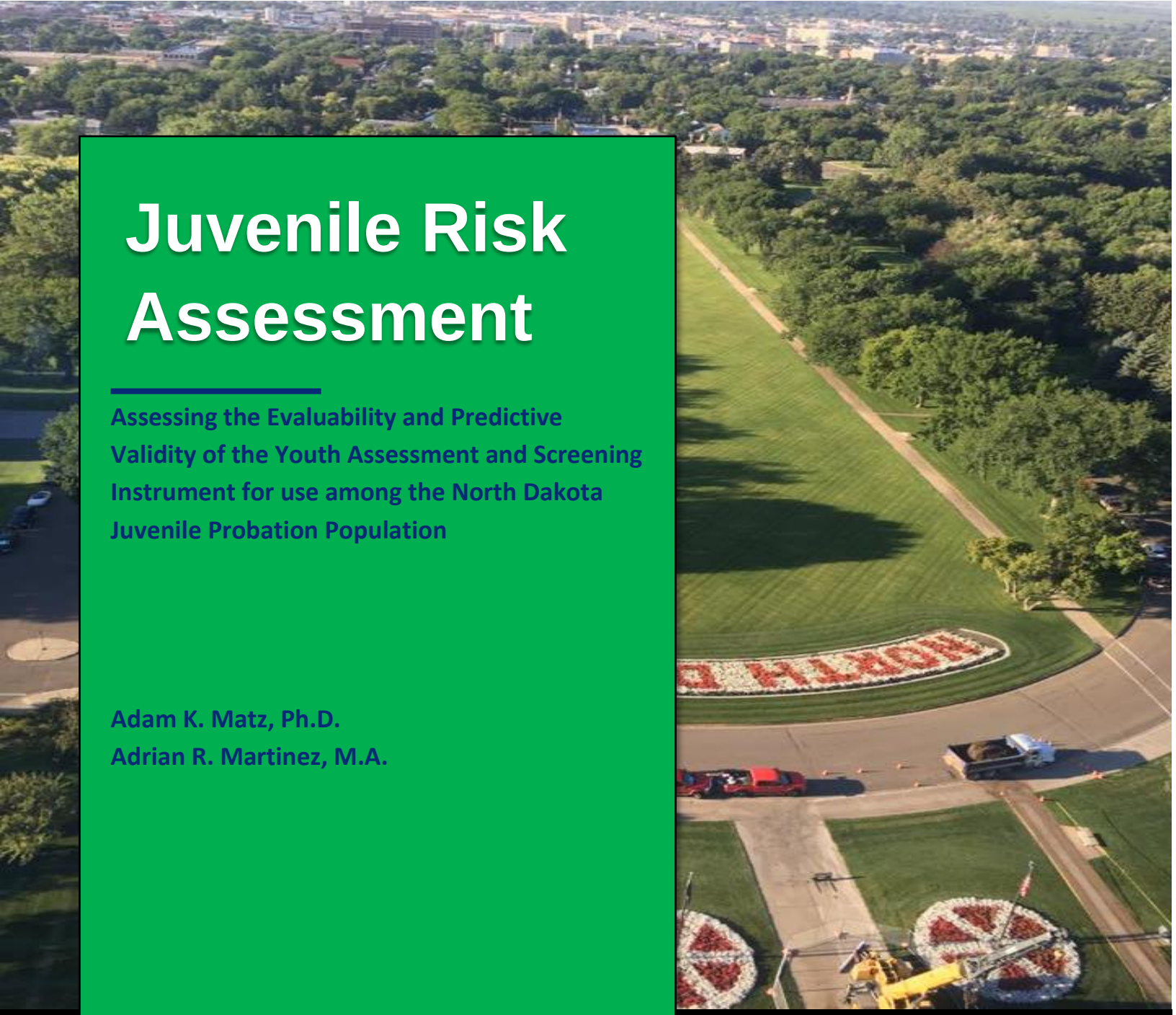
6-26-2019

Juvenile Risk Assessment: Assessing the Evaluability and Predictive Validity of the Youth Assessment and Screening Instrument for use among the North Dakota Juvenile Probation Population

Adam K. Matz
University of North Dakota, adam.matz@und.edu

Adrian R. Martinez



Juvenile Risk Assessment

Assessing the Evaluability and Predictive Validity of the Youth Assessment and Screening Instrument for use among the North Dakota Juvenile Probation Population

Adam K. Matz, Ph.D.
Adrian R. Martinez, M.A.

JUNE 26, 2019

University of North Dakota
College of Arts & Sciences
Department of Criminal Justice

ACKNOWLEDGEMENTS

Special thanks to the North Dakota Juvenile Court and each of the unit directors; including Shawn Peterson of Unit 1 (Northeast and Northeast Central Judicial Districts), Karen Kringlie of Unit 2 (East Central and Southeast Judicial Districts), Cory Pedersen of Unit 3 (South Central and Southwest Judicial Districts), and Scott Hopwood of Unit 4 (Northwest and North Central Judicial Districts). In addition, special thanks to the State Court Administrator, Sally Holewa, and the Family Law Mediation Program Administrator, Cathy Ferderer, for their continued support.

This project was supported by the UND Early Career Award program. It was one of seven, out of 32, proposals selected by the Vice President for Research & Economic Development in fall 2017, with a project period from January 1, 2018 to June 30, 2019. In addition, this research project is in alignment with the Rural Health & Communities Grand Challenge associated with the university's ambitious 2017-2022 Strategic Plan. Special thanks to my colleagues, students, and others at the university that supported this research.

Suggested Citation (APA):

Matz, A. K., & Martinez, A. R. (2019). *Juvenile risk assessment: Assessing the evaluability and predictive validity of the Youth Assessment and Screening Instrument for use among the North Dakota juvenile probation population*. Grand Forks, ND: University of North Dakota, College of Arts & Sciences, Department of Criminal Justice.

For questions or inquiries about this report please contact Dr. Adam K. Matz at 701.777.2010 or by email at adam.matz@UND.edu.

EXECUTIVE SUMMARY

Risk assessment is a central component of juvenile probation work and considered an evidence-based practice by the National Institute of Corrections (NIC). The Youth Assessment and Screening Instrument (YASI) was implemented statewide in 2002. Though subjected to validation in other states and Canadian provinces, it has not been validated in North Dakota.

This research was premised on four guiding questions, with the intent to assess the current evaluability of the instrument and provide preliminary validation estimates. A synopsis of key findings associated with each question is provided below, followed by a list of practice, policy, and research recommendations.

1. *To what extent are the data needed to assess the YASI available and retrievable?*
    - Most of the data needed for assessment exist in CMS, CASEWORKS, and Odyssey. However, it must be manually collected, a time-consuming process. Detailed raw risk and strength scores, individual domain item weighting, and other item adjustments were not retrievable from CASEWORKS.

2. *To what extent does this tool (YASI) accurately predict the likelihood of recidivism among a sample of the state's juvenile probation population?*
    - Results of this study indicate, overall, YASI possesses a statistically significant but small-to-moderate effect in correctly predicting recidivism. Though not optimal, according to Jones, Brown, Robinson, & Frey (2016, p. 189), this is consistent with many prior studies of YASI. Further, evidence of under-classification exists, whereby low and moderate risk youth recidivate at higher levels than expected (see also Jones et al., 2016, p. 185). Due to its low frequency, the high-risk subpopulation could not be adequately assessed.

3. *Does the tool's predictive accuracy differ by sex, race, or region of the state?*
    - The tool performs best for males and whites.
    - There was no evidence to suggest the instrument performs well for females.
    - Due to low frequency within this study sample, YASI's performance in terms of predicting recidivism for African American and Native American youth remains unclear.
    - All units were associated with small-to-moderate effect sizes, though statistical significance varied. Unit 3 possessed the most consistent results.

4. *Are there specific domains (e.g., legal history, peers, alcohol and drugs) of YASI that contribute more (or less) to its predictive accuracy?*
    - We were unable to assess individual domains. Individual item weights and raw risk and strength scores were not retrievable, presumably proprietary property of *Orbis Partners*. However, exploratory multivariate analyses identified five statistically significant variables; *age at first offense* (younger age associated with greater risk), *noncompliance with parental rules* (the higher the rating on this five-point measure the greater the risk), *poor academic performance* (the higher the rating on this five-point measure the greater the risk), *suicidal ideation* (if present less likely to recidivate), and *lack of consequential thinking skills* (those who scored higher on this four-point measure were less likely to recidivate).

Recommendations:

1. Continue to provide ongoing YASI trainings to those tasked with completing assessments. Appropriate training is *essential* to producing valid assessments (Flores, Lowenkamp, Holsinger, & Latessa, 2006; Latessa & Lovins, 2010).

2. Consider implementing a supplemental measure to enhance risk prediction for females. Though in preliminary planning phases and not yet validated, *Orbis Partners* has developed a YASI-Girls (YASI-G) instrument which includes measures concerning relationships, emotional expression, self-efficacy, sexual vulnerability, and early parenthood (Jones et al., 2016, pp. 190-191). Another example would be the Early Assessment Risk List for Girls (EARL-21G) (for further discussion see Van Voorhis, Wright, Salisbury, & Bauman, 2010; Shepherd, Luebbers, and Dolan, 2013).

3. Conduct further research on the YASI's predictive validity as it pertains to female, African American, and Native American youth.

4. Ensure each assessment is documented accurately, reassessments clearly distinguished, and conduct further research on reassessments.

5. Conduct further validation research at least once every five years to continually assess the predictive accuracy of the instrument. Using the current study as a baseline, the next evaluation should demonstrate an improvement in relation to recent changes in practice not captured within this study (including new YASI trainings provided and consolidation of risk assessment to a single person at each unit).

6. In the event subsequent research fails to see an improvement in the YASI's predictive accuracy, consider exploring the re-evaluation and adjustment of item weights and cutoff scores, which may need to be unique for special populations (for further guidance see Georgiou, 2019; Jones et al., 2016).

7. Consider further research on interrater reliability (e.g., Baird et al., 2013; Skeem, Hernandez, Kennealy, & Rich, 2012) and internal consistency of domain measures (e.g., Jones et al., 2016, p. 186).

8. Develop a streamlined research method that includes a standardized sampling frame, method for extraction of samples, and the ability to stratify samples by sex, race, and/or unit. A more automated or pseudo-automated method of data collection that relies less on manual counting and coding would improve efficiency and timeliness, as well as reduce costs, in the future.

TABLE OF CONTENTS

LIST OF TABLES

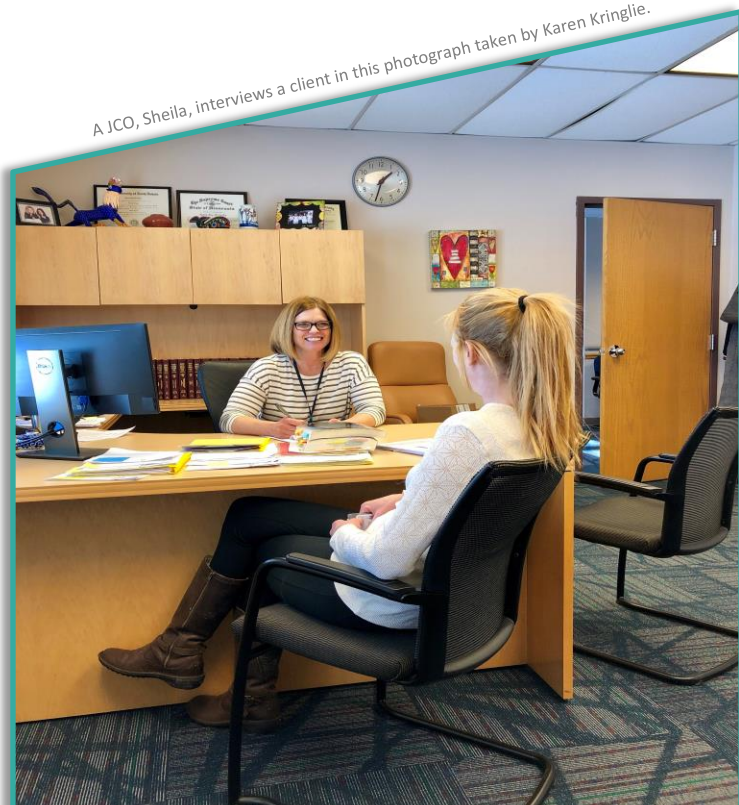
LIST OF FIGURES

INTRODUCTION

Risk assessment is a core component of the judicious prevention of recidivism (i.e., reoffending) in the juvenile justice system (Schwalbe, 2007). Effective assessment has enabled juvenile court officers [JCOs] (a.k.a., juvenile probation officers) to identify and target high risk offenders, more deliberately pooling their resources for those at greatest need of intervention. The use of risk assessment instruments by state juvenile justice systems has increased from 33% in the 1990s to over 85% in the 2000s (Schwalbe, 2007). Though their adoption has been wide-spread, many states have neglected to assess the predictive validity of these instruments. Indeed, confirmation of these instruments' actual utility in predicting recidivism has not received adequate attention.

This report assesses the evaluability of the Youth Assessment and Screening Instrument (YASI), implemented statewide in 2002, using a North Dakota juvenile probation sample. While YASI has been found to possess respectable predictive outcomes in other states and Canadian provinces (Baird et al., 2013; Jones et al., 2016), this is the first attempt to assess its use within North Dakota. We asked four guiding research questions at the outset of the project.

1. To what extent are the data needed to assess the YASI available and retrievable?
2. To what extent does this tool accurately predict the likelihood of recidivism among a sample of the state's juvenile probation population?
3. Does the tool's predictive accuracy differ by sex, race, or region of the state?
4. Are there specific domains or within-domain measures of YASI that contribute more (or less) to its predictive accuracy?



A JCO, Sheila, interviews a client in this photograph taken by Karen Kringlie.

LITERATURE REVIEW

Juvenile Probation in North Dakota

Community supervision is a considerably less expensive alternative to out-of-home placements (Greenwood, 2014). It also allows youth the opportunity to continue attending school and work while abiding by set probation conditions which may include curfew restrictions, community service, restitution, or counseling. YASI assists JCOs in assessing each youth and their likelihood of recidivism, (i.e., risk to public safety) as well as identify criminogenic needs for use in case plan development. It also considers existing protective factors (e.g., positive attitudes towards school, prosocial peers) within its calculation of risk.

As displayed in **Figure 1,** the state of North Dakota has experienced a steady decline in total juvenile referrals over the past eight years (compiled from North Dakota Juvenile Court annual reports from 2010-2018). This is consistent with a larger national trend in which the number of delinquency cases across juvenile courts has dropped 49% from 2005 to 2016 (Hockenberry, 2019). In North Dakota, juvenile probation is administered through the Supreme Court Administrator's Office and seven judicial districts. Probation, historically the most common disposition for delinquency nationally and for the state, has likewise experienced a steady decline from 2,682 juveniles in 2010 to 1,048 in 2018. Beginning in 2016, the use of diversionary programs surpassed probation supervision as the most common sanction in juvenile dispositions.

Note, in recent years the state has experienced an economic recession (Webster, 2016) which contributed to a reduction in the number of JCOs; a decrease from 39 to 36. Despite these changes, the court has been deliberate in maintaining caseload ratios consistent with national recommendations (Burrell, 2006).

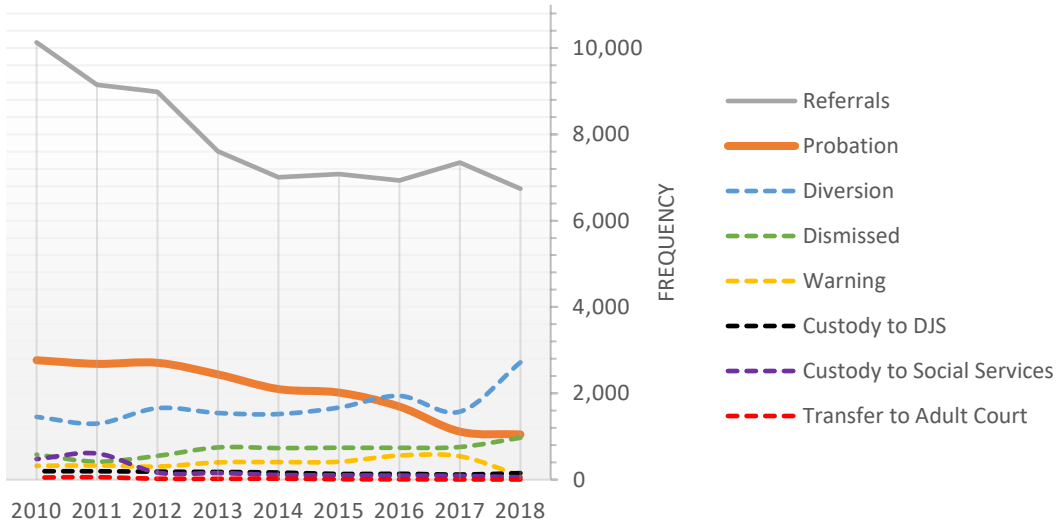


Figure 1. North Dakota Juvenile Court trends in delinquent/unruly referrals and dispositions compiled from 2010-2018 annual reports.

Risk Assessment: An Essential Tool

The National Institute of Corrections (NIC) identified eight evidence-based practices (EBPs) associated with community supervision (Taxman & Belenko, 2012, p. 47);

1. Assess actuarial risk/needs using a standardized instrument(s).
2. Enhance intrinsic motivation.
3. Target interventions.
    a. *Risk Principle*: Prioritize supervision and treatment resources for higher risk offenders.
    b. *Need Principle*: Target interventions to criminogenic needs.
    c. *Responsivity Principle*: Be responsive to temperament, learning style, motivation, culture, and gender when assigning programs.
    d. *Dosage*: Structure 40-70% of high-risk offenders' time for 3-9 months.
    e. *Treatment*: Integrate treatment into the full sentence/sanction requirements.
4. Skill train with directed practice.
5. Increase positive reinforcement.
6. Engage ongoing support in natural communities.
7. Measure relevant processes/practices.
8. Provide measurement feedback.

Risk assessment is a core component of EBP, and also the key to appropriately referring youth to services, ideally evidence-based *programs* (for recommended programs see Greenwood, 2014; Taxman & Belenko, 2012; Weisburd, Farrington, & Gill, 2016) that target dynamic risk factors (Andrews, Bonta, & Hoge, 1990; Latessa, Listwan, & Koetzle, 2015).

Risk assessment involves estimating an individual's likelihood to recidivate. The goal is to identify those at greatest risk and focus rehabilitative services on that specific population. This is not only fiscally responsible, but prior research has demonstrated exposing low-risk youth to intensive services can lead to adverse outcomes (Krysik & LeCroy, 2002; Schwalbe, 2007). A large proportion of delinquent behavior stems from the actions of a small number of juveniles. Accurate assessments allow resources to be directed at this criminogenic population, where the impact is most likely to have a desirable effect. Further, empirical assessments have been shown to be more objective, reliable, and equitable than clinical risk assessments or professional judgements (Duwe & Rocque, 2017; Schwalbe, 2007).

Actuarial risk assessment instruments divide youth into low, medium, or high risk for repeat offending by assigning numerical scores to a series of risk factors known to correlate with subsequent delinquent behaviors (Schwalbe, 2007). Domains of interest include prior criminal history, substance abuse, family relations, peer delinquency, and school-related bonds and performance. The scores associated with these domains, comprising a risk index, are weighted and summed to yield an overall composite risk score. Standardized cut-offs are developed to categorize youth into one of several ordinal classifications ranging from low to high risk. While early-era instruments focused primarily on risk and classification, modern day instruments are also used to guide intervention selections based on need and responsivity considerations (Andrews et al., 1990). This has led to a greater interest in *dynamic* (i.e., alterable) risk factors (e.g., substance abuse, cognitive errors) in addition to *static* risk factors (e.g., criminal history, age, race, gender).

In the past there was resistance to the adoption of actuarial risk assessment instruments, which were viewed as negating one's professional expertise (Schnieder, Ervin, & Snyder-Joy, 1996). However, when implemented well, empirical risk assessments have been shown to be more reliable in predicting actual risk to reoffend (Dawes, Faust, & Meehl, 1989; Duwe & Rocque, 2017; Grove & Meehl, 1996; Schwalbe, 2007; Vincent, Guy, Perrault, & Gershenson, 2016).

The Need for Validation

There are numerous risk assessment instruments available of varying quality and effectiveness (Desmarais & Singh, 2013; Schwalbe, 2007). In more recent years, as the use of risk assessment instruments has permeated the field of probation, more attention has been paid to their predictive validity, especially in comparisons across instruments and between differing populations subjected to a given assessment (Shepherd et al., 2013). Whereas instrument development is based on an estimation sample in which risk factors associated with recidivism are identified and combined to form an index, validation is assessed using a separate sample in which the predictive validity of that index is examined (Krysik & LeCroy, 2002). It must be recognized that not all instruments are developed with the same level of statistical rigor, nor subsequently validated. Examples include adaptations of the Model Risk Assessment Instrument (MRAI) and the North Carolina Assessment of Risk (NCAR) (Howell, 1995; Schwalbe, Fraser, Day, & Arnold, 2004). Though often containing similar risk factors as actuarially developed instruments, such instruments were typically developed through consensus building with juvenile justice professionals (Schwalbe, 2007).

Further, instruments developed more recently have tended to be lengthier, measuring risk factors using large multi-item scales (DeVellis, 2012). The Youth Level of Service/Case Management Inventory (YLS/CMI), for example, examines eight domains of risk using a total of 42 items (Schwalbe, 2007). Others, such as the Washington State Juvenile Court Assessment (WSJCA), utilize alternative matrix scoring procedures and other novel approaches in calculating risk (Barnoski, 2004). Though often measuring similar concepts, their method of development and execution can vary significantly.

Regardless of the approach, the goal is to produce an instrument that possesses high predictive validity (i.e., can accurately predict the risk of recidivism). When instruments possess low levels of predictive validity the information they provide is little better, or even more misleading, than that of subjective professional assessments (Krysik & LeCroy, 2002). Poor predictive validity can be a real concern for agencies that adopt risk assessment instruments from other jurisdictions without subsequent validation (Jones, Harris, Fader, & Grubstein, 2001). By doing so, the agency is making an assumption that what worked in one jurisdiction will work in another. This assumption can be faulty partly due to differences in implementation and fidelity alone (Taxman & Belenko, 2012). One need only to look at the recent debate over attempts to replicate HOPE (Hawaii's Opportunity Probation with Enforcement) in the continental US to see replication without follow-up research can be complicated and risky (Hamilton, Campbell, van Wormer, Kigerl, & Posey, 2016; Lattimore, et al., 2016; O'Connell, Brent, & Visher, 2016).

Systematic reviews and meta-analyses of adult risk assessment validations have generally produced positive outcomes, noting well developed instruments can predict recidivism significantly above chance (Barbaree, Seto, Langton, & Peacock, 2001; Desmarais & Singh, 2013; Gendreau, Goggin, & Smith, 2002; Gendreau, Little, & Goggin, 1996). Similar outcomes have been found for juvenile populations (Schwalbe, 2007). Schwalbe's (2007) systematic review and meta-analysis of juvenile risk assessment instruments included 28 studies. The YLS/CMI was the most commonly researched instrument, present in 11 of the 28 studies. Other instruments examined, for example, included the Wisconsin Juvenile Probation and Aftercare (WJPA) risk instrument (Ashford & LeCroy, 1988), ASSET

(Baker, Jones, Roberts, & Merrington, 2003), the Washington State Juvenile Court Assessment (WSJCA) (Barnoski, 2004), the Psychopathy Checklist-Youth Version (PCL-YV) (Corrado, Vincent, Hart, & Cohen, 2004), the Young Offender Level of Service Inventory (YO-LSI) (Ilacqua, Coulson, Lombardo, & Nutbrown, 1999), the Arizona Risk/Needs Assessment (ARNA), and the Child and Adolescent Functional Assessment Scale (CAFAS) (Quist & Matshuzi, 2000). Schwalbe (2007) found third generation instruments (e.g., YLS/CMI, YO-LSI) that utilized multi-item scales to measure a given construct tended to have improved predictive validity. That said, the YLS/CMI, which was the most heavily researched, possessed some of the highest and lowest effect sizes of all instruments studied, suggesting further research is warranted. With exception to the YLS/CMI, few instruments have been validated across multiple samples. Not only are validation studies needed, periodic re-validation studies were also recommended.

Another issue that has complicated the use of risk assessment instruments has been their applicability to different juvenile populations in terms of gender, race, and ethnicity (Shepherd et al., 2013). Shepherd and colleagues (2013) argue that risk assessment in juvenile justice is still relatively new and the majority of validation research has focused on males because they constitute a large proportion of delinquent and criminal behavior. Their review of the literature concerning the Structured Assessment of Violence Risk in Youth (SAVRY), YLS/CMI, and PCL-YV concluded further investigation is still needed in this regard. Female recidivism, for example, is likely to be influenced by victimization and abuse, depression, self-esteem issues, mental illness, substance abuse, truancy, sexual promiscuity, and relationship or family issues. Variations in coping responses across gender suggests the predictive validity of risk assessment is likely to be impacted by these differing behavioral patterns. Some instruments have been developed specifically to apply to female antisocial behaviors, including the Early Assessment Risk List for Girls (EARL-21G), but little research has been conducted, often due to small sample sizes. Attempts have also been made to examine factors unique to Aboriginal offenders (e.g., chronic criminal histories, pain, anger, and depression passed down through generations) in Australia and Canada, but with no discernable improvement on recidivism prediction. Despite disproportionate representation of minority populations involved in the criminal justice system, little research has attempted to compare the predictive validity of risk assessment instruments across racial and ethnic populations. From the research that does exist, results are often mixed or contradictory (Singh, Grann, & Fazel, 2011; Skeem, Edens, Camp, & Colwell, 2004; Snowden, Gray, & Taylor, 2010).

North Dakota's Risk Assessment Tool: YASI

YASI is a juvenile risk assessment instrument with widespread adoption in the US and UK (Jones et al., 2016). The instrument is derived from the Case Management Assessment Protocol (CMAP) originally used with juveniles in Washington State. Two versions of the YASI exist including a Pre-Screen and Full Assessment version (examples of the Pre-Screen domains and items are contained in the appendices of Jones et al., 2016, pp. 193-194). The Pre-Screen contains 34 items, with the Full Assessment (possessing 90 items) reserved for use with moderate and high risk youths. Whereas the Pre-Screen is primarily used for risk classification, the Full Assessment is intended to provide greater depth to guide case plan development, providing a more substantive needs assessment for prioritizing treatment goals. In either case, the instrument is scored based on the results of semi-

structured interviews, supplemented with information from collateral sources such as input from parents, police files, probation records, school records, and mental health reports. However, only recently has it been subjected to validation research (Jones et al., 2016). Like other actuarial instruments, it was developed to objectively measure indicators associated with future criminal behavior. Predictors of such behavior include eight core domains; antisocial cognitions, antisocial associates, criminal history, antisocial personality factors, substance abuse, family dysfunction, deficits in education, and inappropriate leisure time.

A unique aspect of YASI is the inclusion of what Jones et al. (2016) refers to as "strength-based" items (a.k.a., promotive or protective factors). These are similar to the concept of protective factors associated with adults (e.g., marital status, attachment to employment). For juveniles, specifically, these protective factors include positive temperament, bonds to the school environment (e.g., school attendance), presence of a caring adult mentor in school or in the community, positive peer relationships, educational achievement, positive responses to authority, and effective prosocial uses of leisure time. Though research is very limited, six items pertaining to protective factors on the SAVRY were found to be significantly related to nonviolent recidivism, with no significant impact on the prediction of violent repeat offending (Jones et al., 2016). Several studies have also noted the inclusion of protective factors can enhance and strengthen the accuracy of risk assessment instruments (Jones et al., 2016; Lodewijks, Doreleijers, de Ruiter, & Borum, 2008).

Though YASI is considered a gender-neutral tool, it does include several gender-responsive items adapted from the feminist literature including gender-specific poverty and mental health factors (Jones et al., 2016). In addition to the incorporation of these factors, the YASI also utilizes separate classification cutoff points for delinquent girls. However, no adjustments or additions have been made in terms of racial or ethnic considerations.

Overall, in terms of interrater reliability, Jones and colleagues (2016) report three studies where YASI demonstrates acceptable and, in some cases, very respectable ratings compared to other juvenile risk assessment instruments (Baird et al., 2013; Skeem et al., 2012). Prior research conducted in Illinois and New York State have also provided evidence of "acceptable" predictive validity (i.e., better than random chance) for juveniles placed on probation based on what Rice and Harris (2005) would consider a moderate effect size (Jones et al., 2016). Consistent with prior literature, Jones and colleagues' (2016) study of a Canadian sample of youth found criminal history, community and peer associations, and antisocial attitudes to be the strongest predictors of recidivism. Finally, *Orbis Partners* is currently working on a YASI-Girls adaptation, a tool that will further consider gender-responsive domains and bolster the instrument's predictive accuracy for female populations (with the current instrument shown to be lacking in predictive validity for females compared to males). This instrument, however, has yet to be implemented or evaluated (Jones et al., 2016).

METHODS

Sample

A random sample of 500 cases was extracted from a sampling frame provided by the North Dakota Juvenile Court. The sampling frame contained client identification numbers for 3,754 cases. Archived cases, consisting of another 7,796 cases, were not included in the sampling frame given the requisite data would not be retrievable by external researchers. Further, the court was concerned older cases would not reflect modern practices. These figures represent cases documented within the court's YASI management system known as CASEWORKS. It does not represent all cases contained within the court's case management system (CMS).

Nine cases were immediately removed due to duplication of identifiers. Another 97 cases were omitted due to a lack of retrievable recidivism data. Fourteen cases involved a youth that received an assessment but no subsequent probation supervision. Two cases did not include an assessment or classification outcome. Further, 239 cases possessed supervision end dates that indicated they were either still active or were under less than three years from completion (i.e., insufficient follow-up period). The final sample consisted of 139 clients. In addition, we examined recidivism for one- and two-year follow-up periods that subsequently expanded the sample size to 270 and 209. All cases included in this sample had a probation termination date between 2010 and 2017.

The final sample was predominantly white (66.9%, *n* = 93) and male (64.8%, *n* = 90). The proportion of African Americans (11.5%, *n* = 16) and Native Americans (14.4%, *n* = 20) exceeds that of the state at 3.1% and 5.5%, respectively.

Data Collection, Assessments, and Measures

This study relied solely on secondary data obtained from CASEWORKS, CMS, and Odyssey. Data were manually accessed through a designated terminal at the Grand Forks County Building. Information was recorded into a SPSS database by two graduate research assistants and the principal investigator. Recidivism information was tied to YASI risk assessment data using a commonly shared client identification number. The following is a brief summary of these measures.

Recidivism.

The Juvenile Court possesses an explicit definition for probation recidivism;

> *Youth under community supervision (formal and informal) for a delinquent offense that admit or are adjudicated or convicted within three years of supervision closure/termination and youth under community supervision that are placed with an agency within three years of community supervision closure/termination.*

The court later provided additional clarifications to their probation recidivism definition by adding the following;

Group tracking:

> *Youth supervised (reporting probation) in the community on formal and informal supervision.*

Recidivism events:

> *Informal admission/ adjudication/ convicted – youth admits or are adjudicated as a juvenile or convicted as an adult for a misdemeanor or felony offense within three years of supervision closure/termination.*

Review of probation that receives a disposition of custody to an agency for placement.

Tracking timeframe:

One, two, and three years after supervision closure/termination.

Data sources:

CMS – for juvenile informal admissions, adjudication and commitment dispositions.

Odyssey – for juvenile adjudication and commitment dispositions; for adult convictions for misdemeanors or higher dispositions (retrievable from www.ndcourts.gov/public-access).

Subsequent analyses presented within this report were developed in response to these definitions. However, one additional clarification is needed. For the purposes of this research, we focused solely on recidivism in relation to the first offense that resulted in a term of probation supervision, formal or informal. The research team created multiple dichotomously coded variables to capture the occurrence of a recidivating event at one, two, and three-year intervals. Subsequent analyses focus on the three-year interval, but links to tables within the **APPENDIX** which replicate these analyses at one- and two-year recidivism intervals are also provided. Additionally, the date associated with the recidivating event was recorded. Any offense that led to a formal disposition or conviction was considered a recidivating event. Subsequent punishments may have included another term of probation supervision or other referral. The supervision end date of the original offense was recorded and used to determine the follow-up interval within which recidivism occurred. Offenses that occurred during supervision were documented separately and not considered a recidivating event (occurred prior to supervision termination). Finally, multiple YASI assessments may exist for a single individual within CASEWORKS. The date of the assessment that coincided with the original probationary outcome was utilized.

The YASI Risk Assessment Instrument.

YASI consists of a Pre-Screen and Full Assessment. The Full Assessment includes an additional 50 items that further expand on the 34 items contained within the Pre-Screen. While the Pre-Screen consists of 9 domains (legal history, family, school, community/ peers, alcohol and drugs, mental health, aggression/ violence, attitudes, skills), the Full Assessment adds one additional domain (employment and free time). Items are summed into a total score to classify youth as low, moderate, or high risk. In practice, all referred youth receive the Pre-Screen and in the event the results indicate a moderate or high risk, a Full Assessment is completed. According to Jones and colleagues (2016) cutoff scores for the YASI differ by gender and were derived from a preliminary study conducted in Illinois. The expected range of recidivism for low risk youth is between 10-20%, 30-40% for moderate risk youth, and 50-60% for high risk cases (Jones et al., 2016, p. 185).

The instrument includes a variety of dynamic and static risk factors, but also protective measures. In addition to their contribution to an overall risk score, individuals are also given a protective classification of low, moderate, or high. Protective measures are designed primarily to assist in case planning. As Jones and colleagues (2016) note, it is possible for a single youth to score high on risk and protective factors within a single domain.

RESULTS

Descriptive and Bivariate Analyses

Table 1 provides a descriptive summary of recidivism at three years from probation termination (*N* = 139). The most serious offense associated with a recidivating event was documented and then recoded based on five classifications as recognized by the Juvenile Court; against person, drug related, property, public order, unruly, and traffic. Note, for those that aged out of the juvenile system some offense types no longer apply (as an adult the offense must have been associated with a misdemeanor or felony conviction). The overall recidivism rate was 53% (*n* = 74). Unruly and drug related offenses were the most common cause for recidivism, comprising about 33% of the sample or 66% of the recidivating events. About 12% of the recidivating cases (6% of the total sample) were associated with an offense against a person.

Note, 34% (*n* = 47) of the sample committed a new offense during their initial supervision prior to termination; these were *not* recorded as recidivating events. About 64% (*n* = 30) of these individuals went on to recidivate, while 36% (*n* = 17) desisted. The result is not statistically significant ($\chi^2[1] = 3.20$, $p = .07$, $\varphi_c = .15$), meaning delinquent behavior while on supervision is not necessarily indicative of future recidivism. However, a significant, though weak, result was found at the one- ($\chi^2[1] = 6.09$, $p = .01$, $\varphi_c = .15$) and two-year intervals ($\chi^2[1] = 7.27$, $p = .01$, $\varphi_c = .19$).

In terms of sex, a notably higher proportion of unruly recidivating events were reported for females (35%) compared to males (16%). The overall recidivism rate for females was also 6% higher overall. African American and Native American populations possessed even higher rates of recidivism at 69% and 70% compared to 48% for whites. Native American juvenile recidivism was predominantly characterized by alcohol possession and consumption as reflected by the high percentage of unruly recidivating events. African American youth were more diverse in their recidivating events, comparable to whites but with a higher proportion of public order offenses.

Though not reported in Table 1, note overall recidivism rates were not significantly different, statistically speaking, for males and females ($\chi^2[1] = .46$, $p = .50$, $\varphi_c = .06$), African Americans and whites ($\chi^2[1] = 2.27$, $p = .13$, $\varphi_c = .14$), or Native Americans and whites ($\chi^2[1] = 3.08$, $p = .08$, $\varphi_c = .17$). Finally, recidivism rates did not differ significantly by unit ($\chi^2[3] = 2.65$, $p = .45$, $\varphi_c = .14$). Note, while not statistically significant, unit 1's recidivism rate is notably higher at 66% compared to 47-51% for the other three units. Results remained the same when examining one- and two-year follow-up periods (for a full breakdown of recidivism rates see **Table 6** and **Table 7** of the **APPENDIX**)

Table 2 further differentiates recidivism rates based on the YASI classifications of risk and strength at low, moderate, or high. Similar to **Table 1**, this is broken down by sex, race, and region. Recall the expected range of recidivism based on YASI is 10-20% for low risk, 30-40% for moderate risk, and 50-60% for high risk juveniles (Jones et al., 2016).

Table 1: *Three-Year Recidivism Rates across Sex, Race, and Region by Recidivating Offense*

		Sex		Race			Region			
Recidivating Offense	Overall Sample (N = 139) % (n)	Male (n = 90) % (n)	Female (n = 49) % (n)	African American (n = 16) % (n)	Native American (n = 20) % (n)	White (n = 93) % (n)	Unit 1[a] (n = 32) % (n)	Unit 2[b] (n = 49) % (n)	Unit 3[c] (n = 39) % (n)	Unit 4[d] (n = 19) % (n)
Against Person[e]	6.5 (9)	6.7 (6)	6.1 (3)	12.5 (2)	10.0 (2)	4.3 (4)	9.4 (3)	8.2 (4)	0.0 (0)	10.5 (2)
Drug Related[f]	10.8 (15)	12.2 (11)	8.2 (4)	18.8 (3)	5.0 (1)	11.8 (11)	3.1 (1)	14.3 (7)	12.8 (5)	10.5 (2)
Property[g]	7.9 (11)	10.0 (9)	4.1 (2)	6.3 (1)	10.0 (2)	6.5 (6)	12.5 (4)	6.1 (3)	10.3 (4)	0.0 (0)
Public Order[h]	4.3 (6)	5.6 (5)	2.0 (1)	12.5 (2)	0.0 (0)	4.3 (4)	12.5 (4)	0.0 (0)	5.1 (2)	0.0 (0)
Unruly[i]	22.3 (31)	15.6 (14)	34.7 (17)	18.8 (3)	45.0 (9)	20.4 (19)	25.0 (8)	18.4 (9)	23.1 (9)	26.3 (5)
Traffic[j]	1.4 (2)	1.1 (1)	2.0 (1)	0.0 (0)	0.0 (0)	1.1 (1)	3.1 (1)	2.0 (1)	0.0 (0)	0.0 (0)
Overall	53.2 (74)	51.1 (46)	57.1 (28)	68.8 (11)	70.0 (14)	48.4 (45)	65.6 (21)	50.0 (24)	51.3 (20)	47.4 (9)

Note. Of the 139 cases 47 (33.8%) committed a new offense prior to the termination of their supervision. These events often led to an extension of supervision which altered the original supervision end date. They may have also led to a formal sanction, revocation, or no formal action.
[a] Unit 1 includes the Northeast Judicial District (Benson, Bottineau, Cavalier, McHenry, Pembina, Pierce, Ramsey, Renville, Rolette, Towner, Walsh) and Northeast Central Judicial District (Grand Forks, Nelson). [b] Unit 2 includes the East Central Judicial District (Cass, Steele, Traill) and Southeast Judicial District (Barnes, Dickey, Eddy, Foster, Griggs, Kidder, LaMoure, Logan, McIntosh, Ransom, Richland, Sargent, Stutsman, Wells). [c] Unit 3 includes the Southwest Judicial District (Adams, Billings, Bowman, Dunn, Golden Valley, Hettinger, Slope, Stark) and South Central Judicial District (Burleigh, Emmons, Grant, McLean, Mercer, Morton, Oliver, Sheridan, Sioux). [d] Unit 4 includes the Northwest Judicial District (Divide, McKenzie, Williams) and North Central Judicial District (Burke, Mountrail, Ward).
Recidivating offense classifications were adapted from the 2018 North Dakota Juvenile Court annual report (see p. 11). [e] Against person offenses include all assaults, menacing, harassment, terrorizing, gross sexual imposition, and robbery. [f] Drug related offenses include any form of illegal drug possession excluding tobacco and alcohol possession/consumption. [g] Property offenses include shoplifting, burglary, criminal mischief/vandalism, criminal trespassing, and all thefts. [h] Public order offenses include disorderly conduct, disturbance of a public school, failure to appear, and resisting arrest. [i] Unruly offenses include curfew, runaway, tobacco, truancy, ungovernable behavior, and possession/consumption of alcohol. [j] Traffic offenses include driving without a license, driving without liability, and the leaving the scene of an accident.

Table 2: *Three-Year Recidivism Rates across Sex, Race, and Region by YASI Risk and Strength Classifications*

		Sex				Race								Region					
	Overall Sample (N = 139) % (n)	Male (n = 90) % (n)	Female (n = 49) % (n)	χ^2	φ_c	African American (n = 16) % (n)	χ^2	φ_c	Native American (n = 20) % (n)	χ^2	φ_c	White (n = 93) % (n)	χ^2	φ_c	Unit 1 (n = 32) % (n)	Unit 2 (n = 49) % (n)	Unit 3 (n = 39) % (n)	Unit 4 (n = 19) % (n)	χ^2	φ_c
Risk																				
Low (n = 66)	37.9 (25)	29.7 (11)	48.3 (14)	2.38	.19	66.7 (4)	1.58	.24	55.6 (5)	1.04	.19	31.1 (14)	2.75†	-.20	42.9 (6)	40.7 (11)	36.8 (7)	16.7 (1)	1.54	.15
Moderate (n = 64)	67.2 (43)	67.4 (31)	66.7 (12)	.01	-.01	77.8 (7)	.51	.10	75.0 (6)	.26	.07	65.9 (29)	.10	-.04	87.5 (14)	57.9 (11)	58.8 (10)	66.7 (8)	4.28	.26
High (n = 9)	66.7 (6)	57.1 (4)	100.0 (2)	1.90	.38	100.0 (1)	1.19	-.41	100.0 (3)	2.83†	.55	50.0 (2)	.91	-.32	50.0 (1)	66.7 (2)	100.0 (3)	0.0 (0)	4.87	.65
χ^2	11.91**	12.04**	3.84			2.7			3.07			11.02**			7.23*	1.73	6.00†	5.60†		
φ_c	.29	.36	.25			.40			.34			.34			.46	.19	.35	.51		
Strength																				
Low (n = 23)	65.2 (15)	64.7 (11)	66.7 (4)	.01	.02	66.7 (2)	.00	.00	66.7 (2)	.00	.00	66.7 (10)	.04	.04	71.4 (5)	77.8 (7)	40.0 (2)	50.0 (1)	2.31	.32
Moderate (n = 54)	61.1 (33)	56.4 (22)	73.3 (11)	1.31	.16	88.9 (8)	3.82†	.28	66.7 (6)	.35	.09	55.9 (19)	1.07	-.14	69.2 (9)	64.3 (9)	55.6 (10)	55.6 (5)	.77	.12
High (n = 62)	41.9 (26)	38.2 (13)	46.4 (13)	.42	.08	25.0 (1)	.22	-.07	75.0 (6)	4.17*	.28	36.4 (16)	1.93	-.18	58.3 (7)	30.8 (8)	50.0 (8)	37.5 (3)	3.17	.23
χ^2	5.85†	3.95	3.22			5.28†			.16			5.32†			.46	8.03*	.40	.56		
φ_c	.21	.21	.25			.57			.09			.24			.12	.40	.10	.17		

Note. YASI = Youth Assessment and Screening Instrument. When the assumption of Chi-Square Test is violated (i.e., less than 80% of cells have a count of 5) the Likelihood Ratio Chi-Square statistic is reported **(McHugh, 2013)**.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

The proportion of high-risk youth contained in the sample is small (7%, *n* = 9), with the majority classified as moderate (46%, *n* = 64) or low risk (48%, *n* = 66). The proportion of low risk youth that recidivated (38%, *n* = 25) is high compared to the benchmarks laid out by Jones et al. (2016). The proportion of moderate risk youth that recidivated (67.2%, *n* = 43) is also higher than expected, while the high-risk group is too limited to draw definitive conclusions (66.7%, *n* = 6). Nonetheless, a statistically significant result was obtained, indicating the instrument does discriminate by risk classification ($\chi^2[3] = 11.91$, $p < .01$, $\varphi_c = .29$). Note, Ellis (2010, p. 41) indicates a phi coefficient (φ_c) of .10 is small, .30 medium, and .40 large; indicating a moderate effect. Clearly, those identified as low risk were less likely to reoffend than those deemed moderate or high risk. In relation to sex, we find that the instrument discriminates based on risk more effectively for males ($\chi^2[3] = 12.04$, $p < .01$, $\varphi_c = .36$) than for females ($\chi^2[3] = 3.84$, $p > .10$, $\varphi_c = .25$). Though no statistically significant differences were observed by sex and risk level, the recidivism rate for low risk females was about 18% higher than that of the males.

Results by race should be interpreted with caution due to the low sample sizes. The frequencies indicate that African American and Native American youth are at greater risk to recidivate compared to whites. Similar to males, the instrument effectively discriminates risk level for whites ($\chi^2[3] = 11.02$, $p < .01$, $\varphi_c = .34$) with a moderate effect. In terms of those classified as low risk, whites were significantly less likely to recidivate than other racial groups ($\chi^2[2] = 2.75$, $p < .10$, $\varphi_c = -.20$) but the effect is weak-to-moderate. No significant differences were observed for moderate or high-risk youth by race.

In terms of geographical unit sample size remains a concern at the three-year interval. Units 1, 3, and 4 each possess a statistically significant outcome. However, with exception to low risk youth in unit 4, all units possess recidivism rates for low and moderate risk youth at nearly double the expected rates of 20% and 30%. Though the strength of evidence that the instrument discriminates by unit varies (with effect sizes ranging from .19 [weak-to-moderate] to .51 [strong]), no statistically significant differences were observed when comparing across units.

In terms of strength classification (i.e., protective factors), an overall statistically significant outcome was observed at the three-year interval ($\chi^2[3] = 5.85$, $p < .10$, $\varphi_c = .21$), albeit considerably weaker than that of risk classification. Those with low strength classifications recidivated at a higher rate (65%, *n* = 15) than those with moderate (61%, *n* = 33) or high strength classifications (42%, n = 26). As observed with risk, it appears the instrument does better discriminating among whites ($\chi^2[2] = 5.32$, $p < .10$, $\varphi_c = .24$) than for other racial groups. No significant differences were observed by sex or unit. Only unit 2 was associated with a significant outcome in terms of recidivism by strength classification ($\chi^2[3] = 8.03$, $p < .05$, $\varphi_c = .40$).

Figure 2 visually displays the recidivism rate in relation to risk and strength levels. Of those youth that recidivated within the three-year sample, none were high risk and high strength nor low risk and low strength. Note the higher the protective measure, the lower the recidivism by risk level. The difference is notably small, and the rates were identical for low and moderate strength youth with a high-risk classification. Recall, however, the limited number of high-risk youth contained in the sample.

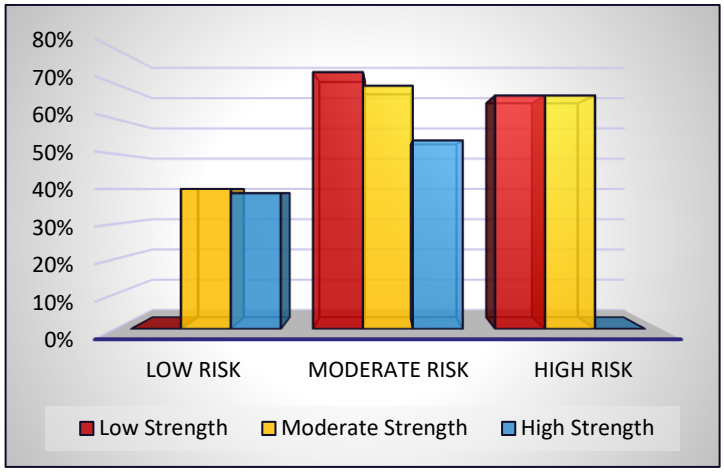


Figure 2. Three-year recidivism by risk and strength classification.

These analyses were subsequently repeated for one- and two-year recidivism intervals. The corresponding figures and tables can be located in the **APPENDIX** (see **Table 8**, **Table 9**, **Figure 6**, and **Figure 7**). Trends remain similar, demonstrating the instrument does discriminate effectively between low, moderate, and high risk to reoffend. However, this appears to hold true for males and whites, but not for females and other specialized populations.

Predictive Validity

The receiver operating characteristics (ROC) area under the curve (AUC) statistic is a diagnostic measure used to assess predictive accuracy (Hamilton, Kowalski, Schaefer, & Kigerl, 2019). It is widely recognized in the literature and has been applied across many fields to a variety of tests in which the probability of a predicted outcome can be verified (Georgiou, 2019). In other words, it produces an effect size representing the likelihood an instrument will correctly predict an outcome. The AUC value ranges from zero to one, with a value of .50 indicating the instrument performed no better than if one were to rely on random chance. Outcomes can be classified into four categories; negligible (<.56), small (.56-.63), moderate (.64-.70), and large (>.71) (Hamilton et al, 2019, p. 8; Rice & Harris, 2005).

In many respects, **Table 3** reflects the mixed findings introduced in **Table 2** and **Figure 2**. This sample provides some evidence of the instrument's predictive validity when applied to the North Dakota population (AUC = .66, p < .01, 95% CI [.56, .75]), with an overall moderate effect size. Notably, each unit demonstrates a moderate to large effect size, but the sample size is small and the confidence intervals wide. Alternatively, a small non-significant effect is observed for females. In terms of race, while the effect sizes are large the confidence intervals are wide and imprecise (as well as non-significant). Any subsample with 50 or fewer cases should be interpreted with caution.

Table 3: *Predictive Accuracy of YASI at Three-Year Recidivism Interval*

Sample	AUC	(95% CI)
Overall sample (*N* = 139)	.66**	(.56, .75)
Male (*n* = 90)	.68**	(.57, .80)
Female (*n* = 49)	.63	(.47, .78)
African American (*n* = 16)	.76	(.43, .99)
Native American (*n* = 20)	.73	(.48, .97)
White (*n* = 93)	.68**	(.56, .78)
Unit 1 (*n* = 32)	.72*	(.51, .93)
Unit 2 (*n* = 49)	.72**	(.58, .87)
Unit 3 (*n* = 39)	.74*	(.58, .90)
Unit 4 (*n* = 19)	.64	(.38, .90)

Note. AUC = Area Under the Curve.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

When repeated for one- and two-year recidivism intervals the predictive accuracy of the instrument drops considerably. At the one-year interval the overall effect becomes small, bordering negligible (see **APPENDIX**; **Table 10**, **Table 11**, **Table 12**, **Table 13**, and **Table 14**). Some notable trends, however, persist. Predictive accuracy is consistently poor for females, and an imprecise and unwieldly outcome continues to be observed for

specialized populations. The instrument clearly performs best for males and whites.

Exploratory Domain Analyses

Though it is not possible to assess the internal consistency of a given domain or its holistic contribution to the prediction of recidivism without access to YASI's raw score data and algorithms (raw case weights are unknown), the following analyses take an exploratory look at select items and their relative predictive power. We use the raw data, consisting of many dichotomously coded variables, to pursue further bivariate and multivariate analyses. In many cases, items are associated with such little variation that they are unusable for subsequent analysis. Any item with variation less than or equal to 10% were removed.

Table 4 displays the frequency or mean and standard deviation for all items with adequate variation and minimal missing data as well as the Pearson correlation in association with recidivism.

At the bivariate level age at first offense possesses the strongest correlation (r = -.294, p < .001), followed by poor academic performance (r = .212, p < .05), noncompliance with parental rules (r = .208, p < .05), prior status offenses (r = .195, p < .05), negative behaviors in school (r = .188, p < .05), and number of runaways (r = .179, p < .05). The Pearson correlation effect sizes all range from small-to-moderate (Ellis, 2010, p. 41). For one- and two-year results see **Table 15** and **Table 16** in the **APPENDIX**. At two years, age at first offense (r = -.147, p < .05) and negative behaviors at school (r = .195, p < .01) are statistically significant. Only prior victimization of property theft (r = -.130, p < .05) was significant at the one-year interval. Note, negative coefficients indicate a reduction in recidivism (i.e., indirect relationship or negative relationship).

Table 4: *Select YASI Domain Items and Pearson Correlations with Three-Year Recidivism*

Item	*N*	*n/M*	*%/SD*	*r*
Legal History				
Prior probation referrals	139	68	48.9%	.081
Age at first offense	139	*M* = 13.23	*SD* = 2.27	-.294***
Prior status offenses	139	57	41.0%	.195*
Prior felony referrals	139	25	18%	-.087
Prior against-person misdemeanor referrals	139	49	35.3%	-.033
Family				
Number of runaways	139	*M* = .40	*SD* = 1.23	.179*
Prior court finding of neglect	139	19	13.7%	.079
Noncompliance with parental rules[a]	139	*M* = 1.73	*SD* = .85	.208*
School				
Truancy in last three months[b]	134	*M* = 1.57	*SD* = 1.09	.065
Negative behaviors in school last three months[c]	134	*M* = 2.73	*SD* = 1.36	.188*
Poor academic performance in last three months[d]	133	*M* = 2.63	*SD* = 1.13	.212*
Community/Peers				
Presence of prosocial peers	139	108	77.7%	-.052
Presence of antisocial peers	139	86	61.9%	.066
Alcohol and Drugs				
Alcohol and drug use in last three months	139	74	53.2%	.104
Mental Health				
Mental health problems in last three months	139	48	34.5%	-.077
Suicidal ideation (thoughts and attempts)	139	33	23.7%	.116
History of physical abuse from parent	139	15	10.8%	.001
Victim of bullying	139	33	23.7%	-.019
Victim of physical assault	139	20	14.4%	.014
Victim of property theft	139	15	10.8%	-.046
Aggression/Violence				
Bullying	139	31	22.3%	.086
Destruction of property	139	17	12.2%	-.002
Assaultive behavior	139	50	36.0%	.011
Attitudes				
Defies accepting responsibility[e]	137	*M* = 2.04	*SD* = 1.01	.158
Skills				
Lack of consequential thinking skills[f]	137	*M* = 2.24	*SD* = .94	.084

[a] Response options of (1) youth usually obeys and follows rules, (2) youth sometimes obeys or obeys some rules, (3) youth often disobeys rules, (4) youth consistently disobeys, and/or is hostile, and (5) no pro-social rules in place.
[b] Response options of (1) attends regularly, (2) some partial day unexcused absences, (3) some full-day unexcused absences, (4) five or more full-day unexcused absences.
[c] Response options of (1) positive behavior, (2) no problems reported, (3), infractions reported, (4) intervention by school administration, (5) police reports filed by school.
[d] Response options of (1) B+ or above, (2), C or better, (3), C- or lower, (4) failing some classes, (5) failing most classes.
[e] Response options of (1) voluntarily accepts full responsibility for delinquent/criminal behavior, (2) recognizes that he or she must accept responsibility, (3) indicates some awareness of the need to accept responsibility, (4) minimizes, denies, justifies, excuses, or blames others.
[f] Response options of (1) acts to obtain good and avoid bad consequences, (2) can identify specific consequences or his/her actions, (3) understands there are good and bad consequences, (4) sometimes confused about consequences of actions.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Table *5* reports logistic regression results for the three-year interval. Nine cases were dropped due to listwise deletion (i.e., missing data). Of the 130 cases included in the model 69 recidivated within three years (53%). Using this specific combination of variables the statistical model was able to correctly predict 80% of those that failed and 70% of those that desisted. McFadden's R^2 indicates these variables collectively explain about 26% of the pseudo-variance in the dependent variable (Pampel, 2000). The model was statistically significant ($\chi^2[25] = 46.011$, $p = .006$). Note, one-year and two-year intervals were associated with non-significant models (therefore, not reported). The Hosmer & Lemeshow goodness of fit test was not significant for the three-year interval, suggesting the model does provide an adequate fit to the theoretical model of perfect prediction ($\chi^2[8] = 8.035$, $p = .430$). Four statistically significant variables were identified;

- *Age at first offense.* The higher the age the lower the likelihood of recidivism. Specifically, with each year higher in age at first offense the odds of recidivism decreases by 41.3%.
- *Noncompliance with parental rules.* Each one-point increase on this five-point measure is associated with a 2.72 times increased likelihood of recidivism.
- *Poor academic performance in last three months.* Each one-point increase on this five-point measure is associated with an 82% increased likelihood of recidivism.
- *Suicidal ideation.* Youth that exhibit signs of suicidal ideation are 81% less likely to recidivate.
- *Lack of consequential thinking skills.* Each one-point increase on this four-point measure is associated with a 61% reduction in the likelihood of recidivism.

This model is provided for demonstrable purposes. It explains little of the pseudo-variance in recidivism. The sample size is limited, suggesting only large effect sizes will be detected (Cohen, 1992; see also Vittinghoff & McCulloch, 2007). Interactions were not considered, and specification is clearly an issue given a majority of the variables/items contained in the instrument were omitted. Further, items were not weighted. There is a heightened risk of a Type II error, meaning some variables that are in fact significant may not be detected, a consequence of low statistical power.

Table 5: *Logistic Regression with Select YASI Domain Items with Three-Year Recidivism*

Item	*B*	S.E.	Wald	*p*	Odds
Legal History					
Prior probation referrals	.344	.508	.460	.497	1.411
Age at first offense	-.532	.151	.125	.001	.587
Prior status offenses	-.167	.516	.104	.747	.847
Prior felony referrals	.743	.600	1.536	.215	2.103
Prior against-person misdemeanor referrals	.804	.632	1.621	.203	2.235
Family					
Number of runaways	.108	.252	.183	.669	1.114
Prior court finding of neglect	.135	.784	.029	.864	1.144
Noncompliance with parental rules	1.001	.410	5.950	.015	2.720
School					
Truancy in last three months	-.328	.258	1.621	.203	.720
Negative behaviors in school last three months	.342	.190	3.251	.071	1.408
Poor academic performance in last three months	.597	.280	4.524	.033	1.816
Community/Peers					
Presence of prosocial peers	-.262	.691	.143	.705	.770
Presence of antisocial peers	-.445	.535	.692	.405	.641
Alcohol and Drugs					
Alcohol and drug use in last three months	.209	.522	.161	.688	1.233
Mental Health					
Mental health problems in last three months	.974	.635	2.354	.125	2.650
Suicidal ideation (thoughts and attempts)	-1.663	.748	4.938	.026	.190
History of physical abuse from parent	.712	.786	.820	.365	2.037
Victim of bullying	.408	.616	.439	.507	1.504
Victim of physical assault	.331	.809	.167	.682	1.393
Victim of property theft	-.035	.776	.002	.964	.966
Aggression/Violence					
Bullying	.077	.622	.015	.902	1.080
Destruction of property	.130	.797	.027	.870	1.139
Assaultive behavior	-.880	.670	1.725	.189	.415
Attitudes					
Defies accepting responsibility	.247	.327	.569	.451	1.280
Skills					
Lack of consequential thinking skills	-.962	.349	7.417	.006	.386

$N = 130$.
$\chi^2[25] = 46.011$, $p = .006$.
McFadden's Pseudo $R^2 = .256$.
Hosmer & Lemeshow goodness of fit $p = .430$.

Failure Analysis

Cox regression represents a proportional hazards model that considers time until a given event occurs (Garson, 2013). A time variable was created using the recidivism date and the termination end date to compute the number of days from the end of supervision to the recidivating event. There were two supervision end dates that appeared in CMS. One was labeled anticipated supervision end date and another was coined case closure date. In many cases these dates were identical. However, on occasion these may differ, in which case we relied on the actual case closure date. Using this time variable, a failure rate can be plotted for the entire sample (not limited to a dichotomous variable at one-, two-, or three-year intervals). In addition, covariates can be introduced to assess their influence on the outcome in terms of temporal immediacy of a recidivating event.

As depicted in **Figure 3**, the probability of a recidivating event occurring is highest shortly after supervision termination. The greatest probability of a recidivating event is within the first 50-100 days. The likelihood of a recidivating event drops considerably after 30-60 days and by 200 days the probability of a recidivating event drops to less than 50%. Stated differently, the odds of survival reaches nearly 75% after the first 365 days. Clearly, the first 100-200 days after supervision, or about 3-6 months, is a critical time period in which the greatest risk of recidivism will occur.

Subsequent failure analyses examine the contribution of the YASI risk and strength classifications (see Figure 4 and Figure 5). All regression analyses produce nonsignificant models, indicating the trend remains comparable regardless of classification. Though not displayed, this holds true when examining time to recidivism by sex, race, and unit as well.

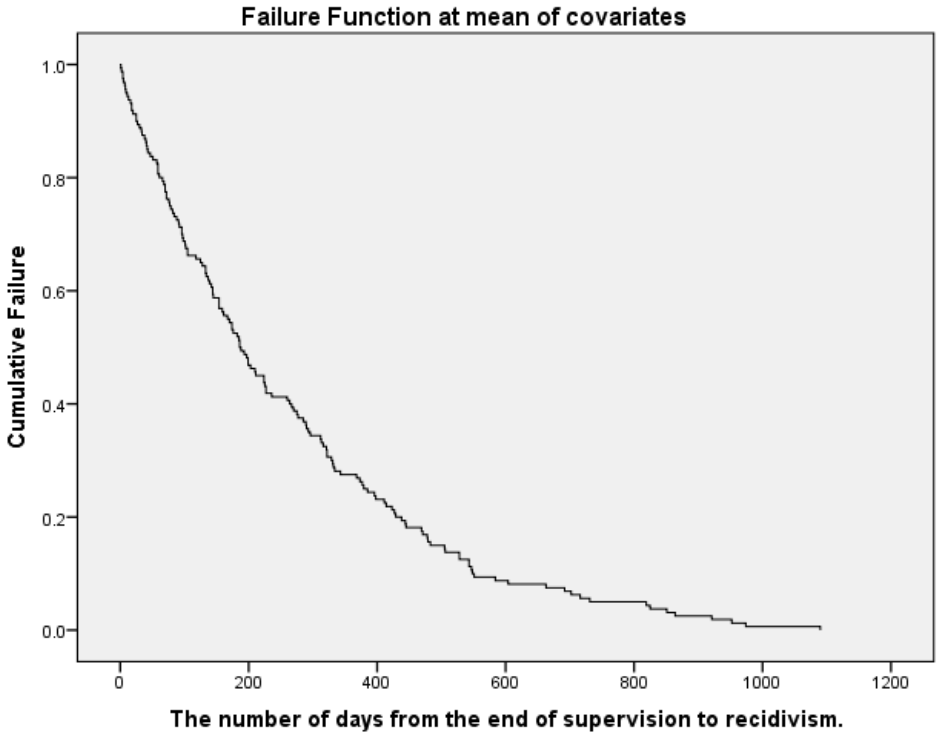


Figure 3. Baseline survival/failure probabilities for recidivism.

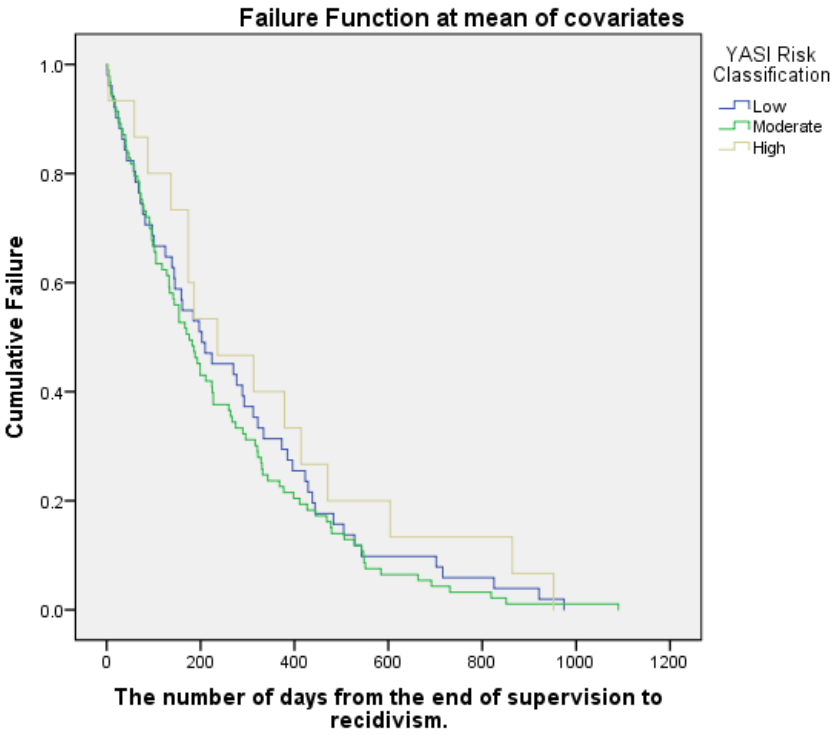


Figure 4. Survival/failure probabilities for recidivism by YASI risk classification.

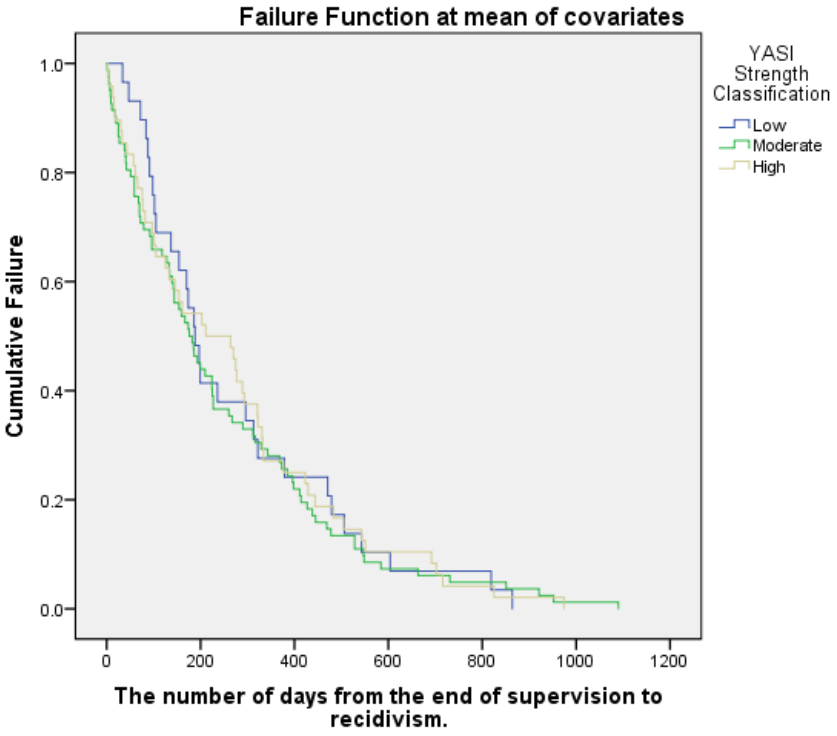


Figure 5. Survival/failure probabilities for recidivism by YASI strength classification.

DISCUSSION

This section starts by responding to each of the four research questions. Limitations and recommendations then follow.

1. *To what extent are the data needed to assess the YASI available and retrievable?*

As described previously, multiple data systems were accessed to retrieve the needed recidivating and assessment information. These sources were linked via a unique identification number. Permissions and online trainings were needed to obtain credentials and these systems could only be accessed by a specific terminal at the closest juvenile probation office, in this case the Grand Forks County Building. Data were manually collected by reviewing CMS, CASEWORKS, Odyssey, and entering the required information into a separate SPSS database. The database contained over 300 variables, due in large part to the number of contingent matrices utilized within YASI (e.g., alcohol and drug abuse matrices). Early in the project several attempts to transcribe a single case were timed and found to range from 20 minutes to one hour depending on the case's complexity, with an average of about 30 minutes. Assuming the information sought was present, collecting the information was straight forward, though time-consuming.

In terms of YASI and CASEWORKS, however, only the data entered by the JCOs could be retrieved. It is clear that YASI employs a weighting system within the computation of its risk assessment scores unique to a given domain (Georgiou, 2019; Jones et al., 2016). The total of these raw scores are associated with cutoff points that are used to determine the respective risk classification (low, moderate, or high). Indeed, Jones and colleagues (2016) indicate that these cutoff scores differ for males and females (p. 185). We did not have access to this weighting information, and to our knowledge it is not publicly available elsewhere. Indeed, it is likely proprietary. As a result, replication was limited to overall outcomes as denoted by risk and strength classifications.

2. *To what extent does this tool accurately predict the likelihood of recidivism among a sample of the state's juvenile probation population?*

Overall, the three-year sample produced the strongest results. While each recidivism interval produced overall significant findings the effect sizes ranged from small to moderate (Rice & Harris, 2005). Jones and colleagues' (2016) Canadian sample, by comparison, achieved a large effect size (note, however, their recidivism measure was based on rearrest at 18 months). By their own admission, however, many prior implementations of YASI were associated with AUCs "in the mid-.60 range" which would be consistent with the results of this study (Jones et al., 2016, p. 189).

As reflected within the bivariate analyses, the high percentage of recidivism among low and moderate risk youth is concerning. It appears youth are being under-classified. In addition, strength measures appear to have limited predictive value. It is unclear if this is due to a lack of emphasis placed on protective factors or if these measures generally contribute less to risk prediction overall. Jones et al. (2016), for example, explains these measures are primarily meant to enhance case plan development. Yet their analyses show greater discrimination amongst strength categories in relation to recidivism. On the other hand, this study demonstrates weak discrimination.

3. *Does the tool's predictive accuracy differ by sex, race, or region of the state?*

In terms of subgroup analyses, several notable patterns emerged. First, under no condition did a significant finding emerge for females. Indeed, the instrument performed best for males and whites. As highlighted in the literature, this has been a consistent problem in juvenile risk assessment (Shepherd et al., 2013). In the case of African Americans and Native Americans results were mixed due in part to the low sample achieved for each respective subpopulation. Finally, all units possessed a moderate-to-large effect size at the three-year interval. Though statistical significance varied, the effect sizes were generally small-to-moderate for each regardless of the time-interval examined.

Taken together, these outcomes indicate the instrument does have the potential to be effective in accurately predicting recidivism among the North Dakota juvenile probation population. That said, its lack of effectiveness for females is concerning, and further research is needed on other specialized populations.

4. *Are there specific domains of YASI that contribute more (or less) to its predictive accuracy?*

We were unable to assess specific domains. The absence of weighting and raw risk score data prohibited the examination of each domain and its unique contribution to overall risk prediction. While new scales could be developed using the collected data for these domains, it would not reflect what was used in determining risk within CASEWORKS (i.e., such analysis wouldn't contribute to the validation of the instrument).

Exploratory analyses were conducted, however, on select YASI items that possessed adequate variation and few missing cases. Such results should be interpreted with caution, but five items emerged as particularly pertinent to three-year recidivism outcomes. First, the younger the onset of delinquency the greater the risk (see Mason & Windle, 2001). Second, high ratings of noncompliance with parental rules increased the likelihood of recidivism. Third, high rating of poor academic performance was also associated with greater risk. Fourth, youth with symptoms of suicidal ideation were significantly less of a risk. No question individuals who display such behaviors have significant needs (Lachal, Massimiliano, Sibeoni, Moro, & Revah-Levy, 2015), but they are not a risk to public safety. Finally, a somewhat contrary finding was discovered in terms of lack of consequential thinking skills. Higher ratings were associated with a reduced likelihood of recidivism. However, this may indicate an opportunity for cognitive correction, whereas the alternative may be indicative of psychopathy (Latessa et al., 2015).

Limitations

First, statistical power is a concern (Cohen, 1992). While this sample size is adequate for conducting an overall assessment of predictive validity, it becomes problematic for subgroup analyses. As a result of low statistical power, there is a greater risk of failing to find a significant outcome when one does actually exist (a.k.a., Type II error). This is of greatest concern for the African American and Native American populations which were notably small even with the larger one-year sample.

Second, the raw risk scores and individual domain item weights were not readily available for retrieval, therefore they could not be examined. As a result, only the final risk and strength classifications could be assessed. In addition, we were unable to replicate the internal consistency ratings as reported in Table 2 of Jones et al. (2016, p. 186). It is unclear which items specifically Jones and colleagues used for each domain when computing their

alpha levels. It is also unknown what weighting was employed or what subsequent recoding or other adjustments may have been performed.

Finally, this study made no attempt to address concerns of interrater reliability. That is, no assessment was conducted on how reliable JCOs were in conducting the YASI consistently. However, several changes in practice have occurred within the past year. These changes include additional trainings, online modules, as well as consolidation of who conducts the assessments. Most of the units have now limited assessments to one or two JCOs with specialized training. These changes were too recent to be reflected within the current study and may lead to considerable improvements in predictive validity and reliability in future validation studies.

Recommendations

First, some concerns exist about the documentation of new assessments relative to reassessment. Inconsistencies were reported indicating that in some cases JCOs replaced data in existing assessments as opposed to creating a new assessment. It is highly recommended that all assessments be documented independently. Of further note, while the current study focused on the initial offense and assessment, subsequent research should examine the impact of reassessments associated with later offenses.

In terms of implementation fidelity, the Juvenile Court should continue to provide YASI trainings for those JCOs that will specialize in its administration. The empirical literature has established that training is critical to extracting the full potential of a risk assessment instrument (Flores et al., 2006; Latessa & Lovins, 2010; Latessa et al., 2015; Taxman & Belenko, 2012)

The data collection process for future research on YASI could be more efficient if streamlined for continual validation checks, at least once every five years. First, a process for obtaining a full sampling frame is needed that includes variables for stratification based on sex, race, and unit. Note, there is an immediate need to examine the instrument's validity further among African American and Native American populations, given the current study performed poorly in this regard.

The current study utilized a manually coded database. By streamlining this process so that individual level data can be extracted and imported through a more automated method larger sample sizes could be obtained, increasing the power of future analyses without increasing the costs associated with manual data collection.

The juvenile court should consider re-weighting items or altering cutoff scores if subsequent validation studies fail to show improved predictive accuracy (Georgiou, 2019). It is not unusual for instruments to be calibrated to reflect actual risk propensities. This involves adjusting raw cutoff scores which dictate whether an individual is low, moderate, or high risk. Tools, such as PACT in Iowa, have been subsequently modified after adoption to improve predictive accuracy (Hamilton et al., 2019). Individual item weighting also matters, Georgiou's (2019) research on the LSI-R has demonstrated correct weighting can improve the predictive validity of an instrument significantly.

Finally, implementation of supplemental measures to enhance risk prediction for females should be explored and considered (Van Voorhis et al., 2010). Though in preliminary planning phases and unvalidated, *Orbis Partners* has developed a YASI-Girls (YASI-G) instrument which includes measures concerning relationships, emotional expression, self-efficacy, sexual vulnerability, and early parenthood (Jones et al., 2016, pp.

190-191). Another example of a female-specific instrument includes the Early Assessment Risk List for Girls (EARL-21G) (Shepherd et al., 2013). As Van Voorhis and colleagues' (2010) research notes, while existing instruments can predict recidivism for females it can be strengthened with the inclusion of gender-responsive factors. One pressing difficulty, however, is the continued lack of validations of the female-specific measures they reference. Indeed, there appears to be an ongoing and intense debate about the validity of gender-neutral measures with female populations, with some showing assessments such as the YLS perform equally well for males and females (Pusch & Holtfreter, 2018). This does not appear to be the case for YASI given the results in this study and in Jones et al.'s (2016) recent research. While there is indication it may be predictive to some extent for females, it is considerably weaker than that observed for males. Further research, deliberation, and long-term planning will be required to address the needs of this distinct population.

REFERENCES

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*(1), 19-52. Retrievable from https://doi.org/10.1177/0093854890017001004

Ashford, J. B., & LeCroy, C. W. (1988). Predicting recidivism: An evaluation of the Wisconsin Juvenile Probation and Aftercare risk instrument. *Criminal Justice and Behavior, 15*(2), 141-151. Retrievable from https://doi.org/10.1177/0093854888015002002

Baird, C., Healy, T., Johnson, K., Bogie, A., Dankert, E. W., & Scharenbroch, C. (2013). *A comparison of risk assessment instruments in juvenile justice.* National Council on Crime and Delinquency. Retrievable from https://nicic.gov/comparison-risk-assessment-instruments-juvenile-justice

Baker, K., Jones, S., Roberts, C., & Merrington, S. (2003). *The evaluation of the validity and reliability of the Youth Justice Board's assessment for young offenders: Findings from the first two years of ASSET.* Oxford: University of Oxford, Centre for Criminological Research. Retrievabe from https://www.ncjrs.gov/App/publications/Abstract.aspx?id=233981

Barbaree, H. E., Seto, M. C., Langton, C. M., & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior, 28*(4), 490-521. Retrievable from https://doi.org/10.1177/0093854801028004006

Barnoski, R. (2004). *Assessing risk for re-offense: Validating the Washington State Juvenile Court Assessment.* Olympia: Washington State Institute for Public Policy. Retrievable from https://www.scribd.com/document/11202103/Assessing-Risk-for-Re-Offense-Validating-the-Washington-State-Juvenile-Court-Assessment

Burrell, W. D. (2006). *Caseload standards for probation and parole.* Lexington, KY: Council of State Governments, American Probation and Parole Association. Retrievable from https://nicic.gov/caseload-standards-probation-and-parole-2006

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. Retrievable from http://dx.doi.org/10.1037/0033-2909.112.1.155

Corrado, R. R., Vincent, G. M., Hart, S. D., & Cohen, I. M. (2004). Predictive validity of the psychopathy checklist: Youth version for general and violent recidivism. *Behavioral Sciences & the Law, 22*(1), 5-22. Retrievable from https://doi.org/10.1002/bsl.574

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical vs. actuarial judgement. *Science, 243*, 1668-1674. Retrievable from https://doi.org/10.1126/science.2648573

Desmarais, S. L., & Singh, J. P. (2013). *Risk assessment instruments validated and implemented in correctional settings in the United States.* New York, NY: Council of State Governments Justice Center. Retrievable from https://csgjusticecenter.org/reentry/publications/risk-assessment-instruments-validated-and-implemented-in-correctional-settings-in-the-united-states/

DeVellis, R. F. (2012). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: Sage. Retrievable from https://us.sagepub.com/en-us/nam/scale-development/book246123

Duwe, G., & Rocque, M. (2017). Effects of automating recidivism risk assessment on reliability, predictive validity, and return on investment (ROI). *Criminology & Public Policy, 16*(1), 235-269. Retrievable from https://doi.org/10.1111/1745-9133.12270

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. New York, NY: Cambridge. Retrievable from https://doi.org/10.1017/CBO9780511761676

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. Retrievable from https://doi.org/10.1016/j.jcrimjus.2006.09.007

Garson, G. D. (2013). *Cox regression.* Asheboro, NC: Statistical Associates Publishers. Retrievable from http://www.statisticalassociates.com/coxregression.htm

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the "unparralled" measure of offender risk? A lesson in knowledge accumulation. *Criminal Justice and Behavior, 29*(4), 397-426. Retrievable from https://doi.org/10.1177/0093854802029004004

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34*(4), 557-607. Retrievable from https://doi.org/10.1111/j.1745-9125.1996.tb01220.x

Georgiou, G. (2019). Weights matter: Improving the predictive validity of risk assessments for criminal offenders. *Journal of Offender Rehabilitation, 58*(2), 92-116. https://doi.org/10.1080/10509674.2018.1562504

Greenwood, P. (2014). *Evidence-based practice in juvenile justice: Progress, challenges, and opportunities.* New York, NY: Springer. Retrievable from https://www.springer.com/gp/book/9781461489078

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Pyschology, Public Policy, and Law, 2*(2), 293-323. Retrievable from http://dx.doi.org/10.1037/1076-8971.2.2.293

Hamilton, Z., Campbell, C. M., van Wormer, J., Kigerl, A., & Posey, B. (2016). Impact of swift and certain sanctions: Evaluation of Washington State's policy for offenders on community supervision. *Criminology & Public Policy, 15*(4), 1009-1072. Retrievable from https://doi.org/10.1111/1745-9133.12245

Hamilton, Z., Kowalski, M. A., Schaefer, R., & Kigerl, A. (2019). Recrafting youth risk assessment: Developing the Modified Positive Achievement Change Tool for Iowa. *Deviant Behavior*. Advance online publication. Retrievable from https://doi.org/10.1080/01639625.2019.1609302

Hockenberry, S. (2019). *Delinquency cases in juvenile court, 2016.* Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention. Retrievable from https://www.ojjdp.gov/pubs/252473.pdf

Howell, J. C. (Ed.). (1995). *Guide for implementing the comprehensive strategy for serious, violent, and chronic juvenile offenders.* Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrievable from https://www.ncjrs.gov/pdffiles/guide.pdf

Ilacqua, G. E., Coulson, G. E., Lombardo, D., & Nutbrown, V. (1999). Predictive validity of the Young Offender Level of Service Inventory for criminal recidivism of male and female young offenders. *Psychological Reports, 84*(3), 1214-1218. Retrievable from https://doi.org/10.2466/pr0.1999.84.3c.1214

Jones, N. J., Brown, S. L., Robinson, D., & Frey, D. (2016). Validity of the youth assessment and screening instrument: A juvenile justice tool incorporating risks, needs, and strengths. *Law and Human Behavior, 40*(2), 182-194. Retrievable from http://dx.doi.org/10.1037/lhb0000170

Jones, P. R., Harris, P. W., Fader, J., & Grubstein, L. (2001). Identifying chronic juvenile offenders. *Justice Quarterly, 18*(3), 479-507. Retrievable from https://doi.org/10.1080/07418820100094991

Krysik, J., & LeCroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism among juvenile offenders. *Research on Social Work Practice, 12*(1), 71-81. Retrievable from https://doi.org/10.1177/104973150201200106

Lachal, J., Orri, M., Sibeoni, J., Moro, M. R., & Revah-Levy, A. (2015). Metasynthesis of youth suicidal behaviours: Perspectives of youth, parents, and health care professionals. *PLoS ONE, 10*(5). Retrievable from https://doi.org/10.1371/journal.pone.0127359

Latessa, E. J., Listwan, S., & Koetzle, D. (2015). *Work works (and doesn't) in reducing recidivism.* Routledge. Retrievable from https://www.elsevier.com/books/what-works-and-doesn-t-in-reducing-recidivism/latessa/978-1-4557-3121-3

Latessa, E. J., & Lovins, B. (2010). The role of offender risk assessment: A policy maker guide. *Victims & Offenders, 5*(3), 203-219. Retrievable from https://doi.org/10.1080/15564886.2010.485900

Lattimore, P. K., MacKenzie, D. L., Zajac, G., Dawes, D., Arsenault, E., & Tueller, S. (2016). Outcome findings from the HOPE demonstration field experiment: Is swift, certain, and fair an effective supervision strategy? *Criminology & Public Policy, 15*(4), 1103-1141. Retrievable from https://doi.org/10.1111/1745-9133.12248

Lodewijks, H. P., Doreleijers, T. A., de Ruiter, C., & Borum, R. (2008). Predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) during residential treatment. *International Journal of Law and Psychiatry, 31*(3), 263-271. Retrievable from https://doi.org/10.1016/j.ijlp.2008.04.009

Mason, W. A., & WIndle, M. (2001). Delinquency risk as a function of number of early onset problem behaviors. *International Journal of Offender Therapy and Comparative Criminology, 45*(4), 436-448. Retrievable from https://doi.org/10.1177/0306624X01454004

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica, 23*(2), 143-149. Retrievable from http://dx.doi.org/10.11613/BM.2013.018

North Dakota Juvenile Court. (2010). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2010.pdf

North Dakota Juvenile Court. (2011). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2011.pdf

North Dakota Juvenile Court. (2012). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2012.pdf

North Dakota Juvenile Court. (2013). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2013.pdf

North Dakota Juvenile Court. (2014). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2014.pdf

North Dakota Juvenile Court. (2015). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2015.pdf

North Dakota Juvenile Court. (2016). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2016.pdf

North Dakota Juvenile Court. (2017). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2017.pdf

North Dakota Juvenile Court. (2018). *Annual report.* Bismarck, ND. Retrieved from https://www.ndcourts.gov/Media/Default/Trial%20Courts/Juvenile%20Court/juvenile-court-annual-report-2018.pdf

O'Connell, D. J., Brent, J. J., & Visher, C. A. (2016). Decide your time: A randomized trial of a drug testing and graduated sanctions program for probationers. *Criminology & Public Policy, 15*(4), 1073-1102. Retrievable from https://doi.org/10.1111/1745-9133.12246

Pampel, F. C. (2000). *Logistic regression: A primer.* Thousand Oaks, CA: Sage. Retrievable from https://us.sagepub.com/en-us/nam/logistic-regression/book10146

Pusch, N., & Holtfreter, K. (2018). Gender and risk assessment in juvenile offenders: A meta-analysis. *Criminal Justice and Behavior, 45*(1), 56-81. Retrievable from https://doi.org/10.1177/0093854817721720

Quist, R. M., & Matshuzi, D. G. (2000). The child and adolescent functional assessment scale (CAFAS): A dynamic predictor of juvenile recidivism. *Adolescence, 35*(137), 181-192. Retrievable from https://www.ncbi.nlm.nih.gov/pubmed/10841305

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior, 29*(5), 615-620. Retrievable from http://dx.doi.org/10.1007/s10979-005-6832-7

Schnieder, J. C., Ervin, L., & Snyder-Joy, Z. (1996). Further exploration of the flight from discretion: The role of risk/need instruments in probation supervision decisions. *Journal of Criminal Justice, 24*(2), 109-121. Retrievable from https://doi.org/10.1016/0047-2352(95)00059-3

Schwalbe, C. S. (2007). Risk assessment for juvenile justice: A meta-analysis. *Law and Human Behavior, 31*(5), 449-462. Retrievable from http://dx.doi.org/10.1007/s10979-006-9071-7

Schwalbe, C. S., Fraser, M. W., Day, S. H., & Arnold, E. M. (2004). North Carolina Assessment of Risk (NCAR): Reliability and predictive validity with juvenile offenders. *Journal of Offender Rehabilitation, 40*(1-2), 1-22. Retrievable from https://doi.org/10.1300/J076v40n01_01

Shepherd, S. M., Luebbers, S., & Dolan, M. (2013). Gender and ethnicity in juvenile risk assessment. *Criminal Justice and Behavior, 40*(4), 388-408. Retrievable from https://doi.org/10.1177/0093854812456776

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513. Retrievable from https://doi.org/10.1016/j.cpr.2010.11.009

Skeem, J. L., Edens, J. F., Camp, J., & Colwell, L. H. (2004). Are there ethnic differences in levels of psychopathy? A metaanalysis. *Law and Human Behavior, 28*(5), 505-527. Retrievable from http://dx.doi.org/10.1023/B:LAHU.0000046431.93095.d8

Skeem, J., Hernandez, I., Kennealy, P., & Rich, J. (2012). *CA-YASI reliability: How adequately do staff in California's Division of Juvenile Justice rate youths' risk of recidivism?* Irvine, CA: University of California. Retrievable from http://risk-resilience.berkeley.edu/

Snowden, R. J., Gray, N. S., & Taylor, J. (2010). Risk assessment for future violence in individuals from an ethnic minority group. *International Journal of Forensic Mental Health, 9*(2), 118-123. Retrievable from https://doi.org/10.1080/14999013.2010.501845

Taxman, F. S., & Belenko, S. (2012). *Implementing evidence-based practices in community corrections and addiction treatment.* New York, NY: Springer. Retrievable from https://link.springer.com/book/10.1007/978-1-4614-0412-5

Van Voorhis, P., Wright, E. M., Salisbury, E., & Bauman, A. (2010). Women's risk factors and their contributions to existing risk/needs assessment: The current status of a gender-responsive supplement. *Criminal Justice and Behavior, 37*(3), 261-288. Retrievable from https://doi.org/10.1177/0093854809357442

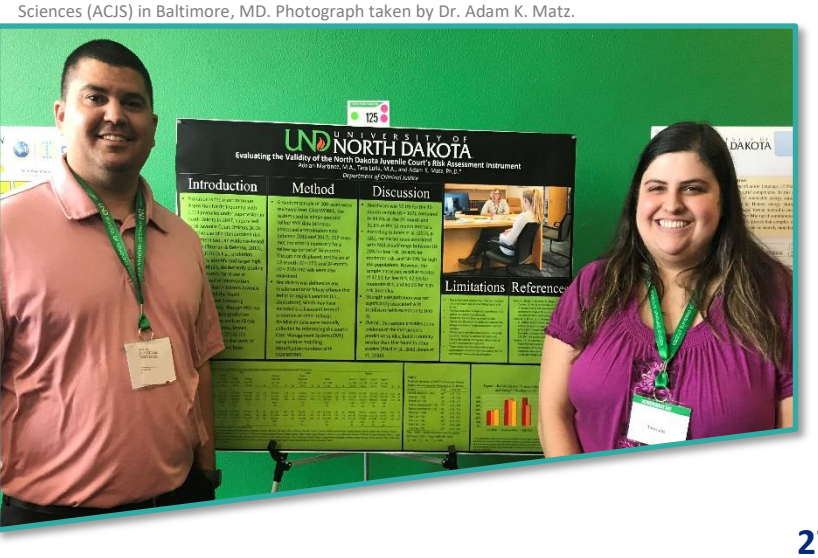
Vincent, G. M., Guy, L. S., Perrault, R. T., & Gershenson, B. (2016). Risk assessment matters, but only when implemented well: A multisite study in juvenile probation. *Law and Human Behavior, 40*(6), 683-696. Retrievable from http://dx.doi.org/10.1037/lhb0000214

Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and cox regression. *American Journal of Epidemiology, 165*(6), 710-718. Retrievable from https://doi.org/10.1093/aje/kwk052

Webster, D. (2016, August 1). North Dakota in recession. *Prairie Public News*. Retrieved from https://news.prairiepublic.org/post/north-dakota-recession

Weisburd, D., Farrington, D. P., & Gill, C. (2016). *What works in crime prevention and rehabilitation: Lessons from systematic reviews.* New York, NY: Springer. Retrievable from https://www.springer.com/gp/book/9781493934751

Doctoral students Adrian Martinez and Tara Lulla presented an early look at this research at UND's 2019 Graduate Research Achievement Day (GRAD). Mr. Martinez received special recognition as one of 10 finalists under the professional, social sciences, humanities, and arts subfields. This research was also presented at the 2019 annual meeting of the Academy of Criminal Justice Sciences (ACJS) in Baltimore, MD. Photograph taken by Dr. Adam K. Matz.



APPENDIX

Table 6: *One-Year Recidivism Rates across Sex, Race, and Region by Recidivating Offense*

		Sex		Race			Region			
Recidivating Offense	Overall Sample (*N* = 270) % (*n*)	Male (*n* = 176) % (*n*)	Female (*n* = 94) % (*n*)	African American (*n* = 36) % (*n*)	Native American (*n* = 32) % (*n*)	White (*n* = 186) % (*n*)	Unit 1 (*n* = 62) % (*n*)	Unit 2 (*n* = 98) % (*n*)	Unit 3 (*n* = 80) % (*n*)	Unit 4 (*n* = 30) % (*n*)
Against Person	3.3 (9)	4.5 (8)	1.1 (1)	5.6 (2)	3.1 (1)	2.7 (5)	4.8 (3)	3.1 (3)	1.3 (1)	6.7 (2)
Drug Related	7.8 (21)	8.0 (14)	7.4 (7)	13.9 (5)	6.3 (2)	7.0 (13)	4.8 (3)	7.1 (7)	10.0 (8)	10.0 (3)
Property	5.2 (14)	6.3 (11)	3.2 (3)	8.3 (3)	0.0 (0)	4.8 (9)	8.1 (5)	5.1 (5)	3.8 (3)	3.3 (1)
Public Order	3.0 (8)	2.8 (5)	3.2 (3)	8.3 (3)	3.1 (1)	2.2 (4)	8.1 (5)	2.0 (2)	1.3 (1)	0.0 (0)
Unruly	11.9 (32)	8.5 (15)	18.1 (17)	11.1 (4)	34.4 (11)	9.1 (17)	16.1 (10)	9.2 (9)	11.3 (9)	13.3 (4)
Traffic	3.3 (9)	2.8 (5)	4.3 (4)	2.8 (1)	3.1 (1)	3.2 (6)	4.8 (3)	3.1 (3)	2.5 (2)	3.3 (1)
Other	1.5 (4)	0.6 (1)	3.2 (3)	0.0 (0)	0.0 (0)	1.6 (3)	1.6 (1)	1.0 (1)	2.5 (2)	0.0 (0)
Overall	35.5 (97)	33.5 (59)	40.4 (38)	50.0 (18)	50.0 (16)	30.6 (57)	48.4 (30)	30.6 (30)	32.5 (26)	36.7 (11)

Note. Of the 270 cases 99 (36.7%) committed a new offense prior to the termination of their supervision.

Table 7: *Two-Year Recidivism Rates across Sex, Race, and Region by Recidivating Offense*

		Sex		Race			Region			
Recidivating Offense	Overall Sample (*N* = 209) % (*n*)	Male (*n* = 136) % (*n*)	Female (*n* = 73) % (*n*)	African American (*n* = 28) % (*n*)	Native American (*n* = 28) % (*n*)	White (*n* = 140) % (*n*)	Unit 1 (*n* = 48) % (*n*)	Unit 2 (*n* = 71) % (*n*)	Unit 3 (*n* = 63) % (*n*)	Unit 4 (*n* = 27) % (*n*)
Against Person	4.8 (10)	5.1 (7)	4.1 (3)	10.7 (3)	3.6 (1)	4.3 (6)	6.3 (3)	5.6 (4)	0.0 (0)	11.1 (3)
Drug Related	11.5 (24)	11.8 (16)	11.0 (8)	14.3 (4)	7.1 (2)	12.1 (17)	2.1 (1)	15.5 (11)	12.7 (8)	14.8 (4)
Property	8.1 (17)	8.8 (12)	6.8 (5)	10.7 (3)	7.1 (2)	7.1 (10)	12.5 (6)	7.0 (5)	7.9 (5)	3.7 (1)
Public Order	2.9 (6)	2.9 (4)	2.7 (2)	7.1 (2)	0.0 (0)	2.9 (4)	8.3 (4)	1.4 (1)	1.6 (1)	0.0 (0)
Unruly	18.2 (38)	13.2 (18)	27.4 (20)	14.3 (4)	39.3 (11)	16.4 (23)	22.9 (11)	14.1 (10)	19.0 (12)	18.5 (5)
Traffic	3.3 (7)	3.7 (5)	2.7 (2)	0.0 (0)	0.0 (0)	4.3 (6)	4.2 (2)	2.8 (2)	3.2 (2)	3.7 (1)
Other	1.4 (3)	0.7 (1)	2.7 (2)	0.0 (0)	0.0 (0)	1.4 (2)	2.1 (1)	0.0 (0)	3.2 (2)	0.0 (0)
Overall	50.2 (105)	46.3 (63)	57.5 (42)	57.1 (16)	57.1 (16)	48.6 (68)	58.3 (28)	46.5 (33)	47.6 (30)	51.9 (14)

Note. Of the 209 cases 77 (36.8%) committed a new offense prior to the termination of their supervision.

Table 8: *One-Year Recidivism Rates across Sex, Race, and Region by YASI Risk and Strength Classifications*

		Sex				Race									Region					
	Overall Sample (N = 270) % (n)	Male (n = 180) % (n)	Female (n = 93) % (n)	χ^2	φ_c	African American (n = 37) % (n)	χ^2	φ_c	Native American (n = 34) % (n)	χ^2	φ_c	White (n = 185) % (n)	χ^2	φ_c	Unit 1 (n = 62) % (n)	Unit 2 (n = 97) % (n)	Unit 3 (n = 83) % (n)	Unit 4 (n = 31) % (n)	χ^2	φ_c
Risk																				
Low (n = 114)	28.1 (32)	22.6 (14)	34.6 (18)	2.03	.13	36.4 (4)	.72	.09	40.0 (4)	1.09	.11	24.1 (20)	2.39	-.15	40.9 (9)	32.6 (14)	18.6 (8)	16.7 (1)	4.59	.20
Moderate (n = 141)	41.8 (59)	39.6 (40)	47.5 (19)	.73	.07	59.1 (13)	4.44†	.19	57.9 (11)	3.58*	.18	34.7 (33)	6.04*	-.21	56.8 (21)	29.2 (14)	47.1 (16)	36.4 (8)	7.20†	.23
High (n = 15)	26.7 (4)	23.1 (3)	50.0 (1)	.58	.21	33.3 (1)	.08	.08	33.3 (1)	.08	.08	25.0 (2)	.02	-.04	0.0 (0)	14.3 (1)	66.7 (2)	50.0 (1)	5.06	.56
χ^2	5.75†	5.64†	1.64			1.90			1.22			2.49			5.50†	1.09	8.74*	1.17		
φ_c	.15	.18	.13			.23			.19			.16			.27	.10	.33	.19		
Strength																				
Low (n = 41)	36.6 (15)	35.5 (11)	40.0 (4)	.07	.04	50.0 (4)	.43	.11	25.0 (1)	.23	-.08	37.0 (10)	.01	.01	42.9 (6)	29.4 (5)	16.7 (1)	75.0 (3)	4.23	.32
Moderate (n = 127)	37.0 (47)	30.7 (27)	51.3 (20)	4.92*	.20	55.0 (11)	3.79†	.19	43.8 (7)	.87	.09	31.8 (27)	3.03†	-.15	52.0 (13)	33.3 (15)	39.0 (16)	18.8 (3)	5.03	.19
High (n = 102)	32.4 (33)	33.3 (19)	31.1 (14)	.06	-.02	37.5 (3)	.61	.09	66.7 (8)	8.02**	.32	24.3 (18)	7.94**	-.28	47.8 (11)	25.0 (9)	27.3 (9)	40.0 (4)	4.06	.20
χ^2	.58	.28	3.53			.71			2.66			1.90			.31	.67	1.97	4.79†		
φ_c	.05	.04	.19			.14			.28			.10			.07	.08	.15	.40		

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 9: *Two-Year Recidivism Rates across Sex, Race, and Region by YASI Risk and Strength Classifications*

		Sex				Race									Region					
	Overall Sample (N = 209) % (n)	Male (n = 136) % (n)	Female (n = 73) % (n)	χ^2	φ_c	African American (n = 28) % (n)	χ^2	φ_c	Native American (n = 28) % (n)	χ^2	φ_c	White (n = 140) % (n)	χ^2	φ_c	Unit 1 (n = 48) % (n)	Unit 2 (n = 71) % (n)	Unit 3 (n = 63) % (n)	Unit 4 (n = 27) % (n)	χ^2	φ_c
Risk																				
Low (n = 92)	37.0 (34)	25.5 (13)	51.2 (21)	6.46*	.27	44.4 (4)	.38	.07	50.0 (5)	.95	.11	33.8 (22)	.92	-.10	45.0 (9)	39.4 (13)	33.3 (11)	16.7 (1)	2.01	.14
Moderate (n = 104)	61.5 (64)	60.8 (45)	63.3 (19)	.06	.02	68.8 (11)	.34	.06	60.0 (9)	.00	-.01	60.9 (42)	.04	-.02	72.0 (18)	54.5 (18)	59.3 (16)	63.2 (12)	1.92	.14
High (n = 13)	46.2 (6)	36.4 (4)	100.0 (2)	3.52†	.46	33.3 (1)	.23	-.16	66.7 (2)	.23	.16	50.0 (3)	.07	.07	33.3 (1)	20.0 (1)	100.0 (3)	50.0 (1)	6.35†	.63
χ^2	11.87**	15.60***	3.29			2.18			.37			9.81**			4.21	3.01	8.69*	4.21		
φ_c	.24	.34	.19			.28			.12			.27			.29	.20	.34	.38		
Strength																				
Low (n = 33)	57.6 (19)	56.5 (13)	60.0 (6)	.04	.03	50.0 (3)	.54	-.14	50.0 (2)	.39	-.13	66.7 (14)	1.95	.24	54.5 (6)	66.7 (8)	33.3 (2)	75.0 (3)	2.42	.27
Moderate (n = 90)	55.6 (50)	49.2 (31)	70.4 (19)	3.43†	.20	73.3 (11)	2.08	.17	50.0 (7)	.03	-.02	52.6 (30)	.54	-.08	63.2 (12)	51.9 (14)	58.1 (18)	46.2 (6)	1.14	.11
High (n = 86)	40.7 (35)	36.0 (18)	47.2 (17)	1.09	.11	28.6 (2)	.20	-.05	70.0 (7)	3.81†	.23	37.1 (23)	1.19	-.12	55.6 (10)	31.3 (10)	38.5 (10)	50.0 (5)	3.24	.19
χ^2	4.84†	3.29	3.41			4.15			1.07			6.38*			.30	5.32†	2.74	1.09		
φ_c	.15	.16	.22			.38			.19			.21			.08	.27	.21	.20		

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

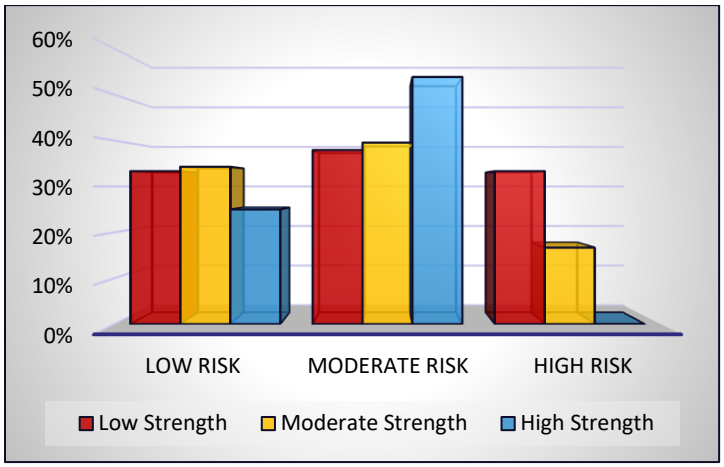


Figure 6. One-Year Recidivism by Risk and Strength Classifications.

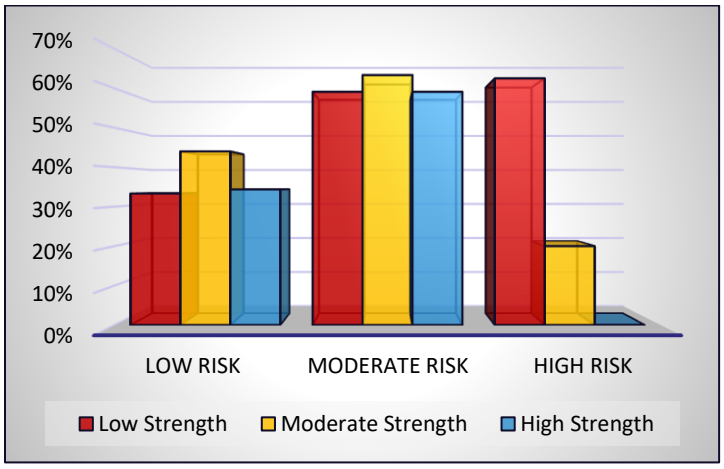


Figure 7. Two-Year Recidivism by Risk and Strength Classifications.

Table 10: *Predictive Accuracy of YASI at Three-Year Recidivism Interval Limited to Assessments within the Past Five Years*

Sample	AUC	(95% CI)
Overall sample (*N* = 76)	.65*	(.54, .78)
Male (*n* = 52)	.68*	(.53, .83)
Female (*n* = 24)	.58	(.34, .81)
African American (*n* = 12)	.89*	(.68, .99)
Native American (*n* = 13)	.71	(.41, .99)
White (*n* = 46)	.64	(.47, .81)
Unit 1 (*n* = 13)	.77	(.50, .99)
Unit 2 (*n* = 28)	.70†	(.51, .90)
Unit 3 (*n* = 23)	.67	(.44, .91)
Unit 4 (*n* = 12)	.64	(.31, .97)

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 11: *Predictive Accuracy of YASI at One-Year Recidivism Interval*

Sample	AUC	(95% CI)
Overall sample (*N* = 270)	.56†	(.49, .63)
Male (*n* = 176)	.58*	(.49, .67)
Female (*n* = 94)	.59	(.48, .71)
African American (*n* = 36)	.54	(.34, .74)
Native American (*n* = 32)	.70†	(.51, .88)
White (*n* = 186)	.57	(.48, .66)
Unit 1 (*n* = 62)	.54	(.39, .68)
Unit 2 (*n* = 98)	.59	(.46, .72)
Unit 3 (*n* = 80)	.72**	(.60, .85)
Unit 4 (*n* = 30)	.60	(.37, .82)

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 12: *Predictive Accuracy of YASI at One-Year Recidivism Interval Limited to Assessments within the Past Five Years*

Sample	AUC	(95% CI)
Overall sample (*N* = 205)	.58†	(.49, .67)
Male (*n* = 136)	.61*	(.51, .71)
Female (*n* = 69)	.61	(.48, .74)
African American (*n* = 31)	.57	(.36, .79)
Native American (*n* = 25)	.77*	(.59, .96)
White (*n* = 138)	.55	(.45, .65)
Unit 1 (*n* = 43)	.60	(.43, .77)
Unit 2 (*n* = 76)	.56	(.40, .71)
Unit 3 (*n* = 64)	.72**	(.57, .87)
Unit 4 (*n* = 22)	.63	(.38, .89)

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 13: *Predictive Accuracy of YASI at Two-Year Recidivism Interval*

Sample	AUC	(95% CI)
Overall sample (*N* = 209)	.61**	(.53, .69)
Male (*n* = 136)	.64**	(.55, .73)
Female (*n* = 73)	.60	(.47, .73)
African American (*n* = 28)	.61	(.38, .85)
Native American (*n* = 28)	.63	(.42, .84)
White (*n* = 140)	.65**	(.56, .74)
Unit 1 (*n* = 48)	.63	(.46, .79)
Unit 2 (*n* = 71)	.68**	(.56, .81)
Unit 3 (*n* = 63)	.69*	(.55, .82)
Unit 4 (*n* = 27)	.66	(.45, .87)

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 14: *Predictive Accuracy of YASI at Two-Recidivism Interval Limited to Assessments within the Past Five Years*

Sample	AUC	(95% CI)
Overall sample (*N* = 144)	.61*	(.51, .70)
Male (*n* = 96)	.65*	(.53, .76)
Female (*n* = 48)	.56	(.39, .72)
African American (*n* = 23)	.64	(.38, .90)
Native American (*n* = 21)	.62	(.38, .87)
White (*n* = 92)	.60	(.48, .71)
Unit 1 (*n* = 29)	.67	(.47, .88)
Unit 2 (*n* = 49)	.59	(.43, .75)
Unit 3 (*n* = 47)	.64	(.46, .81)
Unit 4 (*n* = 19)	.63	(.36, .90)

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 15: *Select YASI Domain Items and Pearson Correlations with One-Year Recidivism*

Item	*N*	*n/M*	*%/SD*	*r*
Legal History				
Prior probation referrals	270	138	51.1%	.038
Age at first offense	270	*M* = 13.74	*SD* = 2.29	-.107
Prior status offenses	270	113	41.9%	.067
Prior felony referrals	270	51	18.9%	-.098
Prior against-person misdemeanor referrals	270	84	31.1%	-.026
Family				
Number of runaways	269	*M* = .41	*SD* = 1.12	.012
Prior court finding of neglect	269	42	15.6%	.025
Noncompliance with parental rules	269	*M* = 1.77	*SD* = .87	.008
School				
Truancy in last three months	247	*M* = 1.53	*SD* = 1.00	.044
Negative behaviors in school last three months	246	*M* = 2.72	*SD* = 1.34	.097
Poor academic performance in last three months	245	*M* = 2.55	*SD* = 1.14	.050
Community/Peers				
Presence of prosocial peers	270	211	78.1%	-.042
Presence of antisocial peers	270	180	66.7%	.077
Alcohol and Drugs				
Alcohol and drug use in last three months	270	176	65.2%	.034
Mental Health				
Mental health problems in last three months	270	101	37.4%	-.073
Suicidal ideation (thoughts and attempts)	270	67	24.8%	.026
History of physical abuse from parent	270	26	9.6%	-.030
Victim of bullying	270	63	23.3%	-.040
Victim of physical assault	270	40	14.8%	-.023
Victim of property theft	270	29	10.7%	-.130*
Aggression/Violence				
Bullying	270	54	20.0%	.078
Destruction of property	270	32	11.9%	-.078
Assaultive behavior	270	85	31.5%	.035
Attitudes				
Defies accepting responsibility	268	*M* = 2.15	*SD* = .99	.018
Skills				
Lack of consequential thinking skills	268	*M* = 2.28	*SD* = .90	.064

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

Table 16: *Select YASI Domain Items and Pearson Correlations with Two-Year Recidivism*

Item	*N*	*n/M*	*%/SD*	*r*
Legal History				
Prior probation referrals	209	108	51.7%	.043
Age at first offense	209	*M* = 13.60	*SD* = 2.30	-.147*
Prior status offenses	209	86	41.1%	.101
Prior felony referrals	209	40	19.1%	-.095
Prior against-person misdemeanor referrals	209	63	30.1%	-.028
Family				
Number of runaways	208	*M* = .38	*SD* = 1.10	-.004
Prior court finding of neglect	208	30	14.4%	.055
Noncompliance with parental rules	208	*M* = 1.74	*SD* = .84	.052
School				
Truancy in last three months	194	*M* = 1.60	*SD* = 1.07	.036
Negative behaviors in school last three months	194	*M* = 2.72	*SD* = 1.35	.195**
Poor academic performance in last three months	193	*M* = 2.58	*SD* = 1.13	.119
Community/Peers				
Presence of prosocial peers	209	167	79.9%	-.050
Presence of antisocial peers	209	140	67.0%	.109
Alcohol and Drugs				
Alcohol and drug use in last three months	209	130	62.2%	.085
Mental Health				
Mental health problems in last three months	209	76	36.4%	-.056
Suicidal ideation (thoughts and attempts)	209	54	25.8%	.068
History of physical abuse from parent	209	21	10.0%	-.014
Victim of bullying	209	51	24.4%	-.053
Victim of physical assault	209	35	16.7%	.015
Victim of property theft	209	23	11.0%	-.044
Aggression/Violence				
Bullying	209	43	20.6%	.014
Destruction of property	209	21	10.0%	-.046
Assaultive behavior	209	67	32.1%	-.027
Attitudes				
Defies accepting responsibility	207	*M* = 2.12	*SD* = 1.00	.059
Skills				
Lack of consequential thinking skills	207	*M* = 2.26	*SD* = .93	.064

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