

Predictive Validity of the MAYSI-2 and PAI-A for Suicide-Related Behavior and Non-Suicidal
Self-Injury among Adjudicated Adolescent Offenders on Probation

Catherine S. Shaffer Erik M. D. Gulbransen Jodi L. Viljoen Ronald Roesch
Kevin S. Douglas
Simon Fraser University

Criminal Justice and Behavior

DOI: <https://doi.org/10.1177/0093854818784988>

Author Note:

Catherine S. Shaffer, Erik M. D. Gulbransen, Jodi L. Viljoen, Ronald Roesch, Kevin S. Douglas, Department of Psychology, Simon Fraser University.

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada, and a Career Investigator Award to Jodi L. Viljoen from the Michael Smith Foundation for Health Research. The views expressed herein are those of the authors and do not necessarily reflect the views or policies of the funding agency. We thank three anonymous reviewers whose comments helped improve and clarify this manuscript.

Correspondence concerning this article should be addressed to Catherine S. Shaffer, Department of Psychology, Simon Fraser University, Burnaby, BC V5A 1S6. Contact: cshaffer@sfu.ca

Abstract

This prospective study evaluated the ability of the MAYSI-2 and PAI-A to predict suicide-related behavior (SRB) and non-suicidal self-injury (NSSI) among adjudicated adolescent offenders on probation. Predictive validity of the MAYSI-2 for SRB and NSSI has generally been postdictively examined among detained adolescents. In addition, no published studies have examined the predictive validity of the PAI-A for SRB and NSSI among adolescent offenders. Neither the MAYSI-2 nor PAI-A added incremental predictive validity above lifetime SRB or NSSI. However, several MAYSI-2 and PAI-A subscales were predictive of SRB or NSSI. With some exceptions, most recommended instrument cut-off scores differentiated between low-risk and high-risk youth. These findings suggest that the MAYSI-2 and PAI-A hold promise for evaluating SRB and NSSI among justice-involved youth. In addition, these findings contribute to more informed decisions regarding the use of these tools and can be used to inform SRB and NSSI prevention efforts.

Keywords: adolescent offenders, MAYSI-2, non-suicidal self-injury, PAI-A, suicide-related behavior

Predictive Validity of the MAYSI-2 and PAI-A for Suicide-Related Behavior and Non-Suicidal Self-Injury among Adjudicated Adolescent Offenders on Probation

Suicide-related behavior (SRB), defined as any behavior undertaken with the deliberate intent to end life (e.g., plans, attempts, completed suicide), and non-suicidal self-injury (NSSI), defined as any self-destructive behavior undertaken to damage or harm oneself without the intention of dying (e.g., cutting, burning, biting; Nock, Joiner, Gordon, Lloyd-Richardson, & Prinstein, 2006), are major health concerns among justice-involved youth (Dixon-Gordon, Harrison & Roesch, 2012; Hayes, 2004). SRB and NSSI among justice-involved youth have been attributed, in part, to elevated rates of mental disorder and emotional distress (Teplin, Abram, McClelland, Dulcan, & Mericle, 2002). Compared to the general population of adolescents, justice-involved youth have higher prevalence rates of depression, anxiety, anger problems, somatic complaints, trauma, substance use, and borderline personality traits. Features of these disorders, such as suicidal ideation, hopelessness, emotional dysregulation, and impulsivity, can lead to SRB (Hayes, 2004) or NSSI (Dixon-Gordon et al., 2012). In addition, comorbid disorders are common (Teplin et al., 2002), with the risk for SRB and NSSI increasing with multiple diagnoses (Fleischmann et al., 2005).

The evaluation of mental health needs on entry to the youth justice system may assist in identification of youth at risk for SRB or NSSI. However, it is unclear how justice agencies should best identify mental health needs in adolescents. Although clinician-administered tools are available for assessing mental health problems in youth (e.g., the Diagnostic Interview Schedule for Children, Version Four, Shaffer et al., 2000), these tools may be inappropriate for routine use in youth justice settings (National Action Alliance for Suicide Prevention, 2013). For instance, comprehensive clinician-rated measures require trained and experienced clinical staff with expertise in scoring and interpretation. Moreover, clinician-rated measures can be time-consuming to administer and score and therefore may be difficult or impractical to implement for every youth. Although briefer assessment tools have been developed (e.g., the Beck Depression Inventory-II; Beck, Steer, & Brown, 1996), these tools often focus on a single problem area and thus are too limited in scope to identify the full range of mental health problems that can occur in adolescence.

To address these concerns, standardized, multidimensional self-report screening measures of mental health have been developed. These measures are intended to identify adolescents in need of a more comprehensive clinician-administered evaluation, or to screen out relatively lower risk adolescents who do not require resource-intensive assessments. In addition, these measures can assist in determining whether monitoring (e.g., suicide monitoring) is required.

Some of these tools, such as the Massachusetts Youth Screening Instrument-Second Version (MAYSI-2; Grisso & Barnum, 2006), a 15-minute screening measure of mental health, were specifically developed for use with justice-involved youth. The MAYSI-2 currently is used in at least 2,000 probation and detention facilities in 47 U.S. states and has demonstrated reliability, validity, and clinical utility (Grisso et al., 2012). Elevated scores (i.e., above “Caution” or “Warning” cut-offs) on the MAYSI-2 are used to identify adolescent offenders in need of further assessment or monitoring. The MAYSI-2 contains 52 items in separate subscales tapping depression and anxiety, anger, thought impairment, somatic complaints, substance use,

and trauma. Thus, it can account for a substantial number of mental health needs related to SRB and NSSI. In addition, the MAYSI-2 includes a subscale tapping suicidal ideation, which provides a starting point on the MAYSI-2 for evaluating SRB potential.

Other self-report measures, such as the Personality Assessment Inventory-Adolescent (PAI-A; Morey, 2007), a downward adaptation of an adult tool, the Personality Assessment Inventory (Morey, 1991), to adolescents, are not designed specifically for use with justice-involved youth, but have aspects that are appealing within youth justice settings, such as validity indices that can be used to detect social desirability, malingering, and other types of response bias. The PAI-A contains measures of depression, anxiety, somatic complaints, suicide ideation, trauma exposure, borderline personality traits, alcohol problems, and drug problems. Like the MAYSI-2, elevated scores (i.e., above “Possible Problems” or “Marked Difficulty” cut-offs) on the PAI-A can be used to identify youth with difficulties in the scale’s content area that warrant further evaluation or monitoring. However, compared to the MAYSI-2, the PAI-A is more comprehensive and time-consuming to administer; it is comprised of 264 items and takes approximately 30 to 45 minutes for a youth to complete.

Despite the arguable utility of the MAYSI-2 and PAI-A for identifying youth at risk to engage in SRB or NSSI, limited empirical research has evaluated the prospective predictive validity of these tools. Such research on the MAYSI-2 is critical given its widespread and recommended use as a screening measure for emergent SRB (National Action Alliance for Suicide Prevention, 2013). In addition, given that the PAI-A incorporates measures of response bias, it is important to establish its predictive utility for SRB and NSSI. For instance, despite its lengthy administration time, the PAI-A may be more effective at screening adolescent offenders who, due to biased responding, may be incorrectly identified on other self-report screening measures such as the MAYSI-2.

To the best of our knowledge, only three published studies have examined associations between MAYSI-2 subscales scores and SRB or NSSI. Wasserman and colleagues (2004) found that boys in custody ($n = 325$) who scored above the Caution cut-off of MAYSI-2 Suicidal Ideation were significantly more likely to self-report prior suicide attempts (odds ratio [OR] = 6.12). In another study, Butler, Loney, and Kistner (2007) found that scores on Angry-Irritable ($r = .40$), Depressed-Anxious ($r = .25$), Somatic Complaints ($r = .38$), and Suicide Ideation ($r = .31$), but not Alcohol/Drug Use, Thought Disturbance, and Traumatic Experience, on the MAYSI-2 significantly predicted placement on suicide watch at 3-month follow-up among adjudicated boys in a residential treatment program ($n = 104$). The authors also examined the classification accuracy of a Caution designation on the MAYSI-2. A Caution designation on MAYSI-2 Suicidal Ideation correctly classified 88% ($n = 92$) of participants, with prediction errors most likely to be false positives. In a more recent study, Archer, Simonds-Bisbee, Spiegel, Handel, and Elkins (2010) found that elevated scores on Depressed-Anxious and Suicidal Ideation on the MAYSI-2 were significantly associated with prior SRB ($r = .17$ to $.28$) and prior NSSI ($r = .19$ to $.30$) among boys ($n = 1,082$) and girls in custody ($n = 110$).

At the time of writing this paper, no published studies have examined the predictive utility of the PAI-A for SRB or NSSI among adolescent offenders. Although studies have found support for the adult PAI in the prediction of prospective SRB (e.g., Wang et al., 1997), due to

differences in item content to modify the PAI for youth (Morey, 2007), these findings cannot be generalized to the adolescent version of the tool. In addition, studies have found a significant association between Borderline Features on the PAI-A and lifetime suicide attempts (e.g., Glenn, Bagge, & Osman, 2013), but these studies were conducted with clinically-referred adolescents (e.g., youth recruited from inpatient psychiatric units).

Preliminary findings on the predictive utility of the MAYSI-2 and PAI-A are promising. However, most investigations used a postdictive approach (but see Butler et al., 2007), which does not reflect the recommended use of the MAYSI-2 (Grisso & Barnum, 2006) or PAI-A in practice (Morey, 2007). One problem with postdictive research designs is that assessment scores may be confounded with the outcome variables (Douglas, Otto, & Borum, 2003). In other words, because the outcomes occurred earlier in time than the predictors, the outcomes might have influenced the predictors (e.g., a failed suicide attempt could have increased scores on measures of depression or emotional dysregulation, a recent incident of NSSI could have increased scores on measures of somatic complaints). As a result, prospective research on the MAYSI-2 and PAI-A is needed.

Beyond conducting this needed research, four additional issues merit research attention. First, although most adolescent offenders serve their sentences in the community rather than in custody (Alam, 2015), past research examining the MAYSI-2 used samples of incarcerated adolescent offenders, which are typically comprised of offenders with severe mental health needs (Fazel, Doll & Långström, 2008) and high rates of SRB and NSSI (Penn, Esposito, Schaeffer, Fritz, & Spirito, 2003). It is unclear whether these results can be generalized to other subsamples of justice-involved youth, such as youth on probation, due to potential differences in base rates of mental disorder and SRB and NSSI across samples. For similar reasons, generalizability of findings on the PAI-A with clinically-referred adolescents to justice-involved youth may also be limited.

Second, research on the MAYSI-2 and PAI-A has primarily focused on the prediction of SRB, rather than both SRB and NSSI simultaneously (but see Archer et al., 2010). While SRB and NSSI appear to be distinct behaviors due to differences in the intent to die, youth who engage in NSSI behaviors are at a greater risk for SRB, suggesting that behaviors are associated or co-occur (Nock et al., 2006). From a primary prevention standpoint, it is important to examine which tools are effective at predicting both outcomes among justice-involved youth. For instance, because justice agencies have limited time and resources, a single tool to evaluate SRB and NSSI would be more cost-effective and efficient than having separate measures for each type of outcome. However, at this point, it is not clear whether a single tool can predict both SRB and NSSI, or a combination of tools is required.

Third, despite the emphasis on cut-off scores in the MAYSI-2 (Grisso & Barnum, 2006) and PAI-A manuals (Morey, 2007) to inform decision-making, only classification accuracy of cut-off scores on MAYSI-2 Suicide Ideation have been examined. Not all youth with suicidal ideation engage in SRB (Fergusson, Beautrais, & Horwood, 2003). In addition, youth may deny suicide ideation to avoid unwanted intervention efforts (Busch, Fawcett, & Jacobs, 2003). Therefore, MAYSI-2 Suicide Ideation is just one of several subscales that could be used to

distinguish between low-risk and high-risk youth, but currently there is inadequate evidence regarding the classification accuracy of other MAYSI-2 and PAI-A subscales.

A final issue is that most research has examined the utility of the MAYSI-2 and PAI-A without consideration of other robust indicators of risk. The assessment of mental health needs may be helpful in identifying youth at risk for SRB and NSSI. However, it might be more efficient to query or triage for other risk factors. As such, it is important to test the incremental utility of the MAYSI-2 and PAI-A. For instance, given the amount of time required to complete the PAI-A, if the PAI-A fails to add predictive value beyond simpler screening approaches then it may be difficult to justify its use as starting point for evaluating youth.

Current Study

To help address the need for further empirical research on these tools, the current study evaluated the prospective predictive utility of the MAYSI-2 and PAI-A for SRB and NSSI among adjudicated adolescent offenders serving community supervision orders. Given the focus of the MAYSI-2 and PAI-A on assessing acute mental health needs (Grisso & Barnum, 2003; Morey, 2007) and the use of a 3-month follow-up period in prior prospective research (e.g., Bulter et al., 2002), a 3-month follow-up period was selected. To the best of our knowledge, this is the first prospective study to examine the predictive utility of the MAYSI-2 and PAI-A for both SRB and NSSI among a non-incarcerated adolescent offender sample. We also extended prior research in two ways. First, we examined the calibration and discrimination of cut-off scores on all MAYSI-2 and conceptually-relevant PAI-A subscales to determine whether they could adequately distinguish between low-risk and high-risk youth. Second, we tested whether the MAYSI-2 and PAI-A added incremental predictive validity to SRB or NSSI history, arguably the strongest risk factors for future SRB and NSSI identified in the literature (Fowler, 2012).

Method

Participants

Study participants were 116 youth recruited between August 2008 and October 2009 from 11 probation offices in a large metropolitan area of Western Canada as part of a larger longitudinal study examining mental health needs, risks, and strengths in adolescents on probation.¹ Of the 508 youth that were approached to participate, 32.1% ($n = 163$) did not meet the following eligibility criteria: between the ages of 12 and 17 ($n = 87$), adjudicated for an offence and placed on probation ($n = 43$), or residing in the Greater Vancouver Regional District ($n = 33$). In addition, 24.8% ($n = 126$) of youth did not wish to participate and 5.1% ($n = 26$) could not be reached for study enrollment. Also, in 5.9% ($n = 30$) of cases parents or guardians could not be reached to obtain informed consent. Of the 163 youth who completed the baseline assessment, 23.3% ($n = 38$) did not complete the 3-month follow-up assessment, 4.3% ($n = 7$) did not complete questions on SRB or NSSI, and 1.2% ($n = 2$) produced invalid PAI-A profiles and were therefore removed from subsequent analyses. The final sample was comprised of 116 adolescent offenders (78 boys and 38 girls). Mean age of participants at the baseline assessment was 16.33 years ($SD = 1.17$, range = 12.93 to 17.89). Of the sample, 37.9% ($n = 44$) identified as Caucasian and 62.1% ($n = 72$) as a racial or ethnic minority group. Specifically, 32.8% ($n = 38$) of youth identified as Aboriginal, 12.1% ($n = 14$) as Asian, 6.9% ($n = 8$) as East

Indian, 5.2% ($n = 6$) as African, 3.4% ($n = 4$) as Hispanic, and 2.6% ($n = 3$) as Middle Easternⁱⁱ. Most of the youth (67.7%, $n = 78$) were first time offenders and had committed a violent offense (62.1%; $n = 72$). The sex and ethnic distribution of the sample mirrored the distribution of justice-involved youth reported in national statistics (Calverley, Cotter, & Halla, 2010). Lifetime rates of SRB and NSSI were 37.1% ($n = 43$) and 25.9% ($n = 30$), respectively. Age did not significantly differ between youth who agreed to participate in the study and youth who declined ($p = .686$). However, girls were more likely to participate in the study compared to boys ($p = .022$). Demographic characteristics, offense characteristics, and lifetime history of SRB and NSSI did not significantly differ between youth retained for analysis and youth who were excluded due to missing follow-up data or an invalid PAI-A profile ($p = .141$ to $.973$). In addition, MAYSI-2 ($p = .073$ to $.885$) and PAI-A subscale scores ($p = .115$ to $.954$) did not significantly differ between youth who completed the 3-month follow-up and youth lost to attrition.

Procedures

All relevant review boards approved this project. Youth were informed about the study at their probation offices via youth probation officers, undergraduate research assistants (RAs), or study flyers. If a youth expressed interest in learning about the study, a RA provided a brief overview of the study. A youth also could self-refer to the study by completing a flyer that was available at his or her probation office. RAs then contacted interested youth to determine eligibility. Following parental or guardian consent and youth assent, youth completed a semi-structured interview and self-report questionnaires, including the MAYSI-2 and PAI-A and questions on prior SRB and NSSI, with a trained undergraduate or graduate RA at the youth's probation office or another community location (e.g., coffee shop). If a youth had difficulty reading or understanding items, questions were read aloud, and the youth circled his or her responses confidentially on the response forms. Following completion of the MAYSI-2 and PAI-A, results were entered into a computer database to obtain subscale scores. In addition, RAs accessed the British Columbia Corrections Network System, an integrated system used for tracking offenders in the community and criminal justice institutions, to obtain background information, including clinician and youth probation officer reports of SRB and NSSI over the youth's lifetime. Three months after the baseline assessment, youth completed a follow-up assessment in which they self-reported whether they had engaged in SRB or NSSI since the baseline assessment and if they had received any treatment services for mental health or behavioral problems. RAs also reviewed justice records to obtain collateral information on SRB and NSSI. Youth were compensated \$20 for completing the baseline assessment and \$15 for completing the 3-month follow-up assessment.

Measures

MAYSI-2. The MAYSI-2 is 52-item self-report screening inventory of mental health. It is designed for use with justice-involved boys and girls aged 12 to 17 years and requires a fifth-grade reading level to complete (Grisso & Barnum, 2006). The MAYSI-2 contains seven subscales for boys and six subscales for girls: Alcohol/Drug Use, Angry-Irritable, Depressed-Anxious, Somatic Complaints, Suicide Ideation, Traumatic Experiences (with different items for boys and girls), and Thought Disturbance (for boys only). Items on the Traumatic Experiences

subscale are answered as yes or no regarding the youth's lifetime and items on the other scales are answered as yes or no regarding the past few months. The yes responses within each subscale are summed to yield a subscale total score. The MAYSI-2 was standardized using data from over 70,000 justice-involved youth. Except for Traumatic Experiences, each subscale has a Caution cut-off score indicating "mental disturbance of possible clinical significance" (Grisso & Barnum, 2006, p. 21) and a Warning cut-off score indicating that "the youth has scored exceptionally high in comparison to other youth in the juvenile justice system" (Grisso & Barnum, 2006, pg. 22). The MAYSI-2 has demonstrated good internal consistency and concurrent validity with other self-report measures of adolescent mental health (Grisso et al., 2012). In the current study, internal consistency on the MAYSI-2 ranged between .58 and .84 (see Table 1). In line with prior research (Grisso et al., 2012), internal consistency of Thought Disturbance ($\alpha = .58$) was lower than the recommended value of .70 (Nunnally, 1978). In addition, internal consistency was low for Somatic Complaints ($\alpha = .65$) and Traumatic Experiences ($\alpha = .63$ and $.59$ for boys and girls, respectively).

--Insert Table 1 about here--

PAI-A. The PAI-A is a 264-item self-report measure of mental health and personality. It is designed for use with boys and girls aged 12 to 18 years and requires a fourth-grade reading level to complete (Morey, 2007). The PAI-A is a youth-adapted version of the Personality Assessment Inventory (Morey, 1991), that uses the same scales but with modified content to reflect developmental differences between adolescents and adults. Each item on the PAI-A is rated as *very true* (1), *mainly true* (2), *slightly true* (3), or *false* (4) in reference to the past few months. In the current study, eight of the 22 subscales on the PAI-A were examined: Depression, Anxiety, Somatic Complaints, Suicide Ideation, Traumatic Stress, Borderline Features, Alcohol Problems, and Drug Problems. These scales were selected due to their conceptual relevance to SRB and NSSI. The PAI-A was standardized using data from 707 school students and 1,160 clinically-referred adolescents. Raw scores on each subscale are converted into a *T* score with a mean of 50 and a standard deviation of 10. A *T* score > 60 suggests that a youth has possible problems in the scales content area and a *T* score > 70 suggests marked difficulty in the scales content area. The PAI-A has demonstrated good internal consistency and concurrent validity with other self-report measures of mental health and personality (Morey, 2007). In the current study, internal consistency of PAI-A subscale ranged between .48 and .80 (see Table 1).

Treatment services. Treatment services for emotional or behavioral problems received since the baseline assessment were examined using a modified version of the Child and Adolescent Services Inventory (CASA; Burns et al., 1992; Mulvey, Schubert, & Chung, 2007). The CASA a 31-item self-report inventory of mental health and other service utilization designed for use with boys and girls aged 8 to 18 years. For each type of treatment, youth are asked to indicate yes or no if a service was received over the past few months. In the current study, an additional item measuring whether the youth had received medication for emotional problems was included. Responses on the CASA were used to create a treatment services total score by totalling yes responses to individual therapy, group therapy, therapy at school, family treatment, drug or alcohol treatment, and medication for emotional problems. Prior research (e.g., Ascher, Farmer, Burns, & Angold, 1996; Mulvey et al., 2007) has indicated that the CASA has good

construct validity and concurrent validity with official measures of service involvement. In the current study, 52.6% ($n = 61$) of youth received individual therapy, 10.3% ($n = 12$) group therapy, 79.0% ($n = 24$) therapy at school, 7.8% ($n = 9$) family treatment, 14.7% ($n = 17$) drug or alcohol treatment, and 14.7% ($n = 17$) medication for emotional problems since the baseline assessment.

SRB and NSSI. Consistent with prior work (e.g., Nock et al., 2006), SRB was defined as any intentional action (e.g., plan, attempt) to end life, and NSSI as any self-destructive behavior undertaken to harm oneself but not to end life. SRB and NSSI were measured using official records and three self-report questions derived from the National Longitudinal Study of Adolescent Health (Carolina Population Centre, 1999) that queried suicide plans, suicide attempts, and whether youth had hurt themselves on purpose. At the baseline assessment these questions were answered regarding the youth's lifetime, and at the 3-month follow-up these questions were answered regarding the previous three months since the baseline assessment. Based on information obtained from official records and youth self-report, a youth was dichotomously coded (i.e., 0 = no, 1 = yes) as having engaged in SRB if they had made suicide plans or attempted suicide. If a youth did not engage in SRB, but had hurt him or herself on purpose, he or she was dichotomously coded as having engaged in NSSI. For instance, at the 3-month follow-up, three youth (2.6%) had engaged in SRB and hurt themselves on purpose and were therefore coded as having engaged in SRB only. To the best of our knowledge, there were no completed suicides in the current study. The base rates of the above outcomes at the 3-month follow-up were 7.8% ($n = 9$) for SRB and 14.7% ($n = 17$) for NSSI.

Data Analytic Plan

MAYSI-2 and PAI-A subscale scores were moderately positively skewed. As such, we used non-parametric approaches to examine the data. First, we computed Spearman's rho coefficients (r_s) in SPSS ©, Version 22 (IBM Corporation, 2013) to examine convergent validity of MAYSI-2 and PAI-A subscale scores. Next, we computed r_s coefficients to examine whether subscale scores and Warning and Caution cut-offs on the MAYSI-2 and subscale scores and Possible Problems and Marked Difficulty cut-offs on the PAI-A were related to SRB and NSSI at the 3-month follow-up. Because correlations are attenuated with low base-rate events, such as SRB, we also conducted Receiver Operating Characteristics (ROC) analysis, which is less sensitive to base rate levels (Rice & Harris, 2005) and commonly reported in prediction studies of other low base-rates events (e.g., adolescent and adult sexual reoffending; Tully, Chou, & Browne, 2013). ROC analyses produced an Area under the Curve (AUC) value which represented the probability that an adolescent who engaged in SRB or NSSI at the 3-month follow-up had a higher MAYSI-2 or PAI-A score than an adolescent who did not engage in SRB or NSSI. AUC values can range between 0 to 1, with an AUC value of .50 representing chance prediction. In addition, Rice and Harris (2005) provided further criteria for interpreting AUCs as follows: AUCs = .56 to .63 represent small effects, AUCs = .64 to .70 represent medium effects, and AUC \geq .71 represent large effects.

To further test the predictive accuracy of the MAYSI-2 and PAI-A, we examined the calibration and discrimination of Warning and Caution cut-offs on the MAYSI-2 and Possible Problems and Marked Difficulty cut-offs on the PAI-A at the 3-month follow-up. To determine

whether cut-off scores in the MAYSI-2 and PAI-A manuals provided optimal classification accuracy for SRB and NSSI, we also examined the calibration and discrimination of median scores on the MAYSI-2 and the PAI-A. Specifically, we examined (a) the probability that an adolescent who engaged in SRB or NSSI at follow-up would score high on the MAYSI-2 or PAI-A (i.e., sensitivity), (b) the probability that an adolescent who had not engaged in SRB or NSSI at follow-up would score low on the MAYSI-2 or PAI-A (i.e., specificity), (c) the proportion of adolescents with high MAYSI-2 or PAI-A scores who engaged in SRB or NSSI at follow-up (i.e., positive predictive value; PPV), and (d) the proportion of adolescents with low MAYSI-2 or PAI-A scores who did not engage in SRB or NSSI at follow-up (i.e., negative predictive value, NPV). Sensitivity or specificity levels of at least 70%, and PPV and NPV values of at least 50% have been deemed acceptable for screening tests (Bujang & Adnan, 2016). Because MAYSI-2 Thought Disturbance is intended for use with boys only, concurrent and predictive validity analyses involving MAYSI-2 Thought Disturbance were conducted with boys only.

To examine whether MAYSI-2 and PAI-A subscale scores added incremental predictive validity to prior SRB or NSSI, we conducted hierarchical penalized logistic regression analyses using the “penalized” package in *R* Version 3.4.2 (R Core Team, 2014). Penalized regression analysis is a recommended approach for analyzing low base rate events because it reduces the risk of bias in the estimation of the OR (Heinze, 2006). SRB or NSSI history was entered in the first block of a penalized logistic regression model, followed by MAYSI-2 or PAI-A subscale scores in the second block. Prior to conducting these models, problematic multicollinearity among the independent variables was examined. Tolerance (.34 to .73) and VIF values (1.38 to 2.92) were in the acceptable range (Cohen, Cohen, West & Aiken, 2003) indicating an absence of problematic multicollinearity in these models. We also conducted power analyses in *G*Power* 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) to determine whether sample size was adequate to test for incremental effects. Regression models with one to nine predictors had sufficient power (i.e., $\beta = .80$) to detect medium or large incremental effects but not small effects. Because power was limited to conduct separate penalized logistic regression models by sex, MAYSI-2 Thought Disturbance was excluded from incremental predictive validity analyses.

Last, given that treatment services received since the baseline assessment may have a confounding effect on the predictive accuracy of the MAYSI-2 and PAI-A, we conducted a set of follow-up analyses in which we re-ran predictive validity analyses (i.e., r_s correlations) controlling for CASA treatment services total score. In addition, given that predictive accuracy of the MAYSI-2 and PAI-A may be stronger among youth with a lifetime history of SRB or NSSI, we also re-ran r_s correlations controlling for these variables.

Results

Descriptive Statistics of MAYSI-2 and PAI-A Subscales

Means, standard deviations, and median scores on MAYSI-2 and PAI-A subscales are presented in Table 1. Mean scores on the MAYSI-2 ranged between 1.03 (Suicidal Ideation) and 5.47 (Angry-Irritable), and mean scores on the PAI-A ranged between 51.07 (Suicidal Ideation) and 68.95 (Drug Problems). Also presented in Table 1 are the proportion of youth classified into

below cut-off (i.e., low/normal), Caution, and Warning categories on the MAYSI-2, and low/normal, Possible Problems, and Marked Difficulty categories on the PAI-A.

Concurrent Validity

Table 2 reports concurrent validity of the MAYSI-2 and PAI-A. Most associations between MAYSI-2 and PAI-A subscales were significant ($r_s = .19$ to $.73$, $p < .001$ to $.038$). The largest associations were observed between MAYSI-2 Alcohol/Drug Use and PAI-A Alcohol Problems ($r_s = .73$, $p < .001$), MAYSI-2 Depressed-Anxious and PAI-A Depression ($r_s = .62$, $p < .001$), and MAYSI-2 Suicidal Ideation and PAI-A Suicidal Ideation ($r_s = .62$, $p < .001$). Large associations (i.e., $r_s \geq .50$) were also found between MAYSI-2 Alcohol/Drug Use and PAI-A Drug Problems ($r_s = .59$, $p < .001$), MAYSI-2 Angry-Irritable and PAI-A Borderline Features ($r_s = .52$, $p < .001$), and MAYSI-2 Depressed-Anxious and Anxiety ($r_s = .54$, $p < .001$), Somatic Complains ($r_s = .51$, $p < .001$), and Borderline Features ($r_s = .52$, $p < .001$) on the PAI-A. Small and non-significant associations were observed between MAYSI-2 Alcohol/Drug Use and Depression, Anxiety, Suicidal Ideation, and Traumatic Stress on the PAI-A ($r_s = .03$ to $.14$, $p = .139$ to $.717$). In addition, associations between MAYSI-2 Somatic Complaints and PAI-A Drug Problems ($r_s = .16$, $p = .097$), MAYSI-2 Suicide Ideation and PAI-A Alcohol Problems ($r_s = .18$, $p = .051$), and MAYSI-2 Thought Disturbance and PAI-A Drug Problems ($r_s = .22$, $p = .056$) were not significant.

--Insert Table 2 about here--

Predictive Validity

Predictive validity of MAYSI-2 and PAI-A subscale and cut-off scores are presented in Tables 3 and 4. Overall, subscale scores had stronger associations with SRB and NSSI than the cut-off scores. For instance, in ROC analysis, subscale scores on Depression (AUC = $.72$, $p = .028$) and Suicidal Ideation (AUC = $.71$, $p = .042$) of the PAI-A were significantly predictive of SRB with large effect sizes (Rice & Harris, 2005). In addition, subscale scores on Suicidal Ideation (AUC = $.72$, $p = .004$) and Thought Disturbance (AUC = $.72$, $p = .028$) of the MAYSI-2 and Traumatic Stress (AUCs = $.74$, $p = .002$) and Borderline Features (AUC = $.69$, $p = .015$) of the PAI-A were significantly predictive of NSSI with moderate to large effect sizes. In contrast, only the Possible Problems cut-off of PAI-A Traumatic Stress was significantly associated with NSSI in ROC analysis (AUC = $.67$, $p = .027$). Notably, some cut-off but not subscale scores were associated with SRB and NSSI in r_s correlation analyses; Marked Difficulty on PAI-A Traumatic Stress was significantly associated with SRB ($r_s = .33$, $p < .001$) and Marked Difficulty on PAI-A Suicidal Ideation was significantly associated with NSSI ($r_s = .19$, $p = .042$). No other significant predictive effects of MAYSI-2 or PAI-A subscales or cut-off scores were found ($r_s = -.05$ to $.18$, AUC = $.44$ to $.68$, $p = .059$ to $.947$).

--Insert Table 3 and 4 about here--

Tables 5 and 6 present the sensitivities, specificities, PPVs, and NPVs of MAYSI-2 and PAI-A median and cut-off scores, respectively. In general, Caution and Warning cut-offs on the MAYSI-2 yielded better calibration and discrimination values than median scores. Caution cut-offs on the MAYSI-2 were better at screening in high-risk youth (i.e., cut-offs yielded higher

sensitivity and NPV, but lower specificity and PPV), whereas Warning cut-offs on the MAYSI-2 were better at screening out low-risk youth (i.e., cut-offs yielded higher specificity and PPV, but lower sensitivity and NPV). In contrast, median scores on the PAI-A were better at screening in high-risk youth, whereas Marked Difficulty cut-offs on the PAI-A were better at screening out lower risk youth. With some exceptions, most cut-off scores on the MAYSI-2 and median and cut-off scores on the PAI-A yielded acceptable values for either sensitivity or specificity (i.e., sensitivity or specificity $\geq 70\%$; range). In addition, NPV was acceptable (i.e., NPV $\geq 50\%$), indicating a low rate of false negatives. However, most PPV values were lower than 50%, indicating relatively high rates of false positives.

--Insert Tables 5 and 6 about here--

Incremental Predictive Validity

Incremental predictive validity analyses are reported in Table 7. Block 1 was significant with the addition of NSSI history, but not SRB history. When MAYSI-2 and PAI-A subscale scores were added in Block 2, only PAI-A Traumatic Stress remained uniquely predictive of NSSI at the 3-month follow-up (OR = 1.09, 95% CI [1.02, 1.18], $p = .008$). However, the addition of PAI-A subscale scores in the NSSI penalized logistic regression model did not significantly improve model fit ($\Delta\chi^2[8] = 12.40$, $p = .134$).

--Insert Table 7 about here--

Predictive Utility of MAYSI-2 and PAI-A Controlling for Confounding Variables

The pattern of findings obtained in the original r_s correlation analyses did not change when CASA treatment services total score was controlled for (i.e., r_s values of MAYSI-2 and PAI-A subscale scores with SRB and NSSI were typically within .00 and .03 of the original values; see Table 8). Similarly, when we reran r_s correlation analyses controlling for lifetime history of SRB the same pattern of results was obtained. However, controlling for lifetime history of NSSI resulted in substantially smaller r_s correlation coefficients than obtained in the original analysis (r_s values of MAYSI-2 and PAI-A subscale scores with NSSI were within .04 and .13 of the original values).

--Insert Table 8 about here--

Discussion

Clinician-administered instruments are available to assess mental health needs conceptually related to SRB and NSSI in youth; however, the time and resources to administer these tools in youth justice settings is problematic. As such, there is a need for valid screening methods to determine when to apply limited resources. Self-report measures of mental health, such as the MAYSI-2, a well-validated tool that is widely used in youth justice settings, or PAI-A a comprehensive tool with promising features, may assist in identifying justice-involved youth at risk of engaging in SRB and NSSI. However, research evaluating their predictive utility among non-incarcerated adolescent offenders is lacking. Thus, this prospective study examined

the predictive utility of the MAYSI-2 and PAI-A for SRB and NSSI among adolescent offenders on probation.

Contrary to prior prospective research on the MAYSI-2 (Bulter et al., 2007), scores on Angry-Irritable, Depressed-Anxious, Somatic Complaints, and Suicide Ideation subscales of the MAYSI-2 were not significantly associated with SRB at 3-month follow-up. Differences in results across studies may be due to differences between samples (i.e., youth on probation versus incarcerated adolescents) or differences in assessment (i.e., SRB versus placement on suicide watch). For instance, significant associations reported in prior research may reflect policies that incarcerated youth scoring high on mental health needs be automatically placed under suicide observation (Hayes, 2004).

Commensurate with results reported in prior postdictive investigations of the MAYSI-2 (Archer et al., 2010), scores on MAYSI-2 Suicide Ideation were significantly predictive of prospective NSSI at 3-month follow-up. Although NSSI is differentiated from SRB by intent to die and therefore suicidal ideation is expected to be unrelated to NSSI, these findings could reflect that SRB and NSSI are associated or co-occur (Nock et al., 2006). Consequently, there is potential for shared risk factors, such as suicidal ideation, underlying SRB and NSSI.

Expanding on prior work, scores on MAYSI-2 Thought Disturbance were significantly predictive of prospective NSSI. Whereas prior research on the association between the MAYSI-2 and NSSI has predominantly focused on predictive utility of Depressed-Anxious and Suicide Ideation subscales on the MAYSI-2 (Acher et al., 2010), the current findings suggest that screening may be more effective if a broader range of problems than depression and suicide ideation are considered.

Importantly, this is the first study to demonstrate that multiple subscales on the PAI-A have utility for identifying justice-involved youth at risk for SRB and NSSI. With respect to SRB, significant predictive effects of PAI-A Traumatic Stress were found. In addition, Suicidal Ideation, Traumatic Stress, and Borderline Features subscales on the PAI-A were significantly predictive of NSSI. However, contrary to prior postdictive research among clinically-referred adolescents (Glenn et al., 2013), PAI-A Borderline Features was not predictive of prospective SRB. It is possible that Borderline Features on the PAI-A has stronger predictive effects among clinically-referred youth due to differences in base rates of borderline personality traits across samples. It is also possible that the significant findings reported in postdictive research were artificially inflated due to confounding.

Given the need to determine classification accuracy of subscales that tap constructs other than suicide ideation, we examined the calibration and discrimination of Caution and Warning cut-offs of all MAYSI-2 subscales and Possible Problems and Marked Difficulty cut-offs on conceptually-relevant PAI-A subscales. Ideally, cut-off scores would have a high degree of sensitivity, a meaningful level of specificity, and a low rate of false positive and false negatives. With some exceptions, most cut-off scores on the MAYSI-2 and PAI-A adequately distinguished between low-risk and high-risk youth, either by “screening in” most high-risk youth or “screening out” most low-risk youth with low rates of false negatives. However, consistent with prior research on MAYSI-2 Suicide Ideation (Butler et al., 2007), rates of false-positives were

relatively high. Because of the serious consequences that can arise from SRB and NSSI (e.g., injuries that require medical attention, death), some authors have argued that high false-positive rates on SRB and NSSI screening measures may be viewed as acceptable (e.g., Wasserman et al., 2003). That said, given that MAYSI-2 and PAI-A subscale scores had stronger associations with SRB and NSSI than cut-off scores in ROC analysis, and median scores on the PAI-A were better at screening in high-risk youth, further research is needed on appropriate cut-off scores on the MAYSI-2 and PAI-A that maximize classification accuracy.

One question that has remained unaddressed in the literature is whether the evaluation of mental health needs improves the prediction of emergent SRB and NSSI beyond other robust indicators of risk. The current study did not find evidence that the MAYSI-2 or PAI-A added incrementally to SRB or NSSI history. Thus, triaging for prior SRB and NSSI through file review or self-report questions included on an intake assessment may be a more effective starting point for evaluating emergent risk. However, lifetime history variables were not significantly associated with prospective SRB in our analytic models. Moreover, measures of more proximal risk factors, such as mental health, can help to supplement initial evaluations, especially when historical information on SRB and NSSI is unavailable or incomplete.

Some research has found that the length of time that has passed since the last incident of SRB, rather than the presence of prior SRB, is associated with future SRB (Fowler, 2012). In addition, repeated SRB may be better predictive of future SRB, as subsequent SRB may require less mental distress to precipitate (e.g., Joiner & Rudd, 2000). As such, an important direction for future work is to test the incremental utility of the MAYSI-2 and PAI-A using more refined measures of SRB and NSSI history. Furthermore, given that SRB and NSSI among adolescents occurs during an age period when risk behaviors can occur in reaction to a distressing event (Dixon-Gordon et al., 2012), research should also test the incremental predictive validity of the MAYSI-2 and PAI-A beyond other risk factors that were not examined in the current research (e.g., interpersonal conflict, victimization).

Overall, the findings obtained indicate that the PAI-A has more promise in predicting SRB and NSSI than the MAYSI-2 (i.e., 37.5% of associations on the PAI-A vs. 14.3% on the MAYSI-2 were significant in correlation or ROC analysis). Although this is to be expected given that the MAYSI-2 is a brief screening measure intended to identify adolescents in need of further assessment (Grisso & Barnum, 2006), it is possible that the MAYSI-2 may not fully capture the extent of mental health needs associated with SRB and NSSI, and therefore more comprehensive measures, like the PAI-A, are required. For instance, despite tapping the same mental health needs, indices of depression, suicide ideation, and trauma on the PAI-A but not the MAYSI-2 were associated with SRB or NSSI at 3-month follow-up. In addition, although indices of calibration and discrimination were largely similar across these subscales, PAI-A subscales were able to flag false negatives that MAYSI-2 subscales had missed. Nevertheless, given the widespread use of the MAYSI-2, as well as the resources needed to routinely administer the PAI-A in youth justice settings, further research is needed to determine the relative predictive value of the MAYSI-2 to the PAI-A.

Given that this study is the first prospective study to examine the predictive utility of the MAYSI-2 and PAI-A for SRB and NSSI among a non-incarcerated adolescent offender sample,

additional research replicating these findings is important. Nevertheless, current findings regarding the use of the MAYSI-2 and PAI-A with adolescent offenders who are not yet deeply entrenched in the justice system are encouraging. Elevated scores on Depression, Suicidal Ideation, and Traumatic Stress of the PAI-A may assist in the identification of youth at risk for SRB. In addition, elevated scores on Suicidal Ideation and Thought Disturbance of the MAYSI-2 and on Suicide Ideation, Traumatic Stress and Borderline Features of the PAI-A may assist in the identification of youth at risk for NSSI. Furthermore, these subscales remained predictive controlling for lifetime history of SRB, suggesting that the use of these scales may aid in the identification of at-risk youth before mental distress can escalate into more harmful behavior. However, predictive validity of these scales was substantially reduced when prior NSSI was controlled for. Therefore, it would be worthwhile to further explore whether the MAYSI-2 and PAI-A are predictive among youth without a lifetime history of NSSI.

Four limitations should be considered when interpreting these findings. First, internal consistency was low (Nunnally, 1978) for seven of the 13 examined MAYSI-2 and PAI-A subscales, which may have attenuated associations between scores on these subscales and the outcome variables (Nunnally & Bernstein, 1994). Second, we relied on a small number of questions to evaluate SRB and NSSI; however, more comprehensive and validated measures are available (e.g., Nock, Holmberg, Photos, & Michel, 2007). Third, incremental predictive validity analyses were underpowered to detect small effects. Fourth, although youth who refused study participation did not differ from youth who agreed to participate on age, boys were less likely to participate in the study than girls. Also, information on offense characteristics and prior SRB and NSSI of youth who refused study participation was unavailable. As such, sample generalizability may be limited.

Despite these limitations, the current research has important clinical implications for youth justice professionals. Although professionals should be aware of potential differences of MAYSI-2 and PAI-A subscales in terms of predictive value for SRB versus NSSI, as well as potential for false-positive and false-negative designations, professionals should prioritize justice-involved youth with high scores on the MAYSI-2 and PAI-A for further assessment or monitoring. In addition, findings suggest that prevention and intervention efforts targeting depression, suicidal ideation, trauma, thought disturbance, and borderline features may be effective in reducing SRB and NSSI. Because mental health difficulties have been linked to increased risk of reoffending (Wibbelink, Hoeve, Stams, & Oort, 2017), addressing the mental health needs of youth upon entry to the justice system may not only help to reduce SRB and NSSI, but also ongoing justice system involvement.

References

- Alam, S. (2015). *Youth court statistics, 2013/2014* (Catalogue no. 85–002-X). Ottawa: Statistics Canada.
- Archer, R. P., Simonds-Bisbee, E. C., Spiegel, D. R., Handel, R. W., & Elkins, D. E. (2010). Validity of the Massachusetts Youth Screening Instrument–2 (MAYSI–2) scales in juvenile justice settings. *Journal of Personality Assessment, 92*, 337-348. doi:10.1080/00223891.2010.482009
- Ascher, B. H. Z., Farmer, E. M., Burns, B. J., & Angold, A. (1996). The child and adolescent services assessment (CASA) description and psychometrics. *Journal of Emotional and Behavioral Disorders, 4*, 12-20. doi:10.1177/106342669600400102
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Bujang, M. A., & Adnan, T. H. (2016). Requirements for minimum sample size for sensitivity and specificity analysis. *Journal of Clinical and Diagnostic Research, 10*(10), e01 – e06. doi:10.7860/JCDR/2016/18129.8744
- Busch, K. A., Fawcett, J., & Jacobs, D. G. (2003). Clinical correlates of inpatient suicide. *The Journal of Clinical Psychiatry, 64*, 14-19. doi:10.4088/JCP.v64n0105
- Burns, B. J., Angold, A., Magruder-Habib, K., Costello, E. J., & Patrick, M. K. S. (1992). *The Child and Adolescent Services Assessment (CASA)*. Durham, ND: Duke University Medical Center.
- Butler, M. A., Loney, B. R., & Kistner, J. (2007). The Massachusetts Youth Screening Instrument as a predictor of institutional maladjustment in severe male juvenile offenders. *Criminal Justice and Behaviour, 34*, 476-492. doi:10.1177/0093854806291711
- Calverley, D., Cotter, A., & Halla, E. (2010). *Youth custody and community services in Canada, 2008/2009*. Ottawa: Statistics Canada.
- Carolina Population Center (1999). *National longitudinal study of adolescent health wave II adolescent in-home questionnaire code book*. Chapel Hill, NC: Carolina Population Center. Retrieved from <http://www.cpc.unc.edu/projects/addhealth/codebooks/wave2>.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Dixon-Gordon, K., Harrison, N., & Roesch, R. (2012). Non-suicidal self-injury within offender populations: a systematic review. *International Journal of Forensic Mental Health, 11*, 33-50. doi:10.1080/14999013.2012.667513
- Douglas, K. S., Otto, R. K., & Borum, R. (2003). Research methods in clinical forensic psychology. In J. A. Schinka, W. F. Velicer, & I. B. Weiner (Eds.), *Handbook of psychology, volume two: Research methods in psychology* (pp. 189-211). New York, NY: John Wiley & Sons, Inc.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160. doi:10.3758/brm.41.4.1149
- Fazel, S., Doll, H., & Långström, N. (2008). Mental disorders among adolescents in juvenile detention and correctional facilities: A systematic review and metaregression analysis of 25 surveys. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*, 1010-1019. doi:10.1097/chi.obo13e31817eefc3

- Fergusson, D. M., Beautrais, A. L., & Horwood, L. J. (2003). Vulnerability and resiliency to suicidal behaviours in young people. *Psychological Medicine, 33*, 61-73. doi:10.1017/s0033291702006748
- Fleischmann, A., Bertolote, J. M., Belfer, M., & Beautrais, A. (2005). Completed suicide and psychiatric diagnoses in young people: A critical examination of the evidence. *American Journal of Orthopsychiatry, 75*, 676- 683. doi:10.1037/0002-9432.75.4.676
- Fowler, J. C. (2012). Suicide risk assessment in clinical practice: Pragmatic guidelines for imperfect assessments. *Psychotherapy, 49*, 81-90. doi:10.1037/a0026148
- Glenn, C. R., Bagge, C. L., & Osman, A. (2013). Unique associations between borderline personality disorder features and suicide ideation and attempts in adolescents. *Journal of Personality Disorders, 27*, 604-616. doi:10.1521/pedi_2013_27_102
- Grisso, T., & Barnum, R. (2006). *Massachusetts Youth Screening Instrument-Version 2: User's manual and technical report*. Sarasota, FL: Professional Resource Press.
- Grisso, T., Fusco, S., Paiva-Salisbury, M., Perrautot, R., Williams, V., & Barnum, R. (2012). *The Massachusetts Youth Screening Instrument-Version 2 (MAYSI-2): Comprehensive research review*. Worcester, MA: University of Massachusetts Medical School.
- Hayes, L. M. (2004). *Juvenile suicide in confinement: A national survey*. Mansfield, MA: National Center on Institutions and Alternatives.
- Heinze, G. (2006). A comparative investigation of methods for logistic regression with separated or nearly separated data. *Statistics in Medicine, 25*, 4216-4226. doi:10.1002/sim.2687
- IBM Corporation (2013). IBM SPSS Statistics for Windows, Version 22.0 [Computer software]. Armonk, NY: IBM Corporation.
- Joiner Jr, T. E., & Rudd, M. D. (2000). Intensity and duration of suicidal crises vary as a function of previous suicide attempts and negative life events. *Journal of Consulting and Clinical Psychology, 68*, 909-916. doi:10.1037/0022-006X.68.5.909
- Morey, L. C. (1991). *Personality Assessment Inventory: Professional Manual*. Tampa: Psychological Assessment Resources.
- Morey, L. C. (2007). *Personality Assessment Inventory-Adolescent (PAI-A)*. Lutz, FL: Psychological Assessment Resources.
- Mulvey, E. P., Schubert, C. A., & Chung, H. L. (2007). Service use after court involvement in a sample of serious adolescent offenders. *Children and Youth Services Review, 29*, 518-544. doi:10.1016/j.childyouth.2006.10.006
- National Action Alliance for Suicide Prevention: Youth in Contact with the Juvenile Justice System Task Force. (2013). *Screening and assessment for suicide prevention: Tools and procedures for risk identification among juvenile justice youth*. Washington, DC: Author
- Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). Self-Injurious Thoughts and Behaviors Interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment, 19*, 309–317. doi:10.1037/1040-3590.19.3.309.
- Nock, M. K., Joiner, T. E., Gordon, K. H., Lloyd-Richardson, E., & Prinstein, M. J. (2006). Non-suicidal self-injury among adolescents: Diagnostic correlates and relation to suicide attempts. *Psychiatry Research, 144*, 65-72. doi:10.1016/j.psychres.2006.05.010
- Nunnally, J.C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Nunnally, J.C. & Bernstein, I. (1994). *Psychometric theory (3rd edition)*. New York: McGraw-Hill.

- Penn, J. V., Esposito, C. L., Schaeffer, L. E., Fritz, G. K., & Spirito, A. (2003). Suicide attempts and self-mutilative behavior in a juvenile correctional facility. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*, 2-769. doi:10.1097/01.CHI.0000046869.56865.46
- R Core Team. (2014). R: A language and environment for statistical computing reference index [Computer software]. Vienna, Austria: The R Foundation for Statistical Computing.
- Rice, M. & Harris, G. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's *d*, and *r*. *Law and Human Behavior, 29*, 615-620. doi:10.1007/s10979-005-6832-7.
- Shaffer, D., Fisher, P., Lucas, C., Dulcan, M., & Schwab-Stone M. (2000). NIMH Diagnostic Interview Schedule for Children, Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 28-38. doi:10.1097/00004583-200001000-00014
- Teplin, L. A., Abram, K. M., McClelland, G. M., Dulcan, M. K., & Mericle, A. A. (2002). Psychiatric disorders in youth in juvenile detention. *Archives of General Psychiatry, 59*, 1133-1143. doi:10.1001/archpsyc.59.12.1133
- Tully, R. J., Chou, S., & Browne, K. D. (2013). A systematic review on the effectiveness of sex offender risk assessment tools in predicting sexual recidivism of adult male sex offenders. *Clinical Psychology Review, 33*, 287-316. doi:10.1016/j.cpr.2012.12.002
- Wang, E. W., Rogers, R., Giles, C. L., Diamond, P. M., Herrington-Wang, L. E., & Taylor, E. R. (1997). A pilot study of the Personality Assessment Inventory (PAI) in corrections: Assessment of malingering, suicide risk, and aggression in male inmates. *Behavioral Sciences & the Law, 15*, 469-482. doi:10.1002/(SICI)1099-0798(199723/09)
- Wasserman, G. A., Jensen, P. S., Ko, S. J., Cocozza, J., Trupin, E., Angold, A., ... & Grisso, T. (2003). Mental health assessments in juvenile justice: Report on the consensus conference. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*, 752-761. doi:10.1097/01.CHI.0000046873.56865.4B
- Wasserman, G. A., McReynolds, L. S., Ko, S. J., Katz, L. M., Cauffman, E., Haxton, W., & Lucas, C. P. (2004). Screening for emergent risk and service needs among incarcerated youth: Comparing MAYSI-2 and Voice DISC-IV. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*, 629-639. doi:10.1097/00004583-200405000-00017
- Wibbelink, C. J., Hoeve, M., Stams, G. J. J., & Oort, F. J. (2017). A meta-analysis of the association between mental disorders and juvenile recidivism. *Aggression and Violent Behavior, 33*, 78-90. doi:10.1016/j.avb.2017.01.005

Endnotes

ⁱ As part of this study, adolescents on probation were evaluated at 3-month intervals over a 1-year period.

ⁱⁱ One participant self-identified as both Aboriginal and Hispanic. Therefore, the total number of youth in each ethnic minority category does not exactly equal 72.

Tables

Table 1. *Psychometric and Descriptive Properties of MAYSI-2 and PAI-A Subscales*

Scale	Items	A	M (SD)	Median	Risk Level		
					% (n) Low/Normal	% (n) Caution/Possible Problems	% (n) Warning/Marked Difficulty
MAYSI-2							
Alcohol/Drug Use	8	.84	4.35 (2.67)	5.00	37.1 (43)	21.6 (25)	41.4 (48)
Angry-Irritable	9	.79	5.47 (2.68)	6.00	32.8 (38)	39.7 (46)	27.6 (32)
Depressed-Anxious	9	.71	2.64 (2.02)	2.00	53.4 (62)	35.3 (41)	11.2 (13)
Somatic Complaints	6	.65	3.34 (1.98)	3.00	37.9 (44)	41.4 (48)	20.7 (24)
Suicidal Ideation	5	.82	1.03 (1.48)	0.00	70.7 (82)	11.2 (13)	18.1 (21)
Traumatic Experiences	5	.63 _B	2.61 (1.57)	3.00	--	--	--
		.59 _G					
Thought Disturbance _B	5	.58	1.08 (1.28)	1.00	33.6 (39)	14.7 (17)	20.7 (24)
PAI-A							
Depression	18	.67	55.03 (10.67)	54.00	73.3 (85)	16.4 (19)	10.3 (12)
Anxiety	18	.78	53.42 (11.57)	51.00	76.7 (89)	14.7 (17)	8.6 (10)
Somatic Complaints	18	.74	54.84 (10.80)	52.00	79.3 (92)	12.1 (14)	8.6 (10)
Suicidal Ideation	8	.65	51.07 (8.45)	50.00	89.7 (104)	6.9 (8)	3.4 (4)
Traumatic Stress	6	.80	54.72 (10.47)	52.00	75.9 (88)	13.8 (16)	10.3 (12)
Borderline Features	20	.69	54.51 (9.94)	55.00	72.4 (84)	22.4 (26)	5.2 (6)
Alcohol Problems	8	.77	62.23 (15.97)	57.00	56.0 (65)	17.2 (20)	26.7 (31)
Drug Problems	8	.48	68.95 (16.53)	66.00	29.3 (34)	30.2 (35)	40.5 (47)

Note. MAYSI-2 = Massachusetts Youth Screening Instrument, Version 2. PAI-A = Personality Assessment Inventory-Adolescent. α = Cronbach's alpha. M = Mean. SD = Standard deviation. B = Boys. G = Girls.

Table 2. *Concurrent Validity of MAYSI-2 and PAI-A Subscale Scores*

MAYSI-2	PAI-A							
	Depression	Anxiety	Somatic Complaints	Suicidal Ideation	Traumatic Stress	Borderline Features	Alcohol Problems	Drug Problems
Alcohol/Drug Use	.10	.14	.23*	.03	.10	.24*	.73***	.59***
Angry-Irritable	.37***	.40***	.43***	.27**	.39***	.52***	.44***	.44***
Depressed-Anxious	.62***	.54***	.51***	.43***	.47***	.52***	.32**	.34***
Somatic Complaints	.36***	.42***	.41***	.19*	.34***	.36***	.24*	.16
Suicidal Ideation	.48***	.34***	.32***	.62***	.38***	.43***	.18	.22*
Traumatic Experiences	.28**	.28**	.30**	.27**	.39***	.37***	.31**	.39***
Thought Disturbance	.25*	.25*	.42***	.33**	.29*	.38***	.33**	.22

Note. Spearman rho correlations. MAYSI-2 = Massachusetts Youth Screening Instrument, Version 2. PAI-A = Personality Assessment Inventory-Adolescent. * $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed test). M = Males.

Table 3. *Predictive Utility of MAYSI-2 Subscale Scores and Caution and Warning Cut-Offs at 3-Month Follow-Up*

Scale	Suicide-Related Behavior			Non-Suicidal Self-Injury		
	r_s	AUC	95% CI	r_s	AUC	95% CI
Alcohol/Drug Use						
Total	-.01	.49	[.36, .63]	.10	.58	[.45, .71]
Caution	.09	.58	[.40, .76]	.17	.61	[.48, .75]
Warning	-.11	.40	[.22, .58]	.05	.53	[.38, .68]
Angry-Irritable						
Total	-.01	.49	[.32, .67]	.18	.64	[.49, .79]
Caution	.07	.56	[.37, .74]	.08	.55	[.41, .70]
Warning	-.04	.47	[.28, .66]	.18	.61	[.46, .77]
Depressed-Anxious						
Total	.17	.68	[.55, .82] ¹	.16	.63	[.50, .76]
Caution	.18	.67	[.50, .84]	.15	.61	[.46, .75]
Warning	-.00	.50	[.30, .70]	.01	.50	[.35, .65]
Somatic Complaints						
Total	.11	.62	[.42, .82]	.11	.59	[.46, .72]
Caution	.09	.59	[.40, .77]	.07	.55	[.41, .70]
Warning	.17	.63	[.42, .84]	.03	.52	[.37, .69]
Suicidal Ideation						
Total	.16	.68	[.50, .87]	.30**	.72**	[.59, .85]
Caution	.17	.64	[.44, .84]	.22*	.64	[.49, .79]
Warning	.12	.58	[.38, .79]	.25**	.63	[.48, .79]
Traumatic Experiences						
Total	.05	.55	[.38, .73]	.14	.61	[.47, .75]
Thought Disturbance						
Total	-.05	.44	[.25, .63]	.26*	.72*	[.54, .90]
Caution	.05	.55	[.29, .81]	.22	.66	[.50, .83]
Warning	-.17	.34	[.15, .53]	.25*	.67	[.48, .86]

Note. MAYSI-2 = Massachusetts Youth Screening Instrument, Version 2. PAI-A = Personality Assessment Inventory-Adolescent. r_s = Spearman rho correlation. AUC = area under the curve. 95% CI = 95% confidence intervals of AUC. * $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed test). ¹Although 95% CI did not include .50 (representing change prediction), AUC values were not significant at $p < .05$ ($p = .070$).

Table 4. *Predictive Utility of PAI-A Subscale Scores and Possible Problems and Marked Difficulty Cut-Offs at 3-Month Follow-Up*

Scale	Suicide-Related Behavior			Non-Suicidal Self-Injury		
	r_s	AUC	95% CI	r_s	AUC	95% CI
Depression						
Total	.21*	.72*	[.58, .86]	.10	.58	[.42, .73]
Possible Problems	.19*	.66	[.46, .85]	.08	.55	[.40, .70]
Marked Difficulty	.13	.57	[.36, .78]	.12	.55	[.39, .70]
Anxiety						
Total	-.03	.47	[.26, .68]	.05	.54	[.40, .68]
Possible Problems	-.01	.49	[.30, .69]	.00	.50	[.35, .65]
Marked Difficulty	.14	.57	[.36, .79]	-.04	.48	[.38, .63]
Somatic Complaints						
Total	.03	.53	[.32, .74]	.17	.64	[.50, .78]
Possible Problems	.09	.57	[.36, .77]	.09	.55	[.40, .71]
Marked Difficulty	.14	.57	[.46, .79]	.05	.52	[.37, .67]
Suicidal Ideation						
Total	.19*	.71*	[.52, .89]	.17	.64	[.48, .79]
Possible Problems	.33***	.69	[.47, .90]	.18	.58	[.58, .74]
Marked Difficulty	.12	.54	[.33, .75]	.19*	.55	[.39, .71]
Traumatic Stress						
Total	.03	.54	[.27, .80]	.29**	.74**	[.61, .86]
Possible Problems	.14	.61	[.41, .81]	.28**	.67*	[.52, .82]
Marked Difficulty	.33***	.69	[.47, .90]	.34***	.65	[.49, .81]
Borderline Features						
Total	.08	.58	[.34, .82]	.23*	.69*	[.55, .82]
Possible Problems	.11	.59	[.39, .79]	.18	.61	[.46, .77]
Marked Difficulty	.08	.53	[.32, .74]	.12	.54	[.38, .69]
Alcohol Problems						
Total	-.03	.46	[.26, .67]	.07	.57	[.41, .70]
Possible Problems	.02	.52	[.32, .72]	.10	.57	[.42, .72]
Marked Difficulty	-.01	.49	[.30, .69]	.12	.57	[.42, .73]
Drug Problems						
Total	-.00	.54	[.35, .73]	.13	.60	[.45, .76]
Possible Problems	.05	.54	[.35, .73]	.11	.57	[.43, .71]
Marked Difficulty	.02	.52	[.32, .72]	.06	.54	[.39, .69]

Note. MAYSI-2 = Massachusetts Youth Screening Instrument, Version 2. PAI-A = Personality Assessment Inventory-Adolescent. r_s = Spearman rho correlation. AUC = area under the curve. 95% CI = 95% confidence intervals of AUC. *? $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed test).

Table 5. Sensitivity, Specificity, PPV, and NPV of MAYSI-2 Caution and Warning Cut-Offs at 3-Month Follow-Up

Scale	Suicide-Related Behavior				Non-Suicidal Self-Injury			
	Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV
Alcohol/Drug Use								
Median	22.22%	57.01%	4.17%	89.71%	47.06%	59.60%	16.67%	13.24%
Caution	77.78%	38.32%	9.59%	95.34%	82.35%	40.40%	19.18%	95.02%
Warning	22.22%	57.01%	4.17%	89.71%	47.01%	59.60%	16.67%	86.76%
Angry-Irritable								
Median	33.33%	55.14%	5.88%	90.77%	64.71%	59.60%	21.57%	9.23%
Caution	77.78%	33.64%	8.97%	94.74%	76.47%	34.34%	16.67%	89.47%
Warning	22.22%	71.96%	6.25%	91.67%	52.94%	75.76%	27.27%	90.36%
Depressed-Anxious								
Median	66.67%	56.07%	11.32%	95.24%	58.82%	56.56%	18.87%	11.11%
Caution	77.78%	56.07%	12.96%	96.78%	64.71%	56.57%	20.37%	90.32%
Warning	11.11%	88.79%	7.69%	92.23%	11.76%	88.89%	15.38%	85.45%
Somatic Complaints								
Median	66.67%	53.27%	10.71%	95.00%	58.82%	53.53%	17.86%	11.67%
Caution	77.78%	39.25%	9.72%	95.45%	70.59%	39.40%	16.67%	88.63%
Warning	44.44%	81.31%	16.67%	94.56%	23.53%	79.80%	16.67%	85.87%
Suicidal Ideation								
Median	75.00%	60.75%	12.50%	97.01%	76.47%	64.29%	27.08%	5.97%
Caution	55.56%	72.90%	14.71%	95.12%	53.94%	74.74%	26.47%	90.24%
Warning	33.33%	83.18%	14.29%	93.68%	41.18%	85.86%	33.33%	89.47%
Traumatic Experiences								
Median	33.33%	67.29%	7.89%	92.31%	41.18%	68.69%	18.42	12.82%
Thought Disturbance								
Median	0.00%	68.00%	0.00%	91.10%	60.00%	74.29%	25.00%	7.14%
Caution	60.00%	49.33%	7.31%	94.87%	80.00%	52.86%	19.51%	94.87%
Warning	0.00%	68.00%	0.00%	91.07%	60.00%	74.29%	25.00%	92.86%

Note. MAYSI-2 = Massachusetts Youth Screening Instrument, Version 2. PPV = Positive Predictive Value. NPV = Negative Predictive Value.

Table 6. *Sensitivity, Specificity, PPV, and NPV of PAI-A Possible Problems and Marked Difficulty Cut-Offs at 3-Month Follow-Up*

Scale	Suicide-Related Behavior				Non-Suicidal Self-Injury			
	Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV
Depression								
Median	88.88%	53.27%	13.79%	98.28%	64.71%	11.32%	47.48%	98.77%
Possible Problems	55.56%	75.70%	16.13%	95.29%	35.29%	74.74%	19.35%	87.06%
Marked Difficulty	22.22%	90.65%	16.67%	93.27%	17.64%	90.90%	25.00%	86.54%
Anxiety								
Median	33.33%	42.99%	4.69%	88.46%	64.71%	10.17%	53.53%	98.61%
Possible Problems	22.22%	76.64%	7.41%	92.13%	23.52%	76.68%	14.81%	85.39%
Marked Difficulty	22.22%	90.65%	16.67%	93.27%	5.88%	90.90%	10.00%	84.91%
Somatic Complaints								
Median	44.44%	47.66%	6.66%	91.07%	70.59%	9.43%	48.48%	98.63%
Possible Problems	33.33%	80.37%	12.50%	93.48%	29.41%	80.80%	20.83%	86.96%
Marked Difficulty	22.22%	92.52%	20.00%	93.39%	11.76%	91.91%	20.00%	85.85%
Suicidal Ideation								
Median	88.88%	38.31%	10.81%	97.62%	76.47%	6.15%	61.62%	98.03%
Possible Problems	44.44%	92.52%	33.33%	95.19%	23.52%	91.91%	33.33%	87.50%
Marked Difficulty	11.11%	97.20%	25.00%	92.86%	11.76%	97.79%	50.00%	86.60%
Traumatic Stress								
Median	55.56%	43.93%	7.69%	92.16%	82.36%	5.56%	51.52%	98.31%
Possible Problems	44.44%	77.57%	14.29%	94.32%	52.94%	80.80%	32.14%	90.90%
Marked Difficulty	44.44%	92.52%	33.33%	95.19%	35.29%	93.93%	50.00%	89.42%
Borderline Features								
Median	66.67%	59.47%	10.17%	94.74%	76.47%	8.00%	46.46%	98.58%
Possible Problems	44.44%	73.83%	12.50%	94.05%	47.06%	75.76%	25.00%	89.29%
Marked Difficulty	11.11%	95.33%	16.67%	92.72%	11.76%	95.95%	33.33%	86.36%
Alcohol Problems								
Median	44.44%	42.99%	6.15%	90.12%	44.44%	7.58%	57.01%	99.04%
Possible Problems	44.44%	58.82%	8.69%	92.31%	44.44%	10.63%	41.76%	99.27%
Marked Difficulty	22.22%	76.47%	7.69%	91.76%	22.22%	22.58%	23.53%	99.71%
Drug Problems								
Median	44.444%	52.34%	7.27%	91.80%	59.94%	14.81%	46.47%	99.01%

Predictive Validity of the MAYSI-2 27

Possible Problems	77.78%	29.90%	8.53%	94.11%	82.35%	4.22%	68.69%	97.42%
Marked Difficulty	44.44%	59.81%	8.51%	92.75%	47.06%	18.7%	39.39%	99.22%

Note. PAI-A = Personality Assessment Inventory-Adolescent. PPV = Positive Predictive Power. NPV = Negative Predictive Power.

Table 7. Incremental Validity of MAYSI-2 and PAI-A Subscale Scores over Lifetime History: Penalized Logistic Regression Models

	Suicide-Related Behavior				Non-Suicidal Self-Injury			
	<i>b</i> (SE)	OR [95% CI]	Wald	<i>P</i>	<i>b</i> (SE)	OR [95% CI]	Wald	<i>P</i>
Block 1								
Lifetime history	0.79 (0.67)	2.21 [0.59, 8.66]	1.41	.235	1.69 (0.54)	5.43 [1.91, 16.19]	10.04	.002
		$\chi^2(1) = 1.41, p = .235$				$\chi^2(1) = 10.04, p = .002$		
Block 2								
MAYSI-2								
Lifetime history	0.73 (0.81)	2.07 [0.35, 11.44]	0.69	.405	1.23 (0.65)	5.41 [0.97, 12.17]	3.65	.056
Alcohol/Drug Use	-0.00 (0.16)	1.00 [0.72, 1.40]	0.00	.980	0.04 (0.13)	1.04 [0.82, 1.40]	0.08	.772
Angry-Irritable	-0.20 (0.19)	0.82 [0.53, 1.20]	1.05	.306	0.04 (0.15)	1.04 [0.77, 1.42]	0.07	.792
Depressed-Anxious	0.08 (0.22)	1.08 [0.69, 1.76]	0.13	.723	-0.16 (0.19)	0.85 [0.58, 1.26]	0.79	.374
Somatic Complaints	0.13 (0.21)	1.14 [0.74, 1.82]	0.34	.562	0.10 (0.16)	1.11 [0.82, 1.54]	0.43	.512
Suicidal Ideation	0.19 (0.25)	1.20 [0.72, 2.04]	0.52	.470	0.22 (0.22)	1.25 [0.82, 1.93]	1.09	.295
Traumatic Experiences	0.11 (0.27)	1.12 [0.66, 2.09]	0.16	.686	0.07 (0.21)	1.07 [0.72, 1.66]	0.13	.723
		$\chi^2(7) = 4.32, p = .742, \Delta\chi^2(6) = 1.94, p = .925$				$\chi^2(7) = 12.00, p = .101, \Delta\chi^2(6) = 2.63, p = .853$		
PAI-A								
Lifetime history	0.12 (0.85)	1.13 [0.17, 6.13]	0.02	.890	0.92 (0.66)	2.50 [0.69, 0.90]	1.98	.159
Depression	0.06 (0.05)	1.06 [0.95, 1.17]	1.15	.284	-0.05 (0.05)	0.95 [0.87, 1.04]	1.29	.256
Anxiety	-0.06 (0.05)	0.94 [0.84, 1.03]	1.64	.199	-0.07 (0.04)	0.93 [0.86, 1.00]	3.56	.059
Somatic Complaints	0.01 (0.05)	1.01 [0.92, 1.11]	0.06	.810	0.03 (0.04)	1.03 [0.96, 1.11]	0.75	.387
Suicidal Ideation	0.03 (0.05)	1.03 [0.94, 1.13]	0.35	.553	0.04 (0.04)	1.04 [0.96, 1.12]	0.98	.322
Traumatic Stress	0.04 (0.05)	1.04 [0.90, 1.11]	0.79	.375	0.09 (0.04)	1.09 [1.02, 1.18]	7.00	.008
Borderline Features	-0.01 (0.05)	0.99 [0.95, 1.10]	0.01	.906	0.04 (0.04)	1.04 [0.96, 1.14]	1.01	.313
Alcohol Problems	-0.01 (0.02)	0.99 [0.93, 1.04]	0.07	.784	0.01 (0.02)	1.02 [0.98, 1.06]	0.66	.417
Drug Problems	-0.01 (0.02)	0.99 [0.95, 1.04]	0.06	.802	-0.02 (0.22)	0.98 [0.94, 1.02]	0.96	.326
		$\chi^2(9) = 6.75, p = .664, \Delta\chi^2(8) = 5.21, p = .735$				$\chi^2(9) = 22.20, p = .008, \Delta\chi^2(8) = 12.40, p = .134$		

Note. MAYSI-2 = Massachusetts Youth Screening Instrument, Version 2. PAI-A = Personality Assessment Inventory-Adolescent. *b* = Unstandardized coefficient. *SE* = Standard error. OR = Odds ratio or standardized coefficient (Exp [*b*]). 95% CI = 95% confidence intervals of OR. 95% CI of significant ORs did not contain 1.

Table 8. *Predictive Validity of MAYSI-2 and PAI-A Subscale Scores Controlling for Potential Confounding Variables*

Scale	Suicide-Related Behavior		Non-Suicidal Self-Injury	
	Partial r_s Controlling for Treatment	Partial r_s Controlling for Lifetime SRB	Partial r_s Controlling for Treatment	Partial r_s Controlling for Lifetime NSSI
MAYSI-2				
Alcohol/Drug Use	-.03	-.03	.07	.06
Angry-Irritable	-.04	-.06	.13	.07
Depressed-Anxious	.16	.14	.13	.04
Somatic Complaints	.11	.10	.10	.07
Suicidal Ideation	.16	.14	.27**	.17
Traumatic Experiences	.02	.02	.09	.08
Thought Disturbance	-.09	-.08	.22*	.21
PAI-A				
Depression	.19*	.18	.06	.01
Anxiety	-.05	-.06	.03	-.02
Somatic Complaints	.01	-.00	.14	.10
Suicidal Ideation	.18*	.16	.16	.05
Traumatic Stress	.01	-.00	.26**	.21
Borderline Features	.04	.04	.17	.13
Alcohol Problems	-.06	-.06	.02	.01
Drug Problems	-.04	-.04	.08	.06

Note. MAYSI-2 = Massachusetts Youth Screening Instrument, Version 2. PAI-A = Personality Assessment Inventory-Adolescent. r_s = Spearman rho correlation. * $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed test).