results of this analysis suggest that the LSI-R predicts recidivism as well for women as it does for men. At the most fundamental level, this indicates that the LSI-R can be applied to help assess women's risk of recidivism. However, the comparable performance between genders was limited to the prediction of recidivism, not the understanding of possible different causes of criminal behavior between women and men. Understanding the nature of the problem arguably is the first step toward developing management and intervention strategies that will maximally reduce risk—for both men and women.

Holtfreter and Cupp (2007) reviewed gender issues not only with respect to the LSI-R but also with respect to criminal behavior more broadly. They made two important points that have implications for constructing risk assessment instruments and treatment programs for women. First, they described several gendered pathways to crime for women that are not reflected in traditional theories of crime (including the social learning theoretical basis of the LSI-R) or in traditional risk assessment instruments that were developed primarily with men in mind. Several of these feature prominently the unique adverse effects that abuse or other harmful relationship characteristics (either during childhood or adulthood) can have on women.

Second, they argued that current assessment instruments such as the LSI-R oversimplify the complex nature of risk factors, both for men and women. Merely assigning a score, for instance, to any given risk factor on a risk assessment instrument fails to reflect the

potential experiential differences between men and women who have the given risk factor. For example, on the LSI-R, the Family/Marital scale is a risk factor for both men and women. However, it may drive the criminal process quite differently for women compared to men. As Holtfreter and Cupp (2007) explained, many women may become involved in crime in the context of a relationship (i.e., becoming involved in drug-related crimes with a partner). Women's aggression, moreover, more often transpires in the context of a violent relationship, compared to men. Even if men and women obtain similar scores on this risk factor, its role and importance in understanding the nature and reasons for crime across men and women may differ qualitatively. To develop and evaluate gender-specific risk assessment procedures and treatments, such qualitative differences in the reasons for crime across gender must first be understood and then must be integrated into assessment and treatment efforts.

Given the steady increase of women in the criminal justice system, correctional systems and parole boards are in need of effective risk assessment and risk management techniques for women. Assessment instruments informed by principles of dynamic risk or criminogenic need have the potential to provide specific guidance for risk reduction. Our evaluation of the LSI-R indicates that this tool is useful for predicting general recidivism with women and includes dynamic scales that uniquely relate to recidivism. With local cross-validation, the LSI-R shows promise in informing release decisions about female offenders that recognize change, protect public safety, and inform supervision strategies to promote successful reentry.

REFERENCES

- Andrews, D. A. (1982). *The Level of Supervision Inventory (LSI): The first follow-up*. Toronto, Canada: Ministry of Correctional Services (Ontario).
- Andrews, D. A., & Bonta, J. L. (1995). *The Level of Service Inventory—Revised*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. L. (2003). *Level of Service Inventory—Revised: U.S. norms manual supplement*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. L. (2006). *The psychology of criminal conduct* (4th ed.). Newark, NJ: LexisNexis.
- Andrews, D. A., Bonta, J., & Wormith, S. J. (2004). *The Level of Service/Case Management Inventory (LS/CMI)*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., Bonta, J., & Wormith, S. J. (in press). The Level of Service (LS) assessment of adults and older adolescents. In R. K. Otto & K. S. Douglas (Eds.), *Handbook of risk assessment tools*. New York: Routledge.
- Austin, J., Bloom, B., & Donahue, T. (1992). *Female offenders in the community: An analysis of innovative strategies and programs*. San Francisco: National Council on Crime and Delinquency.
- Barnoski, R., & Aos, S. (2003). *Washington's Offender Accountability Act: An analysis of the Department of Corrections' risk assessment*. Olympia: Washington State Institute for Public Policy. Retrieved May 3, 2008, from <http://www.wsipp.wa.gov/pub.asp?docid=03-12-1202>
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.
- Belknap, J., & Holsinger, K. (2006). The gendered nature of risk factors for delinquency. *Feminist Criminology*, *1*(1), 48-71.
- Bonta, J., Law, M., & Hanson, C. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, *123*, 123-142.
- Bonta, J., Pang, B., & Wallace-Capretta, S. (1995). Predictors of recidivism among incarcerated female offenders. *The Prison Journal*, *75*, 277-294.
- Brignell, J. (2000). *Sorry, wrong number! The abuse of measurement*. London: Brignell Associates.
- Cantor-Graae, E., & Selton, J.-P. (2005). Schizophrenia and migration: A meta-analysis and review. *American Journal of Psychiatry*, *162*, 12-24.
- Clark, C. (2002). *Addressing histories of trauma and victimization through treatment*. Delmar, NY: National GAINS Center.
- Coleman, C., & Moynihan, J. (1996). *Understanding crime data*. Buckingham, UK: Open University Press.

- Coulson, G., Ilacqua, G., Nutbrown, V., Giulekas, D., & Cudjoe, F. (1996). Predictive utility of the LSI for incarcerated female offenders. *Criminal Justice and Behavior, 23*, 427-439.
- Covington, S., & Bloom, B. (2006). Gender-responsive treatment and services in correctional settings. In E. Leeder (Ed.), *Inside and out: Women, prison, and therapy* (pp. 9-34). Binghamton, NY: Hawthorne.
- Cutler, S., & Nolen-Hoeksema, S. (1991). Accounting for sex differences in depression through female victimization: Childhood sexual abuse. *Sex Roles, 24*, 425-438.
- Daly, K. (1992). Women's pathways to felony court: Feminist theories of lawbreaking and problems of representation. *Southern California Review of Law and Women's Studies, 2*, 11-52.
- DeFrances, C. J. (2001). *State court prosecutors in large districts, 2001*. Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.
- Deschenes, E., Owen, B., & Crow, J. (2007). *Recidivism among female prisoners: Secondary analysis of the 1994 BJS recidivism data set*. Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.
- Douglas, K. S., Guy, L. S., & Weir, J. (2007). *HCR-20 violence risk assessment scheme: Overview and annotated bibliography*. Burnaby, Canada: Simon Fraser University. Retrieved August 1, 2008, from <http://www.sfu.ca/psyc/faculty/hart/Guides.htm>
- Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law, 11*, 347-383.
- Federal Bureau of Investigation. (2005). *Uniform crime reports*. Washington, DC: Author.
- Federal Bureau of Investigation. (2006). *Uniform crime reports*. Washington, DC: Author.
- Folsom, J., & Atkinson, J. (2007). The generalizability of the LSI-R and the CAT to the prediction of recidivism in female offenders. *Criminal Justice and Behavior, 34*, 1044-1056.
- 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.
- Glaze, L. E., & Bonczar, T. P. (2006). *Probation and parole in the United States, 2005*. Washington, DC: U.S. Department of Justice, Office of Justice Programs. Retrieved February 13, 2008, from <http://www.ojp.gov/bjs/pub/pdf/ppus05.pdf>
- Gray, T., Mays, G. L., & Stoher, M. K. (1995). Inmate needs and programming in exclusively women's jails. *The Prison Journal, 75*, 186-202.
- Greenfeld, L., & Snell, T. (1999). *Bureau of Justice Statistics special report: Women offenders*. Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.
- Hare, R. (2003). *The Hare Psychopathy Checklist-Revised*. Toronto, Canada: Multi-Health Systems.
- Harris, M. (1998). *T.R.E.M. Trauma Recovery and Empowerment: A clinician's guide for working with women in groups*. New York: Free Press.
- Heilbrun, K., Douglas, K. S., & Yasuhara, K. (in press). Violence risk assessment: Core controversies. In J. L. Skeem, K. S. Douglas, & S. O. Lilienfeld (Eds.), *Psychological science in the courtroom: Controversies and consensus*. New York: Guilford.
- Hiday, V. A., Swartz, M., & Swanson, J. W. (1998). Male-female differences in the setting and construction of violence among people with severe mental illness. *Social Psychiatry and Psychiatric Epidemiology, 33*, 568-574.
- Hodgins, S., Lapalme, M., & Toupin, J. (1999). Criminal activities and substance use of patients with major affective disorders and schizophrenia: A 2-year follow-up. *Journal of Affective Disorders, 55*, 187-202.
- Hollin, C. R., & Palmer, E. J. (2006). Criminogenic need and women offenders: A critique of the literature. *Legal and Criminological Psychology, 11*, 179-195.
- Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2003). Ethnicity, gender, and the Level of Service Inventory-Revised. *Journal of Criminal Justice, 31*, 309-320.
- Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2006). Exploring the validity of the Level of Service Inventory-Revised with Native American offenders. *Journal of Criminal Justice, 34*, 331-337.
- Holtfreter, K., & Cupp, R. (2007). Gender and risk assessment: The empirical status of the LSI-R for women. *Journal of Contemporary Criminal Justice, 23*, 363-382.
- Holtfreter, K., Reisig, M. D., & Morash, M. (2004). Poverty, state capital, and recidivism among women offenders. *Criminology and Public Policy, 3*, 185-208.
- James, D., & Glaze, L. (2006). *Mental health problems of prison and jail inmates*. Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.
- Junginger, J., Claypoole, K., Laygo, R., & Crisanti, A. (2006). Effects of serious mental illness and substance use on criminal offenses. *Psychiatric Services, 57*, 879-882.
- Kirkpatrick, B. (1999). Exploratory research of female risk prediction and LSI-R. *National Journal of Corrections, 24*(5), 1-17.
- Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior, 28*, 471-489.
- Landenberger, N., & Lipsey, M. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology, 1*, 451-476.

- Langan, P., & Levin, D. (2002). *Recidivism of prisoners released in 1994: Bureau of Justice Statistics special report*. Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.
- Laub, J., & Sampson, R. (2003). *Shared beginnings, divergent lives: Delinquent boys to age 70*. Cambridge, MA: Harvard University Press.
- Lidz, C., Mulvey, E., & Gardner, W. (1993). The accuracy of predictions of violence to others. *Journal of the American Medical Association*, *269*, 1007-1011.
- Little Hoover Commission. (2004). *Breaking the barriers for women on parole*. Sacramento: Author. Retrieved December 9, 2008, from <http://www.lhc.ca.gov/lhcdir/report177.html>
- Lowenkamp, C. T., Holsinger, A. M., & Latessa, E. J. (2001). Risk/need assessment, offender classification, and the role of childhood abuse. *Criminal Justice and Behavior*, *28*, 543-563.
- Lowenkamp, C., Smith, P., Latessa, E., & Cullen, F. (in press). Can 14,737 women be wrong? A meta-analysis of the LSI-R and recidivism for female offenders. *Criminology and Public Policy*.
- Manchak, S., Skeem, J., & Douglas, K. (2008). Utility of the Revised Level of Service Inventory (LSI-R) in predicting recidivism after long-term incarceration. *Law and Human Behavior*, *32*, 477-488.
- McClellan, D., Farabee, D., & Crouch, B. (1997). Early victimization, drug use and criminality. *Criminal Justice and Behavior*, *24*, 455-477.
- Messina, N., & Grella, C. (2006). Childhood trauma and women's health outcomes in a California prison population. *American Journal of Public Health*, *96*, 1842-1848.
- Messina, N., Grella, C., Burdon, W., & Prendergast, M. (2007). Childhood adverse events and current traumatic distress: A comparison of men and women drug-dependent prisoners. *Criminal Justice and Behavior*, *34*, 1385-1401.
- Mihailides, S., Jude, B., & Van den Bossche, E. (2005). The LSI-R in an Australian setting: Implications for risk/needs decision-making in forensic contexts. *Psychiatry, Psychology and Law*, *12*, 207-217.
- Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P., Mulvey, E., et al. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York: Oxford University Press.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Palmer, E. J., & Hollin, C. R. (2007). The Level of Service Inventory-Revised with English women prisoners: A needs and reconviction analysis. *Criminal Justice and Behavior*, *34*, 971-984.
- Poels, V. (2007). Risk assessment of recidivism of violent and sexual female offenders. *Psychiatry, Psychology and Law*, *14*, 227-250.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Raynor, P. (2007). Risk and need assessment in British probation: The contribution of LSI-R. *Psychology, Crime & Law*, *13*, 125-138.
- Raynor, P., Kynch, J., Roberts, C., & Merrington, S. (2000). *Risk and need assessment in probation services: An evaluation* (Research Study 211). London: Home Office.
- Reentry Policy Council. (2005). *Parole status by state*. New York: Council of State Governments Justice Center. Retrieved September 8, 2006, from <http://www.reentrypolicy.org/reentry/Search.aspx>
- Robbins, P. C., Monahan, J., & Silver, E. (2003). Mental disorder, violence, and gender. *Law and Human Behavior*, *27*, 561-571.
- Sabol, W. J., & Couture, H. (2008). *Prison inmates at midyear 2007*. Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.
- Sabol, W. J., Minton, T. D., & Harrison, P. M. (2007). *Prison and jail inmates at midyear 2006*. Washington, DC: U.S. Department of Justice.
- Simourd, D. J. (2004). Use of dynamic risk/need assessment instruments among long-term incarcerated offenders. *Criminal Justice and Behavior*, *31*, 306-323.
- Sorbell, L., Eccleston, L., Ward, T., & Jones, R. (2002). Treatment needs of female offenders: A review. *Australian Psychologist*, *37*, 198-205.
- Tolman, A., & Mullendore, K. (2003). Risk evaluations for the courts: Is service quality a function of specialization? *Professional Psychology: Research and Practice*, *34*, 225-232.
- Trickett, P. K., & McBride-Chang, C. (1995). The developmental impact of different forms of child abuse and neglect. *Developmental Review*, *15*, 311-337.
- Valdez, A., Kaplan, C. D., & Curtis, R. L. (2007). Aggressive crime, alcohol and drug use, and concentrated poverty in 24 U.S. urban areas. *American Journal of Drug and Alcohol Abuse*, *33*, 595-603.
- Washington State Institute for Public Policy. (2003). *Sentences for adult felons in Washington: Identifying changes that could reduce costs without endangering public safety—interim report*. Olympia, WA: Author. Retrieved September 8, 2006, from <http://www.wsipp.wa.gov/rptfiles/sentinterim.pdf#search=%22determinate%20sentencing%22>
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence, Version 2*. Burnaby, Canada: Simon Fraser University.

- Widom, C. S. (1989). Child abuse, neglect, and adult behavior: Research design and findings on criminality, violence, and child abuse. *American Journal of Orthopsychiatry*, 59, 355-367.
- Widom, C. S., & Maxfield, M. G. (1996). A prospective examination of risk for violence among abused and neglected children. In C. F. Ferris & T. Grisso (Eds.), *Understanding aggressive behavior in children* (pp. 224-237). New York: New York Academy of Sciences.
- Wright, D. (2005). Receiver operating characteristics curves. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. 4, pp. 1718-1721). Chichester, UK: John Wiley.

Sarah M. Manchak, MA, is a graduate student in the Department of Psychology and Social Behavior at the University of California, Irvine. She received her MA from John Jay College of Criminal Justice and is now pursuing her PhD with a concentration in psychopathology and a minor focus in psychology and the law. Her research focuses on violence risk assessment and the management and treatment of offenders with mental disorder.

Jennifer L. Skeem, PhD, is an associate professor of psychology and social behavior at the University of California, Irvine (UCI). She is a member of the MacArthur Research Network on Mandated Community Treatment and UCI Centers for Psychology and Law and Evidence-Based Corrections. Her research is designed to inform clinical and legal decision making about individuals with mental disorder. Specific topics include understanding psychopathic personality disorder, assessing and treating violence risk, and identifying factors that influence the outcomes of offenders who are required to accept psychiatric treatment. She received the Saleem Shah Award for Early Career Excellence from the American Psychological Association (Division 41) and the Distinguished Assistant Professor Award for Research from the Academic Senate of UCI.

Kevin S. Douglas, LLB, PhD, is an associate professor and coordinator of the Law and Forensic Psychology Program in the Department of Psychology at Simon Fraser University and a guest professor of applied criminology at Mid-Sweden University, Sundsvall, Sweden. His work focuses on forensic assessment and more specifically on violence risk assessment and psychopathy. In 2005, he received the Saleem Shah Award for Early Career Contributions to Law and Psychology, jointly awarded by the American Psychology-Law Society and the American Academy of Forensic Psychology, as well as a Career Scholar Award from the Michael Smith Foundation for Health Research.

Maro Siranosian received her BA in criminology, law, and society from the University of California, Irvine. As an honors student, she focused her thesis work on violence risk assessment and female offenders. Currently, she is involved in international research focusing on coercion/social support theory.