

# The Impact of Using Documented but Uncharged Offense Data on JSORRAT-II Predictive Validity

Sexual Abuse: A Journal of  
Research and Treatment  
1–17

© The Author(s) 2015

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/1079063215582011

sax.sagepub.com



**Christopher A. Ralston<sup>1</sup>, Amar Sarkar<sup>1,2</sup>,  
Grace T. Philipp<sup>1,3</sup>, and Douglas L. Epperson<sup>4</sup>**

## Abstract

Following the implementation of sexual offender notification laws, researchers have found a drop in the rate of prosecutions and an increase in plea bargains for sexual offenses committed by male juveniles. This type of prosecutorial hesitation has implications for the predictive validity of sexual recidivism risk assessments, such as the Juvenile Sexual Offender Recidivism Risk Assessment Tool-II (JSORRAT-II), that require data from officially adjudicated offenses in the scoring of several items. The present study sought to test the impact of including data from documented but uncharged (DBU) sexual offenses in the scoring of the JSORRAT-II on its predictive validity using an exhaustive sample of 1,095 juveniles who offended sexually from the states of Iowa and Utah. Although sexual recidivists had significantly more DBU data, the inclusion of those data did not improve the predictive validity of the tool. The authors discuss additional reasons why changes in prosecutorial practice might remain confound in risk assessment studies and suggest future research to investigate those hypotheses.

## Keywords

risk assessment, sexual offending, juvenile, JSORRAT-II, documented but uncharged sexual offenses

---

<sup>1</sup>Grinnell College, IA, USA

<sup>2</sup>Oxford University, UK

<sup>3</sup>Columbia University, New York, NY, USA

<sup>4</sup>California Polytechnic State University, San Luis Obispo, CA, USA

## Corresponding Author:

Christopher A. Ralston, Department of Psychology, Grinnell College, 1116 8th Ave., Grinnell, IA 50112-6190, USA.

Email: ralstonc@grinnell.edu

## **Impact of Scoring the Juvenile Sexual Offense Recidivism Risk Assessment Tool-II (JSORRAT-II) Using Documented but Uncharged (DBU) Data**

Highly publicized and especially brutal crimes by individuals previously convicted of sexual offenses, such as the murder of Megan Kanka, have contributed to a dramatic increase in public awareness of sexual offenses in the last two decades. Indeed, crimes against children seem to possess a unique ability to shock, infuriate, and call governments to action, and have led to equally publicized policies designed to restrain the activities of known sexual offenders to reduce the likelihood of recidivism. These include the Jacob Wetterling Crimes Against Children and Sexually Violent Offender Registration Act of 1994 and the Adam Walsh Child Protection and Safety Act of 2006. These measures seek to prevent recidivism and assist police efforts to investigate and arrest perpetrators of such abuse, primarily through placement on registries and community notification. Although well-meaning, the laws may have inadvertently reduced the likelihood of prosecution for a particular class of offenders: juveniles who sexually offended (JSOs; for example, Letourneau, Bandyopadhyay, Sinha, & Armstrong, 2009). Consequently, the validity indices for sexual recidivism risk assessment tools that were derived from JSO samples after the implementation of these laws might be confounded. The purpose of the present study is to investigate the hypothesis that changes in juvenile justice prosecutorial and legal practices affect the predictive validity of one such measure, the JSORRAT-II (Epperson & Ralston, in press-b; Epperson, Ralston, Fowers, DeWitt, & Gore, 2006).

Many have questioned the utility of applying these laws and policies derived from the Wetterling and Walsh Acts to JSOs (e.g., Batastini, Hunt, Present-Koller, & DeMatteo, 2011; Harris, Lobanov-Rostovsky, & Levenson, 2010; Letourneau & Miner, 2005; Trivits & Reppucci, 2002), and others have demonstrated their ineffectiveness in deterring new sexual offending (e.g., Letourneau, Bandyopadhyay, Armstrong, & Sinha, 2010; Letourneau, Levenson, Bandyopadhyay, Armstrong, & Sinha, 2010) or preventing sexual recidivism (e.g., Tewksbury & Jennings, 2010; Tewksbury, Jennings, & Zgoba, 2012). Indeed, the tiered system used to guide registration and community notification duration and intensity does not seem to align with the empirical literature on known risk factors (Zgoba et al., 2012). Perhaps as a consequence of these observations and because such laws are associated with loss of relationships, threats or harassment, feelings of hopelessness, and loss of jobs and homes, at least for adults who offend sexually (Levenson & Cotter, 2005; Levenson, D'Amora, & Hern, 2007; Tewksbury, 2005), some legal practitioners also question the application of these laws to juveniles. For example, the rate of prosecuting JSOs dropped in conjunction with the implementation of registration and notification laws in South Carolina (Letourneau et al., 2009). More specifically, the odds of prosecuting a sexual offense at the felony level fell by approximately 40% in comparison with juveniles charged with nonsexual offenses after the enactment of lifetime registration laws. Letourneau and colleagues (2009) argued that prosecutors seemed to take into consideration the negative consequences of such laws on JSOs in making prosecutorial

decisions. In addition, Letourneau and colleagues (2013), using the same sample, found that the probability of plea bargaining to nonsexual or less severe charges increased across three time periods (1990-1994, 1995-1998, and 1999-2004), corresponding to the passage of registration and notification policies in that state. This pattern was not observed for nonsexual offenses, again signaling hesitation by prosecutors to seek charges that simultaneously capture the true nature of the crime and lead to collateral consequences through the application of registration and notification policies.

It is unclear to what extent prosecutorial practices with JSOs have changed in other states in response to these laws; however, given the possibility and consequences of long-term registration and community notification requirements, it is reasonable to assume that South Carolina prosecutors are not alone in their protective attitude toward JSOs. Although such practices result in fewer negative consequences for the JSOs themselves, this lenience may pose a valid risk to the community. While being spared undue punishments, JSOs may also be eluding much-needed judicial and treatment intervention because of well-intentioned interference by legal agents (Letourneau, Armstrong, Bandyopadhyay, & Sinha, 2013; Letourneau et al., 2009). Ultimately, overly harsh laws defeat the purpose of establishing community harmony, both in the consequences they frame for JSOs and also by the overcompensating actions taken to protect the JSOs from those consequences.

Although the Adam Walsh Act does not provide for tier placement considerations outside of the type of crime committed, the potential negative effects to some JSOs could be mitigated with targeted risk assessment, thereby limiting the application of such laws only to those JSOs most at risk to sexually recidivate. A number of assessments have been developed for adults who offended sexually, including the Static-99 (Harris, Phenix, Hanson, & Thornton, 2003), the Minnesota Sexual Offender Screening Tool-Revised (MnSOST-R; Epperson et al., 2004), and the Sexual Offender Risk Appraisal Guide (SORAG; Quinsey, Harris, Rice, & Cormier, 1998). Empirically derived risk assessment tools for adults have been found to aid decision making regarding level of registration and community notification, placement, and other release-related questions with average effect sizes ranging from moderate to large (Hanson & Morton-Bourgon, 2009). Given the effect sizes, it seems reasonable that the risk level assessed by these tools is a better predictor of the threat posed to communities than tier placement.

There are a number of empirically guided risk assessment tools designed specifically for JSOs. These include the Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR; Worling & Curwen, 2001), the Juvenile Risk Assessment Scale (JRAS; New Jersey Attorney General's Office, 2006), and the Juvenile-Sex Offender Assessment Protocol-II (J-SOAP-II; Prentky & Righthand, 2003), among others. Although the predictive validity of some adult tools is well established, there has been much more limited research exploring the predictive accuracy of juvenile-specific tools.

The JSORRAT-II (Epperson & Ralston, in press-b; Epperson et al., 2006) is the first fully empirically derived sexual offense recidivism risk assessment tool designed for

male JSOs. It consists of 12 historical items that can be reliably scored on the basis of juvenile justice case file review by coders with a 1-day training course. Developed on an exhaustive development sample of 636 juveniles from Utah adjudicated guilty for a sexual offense between 1990 and 1992, the JSORRAT-II successfully predicted sexual recidivism before age 18 at well above chance levels, with the area under the receiver operating characteristic curve (AUC) = .89 (95% confidence interval [CI] = [.85, .92]; Epperson & Ralston, *in press-b*; Epperson et al., 2006). The original authors also successfully cross-validated the tool using exhaustive samples from Utah (AUC = .64, 95% CI = [.58, .71]; Epperson & Ralston, *in press-b*; Epperson et al., 2006) and Iowa (AUC = .70, 99% CI = [.60, .81]; Ralston, Epperson, & Edwards, *in press*). However, a cross-validation attempt by independent researchers found that the JSORRAT-II did not significantly predict sexual misconduct during treatment or sexual recidivism posttreatment for a sample of JSOs admitted to a nonresidential treatment program (Viljoen et al., 2008). Finally, Viljoen, Mordell, and Beneteau (2012) found the average AUC across seven published and unpublished studies, excluding the JSORRAT-II development study, to be .61.

It should be noted that the JSORRAT-II was developed on a sample whose index sexual offense occurred before the passage of the Wetterling Act (1994), whereas the two cross-validation studies conducted by the original authors of the JSORRAT-II used samples whose offenses occurred after its implementation (Utah: 1996-1997, Iowa: 2000-2006). Both the findings by Letourneau and colleagues (Letourneau et al., 2013; Letourneau et al., 2009) discussed above and anecdotal evidence from research and treatment providers across several states suggest that some states currently prefer to handle many sexually deviant behaviors nonjudicially to minimize the negative outcomes associated with the application of adult statutes and policies to JSOs, instead pushing for repeated or violent offending to result in formal charges.

If the passage of such laws did reduce the likelihood a sexually deviant act was prosecuted, the validity estimates of tools such as the JSORRAT-II that utilize information from charged sexual offenses to inform risk levels would be compromised with samples taken after the passage of these laws. The nonjudicial handling of sexually deviant behavior would affect cross-validation studies in two ways. First, many nonjudicially handled JSOs would not receive a judicial charge or adjudication as a consequence of JSOs' generally low risk of reoffense (e.g., Caldwell, 2002). If subsequent offending events are more likely to lead to charges and adjudications, initial detection and nonjudicial handling is likely a sufficient deterrent to most future offending. Because a formal charge for a sexual offense is often the characteristic used to decide whom to include in cross-validation samples (as in the three Epperson and Ralston JSORRAT-II studies), this would effectively remove these presumably low-risk offenders from consideration. Consequently, predictive validity estimates, particularly the AUC values, might be suppressed because of the relatively fewer true negatives in more recent samples compared with older, pre-Wetterling Act samples. Second, nonjudicial handling of prior sexual offenses results in the loss of information relevant for scoring 6 of the 12 JSORRAT-II items because the scoring methodology for those items precludes the use of offending behavior that did not reach a formal charge. The

intent behind this decision was to reduce the likelihood of less valid statements (e.g., uncorroborated self-reports, accusations from others) from artificially inflating the risk level designated by the test. However, protective actions by the different legal agents may have prevented consideration of offense data that would otherwise be admissible in the scoring of the JSORRAT-II at earlier times, potentially leading to a reduction in the predictive accuracy upon cross-validation.

If changes in prosecutorial practices for JSOs occurred between the time data were collected from the JSORRAT-II development sample and the collection of data from the two cross-validation samples by the JSORRAT-II authors, it is possible that it became more difficult for a sexually deviant act to rise to the level of a formal charge—effectively changing the nature of the sample and reducing the data available to score the JSORRAT-II. If, in turn, the uncharged sexually deviant behavior was substantiated through documentation by other official sources, the inclusion of such data in the scoring of the JSORRAT-II has the potential to augment its predictive accuracy in those later samples. The present study investigated the effects of including information from DBU sexual offense data in the scoring of the JSORRAT-II on its predictive accuracy.

## Hypotheses

**Hypothesis 1:** Proportionally more sexual recidivists will have DBU offenses than nonrecidivists.

**Hypothesis 2:** The inclusion of DBU data will result in a significantly greater increase in JSORRAT-II total scores for recidivists when compared with nonrecidivists.

**Hypothesis 3:** The inclusion of DBU data will result in a significant increase in the predictive validity of the JSORRAT-II.

## Method

### Sample

The present study utilized the juvenile justice case files from 1,095 male JSOs aged 11 to 18 years who were adjudicated guilty for a sexual offense. The case files represented an exhaustive sample of male JSOs from the states of Utah and Iowa, whose index offense occurred in 1996 or 1997 ( $n = 566$ ) and in 2000 through 2006 ( $n = 529$ ), respectively. The mean age for the combined sample was 15.1 years ( $SD = 1.5$ ), with a mean age of 15.2 years ( $SD = 1.6$ ) for the Utah sample and 15.0 years ( $SD = 1.4$ ) for the Iowa sample. All JSOs in both samples were followed until age 18.

Consistent with Utah's demographic makeup, the majority of juveniles in that sample were Caucasian/White (76.4%). The remainder of the Utah sample were Hispanic/Latino (7.7%), African American/Black (2.2%), Asian American (1.6%), Native American (1.4%), multiethnic or Other (1.1%), and unspecified (9.6%). The majority of the Iowa sample also was identified as Caucasian/White (79.6%), and the remainder

of the sample was identified as African American/Black (8.7%), Hispanic/Latino (4.5%), Native American (1.1%), Asian American (0.6%), multiethnic or Other (3.1%), and unspecified (2.5%). These two samples, combined for this study, were used to assess the predictive validity of the JSORRAT-II in previous studies (Epperson & Ralston, in press-b; Ralston, Epperson, & Edwards, in press).

### *Materials and Procedures*

*Juvenile judicial and corrections case files.* Juvenile justice case files for all JSOs in the study were located and copied by the staff of the Iowa and Utah Divisions of Juvenile Justice Services. All files were transported to the researchers where they were prepared for scoring. Each case file contained a record of the JSO's criminal involvement in the juvenile justice system up to and including their index sexual offense. Records of criminal involvement typically included arrest, investigation, court, and jurisdictional review reports. From these records, information could be extracted about past and current sexual offenses, including information about events leading up to the sexual offense, the nature of the offense, and information about victims. Additional information found in most files included probation or caseworker reports, psychological evaluations, education reports, and Department of Human Services reports. These additional files provided information about educational history, social functioning, substance use and abuse, mental health issues, treatment history, caregiving stability, and histories of abuse and neglect.

Because the Utah sample was not subjected to a prospective study, all case files were arranged chronologically by two undergraduate research assistants. After chronological arrangement, one of the authors removed all information after one of two time points. First, if the JSO did not recidivate sexually after their 1996 to 1997 index sexual offense, all information found in the case file dated January 1, 2000, or later was removed. This ensured that sufficient information was available to score the JSORRAT-II (i.e., 3-4 years of postadjudication file information), while limiting the size of the file to save coder time. Second, if the JSO was identified as having a recidivating offense, all information was removed from the first mention of that offense onward. If the recidivating offense occurred in 2000 or later, all information dated January 1, 2000, or later was removed. The second step was taken to ensure that the coders were blind to each JSO's recidivism status. No file information was removed from the Iowa case files prior to data extraction and entry into an electronic database. However, any information relating to recidivating offenses was removed from the electronic database prior to the calculation of JSORRAT-II scores.

*Sexual recidivism data.* The sexual recidivism follow-up period ranged from the date of the JSO's index adjudication to the date of his 18th birthday for both Iowa and Utah samples. Because the mean age for JSOs in this sample was 15.1 ( $SD = 1.5$ ) at the time of that adjudication, individuals were followed for 2.9 years on average ( $SD = 1.5$ , range from 14 days to 6.9 years). Unfortunately, we were unable to account for true time at risk (e.g., outside a secure facility). The Iowa and Utah Juvenile Justice

Services conducted electronic searches of the statewide juvenile court/juvenile justice services database to generate a list of charges, adjudications, offense dates, charge dates, and adjudication dates for each JSO. Juvenile sexual recidivists, defined as those JSOs with a formal charge for a new sexual offense prior to age 18, were identified from this list.

*JSORRAT-II.* The JSORRAT-II is a 12-item risk assessment tool for male JSOs, aged 12 to 18 years. The first 6 items reflect sexual offending behavior (e.g., number of victims, offending in a public place), and the scoring of those items requires that the offending behavior result in a formal charge. The remaining items reflect nonsexual offending characteristics or behavior (e.g., frequency of personal sexual abuse victimization, performance in sexual offender-specific treatment).

*Data extraction.* Data were extracted from the case files to a coding form that included the JSORRAT-II variables and several research variables found to be significantly associated with sexual recidivism in the JSORRAT-II development study. These data were used to score the JSORRAT-II. The present study sought to assess the impact of including DBU data on the predictive validity of that tool. DBU data were defined as any sexual offense that was not investigated by the juvenile justice system but was clearly documented by another state agency (e.g., child protective services) or was reported by a credible source (e.g., JSO's legal guardian). Scoring included self-report data only when the existence of such offenses was corroborated by other documentation (e.g., child protective services, report of parent). DBU data were collected along with the data used to score the JSORRAT-II items.

Eleven undergraduate research assistants were trained over the course of several didactic meetings on the procedures for extracting information from the case files. During the training meetings assistants were introduced to and trained on how to use the coding form. All research assistants then extracted data from the same set of practice cases. These practice cases were actual cases from the sample. After completing these, all coders met with the lead researcher to discuss the cases, any discrepancies in scoring, and any other questions pertaining to scoring the cases. This process was repeated until all coders completed the coding forms for several cases in a consistent fashion.

In addition to this primary training, all coders met with one of the authors approximately once every 2 weeks throughout the course of the coding to help prevent coder drift. During these sessions, the coders and the researcher reviewed discrepancies in scoring reliability cases, key scoring issues, and any additional questions that they had pertaining to the coding information from the cases.

Furthermore, the coders were instructed to record all scoring questions in a research coding log. The researcher reviewed this log on a near daily basis and responded to these questions. This log was placed near the coding forms for all coders to review prior to each coding session so that all coders would have any new scoring information. These questions were also used to direct discussion during the secondary training meetings.

*Data entry.* Each coding form was double entered into an SPSS database to assess and correct for data entry error. Once all forms had been entered, the researcher analyzed the entries for inconsistencies. Upon finding inconsistencies, the original coding form was consulted for the appropriate entry.

*Data analysis.* Research assistants coded the same 16 cases from the Utah sample and same 50 cases from Iowa over the course of the research project. To assess interrater reliability, the singular interclass correlation (ICC) for absolute agreement, using a two-way mixed effect model, was calculated for the JSORRAT-II total score. The singular ICC for absolute agreement counts baseline differences between raters as error, so it is a very conservative measure and appropriate as an index of reliability for a risk assessment tool where one is interested in absolute and not just relative agreement. Because of the way it is calculated, this index can also be viewed as a coefficient of generalizability, reflecting the proportion of total variance that is due to true differences between the cases. Coefficient alpha was also calculated, which in this context reflects relative consistency across raters, much like its traditional use reflects inter-correlation of test items. It also reflects the increase in reliability that would result if each risk assessment reflected the average score from all coders rather than just the score from one coder.

The three hypotheses were tested using a combination of analyses. First, chi-square analysis was used to test the hypothesis that proportionally more sexual recidivists would have DBU offenses than nonrecidivists. Second, ANOVA was used to test the hypothesis that the inclusion of DBU data would result in a significantly greater increase in JSORRAT-II total scores for recidivists when compared with nonrecidivists. For that analysis, the number of additional points gained through the inclusion of DBU data was used as the dependent variable, while recidivism status was specified as a fixed factor. Finally, change in the predictive validity of the JSORRAT-II with the inclusion of DBU offense data was tested in two ways. First, a hierarchical Cox proportional hazard regression analysis was used to determine whether DBU status and score increase after the inclusion of DBU offense data predicted juvenile sexual recidivism status above and beyond non-DBU JSORRAT-II scores. Second, AUC values were calculated for the prediction of sexual recidivism using JSORRAT-II scores with and without the inclusion of DBU offense data, and Hanley and McNeil's (1983) critical ratio  $z$  was used to test the differences observed in those two AUC values. We used an alpha level of .05 for all statistical tests.

## Results

### *Reliability Analyses*

The singular ICC for absolute agreement for JSORRAT-II total scores in Utah sample was .96 (95% CI = [.92, .98]) and coefficient alpha was .99. The ICC for the Iowa sample was .97 (95% CI = [.94, .98]) with a coefficient alpha of .97. Individual item reliability coefficients are reported in Epperson and Ralston (in press-b) and Ralston



**Table 1.** Sample Characteristics.

	Utah	Iowa	Combined	Development
Total <i>n</i>	566	529	1,095	636
Recidivating <i>n</i> (%)	72 (12.7%)	34 (6.4%)	106 (9.7%)	84 (13.2%)
DBU <i>n</i> (%)	129 (22.8%)	154 (29.1%)	283 (25.8%)	
DBU recidivist <i>n</i>	21	15	36	
(% total/% recidivists)	(3.7%/29.2%)	(2.8%/44.1%)	(3.3%/34.0%)	
DBU nonrecidivist <i>n</i>	108	139	247	
(% total/% nonrecidivist)	(19.1%/21.9%)	(26.3%/28.1%)	(22.6%/25.0%)	

Note. DBU = documented but uncharged offense data.

et al. (in press). Very high reliability was expected because each research assistant received extensive didactic and experiential training at the beginning of the project, and he or she received additional corrective feedback during the course of the project.

### Description of Sample

Sample characteristics, juvenile sexual recidivism rates, and the percentage of samples that had DBU data are presented in Table 1. Out of the 1,095 JSOs, a total of 106 (9.7%) had a sexually recidivating offense before age 18. In contrast, the JSORRAT-II development sample had a juvenile sexual recidivism rate of 13.2% (Epperson & Ralston, in press-b). The difference in rates between the present cross-validation samples and development samples was statistically significant,  $\chi^2(1) = 5.12, p < .05$ . Individually, a total of 34 out of 529 (6.4%) JSOs in Iowa were identified as having a new, recidivating sexual offense as a juvenile. The proportion of JSOs who recidivated in the Iowa sample was significantly different from the rate observed in the JSORRAT-II development sample,  $\chi^2(1) = 14.59, p < .05$ . However, in Utah, the rate of recidivism (72 out of 566, 12.7%) was not significantly different from the rate observed in the development sample,  $\chi^2(1) = 0.06, p > .05$ . Finally, the rates of recidivism between the Iowa sample and the sample from Utah used in the present study were significantly different,  $\chi^2(1) = 12.39, p < .05$ .

### Tests of Hypotheses

Overall, 25.8% of the combined sample had one or more DBU offenses. However, the rate of JSOs with DBU data in Utah (22.8%) was significantly different from the rate observed in Iowa (29.1%),  $\chi^2(1) = 5.70, p < .05$ . For the combined sample, 34.0% of recidivists had DBU data, whereas 25.0% of nonrecidivists had DBU data. Those rates were significantly different,  $\chi^2(1) = 4.03, p < .05$ . Furthermore, while the rates of DBU data for recidivists significantly differed from rates for nonrecidivists in Iowa,  $\chi^2(1) = 3.96, p < .05$ , those rates were not significantly different in Utah,  $\chi^2(1) = 1.91, p > .05$ .

**Table 2.** Means and Standard Deviations for JSORRAT-II Total Scores With and Without DBU Data by State.

	Utah	Iowa	Combined
	M (SD)	M (SD)	M (SD)
JSORRAT-II without DBU			
Recidivists	5.06 (3.64)	4.76 (2.24)	4.96 (3.25)
Nonrecidivists	3.34 (3.14)	3.19 (2.24)	3.27 (2.73)
Total	3.56 (3.26)	3.29 (2.27)	3.43 (2.83)
JSORRAT-II with DBU			
Recidivists	6.35 (4.42)	6.53 (3.40)	6.43 (4.11)
Nonrecidivists	4.31 (3.92)	4.43 (3.16)	4.37 (3.56)
Total	4.57 (4.04)	4.57 (3.22)	4.57 (3.67)
Score increase with full sample			
Recidivists	1.33 (2.19)	1.76 (2.13)	1.47 (2.17)
Nonrecidivists	0.97 (2.06)	1.24 (2.10)	1.11 (2.08)
Total	1.02 (2.08)	1.28 (2.10)	1.14 (2.09)
Score increase only with DBU			
Recidivists	4.24 (2.02)	3.93 (1.28)	4.11 (1.74)
Nonrecidivists	4.35 (2.16)	4.19 (1.80)	4.26 (1.97)
Total	4.33 (2.13)	4.16 (1.76)	4.24 (1.94)

Note. JSORRAT-II = Juvenile Sexual Offender Recidivism Risk Assessment Tool-II; DBU = documented but uncharged offense data.

This pattern of results lends partial support to the first hypothesis that proportionally more sexual recidivists would have DBU offenses than nonrecidivists; however, because of the different results in Iowa and Utah, the remaining statistical tests investigated the possibility that jurisdiction would interact with other variables.

Means and standard deviations for JSORRAT-II total scores with and without DBU data are presented in Table 2. A factorial ANOVA was used to test the hypothesis that the inclusion of DBU offense data would result in a significantly greater increase in JSORRAT-II total scores for recidivists when compared with nonrecidivists. For that analysis, the increase in total score after the inclusion of DBU offense data was set as the dependent variable with recidivism status set as a fixed factor. Because the proportion of recidivists and nonrecidivists with DBU data was not consistent across Iowa and Utah, jurisdiction was also included as a fixed factor to test for an interaction between recidivism status and jurisdiction. The main effect for jurisdiction,  $F(1, 1091) = 2.39, p > .05$ , and the interaction between recidivism status and state jurisdiction,  $F(1, 1091) = 0.13, p > .05$ , were not significant. However, the main effect for recidivism status was marginally significant,  $F(1, 1091) = 3.75, p = .053$ . The effect size associated with that difference was very small ( $\eta_p^2 = 0.003$ ), however. In other words, approximately 0.3% of the variance observed in score increases was due to the recidivism status of the JSOs, indicating minimal support for the second hypothesis.

We tested the third hypothesis, that the inclusion of DBU data would result in a significant increase in the predictive validity of the JSORRAT-II, in two ways. First, a hierarchical Cox proportional hazard regression analysis was used to predict recidivism status from JSORRAT-II total scores in the first block. The second block included DBU status, increases in scores due to DBU data, and state jurisdiction. If the hypothesis was correct, the DBU status and score increase variables should add significantly to the prediction of recidivism status above and beyond JSORRAT-II total scores. Furthermore, because of the difference in recidivism rates observed in Iowa and Utah, jurisdiction should also add significantly to the prediction of recidivism status. Finally, in the third block, the two-way interactions among the variables specified in the second block were investigated. If the hypothesis was correct, the value of DBU status and DBU score increases should not vary by state, and thus, none of the interactions should be significant.

JSORRAT-II scores significantly predicted juvenile sexual recidivism status—Model:  $\chi^2(1) = 34.73, p < .05$ ; JSORRAT-II scores:  $\beta = .16$ , Wald  $\chi^2(1) = 33.46, p < .05$ , odds ratio [OR] = 1.17, 95% CI = [1.11, 1.23]. The second block that included DBU status, DBU score increases, and state jurisdiction also added significantly to the prediction beyond JSORRAT-II scores,  $\chi^2(3) = 15.46, p < .05$ . However, within that block, only the jurisdiction variable significantly added to the prediction,  $\beta = .78$ , Wald  $\chi^2(1) = 13.73, p < .05$ , OR = 2.19, 95% CI = [1.45, 3.31]. Neither DBU status, Wald  $\chi^2(1) = 0.38, p > .05$ , nor increase in scores due to DBU data, Wald  $\chi^2(1) = 0.86, p > .05$  were significant predictors of recidivism status in the second block. Finally, the third block, consisting of the two-way interactions among the variables in the second block, did not significantly contribute to the prediction of recidivism status,  $\chi^2(3) = 5.17, p > 0.05$ .

The pattern of findings from the hierarchical Cox proportional hazard regression analysis did not confirm the third hypothesis. However, to continue the investigation of that hypothesis, AUC values were calculated for the prediction of juvenile sexual recidivism status by DBU-inclusive and DBU-exclusive JSORRAT-II total scores. AUCs were calculated for the combined sample and for both Iowa and Utah samples. Hanley and McNeil's (1983) critical ratio  $z$  was used to test the differences observed in AUC values. As a point of reference, the AUC for the JSORRAT-II development sample was .89 (95% CI = [0.85 to 0.92]; Epperson & Ralston, in press-b). The AUCs for Utah, Iowa, and the combined sample are presented in Table 3.

The JSORRAT-II significantly predicted juvenile sexual recidivism in the Iowa (AUC = .70, 95% CI = [.62, .78]), Utah (AUC = .65, 95% CI = [.59, .72]), and combined samples (AUC = .67, 95% CI = [.61, .72]). However, the original scoring and the DBU-inclusive scoring yielded very similar AUC values (see Table 3). In fact, the DBU-inclusive scoring did not represent a significant improvement in AUC values over the original scoring JSORRAT-II in each of the three cases when tested using Hanley and McNeil's critical ratio  $z$  test. Thus, it appears that the inclusion of DBU data did not improve the accuracy of the JSORRAT-II in predicting juvenile sexual recidivism in either sample alone or in combination.

**Table 3.** Area Under the Receiver Operating Characteristic Curve and Hanley and McNeil Critical Ratio *z* Statistics by Sample.

	Original (95% CI)	DBU-inclusive (95% CI)	<i>z</i>
Iowa	.70 [.62, .78]	.68 [.59, .77]	0.78, <i>p</i> > .05
Utah	.65 [.59, .72]	.65 [.58, .71]	0.39, <i>p</i> > .05
Combined	.67 [.61, .72]	.65 [.60, .70]	0.92, <i>p</i> > .05

Note. DBU = documented but uncharged offense data; CI = confidence interval.

## Discussion

Accurate risk assessment contributes to community safety by allowing for a more targeted allocation of limited resources (e.g., intensive supervision, treatment) to those most likely to reoffend (Epperson & Ralston, in press-a). Also, by targeting higher risk JSOs for more intensive handling, it is likely to reduce contagion effects to low-risk JSOs by limiting exposure to high-risk JSOs in rehabilitation programs (e.g., Boxer, Guerra, Huesmann, & Morales, 2005). It further benefits the lowest risk JSOs by limiting the application of adult statutes that may result in further negative consequences that may otherwise reduce the likelihood of successful, prosocial reintegration (Levenson et al., 2007). The JSORRAT-II is capable of assisting such decisions.

The present study sought to determine whether changing juvenile justice prosecutorial practices have an impact on the predictive accuracy of the JSORRAT-II. Letourneau and colleagues (Letourneau et al., 2013; Letourneau et al., 2009) and anecdotal evidence from research and treatment providers across several states suggest that some states prefer to handle many sexually deviant behaviors nonjudicially for the purposes of minimizing the negative outcomes associated with the application of adult statutes and policies to JSOs. Instead, formal charges might only be brought forward in cases of repeated or violent sexual offending. If this is the case, newer, post-Wetterling Act cross-validation samples might be different in two ways. First, some JSOs who would have been included in such samples prior to the application of registration and notification statutes and policies to JSOs are now less likely to be included by virtue of their only sexual offense being handled nonjudicially. In effect, this has the potential to deflate validity estimates because those excluded JSOs are likely to have lower scores on risk assessment tools such as the JSORRAT-II and not recidivate sexually. Essentially, their exclusion reduces the number of true negative predictions, on which some standard metrics of predictive accuracy are based (e.g., AUC). Unfortunately, this possibility was untestable in the present study. Second, data from DBU offenses cannot be used to score the first six items of the JSORRAT-II, as the criteria for those items require an official charge. If charging practices have changed, then DBU offenses might have risen to the attention of the court in the past and been counted when scoring the tool with older samples. The effect of such practices might be to artificially deflate scores for those most likely to reoffend, given the tendency to bring formal charges only to those who continued offending after informal and nonjudicial sanction or who offended in a violent manner.

The present study investigated this second possibility by rescoring the first six items using DBU data—data that at other times or in other jurisdictions might have been used in scoring the JSORRAT-II. However, the results do not directly support the hypothesis. In the combined sample, recidivists had proportionally more DBU data than nonrecidivists. Yet, the findings from the factorial ANOVA, hierarchical Cox proportional hazards regression, and the analyses using the AUC statistics suggest that such inclusion does not significantly improve the predictive accuracy of the JSORRAT-II in either state. Essentially, the inclusion of DBU data did not provide a predictive advantage over the standard scoring that uses information from charged sexual offenses only.

### *Limitations and Strengths*

This finding suggests that the inclusion of DBU data is not an effective strategy to improve the predictive accuracy of the JSORRAT-II over the standard scoring procedures. However, the hypothesis that changing prosecutorial practices might impact predictive validity estimates in more recent cross-validation samples cannot be ruled out completely, given the study design. As stated above, changes in informal charging practices might have restricted the sample to those who were repeat offenders or those whose offenses were judged by prosecutors to be sufficiently egregious to warrant court attention. If this is the case, the proportion of truly low-risk JSOs may be lower in more modern samples compared with those taken before the passage of state and federal risk management policies. Second, the present study had no way to account for the possibility that those JSOs with DBU offenses received additional supervision, treatment, or resource allocation than those without a prior DBU offense. If this is the case, JSOs with DBU data might be less likely to reoffend due to limited opportunities created by effective risk management strategies (e.g., intensive supervision, secure facility placement) or because risk was reduced through treatment.

Despite these limitations, there are a few noteworthy strengths of the present study. First, the researchers utilized two large, exhaustive samples of male JSOs adjudicated for a sexual offense in Iowa and Utah at two different points in time. The advantage of using such samples is that they are representative of the full spectrum of male JSOs, unlike smaller samples of convenience reported elsewhere in the literature. Such samples (e.g., JSOs in specific treatment programs or placement settings) are likely to be more homogeneous in terms of risk than the population of JSOs. This homogeneity could potentially lead to a restricted range problem for indices of validity and could jeopardize generalizability beyond a narrower range of JSOs.

At the same time, some concerns about generalizability still remain. The study utilized JSOs exclusively from Iowa and Utah. The two samples were predominantly White. Furthermore, the Utah sample likely includes a larger proportion of JSOs subscribing to the Mormon/Latter Day Saints religion than is found in other states. Although we attempted to code for this possibility, in the vast majority of instances, religious affiliation was not listed for JSOs. Finally, the present samples utilized only male JSOs, and thus, the results cannot generalize to female JSOs.

The methodology utilized is also a strength of the present study. Extensive measures were taken to keep research assistants blind to recidivism status of the JSOs to emulate a prospective study. In addition, every research assistant received extensive training that resulted in a very high level of reliability in case scoring.

### *Implications and Future Directions*

These limitations and strengths warrant several future directions. First, a direct test of the hypotheses would require more general samples of juveniles who engaged in sexual misconduct, whether or not that misconduct resulted in a criminal charge or adjudication. Although more difficult to obtain than samples relying on official court documentation, the inclusion of juveniles whose misconduct never reaches court attention would be an important step toward determining the impact of DBU data on the predictive validity of risk assessment tools.

Second, the predictive validity of the JSORRAT-II has been assessed in limited geographic regions and mostly by the original authors. Although it has performed above chance levels in the majority of cross-validation attempts with a mean AUC of .64 (95% CI = [.54, .74]; Viljoen et al., 2012), studies should be conducted in other states that have different geographic locations, racial and ethnic compositions, and dominant religious affiliations. Similarly, researchers other than the original authors should further explore the predictive validity of the JSORRAT-II with other samples.

Third, the present study did not account for the impact of other forms of time at risk and risk management strategies. It is possible that since the implementation of various laws over the past two decades more JSOs are receiving increased supervision, oversight, and secure facility placement. Although the goal of these laws is to reduce recidivism, they also have the effect of contaminating the criterion by which most risk assessment tools are judged (i.e., sexual recidivism). Future research should directly code for different risk management strategies (e.g., level and type of supervision) and time at risk outside of secure facilities, so as to determine the impact on indices of predictive accuracy.

Future research should assess the long-term predictive accuracy of the risk assessment tools, such as the JSORRAT-II, with exhaustive samples of JSOs. Epperson and colleagues (2006) took initial steps to determine whether the JSORRAT-II predicted sexual recidivism into adulthood. For some individuals in that sample, this represented only a few years into their 20s, while others were in their low 30s. Although the results of that study were not promising for the prediction of adult sexual recidivism, future studies should make efforts to follow samples further into adulthood.

Finally, some may argue that the similar predictive validity coefficients between DBU-inclusive and DBU-exclusive scoring justify the use of either approach. At this time, however, we believe it is prudent to replicate the findings before advocating for DBU-inclusive scoring in clinical practice. The JSORRAT-II was developed using several guiding principles that included a reliance on sources of data that were likely to lead to reliable scoring across cases (Epperson et al., 2006). With specific and explicit coding rules that define what constitutes DBU data and warns against the

inclusion of data from documented but uncorroborated allegations, high reliability in scoring might be attainable; however, replication is needed to confirm both the roughly equivalent predictive validity coefficients and interrater reliability of scoring the tool using those methods.

## Conclusion

The inclusion of DBU data in the scoring of the JSORRAT-II did not augment the tool's observed predictive validity coefficients in two states, and at this time the standard, non-DBU scoring criteria should be followed. Questions remain, however, about the impact of changing prosecutorial and judicial practices on risk assessment research. Despite these unanswered questions, the JSORRAT-II remains a promising tool to predict juvenile sexual recidivism when scored using the standards outlined in its scoring manual. Although high levels of reliability were achieved through the procedures used in our laboratory, field-workers are likely to also achieve satisfactory reliability with sufficient training and practice. In the states where it has been validated, the JSORRAT-II can be used to inform a range of shorter-term decisions related to placement, programming, and even treatment (when accompanied by a psychological and needs assessment). Used in that way, limited resources might be allocated more efficiently, undue negative consequences might be avoided for the lowest risk offenders, and communities might also be safer. However, outside those states and purposes, its use should be considered experimental.

## Acknowledgments

Special thanks to Dave Fowers, John Dewitt, Gary Niles, Tom Southard, and Laura Roeder-Grubb for coordinating the identification and transportation of case file information and the collection of recidivism data.

## Authors' Note

Douglas L. Epperson and Christopher A. Ralston developed the JSORRAT-II (Juvenile Sexual Offender Recidivism Risk Assessment Tool-II), the tool examined in this study. The JSORRAT-II is in the public domain and use is free. This research was completed in collaboration with the Utah Juvenile Court, the Utah Division of Juvenile Justice Services, Utah's Network on Juveniles Offending Sexually statewide organization, the Iowa Juvenile Court, and the Iowa Department of Human Rights, Division of Criminal and Juvenile Justice Planning.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by funding from the Iowa Department of Human Services and the Utah State Juvenile Justice Services Division.

## References

- Adam Walsh Child Protection and Safety Act of 2006, H.R. 4472. (2006). Retrieved from <http://www.govtrack.us/congress/billtext.xpd?bill=h109-4472>
- Batastini, A. B., Hunt, E., Present-Koller, J., & DeMatteo, D. (2011). Federal standards for community registration of juvenile sex offenders: An evaluation of risk prediction and future implications. *Psychology, Public Policy, and Law, 17*, 451-474.
- Boxer, P., Guerra, N. G., Huesmann, L. R., & Morales, J. (2005). Proximal peer-level effects of a small-group selected prevention on aggression in elementary school children: An investigation of the peer contagion hypothesis. *Journal of Abnormal Child Psychology, 33*, 325-338.
- Caldwell, M. F. (2002). What we do not know about juvenile sexual reoffense risk. *Child Maltreatment, 7*, 291-302.
- Epperson, D. L., Kaul, J. D., Goldman, R., Huot, S., Hesselton, D., & Alexander, W. (2004). *Minnesota Sex Offender Screening Tool-Revised (MnSOST-R)*. St. Paul: Minnesota Department of Corrections.
- Epperson, D. L., & Ralston, C. A. (in press-a). Conceptual model of risk versus threat and risk management versus risk reduction: Clarification of terms and implications for risk assessment. In A. Phenix & H. Hoberman (Eds.), *Sexual offenders: Diagnosis, risk assessment, and management*. New York, NY: Springer.
- Epperson, D. L., & Ralston, C. A. (in press-b). Development and validation of the Juvenile Sexual Offense Recidivism Risk Assessment Tool-II (JSORRAT-II). *Sexual Abuse: A Journal of Research and Treatment*.
- Epperson, D. L., Ralston, C. A., Fowers, D., DeWitt, J., & Gore, K. S. (2006). Actuarial risk assessment with juveniles who offend sexually: Development of the Juvenile Sexual Offense Recidivism Risk Assessment Tool-II. In D. Prescott (Ed.), *Risk assessment of youth who have sexually abused: Theory, controversy, and emerging strategies* (pp. 118-169). Oklahoma City, OK: Wood & Barnes.
- Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology, 148*, 839-843.
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). Accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*, 1-21.
- Harris, A. J., Lobanov-Rostovsky, C., & Levenson, J. S. (2010). Widening the net: The effects of transitioning to the Adam Walsh Act's federally mandated sex offender classification system. *Criminal Justice and Behavior, 37*, 503-519.
- Harris, A. J., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *STATIC-99 Coding Rules Revised-2003*. Ottawa, Ontario: Solicitor General Canada.
- Letourneau, E. J., Armstrong, K., Bandyopadhyay, D., & Sinha, D. (2013). Sex offender registration and notification policy increases juvenile plea bargaining. *Sexual Abuse: A Journal of Research and Treatment, 25*, 189-207.
- Letourneau, E. J., Bandyopadhyay, D., Armstrong, K., & Sinha, D. (2010). Do sex offender registration and notification requirements deter juvenile sex crimes? *Criminal Justice and Behavior, 37*, 553-569.
- Letourneau, E. J., Bandyopadhyay, D., Sinha, D., & Armstrong, K. (2009). Effects of sex offender registration policies on juvenile justice decision making. *Sexual Abuse: A Journal of Research and Treatment, 21*, 149-165.
- Letourneau, E. J., Levenson, J. S., Bandyopadhyay, D., Armstrong, K., & Sinha, D. (2010). Effects of South Carolina's sex offender registration and notification policy on deterrence of adult sex crimes. *Criminal Justice and Behavior, 37*, 537-552.



- Letourneau, E. J., & Miner, M. (2005). Juvenile sex offenders: A case against the legal and clinical status quo. *Sexual Abuse: A Journal of Research and Treatment, 17*, 293-312.
- Levenson, J. S., & Cotter, L. P. (2005). Effect of Megan's Law on sex offender reintegration. *Journal of Contemporary Criminal Justice, 21*, 49-66.
- Levenson, J. S., D'Amora, D. A., & Hern, A. L. (2007). Megan's Law and its impact on community re-entry for sex offenders. *Behavioral Sciences and the Law, 25*, 587-602.
- New Jersey Attorney General's Office. (2006). *Juvenile risk assessment scale manual*. Retrieved from <http://www.state.nj.us/lps/dcj/megan/jras-manual-scale-606.pdf>
- Prentky, R. A., & Righthand, S. (2003). *Juvenile sex offender assessment protocol: Manual*. Bridgewater, MA: Justice Resource Institute.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Ralston, C. A., Epperson, D. L., & Edwards, S. R. (in press). Cross-validation of the JSORRAT-II in Iowa. *Sexual Abuse: A Journal of Research and Treatment*.
- Tewksbury, R. (2005). Collateral consequences of sex offender registration. *Journal of Contemporary Criminal Justice, 21*, 67-81.
- Tewksbury, R., & Jennings, W. G. (2010). Assessing the impact of sex offender registration and community notification on sex-offending trajectories. *Criminal Justice and Behavior, 37*, 570-582.
- Tewksbury, R., Jennings, W. G., & Zgoba, K. M. (2012). Longitudinal examination of sex offender recidivism prior to and following the implementation of SORN. *Behavioral Sciences & The Law, 30*, 308-328.
- Trivits, L. C., & Reppucci, N. D. (2002). Application of Megan's Law to juveniles. *American Psychologist, 57*, 690-704.
- Viljoen, J. L., Mordell, S., & Beneteau, J. L. (2012). Prediction of adolescent sexual reoffending: A meta-analysis of the J-SOAP-II, ERASOR, J-SORRAT-II, and Static-99. *Law and Human Behavior, 36*, 423-438.
- Viljoen, J. L., Scalora, M., Cuadra, L., Bader, S., Chaves, V., Ullman, D., & Lawrence, L. (2008). Assessing risk for violence in adolescents who have sexually offended: A comparison of the J-SOAP-II, J-SORRAT-II, and SAVRY. *Criminal Justice and Behavior, 35*, 5-23.
- Wetterling, J. (1994). *Crimes Against Children and Sexually Violent Offender Registration Act*. Retrieved from <http://www.gpo.gov/fdsys/granule/USCODE-1999-title42/USCODE-1999-title42-chap136-subchapVI-sec14071>
- Worling, J. R., & Curwen, T. (2001). "ERASOR": *Estimate of Risk Adolescent Sexual Offense Recidivism* (Version 2.0). Toronto, Ontario, Canada: SAFE-T Program, Thistletown Regional Centre for Children & Adolescents.
- Zgoba, K. M., Miner, M., Knight, R., Letourneau, E., Levenson, E., & Thornton, D. (2012). *Multi-state recidivism study using Static-99R and Static-2002 risk scores and tier guidelines from the Adam Walsh Act* (Research report submitted to the National Institute of Justice). Washington, DC: National Institute of Justice. Retrieved from <http://nij.gov/publications/pages/publication-detail.aspx?ncjnumber=240099>