
Does Risk Assessment Make a Difference? Results of Implementing the SAVRY in Juvenile Probation

Gina M. Vincent, Ph.D.^{*}, Laura S. Guy, Ph.D.[†],
Bernice G. Gershenson, M.P.H.[†] and Patrick McCabe, Ph.D.[†]

An effective approach to reducing recidivism is, first, to identify a youth's risk of reoffending and then to match the intensity of interventions to that risk level. This pre-post quasi-experimental, prospective study compared 247 (pre) with 217 (post) adjudicated youths to examine the implementation of the Structured Assessment of Violence Risk in Youth (SAVRY) and its effects on case management practices in Louisiana's Caddo parish probation office. The results indicated that placement rates dropped by 50%, use of maximum levels of supervision dropped by almost 30%, and use of community services decreased except for high-risk youths, but only after the SAVRY was properly implemented. This shift towards more appropriate allocation of resources that are matched to risk level occurred without a significant increase in reoffending. The implications for implementation and for use of risk/needs assessment in juvenile probation are discussed. Copyright © 2012 John Wiley & Sons, Ltd.

Increasing research evidence demonstrates that punishment and sanctions do not deter juvenile reoffending as intended (e.g., Lipsey & Cullen, 2007; Loughran et al., 2009; MacKenzie, Wilson, & Kider, 2001), and that over-involvement may in fact produce iatrogenic effects (see Gatti, Tremblay, & Vitaro, 2009). For example, a recent meta-analysis of 548 juvenile intervention studies reported that discipline increased recidivism rates by an average of 8% (Lipsey, 2009). Another meta-analysis demonstrated that recidivism rates were higher among juveniles processed into the juvenile justice system than among juveniles who were diverted (except for first-time offenders; Petrosino, Turpin-Petrosino, & Guckenbug, 2010).

Alternatively, evidence suggests that case management practices and programming will increase chances of success by considering specific characteristics of the youths; namely, their risk for reoffending and the factors most likely driving their reoffending (criminogenic risk). The first step in promoting these best practices is the use of developmentally appropriate risk assessment tools that provide an accurate estimate of a youth's level of risk for reoffending as well as identification of risk factors that are amenable to change. Even though the use of risk assessment tools is consistent with best practice (see the Juvenile Justice Delinquency and Prevention Act, 2002), little research has been conducted to examine whether the adoption of risk assessment actually leads to changes in the way young offenders are handled. This study contributes to the small research base (Luong & Wormith, 2011; Vieira, Skilling, & Peterson-Badali,

^{*}Correspondence to: Gina M. Vincent, Ph.D., Associate Professor, University of Massachusetts Medical School, Department of Psychiatry, 55 Lake Avenue North, WSH 8B, Worcester, MA 01655. E-mail: Gina.Vincent@umassmed.edu

[†]University of Massachusetts Medical School, MA.

2009) on this topic, and reports on the impact of putting a valid risk assessment tool in place in one juvenile probation office by examining juvenile justice decisions at multiple levels.

CONTRIBUTIONS OF RISK ASSESSMENT AND RISK-NEED-RESPONSIVITY

A widely utilized framework for case management that incorporates risk assessment is the risk-need-responsivity (RNR) approach (Andrews & Bonta, 2006, 2010). The risk principle suggests that the highest-risk offenders should receive the most intensive programming to reduce risk of reoffending. Research shows that placing low-risk youths into intensive services may actually increase their risk (e.g., Andrews, Bonta, & Hoge, 1990; Lowenkamp & Latessa, 2005). Because low-risk youths have a much lower chance of reoffending even in the absence of services (Andrews & Bonta, 2010; Andrews & Dowden, 2006), they may be better served by minimal attention. The need principle identifies the most worthwhile need areas to be targeted to reduce reoffending (i.e., “criminogenic needs”), which comprise dynamic risk factors that are theoretically amenable to treatment (e.g., poor parental monitoring). The responsivity principle suggests that interventions need to consider the offenders’ specific characteristics that may affect their response (e.g., learning style).

Aspects of these principles have been supported by rigorous research and meta-analyses, primarily with adult populations (Andrews & Bonta, 2006, 2010; Andrews & Dowden, 2006; Andrews et al., 1990; Dowden & Andrews, 1999, 2000, 2004; Gendreau, Smith, & French, 2006) but also with young people (Lipsey, 2009). Only a few studies have examined RNR principles at the individual case level where valid risk assessments were used and matched to services or reoffending rates. Vieira et al. (2009), for example, found that young probationers who received services directly aligned with their criminogenic needs as identified by the Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2006) in their probation file reoffended at a rate of 25% vs. 75% for youths who received services that did not match their needs. Similarly, in an archival study of youth probationers in two departments instructed to use the Saskatchewan Youth Edition of the LSI (LSI-SK; Andrews, Bonta, & Wormith, 2001), Luong and Wormith (2011) reported that recidivism significantly increased as the number of untreated needs increased ($r = 0.28$). For high-risk offenders, the match between an assessed need and an identified intervention was associated with a 38% reduction in reconviction.

IMPORTANCE OF SOUND IMPLEMENTATION OF RISK/ NEEDS ASSESSMENT

The first crucial step for promoting juvenile justice practices that are consistent with the RNR approach is identification of youths’ risk for reoffending and criminogenic needs using a valid, developmentally appropriate risk/needs assessment tool. Surprisingly, there is very little research regarding the actual changes that occur in case management decisions once a tool is put in place. The extent of these changes is likely tied to the quality of the implementation process (Bonta, Bogue, Crowley, & Mottuk, 2001; Latessa & Lovins, 2010). It is unlikely that an assessment tool is going to lead to any

change if staff are not trained how to use the tool in their decision-making and if the work environment does not reinforce use of the tool.

Implementation activities for a risk/needs assessment tool include, at a minimum, staff training on how to conduct the risk assessment, revision of office policies to include the risk assessment process, quality assurance, and development of a sound model for use of the assessment in decision-making. Most risk assessment-related research investigates outcomes of interest, and largely ignores the quality of implementation procedures used (see Seave, 2011). However, a few studies exist that can speak to implementation issues. In criminal justice, this research has primarily focused on the importance of staff training and experience. For example, one study found that risk level (from the Level of Service Inventory-Revised, LSI-R) was positively correlated with re-incarceration of adult probationers (Flores, Lowenkamp, Holsinger, & Latessa, 2006). However, this was only the case when the LSI-R was administered by probation officers (POs) with adequate training and by offices where the tool had been in place for 3 years or more. In another study, POs with intensive training in RNR case management had significantly reduced recidivism rates for their caseloads relative to those without training (Bonta *et al.*, 2011).

There is some evidence that front-line workers undervalue and, consequently, under-use results of risk tools to inform case management strategies. Haas and DeTardo-Bora (2009) found that adult correctional staff surveyed 1.5 years after implementation did not believe the LSI-R was appropriate for the offenders with whom they worked, and only 42% used results of the LSI-R to develop re-entry case plans. Similarly, in a survey of juvenile probation officers (JPOs) and court officials in 12 courts in four states, most reported that they stopped using the tools within 2 years of adoption, only half of the court professionals and JPOs ever used the tools regularly in their decision-making, and over-rides were very common (Shook & Saari, 2007). The authors reported that use of risk assessment tools in decision-making was inhibited by, among other things, lack of appropriate training, lack of structured decision-making procedures, and disappointment that implementation of the tool did not achieve targeted goals.

Luong and Wormith (2011) examined the extent to which supervision and case management strategies were guided by the LSI-SK. The LSI-SK risk level was strongly associated with the supervision level assigned to probationers (low-risk youths were given minimum levels of supervision). Further, there was a strong correspondence between identification of needs as classified by the youth worker and by the tool overall, but youth workers over-identified and over-referred youths in some need areas. For example, only 13% of youths scored medium or high risk on the education/employment scale, yet 81% were assigned an educational intervention. Overall, the LSI-SK risk level significantly predicted reconversion over a nearly 2-year follow-up. This study lends support to the notion that risk assessment is used in some juvenile probation decision-making. However, the study was not designed to determine whether the case management practices were an improvement over practices before the LSI-SK was implemented.

Possibly one of the only published studies designed to indicate whether implementation of a standardized risk assessment tool actually effectuated changes in case management was conducted in five counties in Maryland using a pre-post research design (Young, Moline, Farrell, & Biere, 2006). Average adherence to the policy for administering the risk assessment tool was only 55%. With respect to placements, use of risk assessment increased the weight JPOs put on some dynamic risk factors (e.g., participation in school). With respect to service referrals, similar to findings from Luong and Wormith (2011),

JPOs seemed to follow the risk assessment consistently for some service types (e.g., substance abuse, mental health) but not for others (e.g., education). Unfortunately, data regarding changes in the overall percent of service referrals or placements as a result of implementing the risk assessment were not reported.

THE STRUCTURED ASSESSMENT OF VIOLENCE RISK IN YOUTH (SAVRY)

The SAVRY (Borum, Bartel, & Forth, 2006) is a risk/needs assessment tool that was designed to assess violence risk in adolescents aged 12–18 years. However, several studies have demonstrated that the SAVRY has high predictive accuracy for both violent and non-violent reoffending (e.g., Lodewijks, Doreleijers, & de Ruiter, 2008; Welsh, Schmidt, McKinnon, Chattha, & Meyers, 2008). Its predictive validity has been demonstrated in forensic and young offender populations in at least 10 studies published by independent researchers (see Borum, Lodewijks, Bartel, & Forth, 2010 for a review). A recent meta-analysis reported moderate effect sizes (weighted r) for SAVRY total scores of 0.38 for non-violent reoffending and 0.30 for violent reoffending (Olver, Stockdale, & Wormith, 2009). These effect sizes are comparable to or better than other youth risk assessment tools (see Vincent, Terry, & Maney, 2009, for a review).

The SAVRY uses a structured professional judgment approach (as opposed to a mechanical actuarial approach). This means the final estimation of a youth's overall level of risk for violence or reoffending is based on: the evaluator's professional judgment as informed by a systematic appraisal of factors that are empirically related to risk; and a consideration of the relevance of those factors for the youth being assessed. The final risk judgment typically is communicated using categorical estimates of low, moderate, and high. As is the case with most comprehensive risk assessment tools, the SAVRY includes dynamic risk factors that may be considered targets for treatment to reduce risk for reoffending.

Juvenile probation officers can be effectively trained to utilize the structured professional judgment (SPJ) approach. Results of a rigorous study of a state-wide sample of JPOs in Louisiana found good to excellent field inter-rater reliability (IRR) for the SAVRY with an intra-class correlation coefficient (ICC) for the summary risk rating (SRR) of 0.71 (Vincent, Guy, Fusco, & Gershenson, 2012). This is comparable to findings from researcher and clinician raters. The SAVRY also has support for its field predictive validity with young offenders from a large sample of juvenile detainees in Connecticut (Vincent, Chapman, & Cook, 2011).

CURRENT STUDY

Despite the necessity of adopting risk assessment tools, there has been surprisingly little research addressing whether implementation of a risk assessment tool in juvenile justice actually leads to changes in the way JPOs handle young offenders. The Risk/Needs Assessment in Juvenile Probation: Implementation Study is a pre–post prospective study of the change effectuated by implementing a risk assessment tool in six juvenile probation offices. In this article, we report results from the first probation office to complete the implementation study.

This was a pre–post study of implementation of the SAVRY at the post-adjudication/pre-disposition point of case processing. At this decision point, information is gathered about the characteristics of youths in a manner that guides JPOs' decisions and recommendations in the form of a pre-disposition investigation (PDI). Implementation of a risk tool at this decision point should have the largest impact on pre- and post-dispositional planning. In general, we hypothesized that the use of restrictive monitoring and incarceration would decrease once the SAVRY was implemented and RNR principles were followed. The hypotheses predicted decreases in the following areas of study: rates of more restrictive dispositions; rates of out-of-home placements; and use of more intensive supervision levels. The hypotheses predicted increases in the number of youths completing services as a result of better matching of services to youths' risk levels. For all areas of case management, JPOs' decisions were expected to be related to youths' level of risk after SAVRY implementation. Lastly, the study compared rates of reoffending between the groups. Although research reviewed earlier indicated that adherence to the RNR approach reduces recidivism, such a reduction is, of course, tied to the availability of effective services and programming, and therefore cannot be attributed solely to implementing a risk assessment tool. Thus, the hypothesis was that reoffending would not increase after implementation of the SAVRY.

METHOD

This study used a pre–post implementation design with two groups: a sample of youths adjudicated in Caddo parish juvenile probation during the year prior to SAVRY implementation (pre-implementation group) and a sample of youths adjudicated after full implementation of the SAVRY (post-implementation group). The Caddo parish probation office is locally run and is located in a metropolitan city in Louisiana. The office received funding from the MacArthur Foundation Models for Change Initiative to engage in juvenile justice reform. This is a fairly large probation department with 16 JPOs responsible for PDIs and case management, each with an average caseload of 40 youths per month, and an office average of 40 new adjudicated youths per month.

Samples

Youth cases were selected for inclusion in this study based on their adjudication dates (referred to as the “initial adjudication”) with the aim of obtaining close to 250 youths in each sample. All JPOs received training on the SAVRY in November 2008 and the SAVRY was fully implemented 1 February 2009. The pre-implementation group was drawn from a period immediately prior to SAVRY training (1 April 2008 to 15 October 2008). A random sample of 247¹ youths was drawn from the 339 youths adjudicated for a delinquency or “family in need of services” (FINS) offense in the department during this period. The post-implementation group was drawn from all delinquent or

¹ Initially 250 youths were selected, but three were eliminated due to problems with record-keeping. The reason for attempting to obtain 250 youths in each group was to get a good range of cases with different characteristics and outcomes in addition to needing at least 107 subjects in each group to achieve adequate power for regression models with up to eight predictor variables (Cohen, 1992).

FINS adjudications from 1 March 2009 to 1 September 2009. Because adjudications decreased during this year, it was necessary to use all consecutive adjudications rather than a random sample, resulting in a final sample of 217. The pre-implementation and post-implementation cases were mutually exclusive. Data were coded from case files by a trained research assistant (RA) employed by the probation office. Researchers did not have contact with the youths. Table 1 contains the basic demographic and psychosocial information for the groups.

Measures and Procedures

SAVRY

The SAVRY protocol comprises six items defining protective factors (which may lower risk) and 24 items defining risk factors (which may increase risk). The SAVRY risk items are divided conceptually into three domains: historical (10 items that are “static” in the sense that they are rated in terms of their presence in the past), social/contextual (six dynamic items), and individual/clinical (eight dynamic items). Evaluators also are able to identify additional risk and protective factors, recognizing that some cases may present circumstances that are not included among the SAVRY items. The six protective factors are rated as “present” or “absent,” whereas the risk factor items are rated as low, moderate, or high. Total risk scores were generated by assigning numeric values to

Table 1. Group demographics before and after propensity score matching

	Pre-implementation group		Post-implementation group	
	Before match (<i>n</i> =247)	After match (<i>n</i> =205)	Before match (<i>n</i> = 217)	After match (<i>n</i> = 205)
Gender	27.0%	33.2%	34.1%	29.3%
% African American	88.7%	86.8%	86.2%	86.8%
Age at adjudication (years)	15.06 (1.66)	15.00 (1.69)	15.16 (1.53)	15.10 (1.56)
Age at first offense (years)	13.67 (1.70)	13.73 (1.70)	13.94 (1.70)	13.85 (1.56)
% Violent index offense	29.1%	28.3%	24.0%	25.4%
% Adjudicated FINS	23.4%	22.4%	20.2%	21.0%
Prior state JJ commitment	5.3%	3.41%	3.2%	5.4%
Any Axis I diagnosis	15.4%	15.6%	12.0%	12.7%
Any Axis II diagnosis	8.1%	7.8%	7.8%	8.3%
Prior outpatient mental health treatment	17.3%	17.2%	21.3%	21.1%
Prior outpatient substance abuse treatment	0.4%	0.5%	0.5%	0.5%
Any grade failure	25.9%	27.1%	29.3%	29.9%
Regular school attendance	70.4%	71.4%	66.2%	65.1%
Good school performance	70.5%	73.1%	73.9%	72.8%
Living arrangement				
Both parents	18.2%	16.6%	15.2%	15.6%
Single parent	68.8%	69.8%	71.4%	70.7%
Relative	10.9%	11.7%	12.0%	12.2%
Currently placed out of home	1.3%	1.5%	3.7%	3.4%
History of child welfare involvement	8.1%	8.7%	8.3%	7.3%

FINS, family in need of services; JJ, juvenile justice.

Note. No comparisons listed in the table are statistically significant.

the item level categorical ratings (0 = low, 1 = moderate, 2 = high) and then summing across item scores. Protective factors were assigned a 1 (present) or 0 (absent) rating and then summed.

Juvenile probation officers were trained to rate the SAVRY using file information and data from interviews with the youth, a parent, and the youth and parent together. Semi-structured interview scripts were given to them. After considering the presence and relevance of all risk and protective factors, JPOs assigned a SRR of low, moderate, or high for the youth's overall estimated risk of future violence. Except where noted explicitly, the SRR was used in all analyses of risk level reported in the following. A SAVRY was considered invalid if five or more items (i.e., 17% of items) were not rated. Only one SAVRY was invalid.

Inter-rater reliability was evaluated using independent ratings of 10 JPOs and the trained RA. The RA observed the JPOs' interviews and reviewed the same file information. Eight of the JPOs completed three IRR cases with the RA and two JPOs completed two cases. Based on 28 cases, ICC₁ values for the SRR and total scores were 0.75 and 0.94, respectively. Disagreements about the final risk judgment were never more than one level (i.e., no rater pair ever had a high/low combination).

Implementation Procedures

The probation office followed a standardized method of implementation of the SAVRY that was completed with the assistance of the researchers. A detailed description of the implementation procedures has been reported elsewhere (see Vincent, Paiva, Cook, Guy, & Perrault, 2012) so will be summarized only briefly here. First, the researchers conducted an orientation training with staff about risk/needs assessment and what to expect with the SAVRY implementation. Secondly, researchers worked with the administrators to complete and implement a comprehensive policy about how the SAVRY assessments would be completed, and how findings and opinions would be communicated in their reports and used in case planning. Thirdly, researchers worked with the office to revise case plan forms that were organized according to areas of criminogenic needs (e.g., disruptive behavioral problems, substance abuse, education, family).² A service referral matrix that organized services available in the community according to the need area they addressed and the risk level for which they were appropriate (based on service intensity) also was developed. Finally, JPOs completed a SAVRY training workshop, and three additional practice cases were completed over a 2-month period, as well as training on the RNR principles and new policies. They also received a booster training on the SAVRY 6 months later.

Data Collection Procedures

Given the pre–post design of this study, it was essential to control the time periods over which data were gathered. The activity of youths in both the pre- and post-implementation groups was tracked up to 13 months from the last adjudication date in the sample ($M = 11.55$ months, $SD = 5.82$). The RA was trained to gather data from court and

² The SAVRY dynamic items were grouped into criminogenic need areas based on rational item-selection and internal consistency analyses. This was essential for the JPOs' case management. A review of these procedures is beyond the scope of the present paper.

administrative databases and youth probation files, which contained PDI reports, case management forms, JPOs' notes, and, in some cases, psychological assessments. Data coded for each case included, but were not limited to, psychosocial history; prior offenses; dates of various hearings; dates of all movement within the court system and changes in legal status, dispositions, placement dates and locations; probation violations; and all service referrals including dates when services were received. The RA entered data into an access database and participated in biweekly calls with the investigators to troubleshoot any data coding issues.

Dependent Variables

"Disposition" was coded according to the most restrictive disposition received during the initial adjudication. For example, if the youth spent time in detention awaiting disposition and received a disposition of "detention time served" plus 1 year probation, disposition would have been coded as detention. A disposition of commitment to the state juvenile justice (JJ) agency generally resulted in placement in a correctional facility. "Out-of-home placement" was defined as any removal from the home other than to foster care. This included detention, shelters, group homes, residential facilities, substance abuse and mental health treatment facilities, and the state correctional program. Placement rates were examined in two ways: whether youths received any time in a placement during the 13 months of the study (yes/no); and whether youths were placed immediately following their disposition (yes/no). "Level of supervision" refers to the first level of monitoring to which youths were assigned while on probation. There were four possible levels of supervision: minimum (one face-to-face contact every 60 days), moderate (one face-to-face contact every 30 days), maximum (one face-to-face contact every 2 weeks), and intensive (weekly contact). Each youth's assigned level of supervision was documented by JPOs in the case plan.

"Services" were defined as any service in the community aimed at treatment or rehabilitation (e.g., mentoring programs, functional family therapy, individual counseling, psychological evaluations). Sanctions or punishments, such as community service, curfew, electronic monitoring, and JPO supervision, were not counted as services. The following was recorded from the case plans of each youth for every service: the type of service, agency, start and end dates the youth attended the service, and whether the service was successfully completed. The number of different service referrals made by JPOs for each youth and the number of services successfully completed was calculated over the 13-month period.

"Recidivism" was tracked in the community from both juvenile and adult court records for a minimum of 15 months past the last adjudication date in each group. The average length of follow-up was 18.67 months. Recidivism was defined in two ways: a new petition following the initial adjudication; and a new adjudication after the initial adjudication. New offenses were categorized as: "violent" (offenses related to actual, attempted, or threatened harm to another, including sex offenses); "non-violent" (any non-violent offense, such as drug offenses); (c) "violations" (included probation violations, minor traffic violations, and status offenses); and (d) "any reoffending", which included all types. The first three categories were mutually exclusive.

Time at risk for committing an offense was calculated separately for each offense category using the date of the first petition or adjudication in that category. The starting point was the initial adjudication date or date of first release from a placement for

youths placed immediately after adjudication. The end point for youths who did not reoffend was the date of the follow-up data collection. Youths were excluded from recidivism analyses if they spent the entire follow-up period in a placement with no opportunity to reoffend in the community ($n = 5$).

Data Analysis Procedures

Propensity score matching was used to equate the pre- and post-implementation groups along a number of important youth characteristics (e.g., offense history, current offense, demographic variables, psychosocial history variables). Propensity score matching is a technique commonly used in observational studies to reduce potential bias resulting from differences in relevant characteristics between control and treated groups (Rosenbaum & Rubin, 1983). Matching is done to identify control and treated participants who have a better balance on a range of relevant characteristics. Propensity scores were modeled with logistic regression, with the dependent variable being the odds of being in the post-implementation group. For some of the variables, missing data were coded into a separate category (e.g., history of child welfare involvement). Matching was done using the `psmatch2` (Leuven & Sianesi, 2003) procedure with a one-to-one nearest neighbor (with no replacement and with common support) matching schema using Stata (Statacorp, 2007) software.³

Three types of regressions were used to compare the matched pre- and post-implementation groups along a number of dependent variables. The regression model selected was a function of the nature of the dependent variable. Hierarchical logistic regressions were used for analyses with dichotomous dependent variables (e.g., placement status) and were supplemented with generalized linear modeling (GLM) to derive the groups' marginal means on the dependent variables. Hierarchical linear regression was used for analyses using continuous dependent variables (e.g., number of service referrals). Cox proportional-hazards regression was used to compare differences in the time to reoffending between groups. The Cox regression was selected because it permits inclusion of censored cases (in this context, those who have not yet offended) while comparing groups in their time to a reoffense. Covariates were included in most of the analyses even after balancing the groups using propensity-score matching.⁴ Covariates were identified by testing all demographic and psychosocial history variables to identify those that were significantly different between the pre- and post- implementation groups before matching, and those that were significantly related to the specific dependent variable. Every hierarchical regression model entered group membership (pre- vs. post-implementation) at the last step after factoring out significant covariates at the first step (when applicable). The covariates included in each model are listed in the Results section for each regression model.

To examine whether decisions made for the post-implementation group were related to risk, we calculated Cramer's V or Spearman's correlations between the

³ Three different propensity score matching approaches were tried, including nearest neighbor without common support (which eliminated very few subjects and did not result in as good a match between groups) and Mahalanobis-metric matching (which eliminated too many subjects). The results of group comparisons indicated that the nearest neighbor with common support approach was similar to the nearest neighbor approach, except that statistically significant findings were amplified.

⁴ Analyses were conducted both with and without including covariates after propensity matching. The decision was made to retain covariates in the models because these still affected some of the results, and this was the most conservative approach.

Table 2. SAVRY results for the post-implementation group ($n = 195$)

	Range	Mean	SD
SAVRY total (possible 48)	0–34	11.70	7.86
Protective factors scale (possible 6)	0–6	3.69	2.00
Summary risk rating (SRR)	Low	Moderate	High
n (%)	72 (37%)	95 (49%)	28 (14%)
Mean (SD) SAVRY total score	4.94 (3.75)	13.93 (5.87)	28 (14.4)

SAVRY, Structured Assessment of Violence Risk in Youth.

SRR (low, moderate, or high) and the dependent variable. For recidivism analyses, receiver operating characteristic curves (ROCs) were used to examine the predictive accuracy of the SAVRY. The summative index of accuracy is represented by the area under the curve (AUC), which is an ideal index for research in risk assessment because AUC values are not affected by variations in base rates and selection ratios. In the present study, an AUC value can be interpreted as the probability that a SAVRY SRR (or score) drawn at random from the youths who reoffended will be higher than the SRR drawn at random from the group of youths who did not reoffend.⁵ All analyses examining the correspondence between decisions and risk level were conducted with the entire post-implementation group without eliminating the subjects removed for propensity matching, because balancing the pre- and post-implementation groups was not relevant to these analyses.

RESULTS

The propensity score matching procedure eliminated 42 youths from the pre-implementation group and 12 youths from the post-implementation group. The characteristics for the original and final pre- and post-implementation groups are provided in Table 1.

SAVRY Results and Adherence to the Administration Policy

Before we examined any changes due to implementation of the SAVRY, it was first necessary to check whether the probation office was using the tool in accordance with its policy. Ninety per cent ($n = 195$) of adjudicated youths in the post-implementation group had been administered the SAVRY, indicating good adherence to the policy. Table 2 provides descriptive information for SAVRY scores and ratings for the post-implementation group. Most youths were rated as being at moderate risk and the distribution of SAVRY total scores within each SAVRY risk rating was appropriate, providing some validity for JPOs' ratings (see Table 2). Most (90%) had one or more protective factors rated as present. Only 24.0% ($n = 53$) of the 217 youths in the post-implementation group had a SAVRY completed prior to disposition as part of

⁵ AUC values range from 0 (perfect negative prediction) to 1.0 (perfect positive prediction), with 0.50 indicating chance prediction. Values < 0.70 may be considered to be small; values in the range of 0.70 to 0.75 have been viewed as moderate; and values > 0.75 typically are regarded as large (see, e.g., Douglas, Yeomans, & Boer, 2005). However, such descriptors are tentative, as no definitive classification scheme exists.

the PDI. Rather, the majority of cases received a SAVRY assessment immediately after disposition for case planning purposes.

Dispositions

Overall (before eliminating subjects by propensity score matching), the most restrictive disposition for the majority of youths was probation ($n = 341$, 73.5% of total sample), followed by detention ($n = 74$, 15.9%) and commitment to the state JJ agency ($n = 36$, 6.9%).⁶ To determine if the pre- and post-implementation groups differed significantly in the likelihood of obtaining each disposition, three hierarchical logistic regressions were conducted, one for each type of disposition (yes/no). Youths in the post-implementation group (marginal mean = 85%, SE = 0.03) were significantly more likely than those in the pre-implementation group (marginal mean = 67%, SE = 0.03) to receive a probation disposition ($\beta = 0.99$, SE = 0.25, Exp[B] = 2.69, $p < 0.001$) after factoring out covariates [prior violent charges (yes/no) and total number of prior charges].⁷ Youths in the post-implementation group also were significantly less likely to be given dispositions of detention [marginal means = 8% (SE = 0.02) vs. 20% (SE = 0.03), respectively; $\beta = -0.95$, SE = 0.31, Exp(B) = 0.39, $p = 0.002$] and state commitment [marginal means = 4% (SE = 0.01) vs. 10% (SE = 0.02); $\beta = -0.94$, SE = 0.44, Exp(B) = 0.39, $p = 0.03$].

Examination of the post-implementation group indicated the only disposition type that was significantly related to risk level was committing youth to the state (total $n = 8$), where low-risk youths (12.5%) had a lower likelihood than moderate- (25.0%) or high-risk youths (35.5%) of being committed (Cramer's $V = 0.18$, $p = 0.039$). The majority of youths receiving probation ($n = 180$) or detention ($n = 22$) were at moderate risk. The percentages of youths at each level of risk were as follows – for probation: 33.3%, low; 43.3%, moderate; and 12.8%, high; and for detention: 31.8%, low; 59.1%, moderate; and 9.1%, high for.

Many youths did not have a PDI conducted prior to their disposition, in which case the court would not know the youths' SAVRY risk level. In such cases, one would not expect the dispositions to be strongly related to risk level. Table 3 provides the percentage of post-implementation youths at each risk level who received each type of disposition separately for the 53 youths who received a SAVRY prior to disposition vs. those who did not. Results indicated that the only difference from those who did not receive a SAVRY prior to their disposition hearing was that no low-risk youths were committed to the state (see Table 3).

Out-of-Home Placements

Overall (before matching), 198 youths were put in an out-of-home placement at some point during the 13-month tracking period and 125 youths were put in a placement immediately following disposition. Most placements were detention facilities, group

⁶ A few youths received informal dispositions ($n = 17$, 3.7%) instead, but the number was too low to make meaningful comparisons, so this disposition was not a focus for analysis.

⁷ Which judge the youths saw for disposition was also significantly related to the likelihood of receiving detention ($\chi^2(1) = 7.33$, $p = 0.007$) but this was not factored into the models because of the nominal scaling of this variable. Essentially, both judges in this court had significant decreases in use of detention or commitment to the state.

Table 3. Percentage of post-implementation youths receiving each disposition type and any placement as a function of the SAVRY risk level ($n = 195$)

	Low [% (n)]	Moderate [% (n)]	High [% (n)]	Cramer's V
Dispositions made without SAVRY ¹ ($n = 143$)				
Probation ($n = 122$)	83.6 (51)	84.1 (53)	94.4 (17)	0.11
Detention ($n = 12$)	8.2 (5)	11.1 (7)	0.0 (0)	0.13
JJ state commit ($n = 3$)	1.6 (1)	1.6 (1)	5.6 (1)	0.09
Dispositions with SAVRY in PDI ² ($n = 53$)				
Probation ($n = 40$)	81.8 (9)	78.1 (25)	60.0 (6)	0.18*
Detention ($n = 10$)	18.2 (2)	18.8 (6)	20.0 (2)	0.02
JJ state commit ($n = 3$)	0.0 (0)	3.1 (1)	20.0 (2)	0.30
Out-of-home placements				
Any placement received ($n = 67$)	22.2 (16)	35.8 (34)	60.7 (17)	0.26***
Placed right after disposition ($n = 36$)	12.5 (9)	21.2 (20)	25.0 (7)	0.12
SAVRY pre-disposition ($n = 18$) ¹	27.3 (3)	31.3 (10)	50.0 (5)	0.17
SAVRY not pre-disposition ($n = 18$) ²	9.8 (6)	15.9 (10)	11.1 (2)	0.087

JJ, juvenile justice; PDI, pre-disposition investigation report; SAVRY, Structured Assessment of Violence Risk in Youth.

Note. Cells represent the percentage of youths within each SAVRY risk category who received each disposition type or a placement.

¹Cells include only youths who received a disposition/placement from the court before a SAVRY was conducted.

²Cells include only youths for which the court would have seen SAVRY results in a PDI prior to making their disposition/placement decision. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

homes, and shelters. Only nine youths were placed in an in-patient mental health treatment facility of some sort. Two logistic regressions were conducted to determine whether there was a significant difference before and after the SAVRY implementation in the likelihood of receiving any placement immediately following disposition and at any point during the study period.

Immediately following disposition, the pre-implementation group had a significantly higher rate of placement (marginal mean = 33%, SE = 0.03) than the post-implementation group [marginal mean = 15%, SE = 0.03; $\beta = -0.98$, SE = 0.25, Exp(B) = 0.37, $p = 0.001$] after accounting for significant covariates [prior violent charges (yes/no) and total number of prior charges]. Similarly, the pre-implementation group was significantly more likely to receive any placement during the study period (marginal mean = 46%, SE = 0.04) than the post-implementation group [marginal mean = 33%, SE = 0.03; $\beta = -0.58$, SE = 0.212, Exp(B) = 0.56, $p = 0.006$]. Thus, before the SAVRY was implemented, youths were 2.67 times as likely to be put in placement immediately following their disposition and 1.79 times as likely to be placed at any time during their probation. A hierarchical linear regression indicated the average amount of time spent in a placement facility (dependent variable) did not differ between the groups after factoring out covariates [current offense a non-violent felony (yes/no) and total number of prior charges]. The marginal mean length of placement for the pre-implementation group ($n = 94$) was 5.35 months (SD = 17.96 days) and for the post-implementation group ($n = 68$) was 5.13 months (SD = 21.15 days) ($\beta = -12.56$, SE = 26.74, $t = -0.47$, $p = 0.64$, overall model $R^2 = 0.23$).⁸

Whether youths received any placement during the study period (13 months from adjudication) was significantly related to the SAVRY risk level in the post-implementation

⁸ Results are reported from an ANCOVA after factoring out whether youths were adjudicated for a non-violent felony and the number of prior charges.

group (see Table 3), such that higher-risk youths were more likely to be placed. Further, high-risk youths spent more time in placements, on average (Mdn = 182 days) than low-risk youths (Mdn = 61 days) ($Z = -2.25, p = 0.02$). However, whether youths were placed immediately after disposition was not significantly related to risk (see Table 3). Again, similar to the findings related to disposition, one might expect youths ordered into placement prior to having a SAVRY or PDI conducted would have a weaker association between their risk level and the placement decision than youths for whom the court had access to the SAVRY risk level. Indeed, upon this deeper examination, it appeared that for youths who were administered a SAVRY prior to their disposition, there was a trend towards more restrictive placements for higher-risk youths, whereas placements were unrelated to risk for those disposed before a SAVRY was conducted. Table 3 illustrates the correspondence between risk level and placements for these two groups.

Level of Community Supervision

There were significant differences in the starting level of supervision for youths on probation between the pre-implementation and post-implementation groups at every level of supervision: minimum (23.4% vs. 39.7% respectively), moderate (33.2% vs. 47.2%), maximum (32.1% vs. 12.1%), and intensive (11.4% vs. 1.0%). Specifically, in the post-implementation group, there were significantly more youths on minimum and moderate levels of supervision and significantly fewer on maximum and intensive levels of supervision ($V = 0.32, p = 0.001$). Further, in the post-implementation group, youths were being started at low levels of supervision and they were not simply increased to more restrictive levels later in their probation period.

In the post-implementation sample, supervision level was significantly related to risk level, with high-risk youths (77.8%) more likely than low-risk youths (0%) to be placed on maximum supervision, and less likely to be placed on minimum supervision (3.2% vs. 94.4%, respectively) ($V = 1.23, p = 0.001$). The majority of youths assigned to moderate supervision were at moderate risk (91.9%).

Service Referrals

Overall, JPOs' referrals to structured community services for youths who spent any time on probation ranged from one to five referrals. On average, youths received a mean of 1.5 service referrals, but distributions were highly skewed. Linear regression indicated a significant increase in the number of service referrals JPOs gave youths after the SAVRY was implemented (pre-implementation $M = 1.46, SD = 0.70$; post-implementation $M = 1.82, SD = 1.0$; group $\beta = 0.36, SE = 0.09, t = 3.89, p < 0.001$; overall model $R^2 = 0.04$). For exploratory purposes, χ^2 analyses were conducted to examine the odds of obtaining certain types of services in the post-implementation group, relative to the pre-implementation group. There were significant increases in the use of many service types such that youths in the post-implementation group had greater odds of being referred to multi-systemic therapy (OR = 6.02), a life skills service (OR = 2.66), disruptive behavior-related services (OR = 2.06), peer relations services (OR = 3.57), and family-related services (OR = 2.18). Another linear regression indicated there was also a significant increase in the number of youths successfully completing services (pre-implementation marginal mean = 1.32, SE = 0.07; post-implementation marginal mean = 1.71, SE = 0.06; $\beta = 0.39, SE = 0.10, t = 4.02, p < 0.001$, overall model

$R^2 = 0.07$) after accounting for covariates (number of prior violent charges and total number of prior charges).

For the post-implementation group, the number of service referrals was significantly related to risk level. High-risk youths ($M = 2.36$, $SD = 1.10$) were referred to significantly more services than moderate- ($M = 1.87$, $SD = 0.98$) or low-risk youths [$M = 1.67$, $SD = 1.0$; $F(2, 187) = 4.72$, $p < 0.01$]. However, the number of services completed did not differ significantly across youth risk levels (high, $M = 1.78$, $SD = 1.18$; moderate, $M = 1.73$, $SD = 1.01$; low, $M = 1.80$, $SD = 0.83$).

Recidivism: New Petitions and New Adjudications

Overall, the rates for any type of new petition were 42% for the pre-implementation group and 46% for the post-implementation group. Chi-squared results indicated there were no significant differences between these groups in the rates of violent (marginal means = 20% vs. 20%, respectively), non-violent (27% vs. 34%) or probation violation (10% vs. 15%) reoffenses. Similarly, separate Cox regression models conducted for each type of new petition indicated there were no significant differences between youths in the pre- and post-implementation groups in their time to the commission of a new petition. Results of each regression were as follows: any new petition [covariate was number of prior charges, $\beta = 0.18$, $SE = 0.15$, $\text{Exp}(B) = 1.20$, $p = 0.224$], new violent petitions [covariates were number of prior charges, number of prior violent charges, and whether the current offense was a non-violent felony, $\beta = 0.026$, $SE = 0.22$, $\text{Exp}(B) = 1.03$, $p = 0.905$]; new non-violent petitions [no covariates, $\beta = 0.312$, $SE = 0.181$, $\text{Exp}(B) = 1.37$, $p = 0.084$]; and new probation violations [no covariates, $\beta = 0.42$, $SE = 0.029$, $\text{Exp}(B) = 1.53$, $p < 0.141$].

In the post-implementation sample, ROC curves indicated the SAVRY SRR was significantly related to receiving any new petition and violent petitions but not to non-violent petitions or probation violations (see Table 4). AUC values for the SRR ranged from 0.57 to 0.67. Due to the lack of the SRR's effect for predicting non-violent reoffending, the predictive validity of SAVRY total scores was also examined to determine if score-based decisions may have led to a different result. Indeed, total scores significantly predicted all types of reoffending (see Table 4). AUC values for total scores were somewhat higher than the SRR, ranging from 0.62 to 0.73. ROC curves also found that the sum of the protective factors significantly predicted violent and non-violent petitions, and was significantly related to all types of reoffending.⁹ Youths with more protective factors were less likely to reoffend than those with fewer factors (see Table 5).

Conversely, for new adjudications, there were significant increases in the post-implementation group in the any (28% vs. 37%, respectively), non-violent (15% vs. 23%), and violation (5% vs. 11%) categories. The post-implementation youths were adjudicated earlier on any new adjudication [covariates included number of prior charges, $\beta = 0.38$, $SE = 0.17$, $\text{Exp}(B) = 1.47$, $p = 0.03$; $\chi^2 = 18.29[4]$, $p = 0.000$], non-violent adjudications [no covariates, $\beta = 0.508$, $SE = 0.234$, $\text{Exp}(B) = 1.66$, $p = 0.03$; $\chi^2 = 15.487[4]$, $p = 0.004$] and probation violations [no covariates, $\beta = 0.762$, $SE = 0.367$,

⁹ The researchers also conducted a separate set of hierarchical Cox regressions to explore whether the SRR had incremental predictive validity beyond SAVRY total scores for predicting reoffending. The SRR did not add a significant prediction beyond total scores for any of the reoffending dependent variables.

Table 4. Prediction of recidivism by SAVRY summary risk rating (SRR) and total score ($n = 195$)

	Low [% (n)]	Moderate [% (n)]	High [% (n)]	SRR AUC (CI)	SAVRY Total AUC (CI)
New petitions					
Any petition ($n = 90$)	36.1 (26)	47.4 (45)	67.9 (19)	0.61* (0.53–0.69)	0.66*** (0.59–0.74)
Violent ($n = 35$)	8.3 (6)	18.9 (18)	39.3 (11)	0.67** (0.57–0.77)	0.73*** (0.64–0.82)
Non-violent ($n = 68$)	29.2 (21)	34.7 (33)	50.0 (14)	0.57 (0.49–0.66)	0.62** (0.53–0.70)
Violation ($n = 28$)	8.3 (6)	16.8 (16)	21.4 (6)	0.60 (0.50–0.71)	0.67** (0.57–0.77)
New adjudications					
Any petition ($n = 72$)	23.6 (17)	40.0 (38)	60.7 (17)	0.64** (0.56–0.72)	0.71*** (0.63–0.78)
Violent ($n = 27$)	5.6 (4)	14.7 (14)	32.1 (9)	0.68** (0.57–0.79)	0.73*** (0.63–0.83)
Non-violent ($n = 48$)	16.7 (12)	27.4 (26)	35.7 (10)	0.60* (0.50–0.69)	0.65** (0.57–0.73)
Violation ($n = 21$)	6.9 (5)	10.5 (10)	21.4 (6)	0.61 (0.48–0.75)	0.65* (0.52–0.77)

AUC, area under the curve; SAVRY, Structured Assessment of Violence Risk in Youth.

Note. Cells represent the percentage of youth within each risk category who reoffended. ** $p < 0.05$, *** $p < 0.001$.

Table 5. Differences in SAVRY protective factors [M (SD)] between recidivators and non-recidivators

	Recidivators ($n = 89$)	Non-recidivators ($n = 105$)	t -test $t(192)$	Cohen's d	AUCs (CI)
New petitions					
Any	3.29 (1.97)	4.03 (1.99)	2.76*	0.37	0.61** (0.53–0.69)
Violent	2.83 (2.23)	3.88 (1.92)	2.85**	0.50	0.64* (0.53–0.74)
Non-violent	3.24 (2.02)	3.93 (1.98)	2.30*	0.34	0.60* (0.52–0.68)
Violation	2.93 (2.14)	3.82 (1.97)	2.19*	0.43	0.62* (0.50–0.73)
New adjudications					
Any	3.08 (1.93)	4.05 (1.98)	3.31**	0.50	0.64** (0.56–0.72)
Violent	2.96 (2.16)	3.81 (1.97)	2.04*	0.41	0.61 (0.49–0.73)
Non-violent	3.00 (1.96)	3.92 (1.98)	2.79**	0.47	0.63** (0.51–0.74)
Violation	2.95 (1.88)	3.78 (2.01)	1.79	0.43	0.63** (0.51–0.74)

AUC, area under the curve; SAVRY, Structured Assessment of Violence Risk in Youth.

Note. The sum of the protective factors was multiplied by -1 in order to report AUCs and Cohen's d in a direction consistent with the SAVRY total scores. Thus, AUCs > 0.50 indicate that the more protective factors a youth had, the less likely they were to reoffend.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

$\text{Exp}(B) = 2.14$, $p = 0.04$; $\chi^2 = 4.535[1]$, $p = 0.03$. In the post-implementation sample, ROC curves demonstrated that any, violent and non-violent adjudications were significantly predicted by the SRR risk level (see Table 4). AUC values for the SRR for new adjudications ranged from 0.60 to 0.68. As was the case for new petitions, AUCs for new adjudications were higher for SAVRY total scores, ranging from 0.65 to 0.73. Similarly, youths with more protective factors were less likely to receive new adjudications within all categories except violations (see Table 5).

Importance of Quality Implementation

Because the SAVRY was administered to only half the sample (i.e., the post-implementation group), it is unknown whether JPOs and courts were making decisions that corresponded with youths' risk for reoffending before the SAVRY was implemented. As a result of a surreptitious departure from the research protocol in which Caddo parish began to administer the SAVRY to youths before the implementation process was completed (i.e., after SAVRY training but 5 months before staff were trained on office policies about its use in

their decision-making), we were able to conduct additional analyses to examine the relation between risk level and case processing before the SAVRY was fully implemented. There were 61 youths who received a SAVRY prior to full-implementation who could be compared with youths in the whole post-implementation group.

Placement decisions were not related to risk for youths who were assessed with the SAVRY prior to full implementation (% of low-risk youth placed = 23.8%, Moderate = 26.9%, and High = 41.7%; $V = 0.15$, $p = 0.53$), but there was a significant association after full implementation (low = 11.1%, moderate = 21.2%, and high = 31.4%; $V = 0.18$, $p = 0.03$). It seems implementation was associated with a lower percentage of both low- and high-risk youth being placed.¹⁰ Similarly for services, prior to full implementation, there was not a significant difference between low- ($M = 1.32$, $SD = 0.68$) and high-risk youths ($M = 1.09$, $SD = 1.0$) in service referrals received, on average. With respect to supervision level, prior to full implementation, 42% of cases on intensive supervision, 25.9% of youths on maximum supervision, and 20% of youths on moderate supervision were low risk. This compares with no low-risk youths on intensive supervision, 8.8% on maximum supervision, and 5% on moderate supervision after full implementation.

DISCUSSION

This quasi-experimental study used a pre-post, prospective design to answer many questions related to whether implementing risk assessment at an early decision point in juvenile justice effectuates change in the way youths are handled. The results were very promising. Implementation of the SAVRY at the post-adjudication/pre-disposition stage in this probation office led to: a reduction in non-secure and secure placement rates by approximately 50%; a reduction in the use of maximum and intensive levels of supervision by almost 30%; and a shift in resource allocation of services such that the total number of services being used decreased, but the number provided to high-risk youths increased. Altogether, the results suggest a potential significant savings in resources in a manner that was cost-effective and involved evidence-based practices for the handling of young offenders. Further, these changes were achieved without an increase in rates of new petitions.

On the other hand, there was a significant increase in the proportion of youths who were adjudicated. Discussions with administrators in this office indicated that this was likely related to a change in policies followed by the prosecutor's office. A new district attorney (DA) was appointed during the study period, which led to an increase in youth adjudications for about 1 year as per the DA's policy. This is a prime example as to why it is essential for researchers to maintain feedback and discussions with study sites when conducting implementation research in the field in order to avoid misinterpreting findings. Our results also reinforce the value of evaluating reoffending using multiple operationalizations of recidivism, because petition or arrest data and adjudication data

¹⁰ The rates were compared for placements immediately following disposition only. By the time these youths had received any subsequent placements, the implementation process would have been complete, so it would not be an accurate comparison. There is not a strong difference between youths handled prior to full implementation and those handled afterward, given the association indices were similar. The lack of significance in the prior to full implementation group may have been due to having less power.

can be affected by systemic issues that are unrelated to youths' actual behavior.¹¹ There were no other major changes in this office or jurisdiction during the course of the study. The number of staff, administrators, and caseloads stayed consistent.

The one area of youth processing where the SAVRY seemed to have little impact was on the type of dispositions given. Although there was a significant shift towards use of less restrictive dispositions, aside from commitment to the state, which is a very meaningful change, most decisions were not strongly related to risk level. There are two explanations for this. First, the decision to send a youth to detention (which often occurred post-adjudication/pre-disposition with the disposition of "detention time-served") was more strongly related to the youth's offense history and the judge who decided the case than it was to risk level. Secondly, these decisions were generally made quickly without a pre-disposition investigation or SAVRY in place, so one could not expect the SAVRY risk level to have any impact.

Implications: Does Implementation of Risk Assessment in Juvenile Probation Make a Difference?

This study provides evidence of several positive outcomes when an evidence-based risk assessment tool was implemented to inform pre- and post-dispositional planning. Implementation led to more youths being monitored in the community and fewer youths being incarcerated, suggesting that designating youths as high-risk did not condemn them to incarceration. Moreover, these decisions were related to risk level in that few low-risk youths were given intensive services or incarcerated. In addition, less than half of the high-risk youths were incarcerated. However, a take-home message is that comprehensive implementation practices are essential for the benefits of risk assessment to be utilized. Too often, a risk assessment tool is adopted by justice agencies, completed with youth by staff, and then not actually used for decision-making (Shook & Saari, 2007). As was done in this study, the key is implementation of an assessment system, which requires adopting a comprehensive approach for the use of risk assessment tools in youth processing decisions, ensuring it is stated in the policy, and ensuring the staff and stakeholders know its value (Bonta *et al.*, 2001; Ferguson, 2002). Results from the sample of youth who were administered the SAVRY prior to its full implementation offer some tentative evidence for the notion that merely training staff on the SAVRY did not promote decisions in adherence with RNR principles. Risk-related decision-making did not occur fully until office policies were in place, new case management forms were implemented, and staff members were trained on use of the SAVRY in their decisions.

So why did reoffending not decrease after the SAVRY was implemented as others have found (Bonta *et al.*, 2011; Luong & Wormith, 2011)? There are a few potential explanations for this. First, this study tracked reoffending while many youths were still on probation, at least for part of the follow-up period, so youths were receiving some supervision, which may have led to higher detection rates of violations and other delinquent acts. Secondly, previous studies that have shown a decrease in recidivism following the use of a risk assessment tool did not compare overall reoffense rates before and after a tool was put in place. Instead, prior studies have shown reductions

¹¹ Of course, reliance only on official sources of reoffending data may under-represent the recidivism rate; this can be overcome by the use of collateral sources of outcome data (e.g., interviews with family members).

in reoffense rates for subgroups of probation officers that showed the best adherence to RNR-based case management (Bonta et al., 2011), or for subgroups of youths. Specifically, Luong and Wormith (2011) found a 37.9% reduction in recidivism for high-risk offenders only, and Vieira et al. (2009) reported low recidivism rates for youths who received services that were strongly matched to their criminogenic needs. Examination of subgroups of youths for whom recidivism may have decreased was not a focus of this study. Instead, we studied whether implementing risk assessment made a difference on the whole. We encourage other researchers to examine reoffense rates using a pre–post or randomized control design.

Other contributions of this research are the validation of the SAVRY and use of the structured professional judgment approach by JPOs in the field, and the demonstration of the contribution of protective factors to assessing risk for violence. Not only did the SAVRY maintain adequate reliability in the field (see Vincent et al., 2012), but the SRR of JPOs also predicted violent and any reoffending. The fact that the SRR did not predict non-violent reoffending may be a limitation. However, this also could be the result of an implementation issue. The probation department requested that the SAVRY trainer (a co-author of the tool) and researchers create a separate SRR for general delinquent reoffending risk because they were concerned that youths who may be chronic delinquents would be labeled as being at high risk for violence. It is possible that use of this separate rating diluted the predictive accuracy of the SRR for non-violent offending because the JPOs would have seen the two ratings as intended to predict different outcomes. Unfortunately, due to delays in incorporating this rating into the data management system, we did not receive enough of the separate general delinquency reoffending risk ratings for a sufficient sample ($n = 35$) to examine the predictive validity. These unconventional implementation procedures were another reason for examining the predictive validity of the SAVRY total risk score for non-violent offending, even though the tools' authors do not recommend use of total scores in practice. These findings were significant.

Despite the significant association between the SRR and reoffending in this study, results also suggest there is reason to be cautious. AUC values for the SRR were lower than values for the total score, which is using the SAVRY in an actuarial fashion, across all types of reoffending. This is in contrast to the pattern of results found in the more general literature on structured professional judgment tools (Guy, 2008). In a meta-analysis of all available studies, both published and unpublished, that reported on the predictive validity of an SPJ tool, Guy reported that the mean AUC values for any type of recidivism for the SRR was 0.69 (based on 32 effect sizes) and for the total scores of SPJ tools was 0.68 (based on 93 effect sizes). In addition, in the scientific literature the SRR typically has been found to add incremental predictive accuracy to the total score (see Heilbrun, Yasuhara, & Shah, 2010). It may be that JPOs, compared with mental health professionals who have tended to be evaluators in most SPJ studies, were not as adept at the clinical formulation process required to make the SRR in the present study, despite participating in baseline and booster SAVRY training. Making an accurate estimate of whether a youth is at low, moderate, or high risk for reoffending requires the evaluator to consider not only the presence of risk and protective factors (which is what the total score represents), but also the relevance of each item for the particular youth being evaluated. Case formulation is an understudied area in the violence risk assessment field, although some initial work has been conducted (see Hart & Logan, 2010). Thus, this study is not intended to imply that the SPJ approach is superior

to the actuarial approach of violence risk assessment. Instead, our findings suggest that the SPJ approach can be used somewhat effectively by probation officers.

Limitations and Next Steps

It is possible that the generalizability of findings from this probation office is limited. Caddo parish juvenile probation had substantial buy-in from stakeholders for the use of a risk assessment tool because they were funded specifically to incorporate best practices into their pre-dispositional assessment methods to improve case processing and management. At the time, Caddo parish probation had one of the higher rates of commitment to the state juvenile correctional facility in Louisiana. Thus, it is possible that the striking reduction in placement rates was an artifact of a high starting point. Further, the integrity of the implementation procedures at this site was strong, as demonstrated by interviews with staff (Vincent *et al.*, 2012). Thus, other probation departments should not expect to achieve the same results if they have less buy-in, historically different case management practices, or limited implementation activities. A crucial next step is to study the generalizability of these findings using the other sites.

Other limitations included the dependence on file and administrative data, and tracking reoffending while many youths were still on probation. It is unclear that the latter point was a significant limitation given that the rates of new petitions were still over 40%. Nonetheless, a longer follow-up would be ideal. In addition, the study design and follow-up period did not enable us to examine whether use of risk assessment led to shorter periods of time on probation because 75 youths in the post-implementation group were still on probation, compared with 26 youths in the pre-implementation group. This did not lend itself to an accurate comparison.

Another significant limitation of this study is that we do not know if decisions were based on youths' risk for reoffending prior to the SAVRY being implemented because there was no proxy for measuring risk in the pre-implementation group. One might infer that decisions improved after the SAVRY was implemented because there were significant changes in the expected direction. Further, the comparisons between youths who received a SAVRY prior to full implementation and those who received a SAVRY afterward lend some support to the notion that decisions were more strongly related to risk level after SAVRY implementation, but these conclusions are tentative. The ideal study design would be to randomly assign youths within the same probation office to receive or not receive a risk assessment, but this tends to be impractical (and a violation of best practices). The next best design would be to compare the Caddo parish office to a control office that did not implement risk assessment. Unfortunately, the researchers could not identify a probation office in the state that was willing to delay implementation of the SAVRY for that long. Nonetheless, there are benefits to use of a pre-post design with a single jurisdiction because it does not require consideration of jurisdiction-specific controls that would be necessary in a control-group design.

In addition to the need for replication and multi-site studies, other important directions with these data and in future research are to examine decision-making and reoffending for subgroups of youths. For example, the next focus of this project will be to compare reoffense rates of youths who received services strongly aligned with their criminogenic needs as identified by the SAVRY with those whose needs and services were poorly matched. Another important step for future research using pre-post

designs is to compare the case processing of JPOs who shifted their case management practices with JPOs who did not.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Dr. Debra DePrato, M.D., Associate Clinical Professor of Public Health, and Laura Alderman, Ph.D., at Caddo parish juvenile probation for making this study possible. We also offer a special thank you to Joshua Everett, M.A., for his data collection.

REFERENCES

- Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* 4th ed. Newark, NJ: LexisNexis/Matthew Bender.
- Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, 16, 39–55. DOI: 10.1037/a0018362
- Andrews, D. A., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17(1), 19–52. DOI: 10.1177/0093854890017001004
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2001). *Level of Service Inventory–Saskatchewan Youth Edition*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., & Dowden, C. (2006). Risk principle of case classification in correctional treatment: A meta-analytic investigation. *International Journal of Offender Therapy and Comparative Criminology*, 50(1), 88–100. DOI: 10.1177/0306624X05282556
- Bonta, J., Bogue, B., Crowley, M., & Mottuk, L. (2001). Implementing offender classification systems: Lessons learned. In G. A. Bernfeld, D. P. Farrington, & A. Leschied (Eds.), *Offender rehabilitation in practice* (pp. 227–245). Chichester, UK: Wiley.
- Bonta, J., Bourgon, G., Ruggie, T., Scott, T., Yessine, A. K., Gutierrez, L. & Li, J. (2011). An experimental demonstration of training probation officers in evidence-based community supervision. *Criminal Justice and Behavior*, 38, 1127–1148.
- Borum, R., Bartel, P., & Forth, A. (2006). *Structured Assessment of Violence Risk in Youth (SAVRY)*. Florida: Psychological Assessment Resources, Inc.
- Borum, R., Lodewijks, H., Bartel, P. A., & Forth, A. E. (2010). Structured Assessment of Violence Risk in Youth (SAVRY). In R. K. Otto & K. S. Douglas (Eds.), *Handbook of violence risk assessment* (pp. 63–79). New York: Routledge/Taylor & Francis Group.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Douglas, K. S., Yeomans, M., & Boer, D. P. (2005). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior*, 32(5), 479–510. DOI: 10.1177/0093854805278411
- Dowden, C., & Andrews, D. A. (1999). What works for female offenders: a meta-analytic review. *Crime and Delinquency*, 45(4), 438–452. DOI: 10.1177/0011128799045004002
- Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and reoffending: A meta-analysis. *Canadian Journal of Criminology*, 42(4), 449–467. DOI: 10.1177/0306624X03257765
- Dowden, C., & Andrews, D. A. (2004). The importance of staff practice in delivering effective correctional treatment: A meta-analytic review of core correctional practice. *International Journal of Offender Therapy and Comparative Criminology*, 48, 203–214.
- Ferguson, J. L. (2002). Putting the “what works” research into practice: An organizational perspective. *Criminal Justice and Behavior*, 29(4), 472–492. DOI: 10.1177/0093854802029004007
- Flores, A. W., Lowenkamp, C. T., Holsinger, A. M., & Latessa, E. J. (2006). Predicting outcome with the Level of Service Inventory-Revised: The importance of implementation integrity. *Journal of Criminal Justice*, 34(5), 523–529. DOI: 10.1016/j.jcrimjus.2006.09.007
- Gatti, U., Tremblay, R. E., & Vitaro, F. (2009). Iatrogenic effect of juvenile justice. *Journal of Child Psychology and Psychiatry*, 50, 991–998. DOI: 10.1111/j.1469-7610.2008.02057
- Gendreau, P., Smith, P., & French, S. A. (2006). The theory of effective correctional intervention: Empirical status and future directions. In F. T. Cullen, J. P. Wright, & K. R. Blevins (Eds.), *Taking stock: The status of criminological theory: Vol. 15. Advances in criminological theory* (pp. 419–446). New Brunswick, NJ: Transaction.

- Guy, L. S. (2008). Performance indicators of the Structured Professional Judgment approach to assessing risk for violence to others: A meta-analytic survey. Unpublished doctoral dissertation, Simon Fraser University, Vancouver, BC.
- Haas, S. M., & DeTardo-Bora, K. A. (2009, Spring). Inmate reentry and the utility of the LSI-R in case planning. *Corrections Compendium*, 11–16, 49–52.
- Hart, S. D., & Logan, C. (2010). Formulation of Violence Risk Using Evidence-Based Assessments: The Structured Professional Judgment Approach.
- Heilbrun, K., Yasuhara, K., & Shah, S. (2010). Violence risk assessment tools: Overview and critical analysis. In R. K. Otto & K. S. Douglas (Eds.) *Handbook of Violence Risk Assessment* (pp. 1–18). New York, NY: Routledge.
- Hoge, R. D., & Andrews, D. A. (2006). *Youth Level of Service/Case Management Inventory: User's manual*. North Tonawanda, NY: Multi-Health Systems.
- Juvenile Justice Delinquency and Prevention Act (JJDP). (2002). Public Law 107–273, 42 U.S.C. § 5601 *et seq.*
- Latessa, E. J., & Lovins, B. (2010). The role of offender risk assessment: A policy maker guide. *Victims & Offenders*, 5, 203–219. DOI: 10.1080/15564886.2010.485900
- Leuven, E. & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. <http://ideas.repec.org/c/boc/bocode/s432001.html>. version 4.0.4 10nov2010
- Lipsy, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims & Offenders*, 4, 124–147. DOI: 10.1080/15564880802612573
- Lipsy, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science*, 3, 297–320. DOI: 10.1146/annurev.lawsocsci.3.081806.112833
- Lodewijks, H. P. B., Doreleijers, T. A. H., & de Ruiter, C. (2008). SAVRY risk assessment in violent Dutch adolescents: Relation to sentencing and recidivism. *Criminal Justice and Behavior*, 35(6), 696–709. DOI: 10.1177/0093854808316146
- Loughran, T. A., Mulvey, E. P., Schubert, C. A., Fagan, J., Piquero, A. R., Losoya, S. H. (2009). Estimating a dose–response relationship between length of stay and future recidivism in serious juvenile offenders. *Criminology: An Interdisciplinary Journal*, 47(3), 699–740. DOI: 10.1111/j.1745-9125.2009.00165.x
- Lowenkamp, C. T. & Latessa, E. J. (2005). Increasing the effectiveness of correctional programming through the risk principle: Identifying offenders for residential placement. *Criminology and Public Policy*, 4(2), 263–290. DOI: 10.1111/j.1745-9133.2005.00021.x
- Luong, D. & Wormith, J. S. (2011). Applying risk/need assessment to probation practice and its impact on the recidivism of young offenders. *Criminal Justice and Behavior*, 38(12), 1177–1199. DOI: 10.1177/0093854811421596
- MacKenzie, D. L., Wilson, D. B., & Kider, S. B. (2001). Effects of correctional boot camps on offending. *The Annals of the American Academy of Political and Social Science*, 578, 126–143.
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior*, 36(4), 329–353. DOI: 10.1177/0093854809331457
- Petrosino, A., Turpin-Petrosino, C., & Guckenburg, S. (2010). Formal system processing of juveniles: Effects on delinquency. *Campbell Systematic Reviews*, 2010(1), 1–88. DOI: 10.4073/csr.2010.1
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55. DOI: 10.1093/biomet/70.1.41
- Seave, P. L. (2011). Evidence-based practices reduce juvenile recidivism: Can state government effectively promote implementation among probation departments?. *American Journal of Community Psychology*, 48, 138–140.
- Shook, J. L. & Saari, R. C. (2007). Structured decision making in juvenile justice: Judges' and probation officers' perceptions and use. *Children and Youth Services Review*, 29(10), 1335–1351. DOI: 10.1016/j.childyouth.2007.05.008
- StataCorp. (2007). Stata statistical software Release 10.1. College Station, TX: StataCorp LP.
- Vieira, T. A., Skilling, T. A., & Peterson-Badali, M. (2009). Matching court-ordered services with treatment needs. *Criminal Justice and Behavior*, 36(4), 385–401. DOI: 10.1177/0093854808331249
- Vincent, G. M., Chapman, J., & Cook, N. E. (2011). Risk/Needs assessment in juvenile justice: Predictive validity of the SAVRY, racial differences, and contribution of needs factors. *Criminal Justice and Behavior*, 38(1), 42–62. DOI: 10.1177/0093854810386000
- Vincent, G. M., Guy, L. S., Fusco, S. L., & Gershenson, B. G. (2012). Field reliability of the SAVRY with probation officers: Implications for training. *Law and Human Behavior*, 36, 225–236. DOI: 10.1007/s10979-011-9284-21
- Vincent, G. M., Paiva-Salisbury, M. L., Cook, N. E., Guy, L. S., & Perrault, R. T. (2012). Impact of Risk/Needs Assessment on Juvenile Probation Officers' Decision Making: Importance of Implementation. *Psychology, Public Policy, and Law*. Advance online publication. DOI: 10.1037/a0027186

- Vincent, G. M., Terry, A., & Maney, S. (2009). Risk/Needs tools for antisocial behavior and violence among youthful populations. In J. Andrade (Ed.) *Handbook of Violence Risk Assessment and Treatment for Forensic Mental Health Practitioners* (pp. 337–424). New York: Springer.
- Welsh, J. L., Schmidt, F., McKinnon, L., Chattha, H. K., & Meyers, J. R. (2008). A comparative study of adolescent risk assessment instruments: Predictive and incremental validity. *Assessment, 15*(1), 104–115. DOI: 10.1177/1073191107307966
- Young, D., Moline, K., Farrell, J., & Biere, D. (2006). Best implementation practices: Disseminating new assessment technologies in a juvenile justice agency. *Crime & Delinquency, 52*(1), 135–158. DOI: 10.1177/0011128705281752