Criminal Justice and Behavior

http://cjb.sagepub.com

Assessing Risk for Violence in Adolescents Who Have Sexually Offended: A Comparison of the J-SOAP-II, J-SORRAT-II, and SAVRY

Jodi L. Viljoen, Mario Scalora, Lorraine Cuadra, Shannon Bader, Verónica Chávez, Daniel Ullman and Lisa Lawrence

The online version of this article can be found at:
http://cjb.sagepub.com/cgi/content/abstract/35/1/5

Published by:

$SAGE

http://www.sagepublications.com

On behalf of:

International Association for Correctional and Forensic Psychology

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/cgi/content/refs/35/1/5

ASSESSING RISK FOR VIOLENCE IN ADOLESCENTS WHO HAVE SEXUALLY OFFENDED

A Comparison of the J-SOAP-II, J-SORRAT-II, and SAVRY

JODI L. VILJOEN
Simon Fraser University

MARIO SCALORA
LORRAINE CUADRA
SHANNON BADER
VERÓNICA CHÁVEZ
University of Nebraska–Lincoln

DANIEL ULLMAN
Lincoln Regional Center

LISA LAWRENCE
Western New England College

As the youth justice system has evolved, clinicians have been increasingly asked to make judgments about the likelihood that a youth who has committed a sexual offense will reoffend. However, there is an absence of well-validated tools to assist with these judgments. This study examined the ability of the Juvenile Sexual Offense Recidivism Risk Assessment Tool–II (J-SORRAT-II), Structured Assessment of Violence Risk in Youth (SAVRY), and Juvenile Sex Offender Assessment Protocol–II (J-SOAP-II) to predict violent behavior in 169 male youth who were admitted to a residential adolescent sex offender program. Total scores on the SAVRY and J-SOAP-II significantly predicted nonsexual violence but none of the instruments predicted sexual violence. The J-SOAP-II and SAVRY were less effective in predicting violent reoffending in youth aged 15 and younger than in older youth. The implications of these findings are discussed.

Keywords: risk assessment; violence; sex offending; juvenile; adolescent

Clinicians are commonly asked to make judgments about the likelihood that a youth who has committed a sexual offense will reoffend. Courts often seek clinicians' opinions about a youth's violence risk level in making decisions about whether a youth should be transferred to adult court (*Kent v. United States*, 1966), and in determining appropriate placements and dispositions (Hoge, 2002). In treatment settings, knowledge about a youth's risk is used to assist with treatment planning by identifying risk factors that may be targeted

5

by treatment and to guide decisions about when a youth may be ready for a less restrictive environment.

In addition, an increasing number of states have established laws whereby adolescents who have committed a sexual offense, and are judged to be a high risk of subsequent violence, are listed on sex offender registries, made subject to community notification laws, and/or involuntarily committed to a psychiatric hospital following the completion of formal dispositions (Caldwell, 2002; Trivits & Reppucci, 2002; Zimring, 2004). The extension of these laws to adolescents is controversial for a number of reasons, including concern about whether it is possible to assess youths' risk of subsequent violence with sufficient accuracy.

Experts have noted that it may be challenging to assess risk for violence among adolescents. Because adolescence is a period of enormous change and development, assessing youths' violence risk has been described as akin to assessing "moving targets" (Borum, 2003; Grisso, 1998; Prentky & Righthand, 2003). The difficulty of distinguishing between youth who are high and low risk is underscored by the finding that many adolescents who engage in deviant behaviors desist as they mature (Moffitt, 1993). For instance, when adolescents who have committed sexual offenses are followed for periods of up to 6 years, a relatively small proportion, typically 15% or fewer), are rearrested or reconvicted of sexual offenses, although a more substantial proportion commit nonsexual offenses (Worling & Långström, 2003; Worling & Curwen, 2002). With effective treatment, rates of reoffending may be quite low (Borduin, Henggeler, Blaske, & Stein, 1990; Worling & Curwen, 2000).

Another key factor that makes it difficult to assess risk for violence among adolescents is the absence of well-validated approaches to guide judgments. Given the well-documented limitations of unstructured clinical judgments (e.g., Hanson & Bussière, 1998), several adolescent risk assessment tools have recently been developed to assist clinicians. Some of these tools include the include the Juvenile Sex Offender Assessment Protocol -II (J-SOAP-II; Prentky & Righthand, 2003) and the Juvenile Sexual Offense Recidivism Risk Assessment Tool -II (J-SORRAT-II; Epperson, Ralston, Fowers, & DeWitt, 2005), which were designed to assess risk in adolescents who have committed sexual offenses, and the Structured Assessment of Violence Risk in Youth, which was designed to assess general violence risk rather than specifically sexual violence risk (SAVRY; Borum, Bartel, & Forth, 2003).

Although the development of these tools represents a significant step, there is currently inadequate evidence regarding their predictive validity. At the time of writing this article, no published studies have examined the predictive validity of the J-SOAP-II or the SAVRY with adolescents who have committed sexual offenses. Although an unpublished report noted that the J-SORRAT-II was able to predict reoffending (Epperson et al., 2005), this study used the sample with which the instrument was developed, which tends to result in inflated estimates of predictive validity (Silver, Smith, & Banks, 2000). As such, it is important to determine its psychometric properties in an independent sample.

To address the need for empirical research on these tools, the present study examined the ability of the J-SOAP-II, SAVRY, and J-SORRAT-II to predict whether male adolescents in a residential sex-offender-treatment program engaged in sexual and nonsexual violence during and following treatment. In addition, we examined whether these tools were equally effective in predicting violence among younger and older youth. Although the J-SOAP-II, SAVRY, and J-SORRAT-II were designed for use with adolescents aged 12 to 18 (Borum et al., 2003; Epperson et al., 2005; Prentky & Righthand, 2003), it is possible that they may not function equally well for younger and older adolescents. For example, young adolescents may receive higher scores on items on these tools (e.g., impulsivity, lack of empathy)

because of their developmental stage rather than more stable characteristics that are indicative of reoffense risk.

METHOD

PARTICIPANTS

Participants were 169 male adolescents who were admitted to a nonsecure residential treatment program for sexually abusive adolescents in a medium-sized midwestern city in the United States. To be eligible for admittance to the treatment program, adolescents must be adjudicated by the courts for sexual or nonsexual offenses and ordered to receive treatment. Our sample included all youth who were admitted to the program from 1992 to 2005. To help ensure an adequate follow-up period, all youth had been discharged from the program at least 250 days prior to data collection.

The mean age of youth at the time of admission was 15.37 ($SD = 1.51$). The large majority of youth in the sample were non-Hispanic Caucasian (83.4%, $n = 141$), and 8.3% ($n = 14$) were African American, 4.7% ($n = 8$) were Hispanic, 1.2% ($n = 2$) were American Indian or Alaskan Native, and 2.4% ($n = 4$) were of mixed race/ethnicity.

Youth in our sample had committed a variety of sexual-offense-related behaviors that led to treatment, including genital penetration (37.3%, $n = 63$), anal penetration (34.9%, $n = 59$), digital penetration (13.0%, $n = 22$), oral–genital contact (49.1%, $n = 83$), fondling (61.5%, $n = 104$), and exhibitionism (13.6%, $n = 23$). These numbers do not add up to 100% because some youth had multiple index offenses. The majority of youths' index offenses were perpetrated against victims who were at least 3 years younger than themselves (82.8%, $n = 140$). In most cases, the victims were known to the perpetrator (90.5%, $n = 153$). Also, many of the victims were related (69.8%, $n = 118$), such as being a biological sibling, half-sibling, or stepsibling. Almost one quarter of youth had committed index offenses against male-only victims (23.7%, $n = 40$), whereas 46.2% ($n = 78$) perpetrated against female-only victims, and 28.4% ($n = 48$) perpetrated against both male and female victims. Approximately half of youth in our sample had prior charges for sexual (50.9%, $n = 86$) and nonsexual offenses (49.1%, $n = 83$). In total, youth had a mean of 1.57 prior sexual charges ($SD = 1.41$) and 2.00 prior nonsexual charges ($SD = 3.16$).

PROCEDURE

Trained raters completed the J-SOAP-II, J-SORRAT-II, and SAVRY for each youth based on comprehensive file information. Three PhD-level graduate students in clinical forensic psychology, who had completed coursework and practica on clinical and forensic assessment of youth, rated youth on the J-SOAP-II and SAVRY. As the J-SORRAT-II does not require clinical training to complete, two undergraduate raters, who were psychology majors and had completed coursework in forensic psychology, completed this tool. Prior to commencing coding, all raters received comprehensive training on the tools, which included didactic sessions, assigned readings, and the completion of five practice cases using actual case files. The practice cases were reviewed and discussed among the raters. All raters were blind to youths' subsequent charges and convictions.

Because the J-SOAP-II and SAVRY contain a number of dynamic variables, these two instruments were coded twice, once based on information available at admission (admissions scores) and again based on information available at discharge (discharge scores). In investigating the ability of these tools to predict postdischarge offenses, youths' discharge scores were examined. However, in investigating the ability of these tools to predict aggressive behavior during treatment, youths' admissions scores were examined, as these scores are less likely to be contaminated by the outcome (i.e., aggressive behavior during treatment). The J-SORRAT-II does not contain dynamic variables, so it was coded only once based on all of the available file information.

Information was collected on whether youth engaged in sexual aggression and non-sexual aggression both during and following the treatment program by examining law-enforcement, probation, and treatment records. On average, youth spent approximately 1 year in the treatment period ($M = 389.73$ days, $SD = 232.29$), during which time their aggressive behaviors were examined. Also, youth were followed for an average of 6.58 years following discharge from the treatment program ($SD = 3.49$, range = 280 days to 12.01 years). This is longer than the follow-up periods that have been reported by other studies (see Caldwell, 2002).

DESCRIPTION OF FILE INFORMATION

Files were generally at least 600 pages in length and included the following components: psychiatric assessments, psychological assessments (which included a comprehensive battery of tests administered at intake and every 3 months), nursing records, medical examination information, social work reports, teacher assessments, school records, treatment plans, progress notes, physician orders, and correspondence with courts (e.g., arrest records) and other treatment providers.

To examine whether the risk assessment tools could be reliably coded based on the available information, a random sample of 21.9% ($n = 37$) files were selected. Another rater who had similar training and background separately recoded these files. As described in the Results section, intraclass correlations for all measures were sufficient.

Also, we examined whether the raters believed that the file information was adequate to complete the risk assessment tools. Specifically, raters were asked to rate the quality of each file on a scale of 1 (*extremely poor in quality*) to 10 (*extremely good in quality*). The quality ratings was 7.56 ($SD = 1.22$, median = 8.00, mode = 8.00), and only nine files were rated as having scores of 5 or less. Furthermore, raters did not code any of the items for any youth as unratable because of a lack of information. This indicates that, in general, the raters believed that there was adequate information with which to complete the risk assessment tools. To further ensure that the few files that were rated as having comparatively low quality did not alter the estimates of predictive validity, we examined the predictive validity of the risk assessment tools both having included and excluded the files that received quality ratings of 5 or less. Equivalent results were obtained. Therefore, the full sample of youth was included in the presentation of results.

RISK MEASURES

J-SOAP-II. The J-SOAP-II is a 28-item checklist of risk factors that was designed to assess risk for sexual violence as well as general delinquency (Prentky & Righthand, 2003).

It is intended for use specifically with adolescent boys, aged 12 to 18, who have a history of sexually coercive behavior. In developing this instrument, the authors selected risk factors that had empirical support and/or clinical relevance (Prentky & Righthand, 2003). Then, after preliminary investigations, the pilot version of this scale underwent fairly extensive revisions, during which concrete behavioral anchors were developed, items with limited predictive validity were deleted, and new risk factors were added.

The resulting tool, the J-SOAP-II, consists of four scales. Two scales, Sexual Drive/ Preoccupation (e.g., prior sexual offense charges) and Impulsive/Antisocial Behavior (e.g., past school behavior problems), include static risk factors that generally do not change over time. The other two scales, Intervention (e.g., remorse and guilt) and Community Stability/ Adjustment (e.g., management of sexual urges), focus on dynamic (or potentially changeable) risk factors.

Items on the J-SOAP-II are rated on a 3-point scale, with a higher score representing greater risk. A total score is then obtained by summing the items on the four scales. At the present time, there are no classifications associated with various total scores, and the J-SOAP-II functions as an "empirically informed guide" rather than an actuarial tool (Prentky & Righthand, 2003, p. 8).

Although the J-SOAP-II and its earlier version, the J-SOAP, appear to be fairly widely used (see Burton, Smith-Darden, & Frankel, 2006), relatively little is known about its predictive validity. Prentky, Harris, Frizzell, and Righthand (2000) completed the early version of this tool, the J-SOAP, with 96 youths but found the base rate of violence to be too low after a 12-month follow-up to conduct inferential statistics (i.e., only 3 youths sexually reoffended). Furthermore, no published studies have reported information regarding the psychometric properties of the J-SOAP-II. This is important because the J-SOAP-II differs fairly substantially from the J-SOAP; in forming the J-SOAP-II from the J-SOAP, more than 10 items were added, deleted, moved to a different scale, or revised.

J-SORRAT-II. The J-SORRAT-II is a 12-item actuarial tool designed for assessing risk of violence among male juvenile offenders who were 12 to 18 years old at the time of their index sexual offense (Epperson et al., 2005). A number of items on the J-SORRAT-II focus on the youths' sexual and nonsexual offense history (e.g., number of adjudications as a sex offender, number of victims in sex offenses). Other variables examine youths' treatment history (i.e., completion of sex offender treatment), school history (e.g., special education), and past victimization experiences (e.g., number of physical abuse victimization events).

The J-SORRAT-II has an objective, criterion-based scoring system. A number of J-SORRAT-II items are scored "0" or "1" to indicate whether the risk factor is present or absent. Other items are scored on a 3-point scale (0 to 2) or a 4-point scale (0 to 3) to indicate varying degrees of severity for a particular risk factor.

The J-SORRAT-II was developed by identifying key predictors of sexual offending in a sample of 636 male youths who were adjudicated for a sex offense (Epperson et al., 2005). In this test development sample, the J-SORRAT-II was very effective in predicting recidivism (Epperson et al., 2005). Specifically, receiver operating curve (ROC) analyses indicated that the area under the curve (AUC) was .89 for predicting the likelihood that a juvenile would sexually recidivate as a juvenile and .79 in predicting the likelihood that a youth would sexually recidivate as either a juvenile or adult. However, research using independent samples has not yet been conducted.

SAVRY. The SAVRY is a scheme for assessing violence risk among adolescents aged 12 to 18 (Borum et al., 2003). The SAVRY aims to assess general risk of future violence, rather than specifically sexual violence risk (Borum et al., 2003). The SAVRY consists of 24 risk items, which were derived from the empirical literature on youth violence.

The first set of risk factors are historical factors (e.g., early initiation of violence), which are static and generally immune to change. The remaining risk factors include social and contextual factors (e.g., poor parental management) and individual factors (e.g., risk taking and impulsivity), which are all considered to be dynamic. Each risk factor is coded as *high*, *moderate*, or *low*. In addition to these risk factors, the SAVRY assesses whether six protective factors (e.g., prosocial involvement) are present or absent.

After rating risk and protective factors, evaluators use the SAVRY to make a structured professional judgment about a youth's risk for violence (*high*, *moderate*, *low*). For the present study, we created separate section scores for historical, social and contextual, and individual risk factors as well as protective factors by summing the items from the respective sections. In addition, a total score was created by summing the items from the historical, social and contextual, and individual risk factors sections.

There is some evidence that supports the predictive validity of the SAVRY. Catchpole and Gretton (2003) found that SAVRY risk ratings were significantly related to violent reoffending. Specifically, 40% of youth classified as high risk on the SAVRY violently recidivated in a 1-year period as compared to 6% of youth classified as low risk on the SAVRY. To date, however, research has not yet examined the use of the SAVRY with adolescents who have committed sexual offenses.

OUTCOME MEASURES

To assess violence during treatment, raters coded whether youth engaged in sexual aggression and nonsexual aggression during the treatment program. Interrater reliability was examined for a random sample of 21.9% of the youth ($n = 37$). The kappa coefficient for single raters was .77 ($p < .001$) for sexual aggression during treatment and .64 ($p < .001$) for nonsexual aggression. These kappa coefficients are considered "good" (Cichetti, 1994; Cicchetti & Sparrow, 1981).

To assess sexual and nonsexual violence after discharge, youths' statewide juvenile justice and adult criminal records were obtained through law enforcement and probation sources as well as by reviewing subsequent treatment records. A graduate student rater, who was not involved in completing the risk assessment tools, coded whether the youth was charged with sexual offenses, any nonsexual violent offenses (defined as either violent felonies or violent misdemeanors), serious nonsexual violent offenses (defined as at least one violent felony charge), and/or any offenses (defined as any nontraffic felonies or misdemeanors) following discharge. The primary author reviewed all ratings. These raters only rated two cases differently, and the primary author resolved these cases.

DATA ANALYSIS

ROC analyses. ROC analyses were used to evaluate the ability of risk assessment tools to predict youths' treatment behavior and reoffending. ROC analysis computes an AUC by

plotting the sensitivity of a tool against its specificity (Mossman, 1994). The resulting AUC score can range from 0 to 1 and indicates the probability that a randomly selected individual who recidivated would score higher on the risk assessment tool than a randomly selected individual who did not recidivate. An AUC of 1 indicates that the tool is able to predict the criterion behavior perfectly, whereas an AUC of .5 indicates that the tool is not able to predict any better than chance. ROC analysis is now widely used in risk assessment research (Douglas, Ogloff, Nicholls, & Grant, 1999; Harris et al., 2003). One of its major strengths is that it is less affected by base rates than other approaches.

Odds ratios. Odds ratios (ORs) were calculated to determine the likelihood that an individual who scored at or greater than the median on each risk assessment tool engaged in violence. ORs of 2.50 or higher are generally classified as strong (Fleiss, Williams, & Dubro, 1986). For this study, an OR of 2.50 would mean that youth scoring at or greater than the median would be 2.50 times more likely than youth at or less than the median to engage in a particular type of aggressive or offending behavior.

Kaplan-Meier survival curve analyses. Kaplan-Meier survival curve analyses were used to estimate the length of time (in months) that it took for youth who scored at or greater than the median for each tool to reoffend compared to youth who scored less than the median (Luke & Homan, 1998). This method determines the proportion of youths who "survive" in the analyses, meaning they have not yet reoffended at each month of the follow-up.

Moderated hierarchical logistic regression. To examine whether age moderated the relationship between total scores on the risk assessment tools and accuracy of predictions, moderated hierarchical logistic regression was performed (see Dempster, 2001). The total scores on the tools, which were continuous, were first standardized (Cohen, Cohen, West, & Aiken, 2003). Then the standardized scores (z-scores) of the risk assessment tool and age were multiplied to form product terms that represented the interaction between the predictor (risk assessment tool) and potential moderator (age). For each tool, a hierarchical logistic regression was performed, in which the tool's total score and age (defined as 15 and younger versus 16 and older) were entered into the regression equation in Step 1, followed by the product term in Step 2 (Baron & Kenny, 1986; Holmbeck, 1997).

We chose to compare adolescents aged 15 and younger ($n = 89$) to those aged 16 and older ($n = 80$) because research has suggested that there may be important developmental differences between these age groups. Research has, for instance, reported that adolescents aged 15 and younger score significantly lower than older adolescents and adults on measures of psychosocial maturity, cognitive ability, and judgment skills, whereas adolescents aged 16 to 18 score more similarly to adults (Cauffman & Steinberg, 1999; Grisso et al., 2003). In addition, some test developers have argued that although adolescent risk assessment tools are essential for younger adolescents, there may be some situations in which likely adult instruments might be appropriate for older adolescents if used very cautiously (Hanson & Thornton, 1999). This raises the question of whether adolescent risk assessment tools function equally well for younger and older adolescents.

TABLE 1:   Ratings on the Risk Assessment Tools

Measure	Range	M	SD	SEM	Median	ICC_1
J-SORRAT-II	0-17	6.19	3.60	0.28	5.00	.89
J-SOAP-II	9-53	30.91	9.26	0.71	31.00	.84
SAVRY	5-44	24.05	8.14	0.63	24.00	.91

Note. M = mean; *SD* = standard deviation; SEM = structural equation modeling; ICC = intraclass correlation coefficients; J-SORRAT-II = Juvenile Sexual Offense Recidivism Risk Assessment Tool–II; J-SOAP-II = Juvenile Sex Offender Assessment Protocol–II; SAVRY = Structured Assessment of Violence Risk in Youth.

RESULTS

PROPERTIES OF RISK ASSESSMENT TOOLS

The risk scores of youth at discharge are presented in Table 1. On the SAVRY, 17.2% (*n* = 29) youth were classified as low risk, 68.0% (*n* = 115) youth as moderate risk, and 14.8% (*n* = 25) as high risk. At the present time, the J-SOAP-II and the J-SORRAT-II do not use any risk level categories.

To examine the interrater reliability of the measures, separate raters with similar training and background recoded the files of a random sample of 21.9% of the cases (*n* = 37). Intraclass correlation coefficients for single raters (ICC_1s) were calculated using a two-way random effects model (McGraw & Wong, 1996). The ICC_1s for total scores on the SAVRY, J-SOAP-II, and J-SORRAT-II fell in the *excellent* range, according to the classification system used by Cicchetti and Sparrow (1981; see Table 1).

Bivariate Pearson *r* correlations were used to examine correlations between continuous scores on risk measures. When one of the scores was ordinal (i.e., structured professional ratings on the SAVRY), Spearman-Brown correlations were calculated. As shown in Table 2, nearly all the risk measures were correlated with each other. The correlations between the J-SOAP-II total scores, SAVRY total scores, and structured professional judgments on the SAVRY were classified as large, according to Cohen's (1988) classification system. In contrast, the correlations between the J-SORRAT-II and these measures were significant but small.

TREATMENT BEHAVIOR

Predictive validity of the tools—AUC scores. Based on treatment records, 28 youth (16.6%) engaged in sexual aggression and 51 (30.2%) engaged in nonsexual aggression during treatment. Total scores on the J-SORRAT-II did not significantly predict sexual aggression or nonsexual aggression. Although total scores on the J-SOAP-II at admission did not significantly predict sexual aggression during treatment, they significantly predicted nonsexual aggression during treatment (see Table 3). When specific subscales on the J-SOAP-II were examined, all of the subscales on the J-SOAP-II—except for Sexual Drive/Preoccupation—were associated with nonsexual aggression. However, the Sexual Drive/Preoccupation scale was associated with *sexual* aggression. Like the J-SOAP-II, SAVRY total scores at admission did not significantly predict sexual aggression during treatment, but predicted nonsexual aggression. Based on *z*-tests for dependent groups (Hanley & McNeil, 1982),[1] there were no significant differences in the ability of the SAVRY and J-SOAP-II to predict nonsexual aggression.

TABLE 2: Correlations Between Risk Measures

Measure	J-SORRAT-II	J-SOAP-II Total	Sexual Drive Scale	Impulsive/ Antisocial Scale	Intervention Scale	Stability Scale	SAVRY Total	Historical Section	Social/ Contextual Section	Individual Section	SAVRY Structured Professional Rating
J-SORRAT-II	—	.28***	.42***	.26**	.04	.08	.19**	.24**	.09	.13	.16*
J-SOAP-II Total		—	.60***	.72***	.55***	.79***	.88***	.63***	.74***	.79***	.74***
Sexual Drive			—	.25**	.20**	.20**	.31***	.30***	.25**	.21**	.35***
Impulsive/Antisocial				—	.22**	.47**	.70***	.75***	.51***	.42***	.52***
Intervention					—	.50***	.70***	.17*	.49***	.58***	.65***
Stability						—	.50***	.46***	.70***	.81***	.65***
SAVRY Total							—	.77***	.83***	.84***	.74**
Historical								—	.46***	.36***	.50***
Social/Contextual									—	.68***	.62***
Individual										—	.68***

Note. J-SORRAT-II = Juvenile Sexual Offense Recidivism Risk Assessment Tool–II; J-SOAP-II = Juvenile Sex Offender Assessment Protocol–II; SAVRY = Structured Assessment of Violence Risk in Youth.

*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 3:   Comparison of Risk Assessment Tools in Predicting Treatment Behavior for Overall Sample

Measure	Sexual Aggression During Treatment			Nonsexual Aggression During Treat\ment		
	AUC	SE	95% CI	AUC	SE	95% CI
J-SORRAT-II total	.59	.06	.47-.70	.58	.05	.49-.67
J-SOAP-II total (admission)	.61	.06	.50-.72	.66**	.05	.57-.75
Sexual drive/preoccupation	.65*	.05*	.55-.76	.56	.05	.46-.65
Impulsive/antisocial	.52	.06	.41-.63	.63**	.05	.54-.72
Intervention	.60	.06	.49-.71	.61*	.05	.52-.70
Stability	.57	.06	.45-.68	.67**	.05	.58-.76
SAVRY total (admission)	.52	.06	.41-.63	.69***	.04	.60-.77
Historical	.49	.06	.38-.60	.67***	.05	.59-.76
Social/contextual	.50	.06	.38-.62	.61*	.05	.52-.71
Individual	.61	.05	.50-.71	.67***	.04	.59-.76
Protective	.48	.06	.36-.59	.62*	.05	.52-.70
Structured professional rating	.51	.06	.40-.62	.59	.05	.50-.68

Note. J-SORRAT-II = Juvenile Sexual Offense Recidivism Risk Assessment Tool–II; J-SOAP-II = Juvenile Sex Offender Assessment Protocol–II; SAVRY = Structured Assessment of Violence Risk in Youth; AUC = area under the curve; SE = standard error; CI = confidence interval.
*p < .05. **p < .01. ***p < .001.

Predictive validity of the tools—ORs. When the ORs were examined, we found that youth who scored at or greater than the median on the SAVRY were nearly 3 times more likely to engage in nonsexual aggression during treatment than youth who scored less than the median (OR = 3.83, p < .01). The ORs for the J-SOAP-II and J-SORRAT-II were not significant.

Developmental differences in predictive validity. In general, higher AUCs were obtained for older youth (aged 16 to 18 at admission) than younger youth (aged 12 to 15 at admission) on the J-SOAP-II and SAVRY (see Table 4). However, moderated hierarchical logistic regression indicated that there were no significant differences in the ability of the tools to predict sexual or nonsexual aggression during treatment among youth aged 12 to 15 and youth aged 16 to 18.

POSTDISCHARGE OFFENSES

Rates of postdischarge offenses. For the four categories of postdischarge charges, the base rate was 8.3% (n = 14) for sexual offenses, 12.7% (n = 21) for nonsexual violent offenses, 10.1% (n = 17) for *serious* nonsexual violent offenses, and 42.8% (n = 71) for any offense. Three individuals were under investigation at the time of follow-up, so we were unable to obtain their offense records.

Predictive validity of the tools—AUCs. Total scores on the J-SORRAT-II did not predict reoffending of any type. Total scores on the J-SOAP-II did not *significantly* predict reoffending of any type but nearly reached significance in predicting serious nonsexual violence (p = .076; see Table 5). SAVRY total scores did not significantly predict sexual offenses following discharge or the any offense category but predicted *serious* nonsexual violent

TABLE 4: Comparison of Risk Assessment Tools Using Area Under the Receive Operating Characteristic (ROC) by Age of Youth

Measure	J-SORRAT-II Total			J-SOAP-II Total			SAVRY Total		
	AUC	SE	95% CI	AUC	SE	95% CI	AUC	SE	95% CI
Youth aged 12 to 15									
Treatment behavior									
Sexual aggression	.54	.10	.35-.73	.54	.09	.38-.71	.47	.08	.31-.63
Nonsexual aggression	.60	.07	.47 -.73	.62	.07	.50-.75	.66*	.06	.54-.77
Postdischarge offenses									
Sexual offense	.55	.14	.28-.83	.59	.15	.29-.89	.54	.14	.27-.81
Any nonsexual violent offense	.60	.10	.41-.79	.54	.12	.31-.77	.56	.12	.32-.79
Serious nonsexual violence	.57	.12	.32-.81	.44	.13	.19-.70	.52	.15	.23-.81
Any offense	.54	.08	.39-.70	.46	.08	.31-.62	.44	.08	.28-.61
Youth aged 16 to 18									
Treatment behavior									
Sexual aggression	.59	.26	.45-.73	.67*	.07	.53 -.80	.58	.07	.43-.72
Nonsexual aggression	.56	.07	.43-.70	.69**	.07	.56 -.82	.73**	.06	.61-.85
Postdischarge offenses									
Sexual offense	.53	.11	.31 -.75	.53	.09	.36-.70	.53	.09	.35-.71
Any nonsexual violent offense	.53	.06	.41-.66	.58	.06	.47 -.70	.62	.06	.51-.73
Serious nonsexual violence	.53	.07	.40-.66	.71*	.08	.55-.86	.77**	.06	.66-.81
Any offense	.54	.06	.44-.65	.60	.05	.50-.71	.65**	.05	.55 -.75

Note. J-SORRAT-II = Juvenile Sexual Offense Recidivism Risk Assessment Tool–II; J-SOAP-II = Juvenile Sex Offender Assessment Protocol–II; SAVRY = Structured Assessment of Violence Risk in Youth; AUC = area under the curve; *SE* = standard error; CI = confidence interval.
*$p < .05$. **$p < .01$. ***$p < .001$.

offenses. Structured professional ratings on the SAVRY (i.e., low, moderate, and high categories) did not significantly predict reoffending of any type.

Predictive validity of the tools—ORs. When the ORs were examined for each of these tools, no significant differences were found between youth who scored greater than or at the median on any of the offense categories compared to youth who scored less than the median. However, a trend was noted such that youth who scored at or above the median on the SAVRY were significantly more likely than those who scored below the median to be detected for serious nonsexual violent offenses (OR = 2.86, $p = .059$).

Timing of offenses. Youth who reoffended did so within an average of 100.90 months (*SD* = 58.19, median = 146.13). Using survival analyses, we examined whether youth who scored above the median for each tool were more likely to be detected for reoffending. Survival analyses did not yield any significant results for the J-SOAP-II or J-SORRAT-II. However, there was a trend wherein youth who scored greater than the median on the SAVRY at discharge were more likely to be detected for serious nonsexual

TABLE 5: Comparison of Risk Assessment Tools in Predicting Postdischarge Reoffending for Overall Sample

Measure	Sexual Offense			Any Nonsexual Violent Offense			Serious Nonsexual Violent Offense			Any Offense		
	AUC	SE	95% CI	AUC	SE	95% CI	AUC	SE	95% CI	AUC	SE	95% CI
J-SORRAT-II total	.53	.09	.36–.70	.56	.05	.45–.66	.55	.06	.43–.67	.54	.05	.45–.63
J-SOAP-II total	.54	.08	.39–.68	.56	.06	.45–.66	.63*	.07	.49–.78	.56	.05	.47–.65
Sexual Drive/Preoccupation	.60	.08	.44–.76	.54	.06	.42–.66	.52	.08	.36–.69	.53	.05	.44–.62
Impulsive/Antisocial	.54	.08	.38–.69	.61*	.06	.50–.72	.67*	.08	.52–.82	.64**	.04	.55–.72
Intervention	.52	.07	.39–.66	.51	.06	.40–.62	.56	.08	.40–.72	.49	.05	.40–.58
Stability	.45	.08	.29–.61	.53	.06	.42–.64	.63*	.07	.49–.78	.54	.05	.45–.63
SAVRY total	.53	.08	.38–.67	.58	.05	.48–.68	.69*	.06	.56–.81	.58	.04	.49–.67
Historical	.60	.09	.43–.78	.61*	.05	.51–.72	.65	.07	.51–.78	.62*	.04	.53–.70
Social/contextual	.48	.09	.31–.65	.56	.06	.45–.67	.67*	.07	.54–.80	.58	.05	.49–.67
Individual	.46	.07	.33–.60	.53	.05	.43–.63	.67*	.06	.56–.78	.53	.05	.44–.62
Protective	.48	.07	.35–.62	.52	.05	.41–.62	.54	.08	.39–.69	.54	.05	.45–.63
Structured clinical rating	.51	.08	.35–.66	.51	.06	.40–.62	.56	.08	.42–.71	.50	.05	.42–.59

Note. J-SORRAT-II = Juvenile Sexual Offense Recidivism Risk Assessment Tool–II; J-SOAP-II = Juvenile Sex Offender Assessment Protocol–II; SAVRY = Structured Assessment of Violence Risk in Youth; AUC = area under the curve; SE = standard error; CI = confidence interval.
*p < .05. **p < .01. ***p < .001.

TABLE 6:   Age Differences in False Positive Rates

	Ages 12 to 15	Ages 16 to 19	z
Sexual offense			
J-SOAP-II	.74	.39	2.82**
SAVRY	.79	.35	3.87***
Any nonsexual violent offense			
J-SOAP-II	.75	.32	3.57**
SAVRY	.80	.29	4.17***
Serious nonsexual violent offense			
J-SOAP-II	.77	.33	3.69***
SAVRY	.82	.27	4.80***
Any offense			
J-SOAP-II	.80	.26	4.81***
SAVRY	.75	.17	5.35***

Note. J-SOAP-II = Juvenile Sex Offender Assessment Protocol–II; SAVRY = Structured Assessment of Violence Risk in Youth.
*p < .05. **p < .01. ***p < .001.

violent offenses than those who fell at or less than the median, log-rank statistic ($df = 1$) = 3.68, $p = .055$.

Developmental differences in predictive validity. There were no significant age differences in reoffense rates for any of the offense categories. In general, higher AUCs were obtained for older youth (aged 16 to 18 at discharge) than younger youth (aged 12 to 15 at discharge; see Table 4). Although no significant differences were found in the ability of the tools to predict sexual offenses, the J-SOAP-II was significantly better at predicting serious nonsexual violent offenses in older youth than in younger youth, $b = 1.19$, $SE = .57$, Wald = 4.46, $p = .035$, OR = 3.30, and there was a trend wherein the SAVRY was better at predicting serious nonsexual violence offenses in older youth, $b = 1.11$, $SE = .59$, Wald = 3.60, $p = .058$, OR = 3.05. Also, the SAVRY was significantly better at predicting any offense in older youth, $b = 0.86$, $SE = .36$, Wald = 5.78, $p = .016$, OR = 2.35, and there was a trend wherein the J-SOAP-II was better at predicting any offense in older youth, $b = 0.60$, $SE = .34$, Wald = 3.08, $p = .079$, OR = 1.82.

Although moderated logistic regression indicated that the risk assessment tools made more errors in predicting violence and offending in younger youth than in older youth, it does not provide information on the types of errors made, such as whether the tools overestimate or underestimate violence in younger youth. Therefore, to investigate the nature of errors made about younger youth, we compared the false positive rate and false negative rate of younger and older youth. A false positive occurs when a youth is inaccurately judged to be a high risk of reoffending, whereas a false negative occurs when a youth is inaccurately judged to be low risk of reoffending. A youth was classified as "high risk" if his score on a tool fell more than one *SD* greater than the mean for that tool and as "low risk" if his score on a tool fell more than one *SD* less than the mean on that tool. Youth with moderate scores were not included in these analyses because it is less clear whether they should be considered high or low risk.

Although false positive rates were high for both age groups (see Table 6), the false positive rate was significantly higher for younger youth than for older youth on both the J-SOAP-II and

TABLE 7:   Age Differences in False Negative Rates

	Ages 12 to 15	Ages 16 to 19	z
Sexual offense			
J-SOAP-II	0.00	0.33	−3.91***
SAVRY	0.00	1.00	—
Any nonsexual violent offense			
J-SOAP-II	0.20	0.17	0.29
SAVRY	0.17	0.50	−2.64**
Serious nonsexual violent offense			
J-SOAP-II	0.33	0.00	3.51***
SAVRY	0.25	0.00	2.65**
Any offense			
J-SOAP-II	0.30	0.33	0.27
SAVRY	0.11	0.40	−2.51**

Note. J-SOAP-II = Juvenile Sex Offender Assessment Protocol–II; SAVRY = Structured Assessment of Violence Risk in Youth.
*p < .05. **p < .01. ***p < .001.

SAVRY. This was true for all offense categories, including predictions of any nonsexual violent offenses (defined as felonies or misdemeanors), serious nonsexual violent offenses, sexual offenses, and any offenses. In other words, younger adolescents were more likely to be inaccurately judged to be a high risk. False negative rates were lower and much less consistent (see Table 7). Although the false negative rate was significantly higher among older youth than younger youth for predictions of sexual offenses, any nonsexual violent offenses, and any offense, false negatives were more common among younger youth than older youth for predictions of serious nonsexual violent offenses. Importantly, given the low rates of false negatives in the sample, caution should be used in interpreting developmental differences in false negative rates.

DISCUSSION

Although there are many situations in which clinicians may be required to make judgments about risk for reoffending in adolescents who have committed sexual offenses, there are no well-validated measures for use with this population. As such, this study examined the predictive validity of the J-SOAP-II, J-SORRAT-II, and SAVRY in a sample of male adolescent sex offenders who were admitted to a residential treatment program.

PREDICTION OF SEXUAL AND NONSEXUAL OFFENSES

Comparable to other studies (Worling & Långström, 2003), 8.4% of our sample was rearrested for sexual offenses after following youth for an average of 6.58 years. However, a sizable proportion of youth had committed other types of offenses; our overall base rate of reoffending was 42.8%. This finding emphasizes the need to assess risk for nonsexual offending in addition to risk for sexual violence among adolescents who have committed a sexual offense.

None of the tools that we examined were able to significantly predict which youth sexually reoffended following their discharge. In addition, the instruments did not achieve much success in predicting sexual aggression during treatment. Despite the limited ability of the tools to predict sexual violence, total scores on the SAVRY and J-SOAP-II were able to predict nonsexual aggression during treatment and serious nonsexual violent offenses following discharge.

Also, total scores on the SAVRY, which is designed to predict general violence risk, predicted serious nonsexual violent offenses following discharge. The SAVRY did not predict nonsexual violent offenses when the definition was broadened to include misdemeanors as well as felonies. However, the courts are likely to be more interested in serious violence than minor violent acts. Also, while *total scores* on the SAVRY predicted serious nonsexual violence, structured professional ratings on the SAVRY did not. Given that a number of studies have demonstrated that structured professional judgments made based on the administration of a well-validated risk assessment tool show a high degree of accuracy in predicting future violence (e.g., Douglas, Ogloff, & Hart, 2003; Kropp & Hart, 2000; Worling, 2004), this finding merits further investigation. Total scores on the J-SOAP-II were significant in predicting serious nonsexual violent offenses among youths aged 16 and older, but were not significant when the full sample of adolescents were examined.

The finding that total scores on the SAVRY and J-SOAP-II predicted nonsexual violence, to some extent, but not sexual violence or general offending, suggests that a different set of risk factors may be relevant to sexual and nonsexual offending among adolescents who have committed sexual offenses. In similar fashion, other researchers have argued that—although the risk factors for sexual recidivism and general recidivism for adolescent offenders show overlap—they differ in important ways (Långström, 2002; Långström & Grann, 2000; Worling & Curwen, 2000). For example, Långström and Grann (2000) found that risk factors for general delinquency (e.g., previous criminal behavior, use of threats and weapons) were associated with general offending but not sexual reoffending in young sex offenders. However, factors indicative of both general delinquency and sexual deviance (e.g., early onset of sexually abusive behavior, having more than one victim) and poor social skills were associated with sexual reoffending.

On the other hand, others have argued that differences between individuals who commit sexual and nonsexual offenses may be minimal (Gottfredson & Hirschi, 1990; Spaccarelli, Bowden, Coatsworth, & Kim, 1997). Perhaps in light of a belief that sexual and nonsexual offending has similar risk factors, the existing instruments in this field often aim to measure both sexual and nonsexual violence. The J-SOAP-II, for instance, was designed to measure sexual *and* nonsexual offending among male adolescent sex offenders. Whether the same instrument should be used in assessing risk for sexual and nonsexual violence should be reevaluated as more evidence is accumulated.

DEVELOPMENTAL DIFFERENCES IN PREDICTIVE VALIDITY

The J-SOAP-II and the SAVRY were significantly less effective in predicting reoffending following discharge among younger adolescents. In particular, youth aged 15 and younger were more likely than older youth to be incorrectly judged as being a high risk for sexual and nonsexual violence following discharge. This finding could stem from several

possible sources. First, younger adolescents may receive higher scores on certain items on these tools (e.g., understands risk factors, lack of empathy, impulsivity) because of their developmental stage rather than stable personality features that are indicative of long-term reoffense risk (see Seagrave & Grisso, 2002). In addition, sexual violence committed by younger adolescents may seem more serious and aberrant. Therefore, raters may believe these youth are likely to continue to engage in antisocial behavior and thus tend to score them as being high risk.

The differential predictive validity of these tools for youth of various ages could also suggest that there may be developmental differences among adolescent age groups for reoffending. At this point, little is known regarding possible developmental differences in risk factors. In a recent meta-analysis, Hanson and Morton-Bourgon (2005) reported that the basic *sets* of factors that predict sexual reoffending in adult sex offenders (e.g., deviant sex interests) also predict sexual reoffending among adolescents who have committed sexual offenses. However, other empirical reviews have suggested possible developmental differences in specific risk factors for violence between youth and adults and between youth of various ages. For instance, negative peer relationships appear more strongly linked to later offending among youth aged 9 to 11 than those aged 12 and 14 (Lipsey & Derzon, 1998). Also, male gender of victim is a significant predictor of sexual reoffending among adults but not necessarily in youth (Worling & Långström, 2003). If future research finds evidence of developmental differences in risk factors among younger and older adolescents, it may support the development of different risk assessment and management approaches for adolescents of differing ages and maturity levels.

LIMITATIONS AND FUTURE RESEARCH

Given the limited ability of the J-SOAP-II, J-SORRAT-II, and SAVRY in predicting sexual reoffending in this sample, it is critical that future research continue to examine tools for assessing sexual reoffending in adolescents. Given that many sexual offenses, perhaps particularly those that are committed by youth, do not lead to formal convictions, this research should strive to measure reoffending through multiple sources of information, rather than solely convictions (Barbaree & Marshall, 1988; Doren, 1998). In this study, reoffending was measured through official and unofficial sources, including law enforcement, probation, and treatment records. Also, consistent with recommendations in this field, we examined charges rather than convictions (Doren, 1998). Nevertheless, it is likely that our estimate of reoffending is still an underestimate.

Future research on risk assessment tools should also strive to use prospective designs in large and diverse samples, in which risk assessment tools are coded based on clinical interviews as well as file information. Consistent with other studies, the present study coded risk assessment tools based on clinical file information. Although this type of research provides a good starting point by enabling youth to be followed for much longer periods of time, some variables may be difficult to assess accurately based on file information.

It may be particularly useful for future research to examine the Estimate of Risk of Adolescent Sex Offender Recidivism (ERASOR; Worling & Curwen, 2001), a structured professional judgment tool designed to assess risk of sexual violence among adolescents. Preliminary research on this tool has reported promising findings (Worling, 2004). Also, in addition to evaluating risk assessment tools, research should continue to examine specific

risk factors for sexual violence in adolescents. Although some risk assessment tools were developed by reviewing the available research on youth sexual violence, it may be that this research is not sufficiently advanced to produce tools that are highly accurate, especially as there remain a number of unclear and inconsistent findings in this field (Worling & Långström, 2003). Finally, there is a need for research on assessing risk of violence among adolescent girls who have committed sexual offenses, as we currently know very little about this population (Worling & Långström, 2003).

SUMMARY

As more and more jurisdictions require youth who are judged to be a high risk of future violence to be included on sex offender registries and transferred to adult court, there is a pressing need for valid risk assessment approaches. The present study found that, although total scores on the SAVRY and J-SOAP-II were able to predict nonsexual aggression with some degree of accuracy, the tools that were developed for predicting sexual violence (J-SOAP-II and J-SORRAT-II) did not significantly predict sexual violence in our sample. Furthermore, it was especially difficult to predict violence among adolescents aged 15 and younger compared to older adolescents. These limitations, in what are considered to be leading approaches to assessing youths' violence risk, raises questions about what information clinicians can currently offer courts regarding youths' risk for subsequent sexual violence. Furthermore, it may even raise some questions about the legal system's current focus on placing youth on public sex offender registries and transferring them to adult court based on judgments that they pose a high risk of subsequent violence. There is a critical need for further research to examine developmentally appropriate approaches for assessing violence risk among adolescents who have committed sexual offenses.

NOTE

1. The formula used was $z = (A_1 - A_2)/(SE_1^2 + SE_2^2 - 2rSE_1SE_2)^{1/2}$, where A_1 and A_2 are the AUC for the first and second risk assessment, SE_1^2 and SE_2^2 are the standard errors of these AUCs, and r is the correlation between A_1 and A_2.

REFERENCES

Barbaree, H. E., & Marshall, W. L. (1988). Deviant sexual arousal, offense history, and demographic variables as predictors of reoffense among child molesters. *Behavioral Sciences and the Law*, *6*, 267-280.

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.

Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, *34*, 105-113.

Borum, R. (2003). Managing at-risk juvenile offenders in the community: Putting evidence-based principles into practice. *Journal of Contemporary Criminal Justice*, *19*, 114-137.

Borum, R., Bartel, P., & Forth, A. (2003). *Manual for the Structured Assessment of Violence Risk in Youth*, *Version 1.1*. Tampa: University of South Florida.

Burton, D. L, Smith-Darden, J., & Frankel, S. J. (2006). Research on adolescent sexual abuser treatment programs. In H. E. Barbaree & W. L. Marshall, *The juvenile sex offender* (2nd ed.). New York: Guilford.

Caldwell, M. F. (2002). What we do not know about juvenile sexual reoffense risk. *Child Maltreatment: Journal of the American Professional Society of the Abuse of Children*, *7*, 291-302.

Catchpole, R., & Gretton, H., (2003). The predictive validity of risk assessment with violent young offenders: A 1-year examination of criminal outcome. *Criminal Justice and Behavior*, *30*, 688-708.

Cauffman, E., & Steinberg, L. (2000). (Im)maturity of judgment in adolescence: Why adolescents may be less culpable than adults. *Behavioral Sciences and the Law*, *18*, 741-760.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284-290.

Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, *86*, 127-137.

Cohen, J. (1988). *Statistical power for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Dempster, R. (2001). *Understanding errors in risk assessment: The application of differential prediction methodology*. Unpublished doctoral dissertation, Simon Fraser University, Burnaby, British Columbia, Canada.

Doren, D. (1998). Recidivism base rates, predictions of sex offender recidivism, and the "sexual predator" commitment laws. *Behavioral Sciences and the Law*, *16*, 97-114.

Douglas, K. S., Ogloff, J. R. P., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services*, *54*, 1372-1379.

Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: The HCR-20 violence risk assessment scheme and the Psychopathy Checklist: Screening version. *Journal of Consulting and Clinical Psychology*, *67*, 917-930.

Epperson, D. L., Ralston, C. A., Fowers, D., & DeWitt, J. (2005). *Development of a sexual offense recidivism risk assessment tool—II (JSORRAT-II)*. Unpublished manuscript, University of Iowa, Ames.

Fleiss, J., Williams, J. B. W., & Dubro, A. F. (1986). The logistic regression analysis of psychiatric data. *Journal of Psychiatric Research*, *20*, 145-209.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Palo Alto, CA: Stanford University Press.

Grisso, T. (1998). *Forensic evaluations of juveniles*. Sarasota, FL: Professional Resource Exchange.

Grisso, T., Steinberg, L., Woolard, J., Cauffman, E., Scott, E., Graham, S., et al. (2003). Juveniles' competence to stand trial: A comparison of adolescents' and adults' capacities as trial defendants. *Law and Human Behavior*, *27*, 333-363.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29-36.

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66*, 348-362.

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, *73*, 1154-1163.

Hanson, R. K., & Thornton, D. (1999). *Static 99: Improving actuarial risk assessments for sex offenders* (User report 1999-02). Ottawa: Department of the Solicitor General of Canada.

Harris, G. T., Rice, M. E., Quinsey, V. L., Lalumière, M. L., Boer, D., & Lang, C. (2003). A multisite comparison of actuarial risk instruments for sex offenders. *Psychological Assessment*, *15*, 413-425.

Hoge, R. D. (2002). Standardized instruments for assessing risk and need in youthful offenders. *Criminal Justice and Behavior*, *29*, 380-396.

Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, *65*, 599-610.

Kent v. United States, 383 US 541 (1966).

Kropp, P. R., & Hart, S. D. (2000). The Spousal Assault Risk Assessment (SARA) guide: Reliability and validity in adult male offenders. *Law and Human Behavior*, *24*, 101-118.

Långström, N. (2002). Long-term follow-up of criminal recidivism in young sex offenders: Temporal patterns and risk factors. *Psychology, Crime and Law*, *8*, 41-58.

Långström, N., & Grann, M. (2000). Risk for criminal recidivism among young sex offenders. *Journal of Interpersonal Violence*, *15*, 855-871.

Lipsey, M. W., & Derzon, J. H. (1998). Predictors of violent or serious delinquency in adolescence and early adulthood: A synthesis of longitudinal research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 86-105). Thousand Oaks, CA: Sage.

Luke, D. A., & Homan, S. M. (1998). Time and change: Using survival analysis in clinical assessment and treatment evaluation. *Psychological Assessment*, *10*, 360-378.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*, 30-46.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674-701.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, *62*, 783-792.

Prentky, R. A., Harris, B., Frizzell, K., & Righthand, S. (2000). An actuarial procedure for assessing risk with juvenile sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, *12*, 71-93.

Prentky, R., & Righthand, S. (2003). *Juvenile Sex Offender Assessment Protocol II (J-SOAP-II) manual*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Seagrave, D., & Grisso, T. (2002). Adolescent development and the measurement of juvenile psychopathy. *Law and Human Behavior*, *26*, 219-239.

Silver, E., Smith, W. R., & Banks, S. (2000). Constructing actuarial devices for predicting recidivism: A comparison of methods. *Criminal Justice and Behavior*, *27*, 733-764.

Spaccarelli, S., Bowden, B., Coatsworth, J. D., & Kim, S. (1997). Psychosocial correlates of male sexual aggression in a chronic delinquent sample. *Criminal Justice and Behavior*, *24*, 71-95.

Trivits, L. C., & Reppucci, N. D. (2002). Application of Megan's law to juveniles. *American Psychologist*, *57*, 690-704.

Worling, J. R. (2004). The Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR): Preliminary psychometric data. *Sexual Abuse: A Journal of Research and Treatment*, *16*, 235-254.

Worling, J. R., & Curwen, T. (2000). Adolescent sexual offender recidivism: Success of specialized treatment and implications for risk prediction. *Child Abuse and Neglect*, *24*, 965-982.

Worling, J. R., & Långström, N. (2003). Assessment of criminal recidivism risk with adolescents who have offended sexually: A review. *Trauma, Violence, and Abuse*, *4*, 341-362.

Zimring, F. E. (2004). *An American travesty: Legal responses to adolescent sexual offending*. Chicago: University of Chicago Press.

━━━━━━━━━━━━━━━━━━━━━━━

C|E *Criminal Justice and Behavior (CJB)* is pleased to offer the opportunity to earn continuing education credits online for this article. Please follow the access instructions below. The cost is $9 per unit, and a credit card number is required for payment.

Access for AACFP members:

- Go to the AACFP homepage (www.aa4cfp.org) and log in to the site using your email address
- Click on the Publications heading in the sidebar on the left-hand side of the page or on the Publications icon at the bottom of the page
- Click on the *CJB* cover image located at the bottom of the page. You will be prompted for your user name and password to log in to the journal
- Once you have logged in to *CJB*, select the appropriate issue and article for which you wish to take the test, then click on the link in the article to be directed to the test (*note: only marked articles are eligible for continuing education credits*)

Access for institutional subscribers:

- Log in through your library or university portal to access *CJB*
- Select the appropriate issue and article, then click on the link in the article to be directed to the test (*note: only marked articles are eligible for continuing education credits*)

If you are not a member of the AACFP and your institution does not subscribe to *CJB*, individual articles (and corresponding test links within the articles) are available on a pay-per-view basis by visiting http://cjb.sagepub.com. However, membership in the AACFP provides you with unlimited access to *CJB* and many additional membership benefits! Learn more about becoming a member at www.aa4cfp.org.

Should you have difficulty accessing this article's test or the website, please contact Eddie Santos for further assistance at eddie.santos@sagepub.com or 805-410-7528.