The Evaluation of the Streamlined CA-YASI

Kristy N. Matsuda, Laceé N. Pappas, and Susan F. Turner
Center for Evidence-Based Corrections
University of California, Irvine
October 12, 2022

Executive Summary

In 2016, the Division of Juvenile Justice (DJJ) and Orbis Partners Inc., streamlined the risks and needs assessment instrument - the California Youth Assessment Screening Inventory (CA-YASI). The CA-YASI instrument was specially designed to determine the risks and needs of California's most serious youthful offenders. The purpose of this study is to evaluate the streamlined CA-YASI for its practical use in DJJ. This study answers the most basic question: is there any change in risk over time? Then, it examines the streamlined CA-YASI's ability to predict serious offending both in DJJ and upon release to the community. Specifically, this study tests the following research questions:

1) Does risk (as measured by the CA-YASI) change over time while in DJJ?
2) Does CA-YASI predict recidivism (i.e., rearrest or reconviction) post-release from DJJ?
3) Does CA-YASI predict involvement in serious disciplinary incidents while incarcerated in DJJ?

This study sampled virtually all of the male youth residing in DJJ in 2017 (n=628, 98.7%) and released by January 1, 2020. However, each of the three research questions required a unique analysis, and therefore, resulted in usable samples of various sizes. This study utilizes administrative data provided by DJJ, CA-YASI risk assessment data provided by Orbis, and recidivism data provided by the California Department of Justice (DOJ).

Our first research question examined whether level of risk changes over time in this serious youthful offending population. This question stemmed from previous research that questioned whether staff used the tool to score DJJ youth in way that would actually capture change over time (e.g., Matsuda, Hess, & Turner, 2020). Due to maturation and programming, we would expect youth to reduce their risk over time, if, the tool is valid and can be used reliably by staff. The results show that the level of risk does decline for the sample over time. The decline is not dramatic, but it is statistically significant. The results also suggest that the reduction in risk over time may have a limit. Results show a 0.20 unit decrease in risk until Year 4. For youth who are still incarcerated in DJJ at Years 5 and 6 their risk begins to increase again. The average length of stay in DJJ is approximately 2.5 years. Youth that remain in DJJ for these longer terms (i.e., over 4 years) could be higher risk youth. These youth would remain in the sample as the lower risk youth get released. This trend could be a function of the significant reduction in sample size at the later years of incarceration which is more sensitive to extreme observations. It could also be that relation between risk and incarceration is not a simple linear function. The truth could also be a mix of these possible explanations.

This study shows that the CA-YASI significantly predicts rearrest and reconviction in youth who had two years in the community during the follow-up period. We utilized the last CA-YASI risk assessment level on record for each youth in the sample. Results show that youth who exited DJJ with a higher level of risk were more likely to be rearrested and reconvicted upon release. These results support the notion that CA-YASI can be used as a risk prediction instrument for recidivism in a population of serious youthful offenders.

Our final research question examined whether the tool can be used to predict involvement in serious incidents while in DJJ. We measured "serious incidents" by collecting and coding youth involvement as recorded in DJJ's serious incident reports (SIRs). Serious incident reports were available covering the years 2015 to 2020. Unfortunately, we found that a significant portion of our sample (i.e., 120 youth) was incarcerated prior to 2015 and had to be excluded from this analysis. We conducted a series of comparisons of our usable sample to the excluded sample to determine whether the results could be generalized to the excluded group. We found that the excluded sample was, in general, higher risk and more serious offenders than the usable sample. However, the trend in their behavior was the same as the larger, usable group. Thus, our findings suggest the results of the higher risk youth in the sample could likely be generalized to this excluded group. We conducted a path analysis to understand how well the risk level at intake predicted the involvement in SIRs in the first year. Then, the model examines how well the CA-YASI risk at the start of the second year predicts involvement in SIRs in the second year, and so on, until the end of the commitment. The analysis for this question shows that CA-YASI does predict involvement in serious disciplinary incidents while in DJJ. The results also show that as risk declines over time, so does involvement in serious incidents.

Findings from this study support the use of the CA-YASI as a risk assessment instrument for DJJ's serious youthful offender population. It can be used both to predict risk of offending while in DJJ and upon release. Future studies, however, should examine how well the tool functions for case-planning purposes. This study also does not address whether the tool is effective in sub-groups of offenders. Specifically, this study only included males. Future studies should examine the tool's effectiveness in a female offending population or in youth with specific treatment needs. Finally, this study did not compare CA-YASI to any other risk assessment instrument. While CA-YASI is a predictive tool, it is not yet clear if it is *the most* effective predictive tool.

Contents

Executive Summary i
Contents iii
List of Tables iv
List of Figures v
Introduction 1
Risk and Needs Assessment in the Division of Juvenile Justice 1
The History 2
Orbis Partner Inc.'s YASI Instrument 3
The Original California YASI (version 2008) 4
Evaluation of the Original CA-YASI (version 2008) 4
Reliability 4
Construct Validity 5
Predictive Validity 6
Recommendations from the CA-YASI (version 2008) Evaluation 7
Streamlined CA-YASI (2016) 7
Previous Research on Streamlined CA-YASI (2016) 8
The Current Evaluation 10
Methods 11
Sample Selection 11
Sample Characteristics 13
Data & Measures 13
Independent Measure 13
Dependent Measures 15
Recidivism 15
Serious Incident Reports (SIRs) 15
Covariates 16
Race/Ethnicity 16
Sex Offender Status 16
Analytical Strategy 16
Results 18
Question 1: Does risk (as measured by the CA-YASI) change over time while in DJJ? 18
Question 2: Does CA-YASI predict recidivism (i.e., rearrest and reconviction) post-release from DJJ? 21
Question 3: How well does CA-YASI predict disciplinary problems while incarcerated in DJJ? 22
Comparisons of Missing Sample 24
Limitations and Future Research 27
References 28

List of Tables

Table 1. Study sample characteristics (n=615). 12
Table 2. Proportion of DJJ youth by risk level at admission and discharge (n=615). 18
Table 3. Average risk score by covariates at the first and last CA-YASI in DJJ. 21
Table 4. Logistic regression results predicting recidivism outcomes. 22
Table 5. Descriptive comparisons between the RQ3 Sample and the Excluded Youth. 24
Table 6. Comparison of recidivism results for usable (post-2015) and excluded (pre-2015) samples. 26

List of Figures

Figure 1. Average change in risk level over time. 19
Figure 2. Standardized direct and indirect effects of risk predicting yearly rate of involvement in serious incidents. 23
Figure 3. Change in risk level over time for pre-2015 and post-2015 groups..... 25

Introduction

For almost two decades, the California Department of Corrections and Rehabilitation's (CDCR) Division of Juvenile Justice (DJJ), the state's correctional system for youthful offenders, has been reforming its philosophy, policies, and practices. These reforms come in the wake of the consent decree in the *Farrell*[1] lawsuit, which required DJJ to implement evidence-based policies and practices designed to help treat the youthful offenders in their care. As part of their commitment to evidence-based practices, DJJ has relied on outside evaluations to help them assess the effectiveness of their reforms. Recent evaluations include studies on their treatment model (Matsuda, Hess, & Turner, 2020), data systems (Matsuda & Turner, 2020), gang policies (Maxson, et al., 2012), and risk assessment instruments (e.g., Skeem et al., 2013).

The purpose of the current study is to evaluate the effectiveness and utility of DJJ's primary risk assessment instrument, the California Youth Assessment Screening Instrument (CA-YASI). This is not the first time that CA-YASI has been evaluated for use in DJJ. The reliability, construct validity, and predictive utility of the tool have been evaluated in a previous comprehensive study (Kennealy, Skeem, & Hernandez, 2017; Skeem, Hernendez, Kennealy, & Rich, 2011; Skeem, Kennealy, & Hernandez, 2013; Skeem, Kennealy, Hernandez, Clark & Tatar, 2013; Skeem, Kennealy, Tartar, Hernandez, & Keith, 2017). Since that study, however, DJJ and Orbis Partners Incorporated (hereafter referred to as "Orbis"), the designer of the CA-YASI, have streamlined the tool. The streamlined version of the CA-YASI has not yet been evaluated in independent research.

Risk and Needs Assessment in the Division of Juvenile Justice

DJJ's Integrated Behavior Treatment Model (IBTM) describes and details the Division's theoretical treatment framework (for an evaluation of IBTM see Matsuda, Hess, & Turner, 2020). The Risk Needs and Responsivity (RNR) Model (Andrews & Bonta, 2010) is a core component of IBTM. The RNR model is one of the most widely accepted frameworks for providing effective treatment to offenders (Howell, et al., 2019). According to RNR, the best way to determine the risks and needs of an offender is by utilizing a structured risk assessment instrument (Bonta & Andrews, 2007). A risk and needs assessment instrument (hereby referred to as "risk assessment") is an actuarial tool that seeks to determine offenders' risk for unwanted behaviors (e.g., recidivism, violence, drug use). Once classified, individuals can be targeted with different sanctions or services commensurate with the level of risk they pose. Risk assessment instruments can be used to make treatment and placement decisions, create treatment plans, reduce disparity in decision making, predict recidivism, assist in the determination of successful outcomes, or identify gaps in service (Howell, et al., 2019; Shook & Sarri, 2007). Research shows that actuarial tools consistently outperform

[1] The *Farrell v. Harper* lawsuit (No. RG 03079344, Cal. Sup. Ct., Alameda Cnty) was originally filed on January 6, 2003. The name of the case defendant changes to reflect the current head of DJJ at the time. Thus, for ease, the case is just referred to as *Farrell.* For a comprehensive timeline and additional information on the *Farrell* suit, see materials compiled by the California Department of Corrections and Rehabilitation (www.cdcr.ca.gov) and the Center on Juvenile and Criminal Justice (www.cjcj.org).

clinical judgments (e.g., Grove, Eckert, Heston, Bouchard, Segal, & Lykken, 2000). While actuarial tools may not eliminate all types of influence, they can significantly reduce decision making bias (Heilbrun, 1997).

Risk assessments come in different forms. The earliest risk assessment tools were primarily "static." Static risk assessments use characteristics that do not change over time (e.g., criminal history, demographic information) for prediction and assessment. The more recent instruments are often both static and "dynamic." Dynamic factors are features that can change over time (e.g., drug use, criminal thinking, impulsivity). While research generally shows that static factors are the strongest predictors of future offending (Loeber, Slott, & Stouthamer-Loeber, 2008), these factors cannot be changed, so it is difficult to target them for intervention. Dynamic factors, by their nature, do change, and therefore can be intervened upon with treatment or intervention. Thus, risk assessment tools that include dynamic factors can be used to create individual treatment plans (Heilburn, 1997).

The History

In the mid 2000's while under the *Farrell* consent decree, risk assessment became one aspect of DJJ's extensive reform effort. The consent decree included the requirement that DJJ classify youth based on security risk and treatment needs. Prior to *Farrell,* DJJ relied on comprehensive interviews conducted by staff in the reception centers to determine institutional placements and treatment planning (Krisberg, 2003). A security classification system was not generally used in DJJ, except for determining which youth could serve in forestry camps with minimal risk. However, even this classification system had not been validated. Given the empirical research on actuarial risk assessments' accuracy over clinical judgement, the recommendation by the subject-matter expert in *Farrell* was that DJJ adopt standardized criteria for determining effective custody classification (Krisberg, 2003). Reports of the special master associated with the *Farrell* consent decree indicate that initially, DJJ was quite hesitant to fully implement a classification system as recommended and designed by subject-matter experts (Brorby, 2006).

Despite initial hesitations, in their 2006 *Safety and Welfare Plan,* DJJ had enumerated the purpose and plan for their preliminary risk assessment strategy. There were numerous risk assessment instruments available for use with a youth offender population (e.g., see Baird, et al., 2013 or Wachter, 2015). The report does not detail how DJJ reformers and experts decided which risk assessment strategy to utilize. The 2006 plan indicates that Washington State's Juvenile Rehabilitation Administration's Integrated Treatment Model ("JRA Model") was a consideration (JRA, 2002). In fact, much of DJJ's final IBTM model shares many similarities with the JRA model. However, the *Safety and Welfare Plan* also indicated that for their risks and needs assessment, DJJ was closely considering modeling their tool after the Arizona Department of Juvenile Corrections' assessment tool, which included two more domains than the JRA tool (DJJ, 2006) one of which was the domain for sexual offending for which DJJ was interested in including (DJJ, 2006). The Arizona tool was described as both static and dynamic. The tool could be used for initial assessments and reassessments, thus could measure the degree of progress and identify areas needed for targeted intervention (DJJ, 2006).

DJJ's stated purpose for adopting a risk assessment included both the desire to predict future criminal behavior and for case management planning (DJJ, 2006). According to their reports, the "primary outcome variable" of interest is to prevent future criminal behavior (DJJ, 2006, p. 38). They also, however, indicated

the need to create treatment and rehabilitation goals to develop individual case plans based on youths' strengths, needs, risks, and motivations (DJJ, 2006). In their long view, DJJ anticipated their tool being electronic, and that the organization could use the information gathered from the tool to identify their future needs for services throughout the system (i.e., beyond just individual treatment plans) (DJJ, 2006).

Orbis Partners Inc. was contracted to adapt their Youth Assessment and Screening Instrument (YASI) tool to meet the specific needs of the California DJJ population. Orbis' YASI fit a number of the desired criteria as previously mentioned. Orbis provides a dashboard for electronic data entry. YASI captures risks, needs, and strengths in a number of different domains. It includes both static and dynamic items. The YASI tool had previously been validated in probation samples (e.g., Orbis Partners, Inc., 2007a), but was not used in samples as "serious" as the DJJ population.

Orbis Partner Inc.'s YASI Instrument

The history of Orbis' Youth Assessment and Screening Instrument begins in Washington state. In 1997, the administrators of the juvenile court in Washington developed a model for effective case management for youth under community supervision. As part of the Washington State's Case Management Assessment Process (CMAP), they created a risk assessment instrument (e.g., Barnoski, 2003). The current instrument is called the Washington Association of Juvenile Court Administrators Risk Assessment Tool (WAJCA Risk Assessment Tool) or the Positive Achievement Change Tool (PACT).[2] The tool conforms with risk, need, and responsivity principles. It assesses youth on 12 domains.

Orbis designed the training and quality assurance aspects of CMAP during its development (Orbis, 2007). The company then adapted the WAJCA risk assessment tool and has been implementing it in other states (e.g., New York, Illinois, North Dakota, Vermont, Virginia, and Mississippi) since the early 2000's (Baird et al., 2013). It is now administered in over 70 agencies in North America. The 90-item YASI tool can be revised for changes needed by the client and/or intended population (Orbis, 2007). Each version should be validated or re-validated following any new changes. A recent study comparing YASI[3] (in Virginia) with other risk assessments for youth offenders found that that version of YASI had positive indicators of inter-rater reliability and validity (Baird et al., 2013).

By the time DJJ and Orbis established their partnership, the landscape of DJJ had changed considerably. In 2000, DJJ operated 11 institutions. They had over 7,500 youth in their institutions, and over 5,000 more under parole supervision (LAO, 2000). Over time, the DJJ population began to decrease and the state began closing many of their state youth institutions. In 2008, due to the mounting cost of juvenile justice due to the *Farrell* consent decree, the Little Hoover Commission proposed the state eliminate its state-run juvenile justice system in favor of county-run, regional treatment for high-risk, high-need juvenile offenders (Little Hoover Commission, 2008). The state did not close DJJ at that time, but did continue to reduce the

[2] PACT instruments have been adopted and adapted from the original version in various locations. Comparisons studies of the PACT instrument should be careful to consider any changes to the versions.

[3] This study utilized YASI data in a sample of youth on probation, in facilities, or on parole in Virginia with other risk assessments tools in other states and samples.

population. They began to only accept the most serious young violent and sexual offenders in the state. Thus, DJJ needed a validated risk assessment instrument, not for a general probation population with a wide range of risks and offenses, but for a group of the most serious young offenders from around California.[4]

The Original California YASI (version 2008)

Orbis created the California Youth Assessment Screening Instrument (CA-YASI) in 2008. This California version included 117 static and dynamic items in 12 domains (items): legal history (7), correctional history (11), violence and aggression (24), social influences (10), substance use (5), attitudes (8), social/cognitive skills (10), family (10), education and employment (15), health (7), community linkages (4), and community stability (6). DJJ staff were trained to conduct interviews with youth and complete a file review to enter the value of each indicator that they believe best describes the youth's current state. For example, the indicator "Commitment to criminal lifestyles" included the possible responses categories "Expresses total commitment to a criminal lifestyle," "Indicates the positive benefits of a criminal lifestyle," "Is neutral about the benefits of criminal/pro-social lifestyle," "Expresses a clear preference for a pro-social lifestyle," and "Expresses eagerness to disassociate self with criminal lifestyle." Youth were not asked the questions directly, nor were they asked to choose a category to match their commitment to a criminal lifestyle. Staff used their expertise to choose the category that they believed best matched that youth's commitment given their interview with the youth (or their review of the youth's file).

Responses were then entered into the Orbis dashboard, a risk, need, and strength score was calculated, and the youth were classified into one of five risk groups. The content of the CA-YASI assessment could be used to help DJJ staff determine housing placement and/or intervention assignments youth should be placed.

The original CA-YASI included a 100-page administration guide and in-person training sessions taught by Orbis professionals. The administration guide included cues for interviewers to use to help elicit useful responses by the youth. In addition, one-year post the implementation of CA-YASI, DJJ staff received one and a half days of training on CA-YASI from Orbis (Kennealy, Skeem, & Hernandez, 2017). There was, however, no requirement to demonstrate reliability after the training program (Kennealy, et al., 2017).

Evaluation of the Original CA-YASI (version 2008)

Reliability

A team of researchers conducted a comprehensive evaluation of the original CA-YASI tool. The first aspect of that evaluation focused on the reliability of DJJ staff rater's scoring of the CA-YASI tool (see Skeem, Hernandez, Kennealy, & Rich, 2011; Kennealy, Skeem & Hernandez, 2017). The study assessed the 78 staff members (92% of all eligible staff of agency) trained by Orbis on the instrument. The alpha measures the

[4] In 2020, Senate Bill 823 (SB823) realigned the care of youth offenders in California to local jurisdictions and began a phased closure of DJJ facilities. Senate Bill 92 (2021) set a final closure date of DJJ facilities as June 30, 2023.

internal agreement of the (sub)scales. The researchers found that the 12 subscales varied in their internal consistency. Generally, an alpha greater than or equal to 0.70 is acceptable, 0.60 to 0.69 is questionable, and lower than 0.60 is poor. Two of the CA-YASI scales had alphas below 0.60 (i.e., substance use and mental health), four had alphas in the "questionable" level, and six had acceptable levels. The social-cognitive skill subscales had the highest internal agreement (0.93).

The study also measured staffs' agreement to experts' assessment (i.e., the researchers) as indicated by the intraclass correlation or ICC. The ICC measures whether practitioners can consistently rate a given youth in a similar manner to the criterion when accounting for chance. An ICC of 0.60 to 0.74 is good, while 0.75 or more is excellent (Cicchetti & Sparrow, 1981). Unfortunately, for 12 percent of staff (n=9) ICCs could not be calculated due to missing items and/or scoring mistakes. The evaluation team used Kappa (individual items and not scales) with the same reliability markers to assess these individuals. Researchers could not calculate ICC for the mental health subscale due to lack of information. The study found that five of the 11 subscales (that could be measured) fell below the level of "good" rater reliability (Skeem, et al., 2011; Kennealy, et al., 2017). The weakest subscales were substance use (0.57) and social-cognitive skills (0.57). The total score of the instrument was on the lower "good" range with an ICC of 0.63. Subscales that required less judgment (e.g., legal history, education/employment) had stronger level of reliability than those that required more subjectivity (e.g., aggression and violence, substance use).

Examining staff member's individual performance, only 59 percent of staff achieved "good" reliability. Subscales that required more judgment had less agreement than subscales that required less judgement. For example, 82.7 percent of staff demonstrated good reliability on the legal history subscale, but 38.9 percent demonstrated good reliability on the social-cognitive subscale. Interrater accuracy was not significantly impacted by site or staff gender, but the researchers did find that newer staff (i.e., five years or less of experience) performed better than staff with more experience.

Researchers noted that some of the CA-YASI items were poorly defined, especially those related to aggression/violence, attitudes, and social/cognitive skills. In an attempt to increase reliability, they, in cooperation with Orbis, developed additional definitions. The researchers found that the supplemental definitions did not improve accuracy. The researchers also found no significant relationship between more practice and increased staff accuracy. The researchers concluded that the length and complexity of the tool, the lack of definition of some items, and the insufficient training of staff may have contributed to the reliability issues with the tool (Skeem, et al., 2011; Kennealy, Skeem & Hernendez, 2017).

Construct Validity

The second component of the original CA-YASI (2008) evaluation focused on construct validity (see Skeem, Kennealy, & Hernandez, 2013; Skeem et al., 2017). Construct validity measures how well the tool measures the abstract concept that it intends to measure. This is done by measuring two things: convergent validity and discriminant validity. Convergent validity measures how strongly two measures of similar constructs correlate. Discriminant validity would show weak correlations between measures that are supposedly measuring different constructs. For example, a subscale of propensity toward violence should correlate

closely with other scales that have been shown to predict violence (convergent validity) AND should not be closely associated with a scale that only measures drug use (discriminant validity).

The researchers used only the assessments of staff that had sufficient reliability (i.e., the "best case" scenario) to assess the CA-YASI construct validity (Skeem, Kennealy, & Hernandez, 2013; Skeem et al., 2017). They only used specific CA-YASI items, and not Orbis' weighted scales, to maximize potential validity of the tool. The researchers then assessed a sample of DJJ youth using other empirically validated and reliable scales that measure the same theoretical constructs as the CA-YASI.

Results suggested that even using "best case" circumstances, there was limited evidence that the CA-YASI domains were capturing the intended risk factors. Static individual risk factors had the strongest support. These static factors include legal and correctional history and the static violence and aggression measures were the most valid. There was some support for the substance use and mental health domains. The dynamic domains for violence-aggression, attitudes, and social cognitive skills, however, showed no specific association with anger/hostility, procriminal thinking, and executive function deficits. There was also no support for construct validity in the social influences, family, and education and employment domains (Skeem, Kennealy, & Hernandez, 2013; Skeem et al., 2017).

The authors concluded that the tool was more suitable for a "risk" assessment and less for "needs" assessment because of the quality of the construct validity. They recommended using the tool to assign more services to higher risk youth, but not use it to determine what kinds of services (except maybe substance use and mental health which both had higher validity).

Predictive Validity

The final component to the original CA-YASI (2008) evaluation was predictive validity, or the degree to which the instrument can predict risk of future of offending (Skeem, Kennealy, Hernandez, Clark, & Tartar, 2013). The study utilized a sample of 846 youth in DJJ facilities to examine disciplinary infractions, and then a subset of 364 released into the community for recidivism. The researchers excluded 47 percent of youth from the analysis because the DJJ staff did not score the tool reliably. The study found that the predictive utility of the tool worked equally well for different racial and ethnic groups.

The research concluded that CA-YASI did well predicting institutional infractions, especially serious and violent ones. The grand total and the dynamic risk score did a good job of predicting violent and serious infractions, and an average job predicting any infraction. The total static risk score did a moderate job predicting all infractions (Skeem et al, 2013).

Skeem and colleagues also concluded that the tool moderately predicts any post-release arrest, but performs poorly predicting serious and violent arrests. The ability to predict arrest was driven by the total static risk score. Including the dynamic risk score provided no additional increase in prediction capabilities. Static risk only had a moderate effect on predicting serious or violent arrests. Generally, an AUC value of

0.70 is acceptable for clinical purposes.[5] The CA-YASI AUC for re-arrests was 0.66, but the tool could not predict whether the crimes would be serious or violent. The tool failed to predict future offending when attempting to discriminate between moderate and high-risk classifications (Skeem et al, 2013). In short, the CA-YASI had strong utility in predicting serious and violent infractions while in DJJ. It had moderate utility in predicting any infractions and any arrest. And has weak to no utility predicting serious and violent arrests.

Recommendations from the CA-YASI (version 2008) Evaluation

The comprehensive CA-YASI evaluation yielded three major recommendations. First, Skeem and her team suggested streamlining the risk assessment approach. Previous research has found that risk assessment tools that are short and simple tend to predict recidivism as well as longer tools (Baird et al., 2013; Skeem & Monahan, 2011). In fact, one study of YASI in Virginia (i.e., not the California version) found that using a 10-item version of YASI yielded better results than the full version (Baird et al., 2013). Second, the researchers concluded that there was little evidence that CA-YASI validly measured risk factors. This is only important if DJJ intends to adopt programs to address risk factors and then assign youth to programs that address their individual needs. This is, according to the IBTM framework, what DJJ intends to do. The intention to use CA-YASI to create individual treatment plans hinges on the tool's capability of validly measuring risk and need in each specific risk factor domain. Third, the research team recommended that once a streamlined version of CA-YASI was adopted, sufficient training of staff should be conducted to ensure that it is reliably administered. Staff should then be continually monitored to sustain the quality of CA-YASI administrations.

Streamlined CA-YASI (2016)

The 2008 CA-YASI tool (that was the subject of the previous evaluation) included 115 items, 90 items of the original YASI tool plus additional items specific to the unique high-risk population that DJJ serves. Upon the findings of the CA-YASI evaluation, DJJ asked Orbis to streamline the tool. The streamlining process was intended to assess whether all of the items were strictly necessary to either 1) predict recidivism, or 2) provide necessary information for case planning purposes. The process was also intended to maximize validity and simplify the data collection process (Orbis, 2015). These goals were consistent with the prior evaluation findings that the tool lacked construct validity and could improve predictive validity (Skeem, Kennealy, & Hernandez, 2013; Skeem et al, 2013). The evaluation also found that it took a staff member an average of four hours to complete a CA-YASI per administration (Kennealy, Skeem, & Hernandez, 2017). The streamlining of the CA-YASI was conducted with input from the original CA-YASI evaluation team of Dr. Jennifer Skeem and colleagues.

[5] AUC refers to Area Under the Curve (AUC) which is the common metric to assess predictive utility of a risk assessment measure. AUC values range from 0.50, which indicates that accuracy is not improved over chance. A 1.00 score would indicate perfect accuracy of the tool. According to Zhang, Roberts, & Farabee (2011), an instrument that produces AUC values of 0.70 or above are considered acceptable.

The streamlined tool was designed for two purposes: 1) prediction and classification and 2) case planning. The scales of overall risk and overall strength were to be used to classify youth in low, moderate, and high categories. These scales were designed to be most strongly associated with recidivism and optimize prediction. Additional items were included in the tool for case planning purposes. Though these items may not be included in the calculation for the predictive scales, they were retained to help DJJ staff identify the needs and responsivity to select appropriate treatment for the youth.

Orbis recalibrated the CA-YASI on a sample of 502 youth released in 2009. They utilized recidivism (i.e., rearrest) data for the 12-month post-release period (Orbis, 2015). The 1-year recidivism rate for this DJJ sample was reported to be 46 percent (Orbis, 2015). Orbis selected items to be retained in the streamlined CA-YASI that were theoretically or empirically important. For the prediction scales, items that were strongly correlated with recidivism (and not to other items) were retained. Individual items were weighted using a modified Nuffield (1982) procedure (Orbis, 2015). Differential scoring protocols were developed to reflect differences in significant predictors by age group.

The current (streamlined) CA-YASI (2016) version includes 78 items. The tool covers 11 domains: legal history, correctional response, aggression/violence, social networks, substance use, attitudes, adaptive skills, family, education/employment, health, and community. As stated prior, not all domains or items are included in the calculation of recidivism risk.

Staff administer the CA-YASI at intake, exit, and regular intervals (i.e., every three months) throughout a youth's commitment. Additional CA-YASIs can be administered if youth are involved in serious disciplinary infractions. Staff members still complete the CA-YASI by conducting an in-depth interview and file review. Their assessments are entered into the CA-YASI dashboard, and a report is generated summarizing the youth's risks, needs, and strengths. Staff then use that information to make housing determinations and intervention assignments. An assessment of the data collection, storage process, and overall quality was completed in an earlier evaluation (see Matsuda, Turner, & Hess, 2020).

Previous Research on Streamlined CA-YASI (2016)

Research evaluating the quality of the implementation of DJJ's IBTM framework included interviews with current DJJ staff members of various ranks and positions (Matsuda, Hess & Turner, 2020). Part of the interview protocol included questions specifically about CA-YASI and the current use of the tool. The study found that despite the streamlining and reassessment of the CA-YASI, DJJ staff perceived issues with the tool's reliability and validity (Matsuda, Turner, & Hess, 2020). DJJ staff who were familiar with the tool and regularly administered it, had reservations about CA-YASI's accuracy. In short, staff members suggested that the tool was too subjective, and that any noted differences could reflect variation in staff as opposed to actual change in risk or need. Staff also noted concerns with the wording of the questions or the available response categories. Some staff even admitted being reluctant to use certain response categories, especially those related to the aggression/violence domain, even if they knew that it was probably the most accurate answer (Matsuda, Turner, & Hess, 2020).

These concerns closely mirror the findings of the earlier evaluation of CA-YASI (2008). Though CA-YASI underwent a streamline and a recalibration, the questions and the administration process did not change significantly. Thus, the limitations uncovered by Dr. Skeem's team had not been addressed. The administration still required an in-depth interview and file review, the questions were still long, some still lacked clarification (despite additional instruction), and the dynamic factors of the tool still relies heavily on staff discretion and interpretation. Staff members interviewed in the recent interviews (Matsuda, Turner, & Hess, 2020) may not have been part of Dr. Skeem's original research, but their criticisms of the tool are consistent with earlier findings. Furthermore, there is no evidence to suggest that the Skeem evaluation's findings of lack of rater reliability and construct validity have been improved over the years.

The Current Evaluation

The current study focuses on the predictive validity of the current, streamlined CA-YASI tool (2016). While the tool could benefit from another study on interrater reliability and construct validity, this is beyond the scope of the current study. This study will examine the following research questions:

1) Does risk (as measured by the CA-YASI) change over time while in DJJ?
2) Does CA-YASI predict recidivism (i.e., rearrest or reconviction) post-release from DJJ?
3) How well does CA-YASI predict serious disciplinary problems while incarcerated in DJJ?

DJJ uses CA-YASI as a prediction and a case planning management tool (see Heilbrun, 1997 for assumptions on each type of use). The tool is both static and dynamic, it is administered at regular intervals over time in an attempt to capture change in risk, and it is used to select the types of interventions that youth in DJJ are assigned. The first research question tests one of the primary assumptions of the tool. Can CA-YASI measure change in risk over time? DJJ uses the tool to measure risk to help choose appropriate interventions and programming that would assist youth in reducing their risk factors or increasing their protective factors. In addition, a youth's measured risk can implicate their housing assignment. According to the IBTM treatment model, youth that lower their risk should be able to move to a lower risk housing unit that has more privileges. However, if the tool lacks construct validity or allows for too much staff subjectivity, we would not expect the change in risk to follow any predictable pattern. This would make the tool practically useless for case planning purposes.

In addition to CA-YASI's use as a case planning tool, DJJ desires to use the tool to predict the likelihood of recidivism or institutional violence. The streamlined instrument was recalibrated as a risk prediction instrument, but as the reliability and validity issues raised in the first evaluation had not been extensively addressed, there is not yet sufficient evidence to determine whether the current tool can accurately predict recidivism or institutional misconduct. Research questions 2 and 3, as introduced above, seek to explore the prediction capabilities of the streamlined CA-YASI (2016) on youthful offending behavior. This study will explore CA-YASI's association with the likelihood of post-release offending. However, we acknowledge that one's behavior in the community can be impacted by factors outside of DJJ's control, thus, we also examine the tool's association with serious institutional misconduct while in a DJJ facility to examine more proximal outcomes.

Methods

Sample Selection

The sample for this study was selected from the population of male youth incarcerated in a DJJ institution (not fire camp) on February 28, 2017 (n=651).[6] The few youth who were incarcerated in a DJJ facility for 90 days or less were excluded from this analysis (n=15).[7] Of the remaining youth, eight were still incarcerated at the time of our analysis. They were excluded because their commitment was not yet complete by the time our follow-up period began (i.e., January 1, 2020).[8]

Once the sample of 628 DJJ youth (98.7% of the entire DJJ male population) were selected, we sent a request for the corresponding CA-YASI records to DJJ. DJJ, then, sent our roster to Orbis to provide the full history of CA-YASI records. As discussed in a previous evaluation (see Matsuda, Turner, & Hess, 2020), DJJ staff enter data into Orbis' dashboard for CA-YASI. The dashboard utilizes the provided scores and produces the risk and need levels which DJJ staff then use for case planning purposes. The original data, however, is not stored with DJJ. It is in Orbis' possession. As part of their agreement, Orbis provides DJJ periodic transfer of the raw CA-YASI metrics. For this study, we required the use of CA-YASI data for the entire commitment of each person in the sample. At DJJ's request, Orbis provided all of these data for the use in this research.

Upon receipt of the CA-YASI data, we discovered that 13 youth in our sample had no CA-YASI data in the file provided. A follow-up with the organizations revealed that neither DJJ nor Orbis knew why these youth had no CA-YASI records available. These youth had other DJJ records available.[9] Neither of the agencies could recover the data for these youth. Thus, we had to exclude these 13 youth from all analyses. The final sample used for this study is 615 DJJ youth.

---

[6] At that time, DJJ incarcerated 20 female youth in one facility. Females in DJJ receive a separate treatment program than males. They are not included in the current analyses.

[7] Youth can be sent to DJJ for assessment or short-term temporary housing at the request of the county. These youth would not be subject to the standard DJJ treatment model. In total, 4 youth were excluded for being short-term commitments.

[8] A comparison of the study sample with youth excluded from the sample because they were still incarcerated at the follow-up date found that the eight youth still retained had significantly more days in DJJ (an average of 1729.75 days for not released sample and 987.56 days for the released sample), were younger at admission (16 years old v. 16.93 years old), and their cases were more likely to have originated from the superior court instead of the juvenile court (50% v. 17.7%).

[9] Ideally, these data would be missing at random, but parametric tests are difficult to do with such a small group. An examination of the missing shows that, the average age at admission for the missing youth was 17.6 years old. Over 30 percent of the missing records belonged to white youth (30.8%) which is about three times more than the proportion of white youth in the entire sample (10.5%). The average number of days in DJJ was about 686 for the missing youth compared with 977 for the larger sample.

Table 1. Study sample characteristics (n=615).

	n	% / mean (SD)
Case Type		
DJJ Case	506	82.3
DAI Case	10	17.6
Other	3	.5
Race		
Black	211	34.3
Latinx	321	52.2
White	62	10.1
Other	21	3.4
Commitment Offense		
Homicide/Manslaughter	47	7.6
Robbery	215	35.0
Assault	232	37.7
Sex	80	13.0
Other	41	6.7
Court of Jurisdiction		
Juvenile	506	82.3
Superior	109	17.7
Prior DJJ Commitment		
No	588	95.6
Yes	27	4.4
Enhancements		
No	349	56.8
Yes	266	43.3
Gang Affiliation (official)		
No	576	93.7
Yes	39	6.3
Age at Admission	615	16.92 (1.26)
Age at Discharge	615	19.62 (1.45)
Days in DJJ	615	984.14 (463.83)

Sample Characteristics

A summary of the characteristics of our study sample (n=615) is presented in Table 1. Over 82 percent of the sample had a commitment to DJJ. However, DJJ is an agency that houses youth up to 25 years old, and provides the same programming to youth who are legally "adults" and have been committed to the Division of Adult Institutions (DAI) (i.e., adult prison). At this time in DJJ, over 17 percent of the youth in their male institutions were commitments to adult prison that were being housed and treated in DJJ. The remaining three youth had a different type of commitment either as a contract case from the county or potentially a housing for out-of-state. Regardless of how these youth get to DJJ, once in residence, they receive the same type of programming and intervention.

Over half the sample was Latinx/Hispanic (with no other race indicated), over 34 percent were Black, and 10.1 percent were White. The final 3.4 percent of the sample were Asian, Native American, or any other race.

DJJ generally only accepts the most serious youth cases in the state. Thus, there is not as wide a range of commitment offenses or severity as one might expect in other systems. In this sample about 38 percent were committed for assault, almost 35 percent were committed for robbery, 13 percent were convicted of a sex related offense, 7.6 percent were committed for homicide or manslaughter, and fewer than 7 percent for other crimes (e.g., property crimes, kidnapping). Over 82 percent of youth were sent to DJJ from the juvenile court, and far fewer from the superior court (17.7%). Few of the youth in our sample had served a prior DJJ commitment (4.4%). This is not particularly surprising because DJJ's parole system was abolished in 2010, and post-release violations would be handled in the county. Youth would only be returned to DJJ if they completed their DJJ term and were convicted of another crime that included a new commitment to DJJ. An analysis of other commitment characteristics found that over 43 percent of the youth had an enhancement related to their commitment. An "enhancement" is a legal conviction that increases the total incarceration term due to the nature of the crime or the character of the offender.

Gang affiliation, in this study, is measured by an official designation as a "gang member" by DJJ staff. In this sample, which represents almost all of the incarcerated population at the time, 6.2 percent of the youth were officially designated as a gang member.

The average age at admission for this sample was about 17 years old. The average age of discharge of the youth was 19.6. Which then makes the average number of days spent in DJJ a bit over 2.5 years (or 977.98 days).

Data & Measures

Independent Measure

The purpose of this study is to evaluate the utility the streamlined CA-YASI (2016). Our primary independent measure is the level of risk as measured by the global CA-YASI risk scale (as calculated by an algorithm

determined by ORBIS). The CA-YASI tool produces an "Overall Risk" level.[10] The risk scale has four levels: 1 = low, 2 = low moderate, 4 = high moderate, and 5 = high. We were not provided the weighting or algorithm that Orbis uses to calculate the Overall Risk score. It is not definitively known whether the four levels of classifications can be thought of as a "continuum." However, given the nature of the description that we were provided, and the results of our sensitivity analysis, we treat the risk data as continuous in this study.[11] To do this, we recode the data to be divided in one-point intervals. Thus, in this study, a 1 = low, 2 = low moderate, 3 = high moderate, and a 4 = high. The initial risk of our sample included 4.9 percent low risk, 19.2 percent low moderate risk, 41 percent high moderate risk, and the remaining 35 percent as high risk as determined by their first CA-YASI.

In many cases, youth in the sample had both the original CA-YASI and streamlined CA-YASI administrations in their record. Orbis indicated that they had rescored all prior forms of CA-YASI for youth who were admitted to DJJ prior to the creation of the streamline version.

We found that 10 youth in the sample were missing their first CA-YASI assessment (i.e., their "initial" assessment). Thus, for this analysis, we utilized the CA-YASI assessment recorded closest to their DJJ admission date.

The 615 youth in this sample had 6,993 CA-YASI administrations, or an average of 11.4 risk assessments per youth during their DJJ commitment. DJJ policy indicates that staff are to perform a CA-YASI re-assessment every 90 days and/or if youth are involved in any serious incidents. Because involvement in serious incidents is an outcome that we are attempting to predict using the CA-YASI, we excluded any CA-YASI reassessments done in response to involvement in serious incidents. When predicting involvement in serious incidents in DJJ, we utilized scheduled yearly CA-YASI administrations for each individual. A year between assessments would allow for a sufficient period to capture involvement in serious incidents, which is a relatively infrequent event. There were 2,121 yearly CA-YASI administrations for this sample (average of 3.45 per youth).

---

[10] The CA-YASI also produces an Overall Protective level which reflects 5 different levels. We do not know whether Orbis used overlapping or highly correlated measures in their Risk and their Protective level algorithm, we chose to focus solely on the Risk Level for this analysis. However, we did conduct a separate, but similar, analysis of the Protective level and found similar, but opposite, trends that we found for Risk.

[11] To test our assumption that the CA-YASI data could be used as a continuous measure, we performed a sensitivity analysis on the risk data. We analyzed the data using a multilevel mixed-effects ordered logistic regression, which handles the CA-YASI data structure as ordinal (which we know it is) and found the same pattern of results presented in this report (i.e., handling the risk levels as continuous). The results of the sensitivity analysis found significantly lower odds of being in a higher risk group over time for the sample ($OR = 0.84$, $p < 0.001$). This result is also supported when we examined the change in risk from time point to time point. Given that these analyses reveal the same pattern and are robust in terms of the direction and statistical significance of the effects when treating risk as an ordinal variable and as a continuous variable, we opted to use the results from the continuous models for all three research questions. Doing so provides a more intuitive interpretation that is more concretely applied to the practical needs of DJJ.

Dependent Measures

The dependent measures for this study capture offending, both while incarcerated and upon release. For the purposes of this study, we are interested in the most serious types of offending.

Recidivism – Recidivism was measured post-release using: rearrest and reconviction. Recidivism data was obtained from criminal histories maintained by the California Department of Justice (DOJ). We utilize a two-year follow-up period for all youth. Most of the individuals in this sample would be released as an adult, so the criminal history data includes any rearrest or reconviction post-release, regardless of which system inevitably received the case.

The cohort had staggered released dates, so some youth had multiple years of potential follow-up data available, and others had a shorter follow-up period. In this study, we utilize a two-year follow-up period because that was a time-frame available for all youth in the study. From the DOJ criminal history records we extracted information on rearrest and reconviction. California DOJ matched our sample by their unique Criminal Identification and Information (CII) number or name and birthdate. The effort yielded 588 records matches or 95.4 percent of the sample.[12]

Serious Incident Reports (SIRs) – We measure in-facility offending using DJJ's administrative record of Serious Incident Reports. These are distinct from disciplinary infractions. As indicated by the name, SIRs capture the most "serious" incidents that occur in DJJ. A report is taken in the event of group disturbance, battery on staff, attempted/suicide, attempted/escape, sexual assault, death, or walk away/escape. They can capture incidents with youth and staff injuries and other serious consequences. All available Serious Incident Reports were used in this study and reflect a period of January 2015 through January 1, 2020 (i.e., the sample selection end point).

DJJ does not collect and store SIRs electronically. They are written and stored as report documents, thus they are not often used as "data" by the department. Unfortunately, the department did not have SIRs prior to 2015. This study was limited to collecting all of the retained SIRs (i.e., January 2015 to January 2020). The contents were coded by researchers, and an SIR record was created for every youth that was identified as being a participant in the incidents. We could not reliably distinguish between "perpetrator" or "victim," so these data reflect verified "involvement" in a serious incident. In total, 4,373 individual records were coded that represent 799 incidents that occurred during the timeframe. The SIR data were then merged with the administrative data that we had previously received on our sample. Youth could have multiple incidents attached to their files or zero incidents depending on their involvement in the SIRs. In our final sample, 387 (62.9%) were involved in at least one serious incident during their time in DJJ.

[12] One of the 28 youth not located by DOJ was also one of the 21 youth not located by Orbis for CA-YASI records, so they were already excluded from the analysis.

Covariates

DJJ provided the demographic and administrative data from which we drew our sample. As compared to most studies that seek to predict offending, this study utilizes relatively few covariates. The CA-YASI tool's purpose is to account for the most common predictors of offending. We only included covariates if we believed 1) it was not already included in the CA-YASI algorithm, 2) there were theoretical reasons to believe that behaviors may differ by condition, and/or 3) it is a check on the factor to see whether it has any impact when we would expect it should not. It is important to reiterate that while we have the scoring CA-YASI tool, we do not know the algorithm or the factors and weights that Orbis uses to determine the risk groups. Covariates that we included, in some or all of the analyses, include[13]:

Race/Ethnicity – The focus of these analyses is not to validate the tool by race or other covariates, but we do include race/ethnicity as a covariate because we know that it is often correlated with offending patterns and we do not see indication that it was included in the CA-YASI tool. We divide the sample into four racial/ethnic categories: black, Latinx/Hispanic, white, and other. The groups are mutually exclusive (i.e., the Latinx/Hispanic youth are not also included in another racial category). The categories were determined from DJJ's administrative data.

Sex Offender Status – The CA-YASI does include sex offenses as part of its count of previous violent behaviors or adjudications, but otherwise does not appear to score sex offenders in different risk categories. DJJ, however, has implemented a comprehensive and tailored rehabilitative program and housing specifically for sex offenders. Sex offenders do not, generally, share the same treatment programming as non-sex offender youth in DJJ. Thus, there is good reason to look more closely at membership in this group (and their subsequent treatment in DJJ) as potential explanations for differences in behavior).

Analytical Strategy

The analytical strategies vary by research question. We utilize different strategies to understand how risk changes over time, and whether the CA-YASI tool's determined risk can predict offending behaviors both within DJJ and after release.

[13] Age - We know that the CA-YASI includes a measure about age in prior offense history. We do not know whether the CA-YASI includes any measure of age at assessment or any other measure that might account for the maturation effect. We included age at assessment in all of our models and found it had no additional impact, so we excluded the measure from our final model for parsimony and to retain power.

First, we are interested solely in whether risk changes over time in a predictable way. We examine changes in overall risk from the initial assessment, yearly assessments[14], and the final assessment. Using the yearly intervals allows us to see the overall trend in risk over time. We conducted a longitudinal analysis using a fixed-effects model to assess within-person change in risk over time using these yearly CA-YASI assessments.

In the second research question, we use logistic regression models with the final CA-YASI risk level and the other covariates to predict odds of recidivism. We utilize a separate model to predict rearrest and one for reconviction. Results are presented for the first- and second-years post-release.

In the final research question, we utilized a path model to study the relationship between risk over time and in-facility offending. We utilize the initial CA-YASI risk to predict the in-DJJ offending for their first year. We use the first CA-YASI yearly assessment to predict the offending for the next year, and so on and so forth until the youth is released. We focused specifically on the yearly CA-YASI assessments to test whether CA-YASI predicts involvement in serious incidents while committed in DJJ. We calculated the rate of involvement in an SIR by year. The rate was adjusted to reflect youth that did not spend an entire year in DJJ. For example, if a youth was involved in 2 SIRs for the year he was committed to DJJ, the rate of SIR involvement would be 0.17. If, however, a youth was involved in two SIRs but was discharged at month four in his last year, his rate for that year would be 0.5.

[14] "Yearly" meaning on or around the one-year mark in each individual case. It does not represent one particular date, but is instead specific to each subject's admission date and the CA-YASI follow-up that occurs every 12 months from that admission date.

Results

Question 1: Does risk (as measured by the CA-YASI) change over time while in DJJ?

The first research question examines the basic purpose of a risk assessment instrument. Do the scores of the CA-YASI actually show a continuous change in risk over time? This question stems primarily from comments from staff with familiarity with the tool (Matsuda, Hess, & Turner, 2020). Staff members expressed skepticism that the CA-YASI tool was consistently being scored objectively. Some staff admitted being reluctant to mark a category the reflected a lower risk (despite what the scoring guide instructed) because of a youth's prior offending history. If this is consistently the case with the CA-YASI then we would not expect to see any discernable pattern in the change in risk of youth over time.

Descriptively, the overall sample did reduce their risk from the start of their commitment to the end. Table 2 presents the proportion of the youth in our sample in each CA-YASI determined risk category at their initial assessment and at their final assessment (regardless of how many years it took to discharge). As stated prior, fewer than 5 percent of youth were classified as "low" risk, 19 percent as "low moderate," 41 percent were "moderate high," and the remaining 35 percent were placed in the "high" risk group at the initial assessment. In contrast, by the time youth were discharged, 17.6 percent were considered "low risk," around 24 percent "low moderate," 40 percent "moderate high" and only 18.4 percent "high" risk.

Table 2. Proportion of DJJ youth by risk level at admission and discharge (n=615).

Risk Level	Admission / Initial CA-YASI Percent (n)	Discharge / Final CA-YASI Percent (n)
Low	4.9% (30)	17.6% (108)
Low Moderate	19.2% (118)	23.9% (147)
Moderate High	41.0% (252)	40.2% (247)
High	35.0% (215)	18.4% (113)

Figure 1 shows the average change in risk level over time for the sample. We included summaries across six years in DJJ. It is important to consider these data with the understanding that the average length of stay in DJJ is around 2.5 years. Thus, the number of people in the sample in the earlier years (i.e., 1, 2, and 3) is more than those still in the sample at years 4, 5, or 6. The sample size is provided at each year marker. The more people at each time point, the more stable the finding. Thus, we omitted anything past year 6 when there were fewer than 10 people in DJJ. Even with the omission, the 40 people remaining in DJJ in their 6th year is far fewer than the 615 youth that started in the sample.

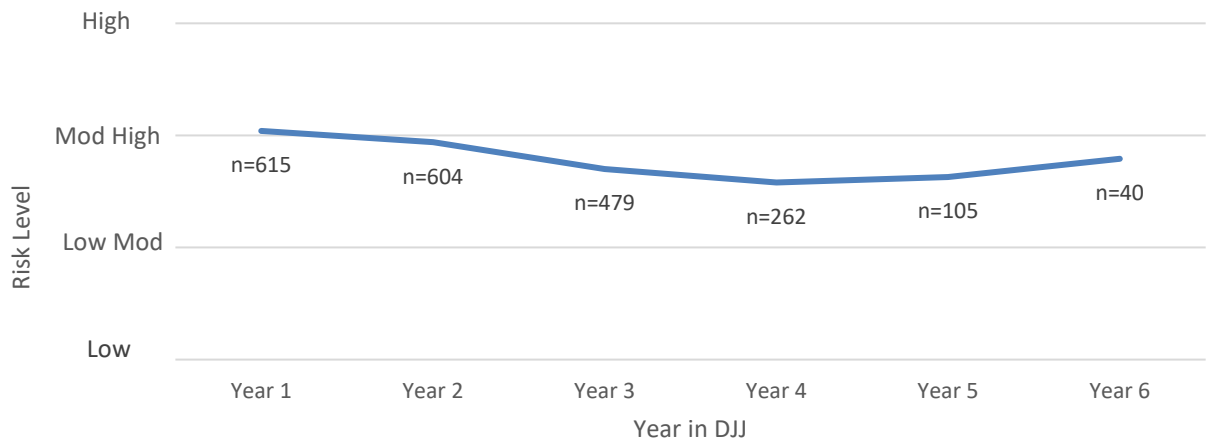

Figure 1. Average change in risk level over time.

We conducted a longitudinal analysis using a fixed-effects model to assess within-person change in risk over time. We utilized the initial CA-YASI administration and the subsequent assessments at one-year intervals until each of the 615 youth were released. Overall, the analysis found a statistically significant but modest reduction in risk over time (b = -0.20, p < 0.001). For every year someone is in DJJ, there is a .20 unit decrease in risk. It is important to note that given that this is a fixed-effects model of time, the coefficient of -0.20 reflects the average level of change between each risk assessment for the youth. Therefore, three years in DJJ would decrease risk an average of 0.20 each year for a total of a 0.60 decrease over the commitment. It is also important to keep in mind that the CA-YASI global risk scale is a four-point scale, so even a partial unit decrease in risk can be substantively important.

Figure 1 is a simple depiction of the change in mean risk of the sample over time. The figure shows that, in general, the risk level of youth in DJJ generally decreases over time until year 4 and then it appears to trend back up. The average risk approaches, but never drops below, what the CA-YASI would consider a "Low Moderate" risk level.

The graph also suggests that the reduction in risk over time may have a limit. In other words, the 0.20 unit decrease in risk stops after Year 4, and for youth who are still incarcerated in DJJ at Years 5 and 6 their risk begins to increase again. Had we restricted our sample to only include four years in DJJ, the magnitude of the coefficient would be slightly more prominent (b= -0.24, p<0.001). However, instead of continuing at an average reduction in risk of 0.24 per year, the overall model produces only a 0.20 reduction because of the uptick in the final years of the analysis. When considering this risk reversal, it is important to remember that we are losing sample size in these later years, so the movement might reflect the loss of power. Another explanation is that higher risk youth are committed to DJJ for a longer period of time, so as the less serious youth are released, the more serious youth remain and drive up the average risk. A final alternative explanation might be that increased time in DJJ does not equate to a linear decrease in risk reduction, and that there may be a plateau. Future research should examine this trend more closely to understand the specific contribution of maturation, selection effects, and treatment on the trajectory of risk over time.

Table 3 examines the overall change in risk by offender characteristics that, to the best of our knowledge, were not included in CA-YASI's calculation of risk. This is a descriptive table to further explore for whom risk is changing over time. The table shows that for every group of offender characteristic, the average risk score decreased from the initial CA-YASI assessment to the last CA-YASI assessment prior to release. Though the magnitude of the change may differ depending on the characteristic. Significance testing showed that almost all of the mean differences from first assessment to last assessment were significant. However, there was no significant difference in the change in risk for superior court commitments over time. Youth who were committed to DJJ by the juvenile court had more than twice the average reduction in risk over their stay compared to youth that were committed by the superior court.

Youth of any race reduced their risk over time. However, white youth, on average, demonstrated a larger reduction in risk over time (-0.75) compared to youth of other racial and ethnic groups. Youth convicted of a sex offense, on average, begin with a lower risk score than non-sex offenders. The youth in the sex offender group also had a larger mean decline in risk over time compared to youth that were not convicted of a sex offense (-0.72 v. -0.46).

This descriptive table does not control for any other extra-legal factors that may explain a decrease in magnitude of risk over time (e.g., time committed in DJJ, programming), but it does highlight that reduction in risk may not be equally distributed across all types of offenders.

Is it possible that staff behavior in scoring still has an impact on these results? Yes, that is possible. In these data, 42 different DJJ assessors scored the CA-YASI administrations. It is possible that the youth in DJJ would be scored differently but for bias by DJJ staff. The reliability of staff scoring was not a part of this evaluation, even though the previous CA-YASI evaluations found that it was a problem.

Table 3. Average risk score by covariates at the first and last CA-YASI in DJJ.

	n	First Assessment		Last Assessment		Diff. / Sig.
		Mean	SD	Mean	SD	
Age at Admission						
≤ 17	409	3.09	0.86	2.59	0.99	-0.50***
≥ 18	206	3.00	0.86	2.47	0.92	-0.53***
Race						
Black	211	2.94	0.84	2.49	0.89	-0.45***
White	62	2.73	0.85	1.98	0.97	-0.75***
Latinx/Hispanic	321	3.21	0.86	2.72	0.96	-0.49***
Other	21	2.96	0.74	2.37	1.07	-0.59*
Court of Commitment						
Juvenile	506	3.08	0.87	2.52	0.97	-0.56***
Superior	109	2.95	0.84	2.72	0.95	-0.23
Offense						
Not Sex Offender	528	3.11	0.85	2.65	0.93	-0.46***
Sex Offender	87	2.74	0.88	2.02	1.01	-0.72***

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

Question 2: Does CA-YASI predict recidivism (i.e., rearrest and reconviction) post-release from DJJ?

To understand whether CA-YASI risk scores significantly predicted post-release behavior, we conducted a series of multivariate logistic regressions. The CA-YASI tool should already account for most of the predictors of offending (as is the intention of the tool), so we include only a few covariates that may be significant but (to the best of our knowledge) are not included in the algorithm. Analyzing the criminal histories maintained by the California Department of Justice, we found that 35.3 percent of the sample (n=588) were rearrested by the end of their first year out of DJJ. About 52 percent (52.4%) were rearrested by the end of year two. After their first year, 20.9 percent of the sample had been reconvicted and 30.4 percent after the second year of follow-up. Table 4 reports the results of the analysis. The coefficients listed in each column report the odds ratio of the variable. Coefficients greater than one reflect a greater odd of recidivism; coefficients that are less than one indicate reduced odds of recidivism.

In this model we include the CA-YASI level determined prior to release, the race/ethnicity (with white as the reference group) and sex offender status as covariates. When controlling for the other factors, risk level at release was the strongest predictor of rearrest and reconviction and significantly predicted rearrest and reconviction in both years of follow-up.

Table 4. Logistic regression results predicting recidivism outcomes.

	Recidivism Outcomes			
	Rearrest		Reconviction	
	Year 1 (SE)	Year 2 (SE)	Year 1 (SE)	Year 2 (SE)
CA-YASI Level at Release	1.36 (0.09)***	1.36 (0.09)***	1.48 (0.13)***	1.50 (0.11)***
Race / Ethnicity				
Black	1.26 (0.46)	2.27 (0.77)**	1.01 (0.50)	1.43 (0.58)
Latinx/Hispanic	1.40 (0.50)	1.48 (0.49)	1.30 (0.59)	1.39 (0.55)
Other	1.06 (0.59)	1.75 (0.95)	1.57 (1.06)	1.75 (1.06)
Sex Offender	0.50 (0.15)*	0.46 (0.13)**	0.41 (0.18)*	0.52 (0.18)

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

The analyses show that the odds of arrest in year one increase 36 percent for each increase in level of risk. The increased risk of rearrest is similar for both years of follow-up. The odds of reconviction one-year post-release increase 48 percent for each level of risk. The risk increases slightly in year two; the odds of reconviction in the second year increase 50 percent for each level of risk. When controlling for their risk level at their release, sex offenders were still significantly less likely to be rearrested and reconvicted (in year 1) than non-sex offenders. This finding supports other studies that have found that sex offenders are less likely to recidivate than non-sex offending youth (e.g., Calleja, 2015).

Race/ethnic status was generally not significant in the models, with one exception. The odds that a black youth was rearrested in their second year was 127 percent higher than compared to white youth. The data show that black youth were not more likely to be reconvicted.

Question 3: How well does CA-YASI predict disciplinary problems while incarcerated in DJJ?

This research question tests the ability of the CA-YASI to predict involvement in serious disciplinary events while incarcerated in DJJ. As stated earlier, DJJ's serious incident reports (SIRs) were available from January 2015 until December 31, 2019. Due to our sampling parameters and the fact that the average length of stay in DJJ is around 2.5 years, we had hoped that most of the study youth would have been in DJJ during the period in which we have complete SIRs to measure carceral behavior. Unfortunately, that did not turn out to be the case. We had to omit 120 youth from this research question who began their stay prior to 2015 (i.e., did not completely overlap with the SIR record period), leaving a total of 495 youth in the inferential analysis model. In this section, we will present the analysis with the 495 youth and then the results of a supplemental analysis that investigates whether the results of the 495 youth can reasonably be extrapolated to the youth that were not included in this analysis (i.e., the 120).

We conducted a path analysis to understand how well the risk level at intake predicted involvement in SIRs in the first year of a youth's stay. Then, the model examines how well the CA-YASI risk at the beginning of

their second year predicts involvement in SIRs in their second year, and so on, until the end of the commitment. We evaluated model fit using four traditional model fit indices: chi-square statistics, comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA). For the purposes of this report, only the standardized results for fit and the path model are reported and interpreted. Overall, the model fit was excellent: $\chi2(36) = 89.368$, $p < .001$, CFI = 0.96, TLI = 0.95, RMSEA = 0.055 according to the acceptable thresholds (Kline, 2009). Traditionally, a significant $\chi2$ signals a poor model fit. However, this estimate should be interpreted alongside other measures of model fit because it is largely influenced by sample size and does not provide much information alone. Both CFI (0.94) and TLI (0.90) are high compared to the acceptable threshold (>.90). RMSEA (0.079) is just at the accepted threshold of less than 0.08. Taken together, these indices reflect a well-fitting model.

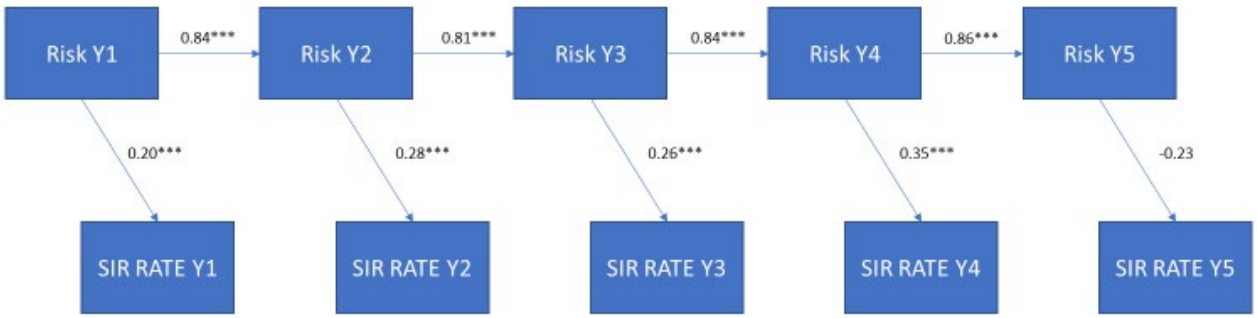

Figure 2. Standardized direct and indirect effects of risk predicting yearly rate of involvement in serious incidents.

Figure 2 shows the results of the path mode with the standardized beta coefficients and indication of significance. The model shows that risk level at each year is significantly correlated with the rate of involvement in serious incidents for the rest of that year. Youth who have a higher risk level at admission are significantly more likely to have a higher rate of involvement in a serious incident in their first year ($\beta = 0.20$, $p < 0.001$, n = 495). Youth who have a higher risk level at year 2 are significantly more likely to have a higher rate of involvement in serious incidents in year 2 ($\beta = 0.28$, $p < 0.001$, n=483). The same pattern held for year 3 ($\beta = 0.26$, $p < 0.001$, n=357) and year 4 ($\beta = 0.35$, $p < 0.001$, n=141). Significance did not continue into year 5 which is not surprising given the small sample who were imprisoned for five years in the sample. Only 31 youth remained in DJJ after five years.

The path analysis also shows that risk level at each year is significantly related to the risk level at the subsequent year, which is expected. Youth who have a higher risk level at intake are more likely to have a higher risk level at year 2 ($\beta = 0.84$, $p < 0.001$). The same pattern continues throughout all the years of the commitment. We know from the results of Question 1 analysis that risk does decline over time for all risk levels, but this model suggests that the decline is still going to be dependent on the previous level, and by extension, one's starting risk level. This analysis suggests that youth with higher risk levels will have higher

risk levels in subsequent years (though it does decline over time), and that higher risk youth will be more likely to be involved in serious incidents while in a DJJ facility.

Comparisons of Missing Sample

Despite the exclusion of 120 youth from this analysis sample, our finding that CA-YASI does predict involvement in serious and violent incidents is consistent with the previous evaluation of the full CA-YASI conducted by Skeem and colleagues (Skeem et al., 2013). However, to better understand the impact of removing 120 youth from the sample, we did descriptive comparisons between the groups and analytical comparisons between the groups' risk and recidivism trends. Table 5 shows a comparison of characteristics of the sample used to analyze question 3 (i.e., Post-2015) with those who were excluded from the sample (Pre-2015) because of lack of SIRs. The comparisons show that the groups are not significantly different in racial/ethnic distribution. They are, however, significantly different on other background characteristics. The analysis sample is more likely to be committed for robbery and assault than murder or manslaughter. They are more likely to be committed from the superior court than the pre-2015 sample, and therefore significantly older upon admission.

Table 5. Descriptive comparisons between the RQ3 Sample and the Excluded Youth.

	RQ3 Sample (Post-2015)	Excluded Sample (Pre-2015)	*p*
n	495	120	
Race / Ethnicity			
Black	34.3%	34.2%	
Hispanic	51.9%	53.3%	
White	3.6%	2.5%	
Other	10.1%	10.0%	
Commitment Offense			
Murder/Manslaughter	5.1%	18.3%	***
Robbery	37.2%	25.8%	
Assault	39.4%	30.8%	
Sex	12.7%	14.2%	
Other	5.7%	10.8%	
Court of Commitment			
Juvenile	80.0%	91.7%	**
Superior	20.0%	8.3%	
Mean Age at Admission (SD)	17.15 (1.22)	16.21 (2.41)	***

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

In addition to the descriptive comparison, we also examined whether the two groups of youth differed in their trajectory of risk (Figure 3). It is important to note that the pre-2015 sample youth were incarcerated for longer (on average). There were 18 youth in the pre-2015 sample that were in their seventh year, but the post-2015 sample only had 4 youth that were committed for six years. Because the sample size is quite small for both groups after year 4, we include only years 1 through 4 in the figure.

Figure 3 shows that the trend of risk in both groups run parallel. Risk scores for both groups significantly decline over time. However, the pre-2015 group, on average, begins and continues to have higher risk scores over their years of incarceration. The differences in means for the two groups is statistically significant at each time point. In short, the pre-2015 group is a higher risk group than the usable sample, but the trajectory over time is consistent.

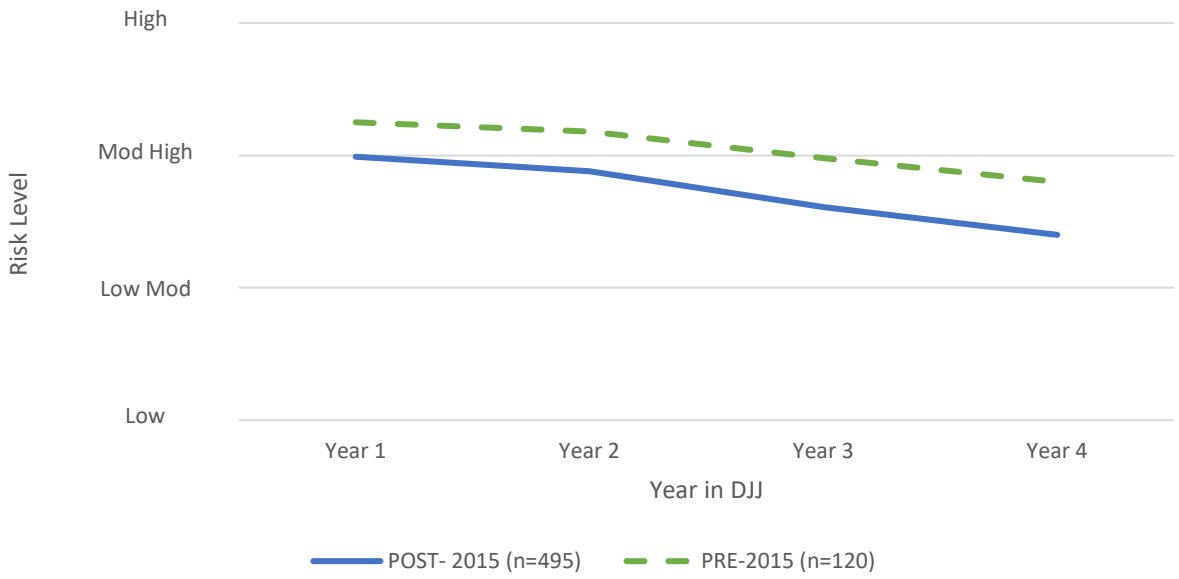

Figure 3. Change in risk level over time for pre-2015 and post-2015 groups.

While they are statistically different in level of risk, should we expect that their recidivism outcomes would be different? If their outcomes are different, then it may be inappropriate to generalize any findings of serious offending in DJJ to the pre-2015 DJJ sample. Because we do not have the SIRs to directly test that question in the pre-2015 sample, we examine the trends in recidivism for the two groups to see if the samples behave similarly in regards to this measure of offending. It is important to keep in mind that the sample sizes of the groups for this analysis are much smaller than in the previous recidivism analysis (Pre-2015 sample = 115, Post-2015 sample = 473). Reliable estimates of regression coefficients are less likely if there are too few events per parameter (Hosmer & Lemeshow, 2000). This was the case with the sex offender indicator, thus it was removed. We limited our two-group recidivism logistic regressions to one covariate (i.e., the ending CA-YASI risk level).

The results of the recidivism analysis are included in Table 6. The analysis for the larger, usable sample (i.e., post-2015) is consistent with our findings of recidivism for the entire sample. The level of CA-YASI

determined risk at their last test administration predicts rearrest and reconviction during the two-year post-release period. For the excluded sample (i.e., pre-2015), the trend of results is consistent with the usable sample, but risk does not predict recidivism until the second year after release. The failure to reach significance in year 1, could be explained with the size of the sample and the smaller proportion of youth that recidivate in the first year. In other words, it could be due to trying to predict a rarer event in a smaller sample (Hosmer & Lemeshow, 2000).

Table 6. Comparison of recidivism results for usable (post-2015) and excluded (pre-2015) samples.

	Usable Sample (post-2015)				Excluded Sample (pre-2015)			
	Arrest Y1	Arrest Y2	Convict Y1	Convict Y2	Arrest Y1	Arrest Y2	Convict Y1	Convict Y2
CA-YASI at	1.46	1.42	1.59	1.61	1.29	1.41	1.36	1.30
Release	(.11)***	(.10)***	(.15)***	(.13)***	(.20)	(.21)*	(.26)	(.21)***

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

Our two-sample comparison suggests that the usable sample in question 3 is significantly different than the excluded sample. However, the behavior of the two groups follows a similar trend, the excluded group is just higher risk and more serious than the included group. Thus, we would expect the CA-YASI to operate similarly in the excluded group, but that the magnitude of the effect might be different. We would expect that, had we had the data, the CA-YASI would significantly predict involvement in serious incidents in DJJ, but that the decline in risk would be less pronounced and that the group would be involved in more serious incidents over time.

Limitations and Future Research

This study is an important first step in exploring the use of the streamlined CA-YASI in this high risk, DJJ population. It should not, however, be the last step. This study found that risk level of youth in DJJ decreases over time (to a point) and that the CA-YASI tool can be used to predict recidivism as well as serious incidents while in DJJ. This study, also however, focused solely on male youth in DJJ. DJJ does house female youthful offenders, generally fewer than 30 at any given time. The current sample size of female offenders is not large enough to provide an analysis of whether the CA-YASI works equally well for male and female youth offenders. We know, from Orbis' documentation of their streamlining item selection, differential scoring protocols were created for different age groups. We do not know if the tool is calibrated differently for female and male offenders. We know that Orbis has an instrument specifically for girls (Orbis Partners, 2007b), but it is not in use at DJJ. It is not clear whether a gender-specific instrument would be more informative. Future research should examine this question.

This evaluation did not examine the extent to which the CA-YASI tool was used to assign programing and interventions. Previous evaluations of the (original) CA-YASI (2008) recommended against using that tool to determine program assignments (Skeem, Kennealy, & Hernandez, 2013; Skeem et al., 2017). Researchers recommended using the tool to determine the amount of intervention, but not the type of intervention. In reality, DJJ offers a limited selection of formal intervention programs (see Matsuda, Hess, & Turner, 2020). This is likely due to the fact that DJJ houses only serious youth offenders. DJJ offers programs in anger management and conflict resolution, substance abuse, counseling, interpersonal skill building, and a tailored program for sex offenders and female offenders. In short, a male youth in DJJ can expect to get all of the programs that they offer because of the limited scope and the nature of their population. Regardless, future evaluations should more closely examine the link between the CA-YASI and its use as a case planning tool specifically because DJJ had stated that it intended to use the tool in that capacity.

Finally, while the CA-YASI tool does significantly predict behaviors in DJJ and upon release, we don't know if it is the "best" tool available for this purpose. This study did not examine whether there is a tool that is more accurate for this population. There are numerous risk assessment tools, and an evaluation should compare the CA-YASI with other options to ensure it is the best fit for DJJ's needs and its population.

References

Andrews, D.A. & Bonta, J., (2010). The psychology of criminal conduct (5th ed.) New Providence, NJ: Matthew Bender Lexis Nexis.

Baird, C., Healy, T., Johnson, K., Bogie, A., Dankert, E.W., Scharenbrock, C. (2013). A comparison of risk assessment instruments in juvenile justice. Report for the National Council on Crime and Delinquency. https://www.ncjrs.gov/pdffiles1/ojjdp/grants/244477.pdf

Barnoski, R. (2003). Changes in Washington State's jurisdiction of juvenile offenders: Examining the impact. Olympia, WA: Washington State Institute of Public Policy.

Bonta, J., & Andrews, D.A. (2007). Risk-Need-Responsivity Model for offender assessment and rehabilitation. Canada: Her Majesty the Queen in Right of Canada.

Brorby, D. (2005). Farrell v. Hickman (No. RG03079344) First Report of Special Master: Compliance with interim measures provisions of consent decree and January 31, 2005 stipulations https://www.clearinghouse.net/chDocs/public/JI-CA-0013-0015.pdf

Brorby, D. (2006). Farrell v. Hickman (No. RG03079344) Second Report of Special Master: Compliance statue interim measures, disability, education, sexual behavior treatment. https://www.clearinghouse.net/chDocs/public/JI-CA-0013-0016.pdf

Calleja, N.G. (2015). Juvenile sex and non-sex offender: A comparison of recidivism and risk. Journal of Addictions & Offender Counseling, 36, 2-12.

CDCR DJJ (2006). Safety and Welfare Remedial Plan: Implementing reform in California. Sacramento, CA: California Department of Corrections and Rehabilitation. https://www.clearinghouse.net/chDocs/public/JI-CA-0013-0005.pdf

Grove, W.M., Eckert, E.D., Heston, L., Bouchard, T.J., Segal, N., & Lykken, D.T. (2000). Clinical vs. mechanical prediction: A meta-analysis. Psychological Assessment, 12, 1, 19-30.

Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. Law and Human Behavior, 21, 347-359.

Hosmer, D.W. & Lemeshow, S. (2000). Applied Logistic Regression (2nd Edition). New York, NY: John Wiley & Sons, Inc.

Howell, J.C., Lipsey, M.W., Wilson, JJ., Howell, M.Q., & Hodges, N.J. (2019). A handbook for evidence-based juvenile justice systems: Revised edition. Lanham, Maryland: Lexington Books.

Juvenile Rehabilitation Administration (2002). Integrated Treatment Model Report. Olympia, WA: Washington State Department of Children, Youth, and Families. https://www.dcyf.wa.gov/sites/default/files/pdf/ITM_Design_Report.pdf

Kennealy, P. J., Skeem, J. L., & Hernandez, I. R. (2017). Does staff see what experts see? Accuracy of front-line staff in scoring juveniles' risk factors. Psychological Assessment, 29, 1, 26.

Kline, R. B. (2005). Principles and practice of structural equation modeling (2nd ed.). New York: Guilford.

Krisberg, B. (2003). General corrections review of the California Youth Authority. Report submitted to the court in Farrell v. Harper. Downloaded at https://www.clearinghouse.net/chDocs/public/JI-CA-0013-0025.pdf.

Legislative Analyst's Office (2000). Analysis of the 2001-02 Budget Bill: Department of the Youth Authority (5460). https://lao.ca.gov/analysis_2001/crim_justice/cj_09_5460_anl01.htm.

Legislative Analyst's Office (2005). Analysis of the 2005-06 Budget Bill: Department of the Youth Authority (5460). https://lao.ca.gov/analysis_2005/crim_justice/cj_08_5460_anl05.htm .

Little Hoover Commission (2008). Juvenile justice reform: Realigning responsibilities. Sacramento, CA: Little Hoover Commission. https://lhc.ca.gov/sites/lhc.ca.gov/files/Reports/192/Report192.pdf

Loeber, R., Slott, N.W., & Stouthamer-Loeber, M. (2008). A cumulative development model of risk and protective factors. In R. Loeber, H.M. Koot, N.W. Slott, P.H. Van der Laan, & M. Hoeve (EDS). Tomorrow's criminals: The development of child delinquency and effective interventions. (pp.3-17). Hampshire, England: Ashgate.

Matsuda, K.N., Hess, J., & Turner, S.F., (2020). Division of Juvenile Justice Treatment Model Process Evaluation. Irvine, CA: Center for Evidence-Based Corrections at University of California, Irvine.

Matsuda, K.N., Turner, S.F. & Hess, J. (2020). Division of Juvenile Justice Data Systems Evaluation. Irvine, CA: Center for Evidence-Based Corrections at University of California, Irvine.

Maxson, C.L., Bradstreet, C.E., Gascon, D., Gerlinger, J., Grebenkemper, J., Haerle, D., Kang-Brown, J., Mesinas, A.M., Omori, M., Reid, S., and Scott., D. (2012). Gang and Violence in California's Youth Correctional Facilities: A Research Foundation for Developing Effective Gang Policies. Irvine, CA: University of California, Irvine.

Nuffield, J. (1982). Parole decision-making in Canada: Research toward decision guidelines. Ottawa, ON: Solicitor General of Canada.

Orbis Partners Inc. (2007a). Long-term validation of the Youth Assessment and Screening Instrument (YASI) in New York State Juvenile Probation. Ottawa, Ontario: Orbis Partners Inc.

Orbis Partners (2007b). Youth Assessment and Screening Instrument: Girls (YASI-G). Ottawa, Ontario, Canada: Author.

Orbis Partners Inc. (2015). CA-YASI Streamlining Methodology. Internal document provided to CDCR's DJJ. 1-3.

Shook, J.J., & Sarri, R.C. (2007). Structured decision making in juvenile justice: Judges' and probation officers' perceptions and use. Children and Youth Services Review, 29, 1334-1351.

Skeem, J.L., Hernandez, I., Kennealy, P.J., & Rich, J. (2011). CA-YASI reliability: How adequately do staff in California's Division of Juvenile Justice rate youths' risk of recidivism? Prepared for the California Division of Juvenile Justice. Irvine, CA: University of California, Irvine.

Skeem, J.L., Kennealy, P., & Hernandez, I. (2013). CA-YASI construct validity: To what extent do the domains measure the risk factors they're supposed to measure. Prepared for the California Department of Juvenile Justice. Irvine, CA: University of California, Irvine.

Skeem, J. L., Kennealy, P. J., Tatar II, J. R., Hernandez, I. R., & Keith, F. A. (2017). How well do juvenile risk assessments measure factors to target in treatment? Examining construct validity. Psychological Assessment, 29(6), 679.

Skeem, J.L., Kennealy, P.J., Hernandez, I., Clark, S., & Tatar, J., (2013). CA-YASI predictive utility: How well do scores and classifications predict youths' infractions and re-arrest? Prepared for the California Department of Juvenile Justice. Irvine, CA: University of California, Irvine.

Skeem, J.L. and Monahan, J. (2011). Current directions in violence risk assessment. Current Directions in Psychological Science, 20, 38-42.

Wachter, A. (2015) Statewide risk assessment in juvenile probation. JJGPS StateScan. Pittsburgh, PA: National Center for Juvenile Justice.

Zhang, S.X., Roberts, R.E., 7 Farabee, D. (2014). An analysis of prisoner reentry and parole risk using COMPAS and traditional criminal history measures. Crime & Delinquency, 60, 2, 167-192.