

Using the MAYSI-2 to Identify Mental Disorder Among Latino Juvenile Offenders

Jennifer Eno Louden, Tamara Kang, Elijah P. Ricks, and Lorae Marquez

The University of Texas at El Paso

©American Psychological Association, 2017. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Please do not copy or cite without author's permission. The final article is available, upon publication, at: doi: 10.1037/pas0000399

Author Note

Jennifer Eno Louden, Department of Psychology, The University of Texas at El Paso; Tamara Kang, Department of Psychology, The University of Texas at El Paso; Elijah P. Ricks, Department of Psychology, The University of Texas at El Paso; Lorae Marquez, Department of Psychology, The University of Texas at El Paso.

Elijah P. Ricks is now at Department of Psychology, Roosevelt University, Chicago.

The authors would like to thank the staff at the El Paso County Juvenile Probation Department, particularly Marc Marquez and Linda Garcia. The authors also thank Dr. James M. Wood for his assistance with interviewer training and project coordination.

Correspondence regarding this article should be addressed to Jennifer Eno Louden, Department of Psychology, The University of Texas at El Paso, 500 W. University Ave., El Paso, TX 79968. Email: jlenolouden@utep.edu

Abstract

Many juvenile justice agencies have adopted the Massachusetts Youth Screening Inventory-Second Version (MAYSI-2; Grisso & Barnum, 2006) to facilitate appropriate programming for young offenders with mental illness. Although Latinos are the fastest-growing ethnic group in the criminal justice system, there is scant research on the utility of the MAYSI-2 among Latino adolescents. The present study examined the utility of the MAYSI-2 in detecting diagnosable mental illness among 398 Latino and 60 European American adolescents in a juvenile justice agency. In addition to testing the scoring configuration used by the agency to identify adolescents in need of further attention, we tested two additional scoring configurations of the MAYSI-2. We found that the MAYSI-2 had similar utility at identifying serious mood and anxiety disorders for both ethnic groups, but was less sensitive to behavioral and substance use disorders among Latinos than it was among European Americans. In addition, the MAYSI-2 overall was less sensitive to mental illness among Latino boys compared to Latina girls. We discuss these findings within the context of best practices for identifying adolescents with mental illness in juvenile justice agencies.

Keywords: screening, psychopathology, adolescents, criminal justice, Latinos

Public Significance Statement: This study suggests that the Massachusetts Youth Screening Inventory-Second Version (MAYSI-2), a tool used by many juvenile justice agencies in the U.S. to identify adolescents with mental illness, works similarly among Latino adolescents as it does with European American adolescents. This suggests that agencies using this tool are likely not under-identifying mental illness among Latino offenders.

Using the MAYSI-2 to Identify Mental Disorder Among Latino Juvenile Offenders

Of the more than 1.4 million adolescents processed through the juvenile justice system annually (Puzzanchera & Robson, 2014), most have some form of mental illness—around 60% of these youthful offenders have a diagnosable mental illness, such as mood, anxiety, psychotic, and behavioral disorders (Teplin, Abram, McClelland, Dulcan, & Mericle, 2002). A sizeable number of these adolescents (13.7%) have a “serious mental illness” (mood, anxiety, or psychotic disorder; Kang, Wood, & Eno Loudon, 2016). Compared to their non-disordered counterparts, adolescents with mental illness tend to become involved in the juvenile justice system for less serious crimes (Skowrya & Cocozza, 2007) and, once involved in the system, are at increased risk of re-offense and further entanglement with the juvenile and adult criminal justice systems (Copeland, Miller-Johnson, Keeler, Angold, & Costello, 2007; Cottle, Lee, & Heilbrun, 2001; Wareham & Boots, 2012). To meet the needs of these juvenile offenders, many agencies have developed specialized interventions, such as mental health courts, which hold promise for improving adolescents’ criminal justice outcomes (Evans Cuellar, McReynolds, & Wasserman, 2006; Pullman et al., 2006; Redding, Lexcen, & Ryan, 2005).

Several research reports and policy statements have advocated for systematic screening for mental health problems by juvenile justice agencies as a first step towards effective service brokerage; this includes both provision of treatment for any individual who needs it, as well as the more specific goal of identifying those who meet the inclusion criteria for specific programs aimed at juvenile offenders with mental illness (see Grisso, 2007; Wasserman et al., 2003). Screening measures are typically administered to all adolescents entering an agency to identify those who possibly have psychological problems and therefore are in need of further in-depth assessment (Grisso, 2007).

Systematic screening is particularly important for members of ethnic minority groups, who represent the majority of adolescents in the juvenile justice system (Hockenberry, 2013), but are less likely than European Americans to receive mental health treatment both in the community and in juvenile justice settings (Alegria et al., 2002; Dalton, Evans, Cruise, Feinstein, & Kendrick, 2009; Rawal, Romansky, Jenuwine, & Lyons, 2004; Vega & Lopez, 2001). Of particular concern are Latino¹ adolescents, who are the fastest growing ethnic group in the criminal justice system and have a host of complex needs (Schuck, Lersch, & Verrill, 2004; Sickmund & Puzzanchera, 2014). In the present study, we examined the identification of mental illness among Latino adolescent offenders using the Massachusetts Youth Screening Inventory-Second Version (MAYSI-2; Grisso & Barnum, 2006), the most widely used mental health screening tool among U.S. juvenile justice agencies (Grisso, 2007).

The MAYSI-2

The MAYSI-2 is a self-report measure designed to be administered upon intake to juvenile justice agencies. It was developed to identify adolescents with a range of problems that are of concern to agencies, including mental illness, substance abuse, suicide risk, and custodial concerns (e.g., risk of violence towards staff; Grisso & Barnum, 2006). Like any screening tool, the MAYSI-2 is not intended to render diagnoses, but rather to identify adolescents who need further evaluation or immediate intervention (e.g., because of suicide risk; Grisso, Barnum, Fletcher, Cauffman, & Peuschold, 2001). It contains seven scales for boys and six for girls: Alcohol/Drug Use, Angry-Irritable, Depressed-Anxious, Somatic Complaints, Suicide Ideation,

¹ Researchers vary in their usage of terms to describe the ethnic groups discussed here and there is lack of consensus regarding which terms are preferred. We use the term “Latino” to reference anyone with Mexican, Central or South American, Cuban, or Puerto Rican ancestry; this term is often synonymous with “Hispanic” (Bean & Tienda, 1987). We use the term “European American” to refer to non-Hispanic Whites or Caucasians.

Traumatic Experiences, and Thought Disturbance (the Thought Disturbance scale is not interpretable for girls due to poor psychometric properties; Grisso et al., 2001).

The MAYSI-2 manual does not specify one pattern of scores that indicates which individuals are in need of services; instead, its creators offer suggestions for potential decisions based on an agency's goals. Scores on the MAYSI-2 are based on two cutoff scores that sort individuals into one of three possible categories for each scale: *normal*, *caution* (indicating some problems in this domain), and *warning* (indicating serious problems in this domain). The caution cutoffs are designed to identify adolescents who would score within the clinical range on a more comprehensive measure of psychopathology. The warning cutoff is designed to identify adolescents with the highest level of need in that domain, meaning that they scored in the top 10% of adolescents in the norm group (Grisso et al., 2001). An agency wanting to prioritize adolescents most in need of services may focus on those who score above the warning cutoff on one or more scales. Alternatively, agencies that want to provide services to all juveniles in need of them may institute some type of additional screening by juvenile justice or mental health personnel for any adolescent who scores above warning on one scale or above caution on two or more scales (see Grisso & Barnum, 2006), which is the criterion used by the agency where the present study was conducted.

The MAYSI-2 demonstrated promising psychometric properties in its development study, where the sample was about half (44%) European American (22.7% African American, 22.5% Latino; Grisso & Barnum, 2006). For example, researchers found alpha coefficients ranging from .61 to .86 (averaging .75), with the exception of Traumatic Experiences among boys at one study site ($\alpha = .51$). However, there is limited research regarding the MAYSI-2's ability to "flag" adolescents who have a diagnosable mental disorder. In other words, what is the

likelihood that a juvenile offender with a diagnosable mental illness will score above an established cutoff on the MAYSI-2? In statistical terms, this refers to the measure's sensitivity (proportion of true positives identified; Zhou, McClish, & Obuchowski, 2002). Ideally, the MAYSI-2 would have high sensitivity (meaning most adolescents with mental illness would be flagged for further assessment) and a low rate of false positives (high specificity; Zhou et al., 2002), although false positives are to be expected in screening measures and can be weeded out with further assessment (Grisso, 2007). Most information regarding the sensitivity of the MAYSI-2 comes from comparisons of scores on the MAYSI-2 to scores on more comprehensive self-report measures, such as the Millon Adolescent Clinical Inventory (Millon, 1993).

Only a few studies have examined the MAYSI-2's ability to identify adolescents with diagnosable mood, anxiety, and psychotic disorders. For example, Archer, Simonds-Bisbee, Spiegel, Handel, and Elkins (2010) compared MAYSI-2 scores for 1,192 juvenile offenders (66.8% African American, 27.3% European American) to diagnoses rendered by staff psychologists as coded from file information. They found only a few meaningful correlations between MAYSI-2 scores and diagnoses. For example, the MAYSI-2 Depressed-Anxious scale was weakly correlated at .16 with a diagnosis of a DSM-IV depressive disorder among boys (but this correlation was negligible among girls, $r = .01$), whereas some scales had little relationship to diagnosis ($r = .05$ for Thought Disturbance and presence of a psychotic disorder among boys; Archer et al., 2010).

A study conducted by Wasserman and colleagues (2004) compared MAYSI-2 scores to DSM-IV diagnoses as determined by the Voice DISC, a computer-administered structured clinical interview, for 325 adolescents in juvenile justice settings in New Jersey and South Carolina (58.2% African American, 20.3% European American). They found that the utility of

the MAYSI-2 at detecting mental disorder varied depending on both which MAYSI-2 scale was used, as well as which type of disorder was being targeted. For example, the Depressed-Anxious scale identified 73% (sensitivity) of adolescents with an affective disorder, whereas the Angry-Irritable scale had only 57% sensitivity for disruptive disorders. A similar study by this research team focusing on 246 juvenile offenders in South Carolina found similar results (see Hayes, McReynolds, & Wasserman, 2005). In sum, there have been only a few examinations of the utility of the MAYSI-2 in identifying individuals with mental illness, and the findings have been mixed.

Ethnic Differences in the MAYSI-2

Best practices for mental health assessment include ensuring that a tool is equivalent for individuals across cultures (Groth-Marnat, 2009). If the MAYSI-2 has less utility at identifying mental illness among ethnic minority groups, this may further disadvantage these already vulnerable groups because they will be less likely than European Americans to be identified by juvenile justice agencies, and therefore less likely to receive treatment referrals and other targeted programming. Among Latinos in particular, there are a number of cultural factors that affect the way that psychopathology is expressed and the way items on self-report measures are interpreted (Cuellar, 1998; Groth-Marnat, 2009; Prieto, McNeill, Walls, & Gomez, 2001). As such, an examination of the utility of the MAYSI-2 among Latinos is warranted.

Prior research on ethnic differences on the MAYSI-2 has revealed differences in patterns of scores for ethnic minority groups as compared to European Americans. For example, studies have found that European American youths score higher than other groups on most MAYSI-2 scales, with the exception of the Depressed-Anxious and Thought Disturbance scales, upon which Latino youths score higher than both European American or African American youths

(Cauffman & MacIntosh, 2006; Grisso et al., 2001; Vincent, Grisso, Terry, & Banks, 2008). A meta-analysis of MAYSI-2 data from 283 agencies concluded that, in general, ethnic differences in MAYSI-2 scores are small, but there is considerable variation in the magnitude of ethnic differences by site. What is unclear is whether these ethnic differences in MAYSI-2 scores are due to differences in the measure's utility in identifying psychopathology by ethnicity. It is unlikely that differences in MAYSI-2 scores reflect true differences in psychopathology, because research suggests that Latino offenders have rates of mental illness at least as high as European Americans (Wasserman, McReynolds, Schwalbe, Keating, & Jones, 2010). It is possible, then, that the MAYSI-2 has more utility among European Americans than with adolescents from ethnic minority groups.

The Present Study

The present study examined the utility of the MAYSI-2 at identifying individuals with serious mental illness in a juvenile justice agency. The agency operates specialty programs for juvenile offenders, including a mental health court that serves to divert offenders with mental illness into treatment and out of the juvenile justice system. A primary criterion for eligibility for the mental health court is a diagnosed mood, anxiety, or psychotic disorder. The agency relied on the MAYSI-2 as part of a comprehensive assessment procedure to identify juvenile offenders with mental illness who may be eligible for this program, as well as to identify those in need of referrals to treatment services in the community. Because the agency serves a population that is largely Latino, we sought to examine whether the MAYSI-2 performed similarly among Latino and European American adolescents.

We examined this question with two specific aims. First, we examined the utility of the MAYSI-2 at identifying mental illness among Latino juvenile offenders compared to its utility

among European American juvenile offenders. We examined the scoring configuration used by the agency to identify need for any type of further assessment (e.g., secondary screening by juvenile justice staff or assessment by a mental health professional), which is either any one scale above warning or any two scales above caution. In addition, we examined a more conservative configuration described in the MAYSI-2 manual, where an individual screens in only if he or she scores above warning on any scale. Finally, we tested one more liberal criterion not mentioned in the MAYSI-2 manual, where an individual could screen in on the MAYSI-2 by scoring above caution on any one scale; we did so to inform practices for agencies aiming to identify as many adolescents with mental disorder as possible, which was a stated goal of the agency where the study was conducted.

Second, we examined the utility of each individual MAYSI-2 scale at identifying mental illness among Latino juveniles. We did so because, although the scales are not necessarily intended to neatly align with specific diagnostic categories (Grisso & Barnum, 2006), it is useful to determine how well the domains assessed by each subscale align with conceptually similar disorders.

Method

We examined the utility of the MAYSI-2 in identifying mental illness among juvenile offenders using two forms of data stemming from a routine intake procedure in a juvenile justice agency in the Southwestern U.S. As described below, the first component was screening via the MAYSI-2, which was administered by intake officers in conjunction with other routine intake measures (e.g., risk assessment tools) that are administered to all adolescents entering the agency. The second component was a structured clinical interview administered by trained graduate students. This clinical interview was part of standard intake for adolescents in one

specific program over a period of 19 months as part of a training program for psychology graduate students; the students were employed by the juvenile probation agency and administered the interviews on the agency's behalf, but as described below, received training and supervision from university faculty. The use of structured interviews during routine intake was part of an initiative by the agency to ensure that all juvenile offenders with mental illness received services, and that those who were eligible for specialized programming such as a mental health court were identified. The analyses presented below reflect secondary analyses of these standard intake procedures, which were approved by the IRB at the primary author's university.

Participants

The sample was composed of two subsets of juvenile offenders, both of which stemmed from low-risk programs in the juvenile justice agency. The first subset included juvenile offenders in a specific diversion program who completed intake procedures at the agency during the time period of the study. By definition, these juvenile offenders were first-time offenders who had committed non-violent misdemeanors. Because these juvenile offenders were required to receive a mental health evaluation as part of intake, the interviewers attempted to interview all juvenile offenders in the diversion program ($N = 456$) during the time period of the study (May 2011 to December 2012). Three hundred fifty-one (76.9%) of these juvenile offenders were interviewed. The remaining 105 (23.1%) were unable to be interviewed for various reasons (e.g., the juvenile offender spoke only Spanish, the intake officer failed to schedule the juvenile for the clinical interview).

Demographic characteristics of the juvenile offenders from the low-risk diversion program who were interviewed are presented in the first column of Table 1. There were no statistically significant differences between the juvenile offenders from this program who were

interviewed and those who were not interviewed with respect to gender, $\chi^2(1, N = 456) = 0.93, p = .334, \phi = 0.04, 95\% \text{ CI } [-.05, .15]$, ethnicity, $\chi^2(5, N = 456) = 4.70, p = .454, \phi = 0.10, 95\% \text{ CI } [-.06, .11]$, age, $t(453) = 0.06, p = .950, d = 0.01, 95\% \text{ CI } [-.34, .36]$, or likelihood of scoring above caution, $\chi^2(1, N = 456) = 0.02, p = .871, \phi = 0.01, 95\% \text{ CI } [-.09, .11]$, or warning, $\chi^2(1, N = 456) = 0.05, p = .814, \phi = 0.01, 95\% \text{ CI } [-.08, .10]$, on the MAYSI-2 overall. The only difference between these groups was that those who were interviewed scored higher on the Suicide Ideation scale compared to those who were not interviewed, $t(418) = -2.09, p = .037, d = 0.20, 95\% \text{ CI } [-.60, -.02]$.

[INSERT TABLE 1]

The second subset of juvenile offenders included 207 individuals who were not yet assigned to a particular program at intake and were ultimately assigned to a program other than the specific diversion program described above. Most commonly, these adolescents were assigned to other specific programs for non-violent first-time offenders ($n = 45$) or their cases were dismissed before assignment to any program ($n = 18$). Demographic characteristics of this group of offenders are presented in the second column of Table 1. There were no statistically significant differences between the interviewed juvenile offenders from the diversion program compared to the juvenile offenders from the non-diversion programs with respect to gender, $\chi^2(1, N = 558) = 0.37, p = .544, \phi = 0.02, 95\% \text{ CI } [-.05, .10]$, or ethnicity $\chi^2(5, N = 558) = 8.36, p = .137, \phi = 0.12, 95\% \text{ CI } [.00, .13]$; but they were different in age, $t(556) = 4.46, p < .001, d = 0.37, 95\% \text{ CI } [.31, .81]$. Among juvenile offenders who were interviewed, those from the non-diversion programs were as likely to score above caution, $\chi^2(1, N = 558) = 1.66, p = .198, \phi = 0.05, 95\% \text{ CI } [-.03, .14]$, and above warning, $\chi^2(1, N = 558) = 0.27, p = .603, \phi = 0.02, 95\% \text{ CI } [-.05-.09]$, on the MAYSI-2 compared to the interviewed juvenile offenders in the diversion

program. Juvenile offenders from the non-diversion programs compared to juveniles in a diversion program scored similarly on the MAYSI-2, with the exception of the Angry-Irritable subscale, $t(524) = -2.65$, $p = .008$, $d = 0.23$, 95% CI [-1.02, -.15]. Given the high degree of similarity between the juvenile offenders in the non-diversion programs and those in the diversion program, we concluded that it was appropriate to combine the two groups for analyses.

The combined sample included a total of 558 adolescents. From this sample, we excluded individuals for the following reasons: the MAYSI-2 was completed in Spanish ($n = 30$), file notes indicated that the juvenile offender could not understand the questions ($n = 2$), the individual was under the age of 12 (the lowest age recommended by the manual; $n = 26$), the individual was not either Latino or European American ($n = 22$), the individual responded “no” to all 52 MAYSI-2 items—a pattern indicative of invalid responding (Maney, 2012; $n = 14$), or the only available MAYSI-2 was completed well *after* the structured interview, such that the timeline for the two measures was not aligned ($n = 6$). This yielded a final dataset with 398 Latino adolescents and 60 European American adolescents ($N = 458$), who had an average age of 14.4 years ($SD = 1.3$) and 67.5% of whom were male, as shown in the third column of Table 1.

Procedure and Measures

The MAYSI-2. As described earlier, the Massachusetts Youth Screening Instrument-Second Version (MAYSI-2) is a self-report questionnaire consisting of 52 items, each answered “yes” or “no” based on whether the item is true for the respondent “within the past few months” (Grisso & Barnum, 2006). Its manual describes standardized administration procedures for both the paper-and-pencil and computerized versions. All but two of the juvenile offenders in our sample were administered the MAYSI-2 in a paper-and-pencil format. Most completed the MAYSI-2 only once, but forty (6.4% of the total sample) had MAYSI-2 scores from multiple

administrations. In these cases, we used the MAYSI-2 that was administered most recent to the date of the clinical interview in our analyses. The average time between the MAYSI-2 and clinical interview was 58.3 days ($SD = 96.5$).

We calculated coefficient alpha (Cronbach, 1951), an index of the average inter-item association of scale items, for all MAYSI-2 subscales, separated by the respondents' ethnicity (see Table 2). The adolescents in our sample appear to have responded to most scale items in a reliable manner, although there are some scales where alpha values are lower, such as the Traumatic Experiences scale for boys. Inter-item correlations, an alternate index of scale reliability, are also presented in Table 2.

[INSERT TABLE 2]

Clinical interview. Juvenile offenders were interviewed using the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS), which conforms to the DSM-IV-TR's (American Psychiatric Association, 2004) diagnostic criteria for Axis I mental disorders (Kaufman, Birmaher, Brent, & Ryan, 1996). In its development study, the K-SADS demonstrated promising psychometric properties; for example, children with K-SADS diagnoses scored statistically significantly higher on self-report measures of psychopathology than did children without diagnoses, suggesting good concurrent validity (Kaufman et al., 1997). Both current and lifetime diagnoses were coded with the K-SADS. We used lifetime diagnoses in our analyses to conform to the criteria used in the agency. This also made sense from a methodological standpoint because (a) current diagnoses may underestimate the utility of the MAYSI—symptoms may be present within the past few months but absent at the time of the interview, and (b) many disorders are episodic, such that even individuals who are currently asymptomatic may indeed have serious mental illness (American Psychiatric Association, 2013).

Interviewers were blind to the results of interviewees' MAYSI-2 results at the time of the interview. Base rates of diagnoses yielded by the K-SADS are presented in Table 3. The most frequent diagnosis (present at time of assessment) for Latino boys was substance abuse ($n = 48$, 17.9%), followed by substance dependence ($n = 26$, 9.7%); European American boys were most frequently diagnosed with a behavior disorder ($n = 8$, 19.5%), followed by ADHD ($n = 7$, 17.1%); Latina girls were most frequently diagnosed with substance dependence ($n = 20$, 15.4%), followed by major depressive disorder ($n = 11$, 8.5%); European American girls' most frequent diagnosis was Generalized Anxiety Disorder ($n = 3$, 15.8%), although there were only 19 people within this category.

[INSERT TABLE 3]

Interviewer training and inter-rater reliability. Interviews were conducted by masters- and doctoral-level graduate students who completed a 3-day didactic training before administering the K-SADS. This training was informed by the Training and Quality Assurance Program for SCID interviewers (Ventura, Liberman, Green, Shaner, & Mintz, 1998), a training program for another structured clinical interview (see also Eno Loudon, Skeem, & Blevins, 2013). The training was conducted by a doctoral-level licensed clinical psychologist and a doctoral-level researcher with expertise in assessing mental illness in offenders (the primary author). The training included sections on general clinical interviewing skills, specific procedures for administering and scoring the K-SADS, and role-plays. Interviewers received clinical supervision on every K-SADS with at least one of the two trainers.

To compute inter-rater reliability, interviewers recorded random interviews of juvenile offenders with permission from the adolescent and his or her parent. Secondary raters who were blind to the original interviewer's ratings listened to each recording and provided their

assessment for each item on the K-SADS. Over the course of the study, interviewers ($N = 6$) recorded 14 interviews. Two to four other interviewers listened to and scored each recording using K-SADS criteria, yielding a total of 50 inter-rater comparisons. A weighted kappa statistic was computed for each comparison; weighted kappa is an index of chance-corrected agreement for ratio scales, and weights disagreements by the magnitude of the difference between ratings (i.e., a rating of 3 disagrees more with a rating of 1 than 2; Cohen, 1968). The mean weighted kappa for all 50 comparisons was .82 ($SD = .14$). A kappa of .70 is conventionally considered adequate inter-rater reliability (Banerjee, Capozzoli, McSweeney, & Sinha, 1999), and also fit our reliability aims.

Results

To assess the MAYSI-2's utility at identifying individuals with mental illness, we computed indices of diagnostic accuracy: sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV; see Zhou et al., 2002). Although each of these indices describes a different aspect of the MAYSI-2's utility, which indices should be emphasized depends on the consequences of different types of incorrect classifications made by the measure (Zhou et al., 2002), as well as the context of the classifications. Sensitivity and specificity are most helpful when examining a measure's utility at the group level, as they indicate which proportion of individuals with mental illness in an agency are correctly classified, whereas PPV and NPV are more useful when making decisions regarding an individual's scores on a measure. As a screening tool, the primary goal of the MAYSI-2 is to identify individuals in need of further assessment, and false positives can be weeded out with further assessment. As such, although we present the full range of utility indices, we focus our discussion primarily on the measure's sensitivity and specificity in an effort to inform agency administrators in selecting decision

criteria based on the MAYSI-2 cutoff scores. In addition, we present area under the curve (AUC; Hanley & McNeil, 1982) indices and their corresponding confidence intervals as an overall measure of the accuracy of classifications made by the MAYSI-2.

Aim 1: Utility of the MAYSI-2 Among Latino and European American Adolescents

To examine the MAYSI-2's utility at detecting mental illness in general as well as the specific disorders targeted by the mental health programming within the agency, we tested the utility of the MAYSI-2 against two different categories of diagnoses. In the first five rows of Table 4, we present the utility of the MAYSI-2 at identifying adolescents with any DSM-IV-TR Axis I disorder, as based on the K-SADS clinical interview. We refer to this criterion as "any disorder." The second set of five rows presents the utility of the MAYSI-2 at identifying serious mental illness, this time categorizing an adolescent as having a mental disorder only if he or she received a diagnosis of a DSM-IV-TR mood, anxiety, or psychotic disorder. Here, adolescents who have a diagnosis of *only* a behavioral disorder (e.g., oppositional defiant disorder) or a substance use disorder were categorized as non-disordered. The columns in Table 4 represent the three different scoring configurations of the MAYSI-2 we examined. As described earlier, these include: one scale above caution, *either* one scale above warning or two scales above caution cutoff ², or one scale above warning.

[INSERT TABLE 4 HERE]

We computed the utility indices separately for Latinos and European Americans, but combined genders because there were not enough European Americans in our sample to render

² We also tested a second scoring configuration suggested by the MAYSI-2 manual where an individual screens in if he or she scores above caution on any two scales. We do not report these results here because they are virtually identical to the results using the agency's scoring configuration of either one scale above warning or two scales above caution. Specifically, all but four individuals in the entire sample were classified in the same manner by these two configurations.

precise estimates by gender. As shown in Table 4, most scoring configurations had similar sensitivity for both ethnic groups. Of primary concern to the agency, the MAYSI-2 flagged 69% of Latino adolescents and 71% of European American adolescents with serious mental illness. Although the MAYSI-2 was more sensitive to both any disorder and serious mental illness at the more liberal one-caution cutoff we tested, this came with decreased specificity, which is to be expected. For example, specificity was 0.70 among Latino adolescents in detecting serious mental illness at the agency's cutoff, but 0.45 when using the one-caution cutoff; this means that the agency's cutoff generated a 30% false positive rate whereas the more liberal cutoff generated a false positive rate of 55%. Comparing the two ethnic groups, there were a few notable differences. First, the MAYSI-2 was more sensitive to serious mental illness at the one-warning cutoff for Latinos than for European Americans (0.51 versus 0.35, respectively). In addition, there was higher specificity for European Americans than Latinos when detecting serious mental illness using the agency's cutoff (where an individual screens in if they score above warning on one scale or above caution on two scales). The AUC values, which reflect both sensitivity and specificity, differed by ethnicity for the agency's cutoff, where the AUC value for European Americans equaled the upper bound of the confidence interval for that of Latinos. This was the case for both any disorder and serious mental illness. However, this appears to be more of a function of higher specificity for European Americans for the MAYSI-2 rather than differences in sensitivity. The AUC values for the other scoring configurations were quite similar between the ethnic groups.

Because there are different rates of mental disorder for boys and girls (Teplin et al., 2002) and a different number of MAYSI-2 scales for boys and girls (Grisso & Barnum, 2006), we present these indices separately by gender for Latinos in Table 5. Differences can be seen in the

MAYSI-2's utility by gender across most of the scoring configuration/mental disorder criteria we tested, where the MAYSI-2 was more sensitive to mental disorder among girls than among boys. For example, although 79% of Latina girls with serious mental illness were flagged by the MAYSI-2 using the agency's cutoff (one scale above warning or two above caution), only 61% of Latino boys were. The more liberal cutoff (one scale above caution) improved identification for both genders, but as expected, this led to lower specificity, which was particularly problematic for girls. For example, of girls who screened in with one scale above caution, 65% had no disorder, and this low specificity can be seen in the difference in AUC values for the two genders. Further, Latino boys are at particular risk of not being identified by the most conservative cutoff of one score above warning, which is apparent both from the low sensitivity of this configuration and the lack of overlap between the confidence intervals in AUC compared to girls.

[INSERT TABLE 5 HERE]

Aim 2: Utility of Individual MAYSI-2 Scales Among Latino Adolescents

To address the second aim, we computed utility indices for each of the MAYSI-2 scales individually, both at the caution cutoff and the warning cutoff. We tested each of these scales' utility at identifying adolescents with categories of disorders we assessed in the clinical interview: mood, anxiety, behavioral, and substance use. We did not assess the utility of the MAYSI-2 in detecting psychotic disorders, because the low base rates of these disorders (1.5% overall) would yield imprecise estimates. Although each individual scale has varying degrees of overlap with each of these disorders, it is informative to examine the extent to which each scale is sensitive to each type of psychopathology. Although we computed indices for every combination of scale/cutoff and disorder category, we focus our discussion primarily on the

utility of each scale to detect the disorder category (or categories) with which it overlaps conceptually—these figures are presented in bold in Table 6.

[INSERT TABLE 6 HERE]

As shown, there were generally low levels of sensitivity for many MAYSI-2 scales among Latino boys. This was true even for the relatively common behavioral and substance use disorders; for example, the Alcohol-Drug scale at the caution cutoff correctly identified only 37% of boys who qualified for a diagnosis of a substance use disorder. Among boys, it appears that the Somatic Complaints scale was more sensitive to anxiety and depression than was the Depressed-Anxious scale. For example, at the caution cutoff, the Somatic Complaints scale correctly identified 70% of Latino boys with a mood disorder and 77% of Latino boys with an anxiety disorder; the sensitivity of the Depressed-Anxious scale was 48% and 35%, respectively, for these disorders.

The MAYSI-2 was somewhat more sensitive to behavioral and substance use disorders among Latina girls than among Latino boys; within our sample, 65% of girls with behavioral disorders were correctly identified with the Angry-Irritable scale and 60% of girls with substance use disorders were correctly identified with the Alcohol-Drug scale. The Alcohol-Drug scale in particular had excellent specificity, which makes sense given that it is unlikely for individuals without substance abuse problems to report them in a juvenile justice setting. The Depressed-Anxious scale was more sensitive among Latina girls as well: it correctly identified 74% of girls with a mood disorder and 77% of girls with an anxiety disorder. As we found with boys, the Somatic Complaints scale was also sensitive to mood and anxiety disorders among Latina girls (67% and 88% correctly identified, respectively), although the specificity was lower for this scale than for the Depressed-Anxious scale.

Discussion

The present study provides evidence for the utility of the widely used Massachusetts Youth Screening Inventory-Second Version (MAYSI-2; Grisso & Barnum, 2006) among Latino adolescent offenders. Specifically, we examined the utility of the MAYSI-2 in identifying adolescents with diagnosable mental illness in a juvenile justice agency with a primarily Latino population and had two overarching findings. First, our comparison of the MAYSI-2's utility with Latinos versus European Americans revealed similar sensitivity across the two ethnic groups. Second, closer examination of the MAYSI-2's utility by gender demonstrated that the measure had less utility among Latino boys than among Latina girls, particularly at identifying behavioral and substance use disorders. These findings are discussed in greater detail below.

The MAYSI-2 Does Not Under-Identify Latinos

We found that the MAYSI-2 was similarly sensitive to all forms of mental illness among Latinos and European Americans. This is encouraging, given that Latinos are the fastest growing ethnic group in the criminal justice system (Schuck et al., 2004), yet Latino offenders are at disproportionate risk of not receiving mental health treatment (Dalton et al., 2009; Rawal et al., 2004). Although prior research has found different patterns of scoring on the MAYSI-2 when comparing Latinos to European Americans (Vincent et al., 2008), the present study suggests that this is not likely related to ethnic differences in sensitivity of the measure. Our results suggest that the MAYSI-2 adequately identified individuals with mental illness in the agency where this research was conducted, and more generally, indicates that the MAYSI-2, the most popular mental health screening tool among juvenile justice agencies (Grisso, 2007), likely does not systematically under-identify mental illness among Latino adolescents. If anything, it is possible that we underestimated the sensitivity of the MAYSI-2, given that it is designed to

detect recent and ongoing symptoms (Grisso & Barnum, 2006) and we tested its ability to detect lifetime diagnoses. We likely would have found even higher levels of sensitivity for the MAYSI-2 had we confined our analyses to present diagnoses.

As expected, we found different levels of sensitivity for the MAYSI-2 when testing different scoring configurations, and these differences existed for Latino boys and Latina girls as well as European Americans. As noted earlier, the MAYSI-2 manual provides multiple possible scoring configurations for agencies to consider based on the agency's goals (Grisso & Barnum, 2006). The scoring criteria used by the agency where this research was conducted, where an individual screens in if he or she either scores above warning on one scale or above caution on two scales, identified about 70% of individuals with mood, anxiety, and psychotic disorders. A more conservative cutoff (one scale above warning) suggested in the MAYSI-2 manual for agencies most concerned with conserving resources would identify far fewer adolescents with these disorders—specifically, only about half (51%) of Latino and about a third (35%) of European American adolescents with diagnosable serious mental illness scored above warning on one MAYSI-2 scale.

The more liberal scoring configuration that we tested, where an individual screens in if he or she scores above caution on one scale, identified the vast majority of adolescents with serious mental illness (90% of Latinos and 88% of European Americans). Notably, this criterion is not part of the “best practices” suggested by the MAYSI-2 manual (Grisso & Barnum, 2006), but agencies are given discretion to determine the scoring criteria they would like to use. The increase in sensitivity associated with this more liberal cutoff score came at the cost of an increased rate of false positives—around half of adolescents who scored above caution on one scale did not have a diagnosable mental illness. However, this high rate of false positives does

not necessarily outweigh the benefits of the increased sensitivity of this cutoff. There are several low-cost methods of providing a second assessment to adolescents who are flagged by the MAYSI-2. For example, the MAYSI-2 manual suggests that a probation officer can provide a secondary assessment by asking follow-up questions to clarify “yes” responses on the MAYSI-2 (Grisso & Barnum, 2006). Agencies wanting more thorough mental health assessments could follow the model used by the agency where the present study was conducted, where graduate students supervised by a faculty member licensed as a psychologist conducted assessments as part of their clinical training.

Further, the absence of a diagnosable mental illness does not indicate the lack of need for services. As noted earlier, the MAYSI-2 is designed to detect a wide range of symptoms and distress, not just diagnosable mental illness (Grisso & Barnum, 2006). It is likely that some of the adolescents who screened in on the MAYSI-2 in this study who do not have a diagnosable mental illness have behavioral problems that may be of concern to the agency. In addition, some adolescents who screened in on the MAYSI-2 likely have subclinical distress, meaning that they have some symptoms of mental illness, but either not enough in number or of insufficient severity to meet diagnostic criteria. Recent research suggests that juvenile offenders with subclinical depression (Kang, Eno Loudon, Ricks, & Jones, 2015) or subclinical symptoms of bipolar disorder (Kang, Perez, & Eno Loudon, 2014) have rates of problematic symptoms such as aggression and substance abuse similar to those of diagnosed offenders. As such, further assessment and even treatment are likely appropriate for many adolescents who screen in on the MAYSI-2 regardless of whether a diagnosable mental illness is present.

Concerns With Using the MAYSI-2 With Latino Boys

Although we found promising evidence for the utility of the MAYSI-2 among Latinos as a group, we found meaningful differences among Latinos by gender. Specifically, for all except the most liberal scoring configuration we tested, the MAYSI-2 was less sensitive among Latino boys than among Latina girls. This was especially true when using the MAYSI-2 to identify behavioral and substance use disorders; the MAYSI-2 at best correctly identified 67% of Latino boys (compared to 88% of Latina girls, at the most liberal cutoff) and at worst correctly identified only 22% of Latino boys (compared to 51% of Latina girls at the most conservative cutoff). This is in contrast to other research finding higher sensitivity for these disorders. For example, Wasserman and colleagues (2004) found that the Angry-Irritable scale had sensitivity of 0.57 for behavioral disorders in their sample whereas in our sample the rate for Latino boys was 0.44. Importantly, the Angry-Irritable scale is not designed specifically to identify adolescents with behavioral disorders, which are defined by a range of symptoms, including rule violations (American Psychiatric Association, 2013; Grisso & Barnum, 2006); rather, the scale assesses constructs such as tendencies towards anger and aggression (Grisso & Barnum, 2006).

Similarly, we found that the MAYSI-2 had low sensitivity for identifying substance use disorders among Latino boys. Even at the most liberal scoring cutoff, the Alcohol-Drug scale correctly identified only 37% of Latino boys who had a diagnosable substance use disorder. This is in contrast to Wasserman and colleagues' (2004) finding of 0.65 sensitivity for substance use disorders in their sample. Research suggests that there are no meaningful differences in rates of self-reporting substance abuse for Latino juvenile offenders compared to European American juvenile offenders (Kim, Fendrich, & Wislar, 2000). Although there are lower rates of endorsement of substance use in juvenile probationers compared to juvenile offenders in residential placements (Coker et al., 2013), this does not explain the ethnic differences we found.

A meta-analysis on the MAYSI-2 (Vincent et al., 2008) found that, on average, Latinos scored lower than European Americans on the Alcohol-Drug scale. Whether this reflects differences in rates of substance use or differences in reporting is a question that warrants further research.

Although the MAYSI-2 scales were overall less sensitive among Latino boys, the Somatic Complaints and Angry-Irritable scales were helpful at identifying some types of psychopathology in this group. Specifically, the Somatic Complaints scale had higher sensitivity to mood and anxiety disorders among Latino boys than did the Depressed-Anxious scale. There is little research on this scale specifically (Wasserman et al., 2004, did not report indices for it), but it was designed to identify physical symptoms consistent with depression and anxiety. Latinos in general are more likely to experience and report mood and anxiety symptoms as physical symptoms than are European Americans, and may not recognize an underlying mood disturbance (Grames, 2006; Katon, Kleinman, & Rosen, 1982). Further, Latinos may confound the concept of anxiety with the culture-specific “*nervios*,” which refers to general distress accompanied by multiple physiological symptoms (American Psychiatric Association, 2013; Grames, 2006; Lopez & Guarnaccia, 2000). Our findings may indicate lack of conceptual equivalence among Latinos for the Depressed-Anxious and Somatic Complaints scales, meaning that the constructs of anxiety and depression are different among Latinos such that measures designed to tap these constructs among European Americans are insufficient for Latinos (see Okazaki & Sue, 1995). Similarly, we found that the Angry-Irritable scale was more sensitive to mood disorders among Latino boys than was the Depressed-Anxious scale. This could be related to gender differences in the manifestation of depression; men with depression report symptoms of anger, aggression, and risk-taking behaviors, which some researchers have suggested is a gender-specific expression of depression (Martin, Neighbors, & Griffith, 2013). Further, there is

some evidence that Latino boys express internalizing symptoms as anger because the expression of emotion, such as depression, is perceived as a threat to masculinity, which is particularly important among Latinos (see Anderson & Mayes, 2010). Notably, the decision-making criteria used by the agency in this study calls for increased staff vigilance for any individual scoring above warning on the Angry-Irritable scale due to increased likelihood of problematic behaviors, but this would not necessarily lead to further assessment. It may be worthwhile to assess the utility of using this scale to identify mood disorders among Latino boys. Given the flexibility allowed in scoring the MAYSI-2, particularly when using the MAYSIWARE software, it is possible that future research could uncover specific combinations of scales and cutoffs that maximally identify mental disorder among different subgroups of adolescents.

Limitations

As with any research, there are some limitations to the present study that should be borne in mind when interpreting its results. First, our data were collected among primarily Mexican American adolescents in the Southwestern United States in an agency that predominantly serves Latino offenders. Given the heterogeneity among Latinos as a group (Bean & Tienda, 1987), further research should assess the utility of the MAYSI-2 among other Latino subgroups. Similarly, there is heterogeneity within Latino subgroups, and although the vast majority of Latino adolescents in our sample were of Mexican American origin, we did not assess acculturation (degree of identification with Mexican versus mainstream U.S. culture), which is known to affect psychological assessment (see Okazaki & Sue, 1995). It is likely that our sample was different in some ways from Mexican Americans in other regions given that Latinos are the majority ethnic group in the region where this study was conducted. Second, our comparison group of European Americans was small, so we were unable to examine gender

differences in the MAYSI-2 within this group, and this may limit the precision of our analyses with them. As such, our indices for the utility of the MAYSI-2 among European Americans should be interpreted cautiously. Third, although about 10% of the adolescents in the agency where we conducted this research completed the MAYSI-2 in Spanish, those adolescents were not included in the study. Future research is needed to determine the utility of the translated MAYSI-2. Fourth, there is limited research examining the utility of the K-SADS among Latino youth, although this is also the case for other comparable semi-structured interviews (see Polo & Lopez, 2008). We feel that this tool was an appropriate choice for the criterion in this research given that it is specifically designed to elicit symptoms that align with DSM diagnostic criteria, which have been shown to apply reasonably well to Latino adolescents (Green et al., 2012). Finally, most of our sample was low-risk, first-time offenders in community-based programs, so the extent to which the pattern of results we obtained would generalize to offenders in detention settings is unknown. The MAYSI-2 development study (Grisso & Barnum, 2006) identified differences in patterns of scores between offenders in different types of legal settings (e.g., detention versus probation); we suggest that these differences are likely due to actual differences in symptoms across groups (e.g., higher risk offenders have more substance abuse problems). Additional research is needed to examine whether the MAYSI-2 has similar utility among Latino juvenile offenders across different risk levels.

Conclusions

The present study adds to the limited research on the utility of the MAYSI-2 among Latino juvenile offenders, but more research is needed to uncover the nuances involved in the assessment of this group. Assessment with bicultural and bilingual individuals is complex (Malgady, Rogler, & Costantino, 1987); issues such as the extent to which different forms for

Latinos on the MAYSI-2 are needed should be examined. More research is needed to uncover the underlying causes of the differences we found between Latinos and European Americans on the MAYSI-2 and to assess the utility of scoring configurations beyond those we tested here, including those used by the MAYSIWARE software.

The present findings may be distilled into some preliminary recommendations for juvenile justice agencies. First, because we found that the MAYSI-2 demonstrated similar sensitivity to serious mental illness among Latino juvenile offenders as it did for European Americans, it can be confidently included in standard intake procedures without concerns of systematic misidentification of Latinos with mental illness. However, because we found that Latinos with mood and anxiety disorders frequently screened in on the Somatic Complaints and Angry-Irritable scales, agencies may need to attend to these scales in addition to the Depressed-Anxious scale when assessing for mood and anxiety disorders among Latino juvenile offenders. In addition, agencies may need to supplement the MAYSI-2 with another measure of substance abuse, particularly for Latino boys, to more accurately identify these issues.

Decisions regarding which cutoff score to use in determining which juvenile offenders should be referred for secondary screening or treatment need to be made taking into account agency goals and consequences of incorrect classifications (Grisso & Barnum, 2006). Agencies that seek to identify the majority of individuals with mental illness may consider using a more liberal cutoff to maximize sensitivity and communicate with staff regarding the expected rates of false positives. The most important issue for agencies is ensuring that referrals to treatment and further assessment are made systematically—some research indicates that referrals are disproportionately made for girls and European Americans, regardless of mental health status (Redding et al., 2005). Because Latinos are at particular disadvantage when it comes to

receiving mental health treatment in juvenile justice settings (Dalton et al., 2009; Rawal et al., 2004), it is imperative that a proper screening procedure reliably inform mental health treatment and specialized programming designed to improve justice system outcomes for young offenders.

References

- Alegria, M., Canino, G., Rios, R., Vera, M., Calderon, J., Rusch, D., & Ortega, A. N. (2002). Mental health care for Latinos: Inequalities in use of specialty mental health services among Latinos, African Americans, and non-Latino whites. *Psychiatric Services, 53*, 1547-1555. doi:10.1176/appi.ps.53.12.1547
- American Psychiatric Association. (2004). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.
- Anderson, E. R., & Mayes, L. C. (2010). Race/ethnicity and internalizing disorders in youth: A review. *Clinical Psychology Review, 30*, 338-348. doi:10.1016/j.cpr.2009.12.00
- 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
- Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *Canadian Journal of Statistics, 27*, 3-23. doi:10.2307/3315487
- Bean, F. D., & Tienda, M. (1987). *The Hispanic population of the United States*. New York, NY: The Russell Sage Foundation.
- Cauffman, E., & MacIntosh, R. (2006). A Rasch differential item functioning analysis of the Massachusetts Youth Screening Instrument: Identifying race and gender differential item

- functioning among juvenile offenders. *Educational and Psychological Measurement*, 66, 502-521. doi:10.1177/0013164405282460
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70, 213-220. doi:10.1037/h0026256
- Coker, K. L., Wernsman, J., Ikpe, U. N., Brooks, J. S., Bushell, L., & Kahn, B. A. (2013). Using the Massachusetts Youth Screening Instrument-Version 2 on a community sample of African American and Latino/a juvenile offenders to identify mental health and substance abuse treatment needs. *Criminal Justice and Behavior*, 4, 492-511. doi:10.1177/0093854813505565
- Copeland, W. E., Miller-Johnson, S., Keeler, G., Angold, A., & Costello, E. J. (2007). Childhood psychiatric disorders and young adult crime: A prospective, population-based study. *American Journal of Psychiatry*, 164, 1668-1675. doi:10.1176/appi.ajp.2007.0612202
- Cottle, C. C., Lee, R. J., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles: A meta-analysis. *Criminal Justice and Behavior*, 28, 367-394. doi:10.1177/0093854801028003005
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. doi:10.1007/BF02310555
- Cuellar, I. (1998). Cross-cultural clinical psychological assessment of Hispanic Americans. *Journal of Personality Assessment*, 70, 71-86. doi:10.1207/s15327752jpa7001_5
- Dalton, R. F., Evans, L. J., Cruise, K. R., Feinstein, R. A., & Kendrick, R. F. (2009). Race differences in mental health service access in a secure male juvenile justice facility. *Journal of Offender Rehabilitation*, 48, 194-209. doi:10.1080/1050967090276657

- Eno Louden, J., Skeem, J. L., & Blevins, A. (2013). Comparing the predictive utility of two screening tools for mental disorder among probationers. *Psychological Assessment, 25*, 405-415. doi:10.1037/a0031213
- Evans Cuellar, A., McReynolds, L. S., & Wasserman, G. A. (2006). A cure for crime: Can mental health treatment diversion reduce crime among youth? *Journal of Policy Analysis and Management, 25*, 197-214. doi:10.1002/pam.20162
- Grames, H. A. (2006). Depression, anxiety, and *ataque de nervios*: The primary mental health care model in a Latino population. *Journal of Systemic Therapies, 25*, 58-72. doi:10.1521/jsyt.2006.25.3.5
- Green, J. G., Gruber, M. J., Kessler, R. C., Lin, J. Y., McLaughlin, K. A., Sampson, N. A., . . . Alegria, M. (2012). Diagnostic validity across racial and ethnic groups in the assessment of adolescent DSM-IV disorders. *International Journal of Methods in Psychiatric Research, 21*, 311-320. doi:http://0-dx.doi.org.lib.utep.edu/10.1002/mpr.137
- Grisso, T. (2007). Progress and perils in the juvenile justice and mental health movement. *Journal of the American Academy of Psychiatry and the Law, 35*, 158-167.
- Grisso, T., & Barnum, R. (2006). *Massachusetts Youth Screening Instrument-2 (MAYSI-2): User's manual and technical report, 2006 revised edition*. Sarasota, FL: Professional Resource Press.
- Grisso, T., Barnum, R., Fletcher, K. E., Cauffman, E., & Peuschold, D. (2001). Massachusetts youth screening instrument for mental health needs of juvenile justice youths. *Journal of the American Academy of Child Adolescent Psychiatry, 40*, 541-548. doi:10.1097/00004583-200105000-0001
- Groth-Marnat, G. (2009). *Handbook of psychological assessment* (5th ed.). Hoboken, NJ: Wiley.

- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29-36.
doi:10.1148/radiology.143.1.706374
- Hayes, M. A., McReynolds, L. S., & Wasserman, G. A. (2005). Paper and voice MAYSI-2 format comparability and concordance with the voice DISC-IV. *Assessment*, *12*, 395-403.
doi:10.1177/1073191105280359
- Hockenberry, S. (2013). *Juveniles in residential placement, 2010*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved from
<http://www.ojjdp.gov/pubs/241060.pdf>
- Kang, T., Eno Loudon, J., Ricks, E. P., & Jones, R. L. (2015). Aggression, substance use disorder, and presence of a suicide attempt among juvenile offenders with subclinical depression. *Law and Human Behavior*, *39*, 593-601.
<http://dx.doi.org/10.1037/lhb0000145>
- Kang, T., Perez, E., & Eno Loudon, J. (2014, March). *Correlates of subclinical bipolar disorders in a juvenile justice setting*. Poster presented at the meeting of the American Psychology-Law Society, New Orleans, LA.
- Katon, W., Kleinman, A., & Rosen, G. (1982). Depression and somatization: A review. *The American Journal of Medicine*, *72*, 127-135. doi:10.1016/0002-9343(82)90599-
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P... Brent, D. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 980-988.
doi:10.1097/00004583-199707000-00021

- Kaufman, J., Birmaher, B., Brent, D., Rao, U., & Ryan, N. (1996). *Kiddie-SADS-Present and Lifetime Version (K-SADS-PL)*. Pittsburgh, PA: University of Pittsburgh, School of Medicine.
- Kim, J. Y. S., Fendrich, M., & Wislar, J. S. (2000). The validity of juvenile arrestees' drug use reporting: A gender comparison. *Journal of Research in Crime and Delinquency*, *37*, 419-432. doi:10.1177/0022427800037004004
- Lopez, S. R., & Guarnaccia, P. J. J. (2000). Cultural psychopathology: Uncovering the social world of mental illness. *Annual Review of Psychology*, *51*, 571-598.
doi:10.1146/annurev.psych.51.1.57
- Malgady, R. G., Rogler, L. H., & Costantino, G. (1987). Ethnocultural and linguistic bias in mental health evaluation of Hispanics. *American Psychologist*, *42*, 228-234.
doi:10.1037/0003-066X.42.3.22
- Maney, S. (2012, March). Violent and nonviolent juvenile offenders: Comparisons concerning self-report of mental health symptoms and response styles. In T. Grisso (Chair), *A national investigation of the mental health symptoms reported by juvenile offenders*. Symposium conducted at the meeting of the American Psychology-Law Society, San Juan, Puerto Rico.
- Martin, L. A., Neighbors, H. W., & Griffith, D. M. (2013). The experience of symptoms of depression in men vs. women: Analysis of the national comorbidity survey replication. *JAMA Psychiatry*, *70*, 1100-1106. doi:10.1001/jamapsychiatry.2013.198
- Millon, T. (1993). *Millon Adolescent Clinical Inventory: Manual*. Minneapolis, MN: National Computer Systems.

- Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment, 7*, 367-375. doi:10.1037/1040-3590.7.3.36
- Pullmann, M. D., Kerbs, J., Koroloff, N., Veach-White, E., Gaylor, R., & Sieler, D. (2006). Juvenile offenders with mental health needs: Reducing recidivism using wraparound. *Crime & Delinquency, 52*, 375-397. doi:10.1177/0011128705278632
- Prieto, L. R., McNeill, B. W., Walls, R. G., & Gomez, P. (2001). Chicanas/os and mental health services: An overview of utilization, counselor preference, and assessment issues. *The Counseling Psychologist, 29*, 18-54. doi:10.1177/001100000129100
- Puzzanchera, C., & Robson, C. (2014). *Delinquency cases in juvenile court, 2010*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved from <http://www.ojjdp.gov/pubs/243041.pdf>
- Rawal, P., Romansky, J., Jenuwine, M., & Lyons, J. S. (2004). Racial differences in the mental health needs and service utilization of youth in the juvenile justice system. *The Journal of Behavioral Health Services and Research, 31*, 242-254. doi:10.1007/BF02287288
- Redding, R. E., Lexcen, F., & Ryan, E. P. (2005). Mental health treatment for juvenile offenders in residential psychiatric and juvenile justice settings. In K. Heilbrun, N. Goldstein, & R. Redding (Eds.), *Juvenile delinquency: Prevention, assessment, and intervention* (pp. 282-309). New York, NY; Oxford.
- Schuck, A. M., Lersch, K. M., & Verrill, S. W. (2004). The “invisible” Hispanic? The representation of Hispanics in criminal justice research. *Journal of Ethnicity in Criminal Justice, 2*, 5-22. doi:10.1300/J222v02n03_02
- Sickmund, M., & Puzzanchera, C. (2014). *Juvenile offenders and victims: 2014 National Report*. Pittsburgh, PA: National Center for Juvenile Justice.

- Skowrya, K. R., & Coccozza, J. J. (2007). *Blueprint for change: A comprehensive model for the identification and treatment of youth with mental health needs in contact with the juvenile justice system*. Delmar, NY: National Center for Mental Health and Juvenile Justice.
- Retrieved from http://www.ncmhjj.com/wp-content/uploads/2013/07/2007_Blueprint-for-Change-Full-Report.pdf
- 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
- Vega, W. A., & Lopez, S. R. (2001). Priority issues in Latino mental health services research. *Mental Health Services*, *3*, 189-200. doi:10.1023/A:101312503071
- Ventura, J., Liberman, R. P., Green, M. F., Shaner, A., & Mintz, J. (1998). Training and quality assurance with the structured clinical interview for DSM-IV (SCID-I/P). *Psychiatry Research*, *79*, 163-173. doi:10.1016/S0165-1781(98)00038-9
- Vincent, G. M., Grisso, T., Terry, A., & Banks, S. (2008). Sex and race differences in mental health symptoms in juvenile justice: The MAYSI-2 national meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*, 282-290.
- doi:10.1097/CHI.0b013e318160d516
- Wareham, J., & Boots, P. (2012). The link between mental health problems and youth violence in adolescence: A multilevel test of DSM-oriented problems. *Criminal Justice and Behavior*, *39*, 1003-1024. doi:10.1177/0093854812439290
- Wasserman, G. A., Jensen, P. S., Ko, S. J., Coccozza, 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.chi0000046873.56865.4B

Wasserman, G. A., McReynolds, L., Ko, S., Katz, L., Cauffman, E., Haxton, W., & Lucas, C. (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:10.1097/01.chi00000116742.71662.b1

Wasserman, G. A., McReynolds, L. S., Schwalbe, C. S., Keating, J. M., & Jones, S. A. (2010). Psychiatric disorder, comorbidity, and suicidal behavior in juvenile justice youth. *Criminal Justice and Behavior*, 37, 1361-1376. doi:10.1177/0093854810382751

Zhou, X.-H., McClish, D. K., & Obuchowski, N. A. (2002). *Statistical methods in diagnostic medicine, Part 1*. New York, NY: John Wiley & Sons Inc.

Table 1

Characteristics of All Interviewed Adolescents From Diversion and Non-Diversion Programs

| | Diversion Program (Interviewed) | Non-Diversion Program (Interviewed) | Final Sample (All interviewed adolescents used for analyses) |
|-------------------------|---------------------------------------|---|---|
| | <i>n</i> = 351 | <i>n</i> = 207 | <i>N</i> = 458 |
| Ethnicity | | | |
| Latino | 81.20% | 88.90% | 86.90% |
| European American | 13.40% | 9.70% | 13.10% |
| Other | 5.40% | 1.40% | 0.00% |
| Gender | | | |
| Female | 33.90% | 31.40% | 32.50% |
| Male | 66.10% | 68.60% | 67.50% |
| Age | | | |
| Mean (<i>SD</i>) | 14.5 (1.5)** | 15.1 (1.2)** | 14.4 (1.3) |
| Range | 10.0-17.0 | 12.0-17.0 | 12.0-17.0 |
| MAYSI 2 Subscale Scores | Mean (<i>SD</i>) | Mean (<i>SD</i>) | Mean (<i>SD</i>) |
| Alcohol-Drug | 1.3 (1.8) | 1.4 (1.8) | 1.5 (1.9) |
| Angry-Irritable | 3.1 (2.4)* | 2.5 (2.5)* | 3.0 (2.5) |
| Depressed-Anxious | 1.9 (1.7) | 1.7 (1.9) | 1.8 (2.5) |
| Somatic Complaints | 1.3 (1.8) | 1.4 (1.8) | 1.5 (1.9) |
| Suicidal Ideation | 0.7 (1.4) | 0.6 (1.2) | 0.7 (1.3) |
| Thought Disturbance | 0.7 (1.0) | 0.5 (0.9) | 0.5 (0.9) |
| Traumatic Experiences | 1.2 (1.2) | 1.1 (1.2) | 1.2 (1.3) |

Note: The first two columns report characteristics of all interviewed adolescents from diversion program v. non-diversion program. The last column titled “Final Sample” is the final sample with all exclusions that was used for the analyses (see methods section for details on exclusions). Therefore, column 1 (interviewed juvenile offenders in a diversion program) and column 2 (interviewed juvenile offenders in a non-diversion program) do not include exclusions, and as a result, do not add up to the final sample provided in the 3rd column (**p* < .05, ***p* < .001).

Table 2

Reliability Coefficients for MAYSI-2 Subscales

| Subscale | # Items | <i>n</i> | Latino Youth | | | European Amer. Youth | | | |
|---------------------------------|---------|----------|--------------|------------|----------|----------------------|----------|------------|----------|
| | | | α | 95% CI | <i>r</i> | <i>n</i> | α | 95% CI | <i>r</i> |
| Alcohol/Drug Use | 8 | 398 | .79 | [.76, .82] | .29 | 60 | .68 | [.55, .79] | .21 |
| Angry-Irritable | 9 | 398 | .79 | [.75, .82] | .29 | 60 | .77 | [.68, .85] | .28 |
| Depressed-Anxious | 9 | 398 | .70 | [.66, .74] | .21 | 60 | .71 | [.60, .81] | .21 |
| Somatic Complaints | 6 | 398 | .66 | [.61, .71] | .25 | 60 | .59 | [.41, .73] | .20 |
| Suicide Ideation | 5 | 398 | .86 | [.83, .88] | .55 | 60 | .84 | [.76, .89] | .54 |
| Thought Disturbance (boys only) | 9 | 268 | .68 | [.62, .74] | .21 | 41 | .72 | [.58, .84] | .19 |
| Traumatic Experiences (boys) | 5 | 268 | .49 | [.38, .58] | .16 | 41 | .55 | [.29, .74] | .20 |
| Traumatic Experiences(girls) | 5 | 130 | .61 | [.49, .71] | .24 | 19 | .78 | [.57, .90] | .40 |

Note. Correlations presented are mean inter-item correlations. The Thought Disturbance subscale is not recommended for use with girls. The Traumatic Experiences subscale uses different items for boys and girls.

Table 3

Base Rates of Diagnoses by Ethnicity and Gender

| Lifetime Prevalence | Boys | | | | Girls | | | |
|------------------------|----------------|------|---------------|------|----------------|------|---------------|------|
| | LA | | EA | | LA | | EA | |
| | <i>N</i> = 268 | | <i>N</i> = 41 | | <i>N</i> = 130 | | <i>N</i> = 19 | |
| | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % |
| Mood Disorder | 23 | 8.6 | 8 | 19.5 | 27 | 20.8 | 3 | 15.8 |
| Anxiety Disorder | 17 | 6.3 | 4 | 9.8 | 17 | 13.1 | 4 | 21.1 |
| Behavioral Disorder | 58 | 21.6 | 9 | 22.0 | 23 | 17.7 | 2 | 10.5 |
| Substance Use Disorder | 96 | 35.8 | 10 | 24.4 | 37 | 28.5 | 4 | 21.1 |
| Psychotic Disorder | 4 | 1.5 | 2 | 4.9 | 2 | 1.5 | 1 | 5.3 |
| Adjustment Disorder | 6 | 2.2 | 0 | 0.0 | 2 | 1.5 | 2 | 10.5 |

Note. “LA” denotes Latinos, “EA” denotes European Americans. Given percentages indicate column-wise prevalence and do not sum to 100 as many juveniles had no diagnosis, whereas some had more than 1.

Table 4

Utility of MAYSI-2 in Detecting Lifetime Psychopathology Among Latino and European American Adolescents

| Criterion | | MAYSI-2 Scoring | | | | | |
|------------------------|-----------------|------------------------------|-----------------------|---|-----------------------|------------------------------|-----------------------|
| | | Above caution on any 1 scale | | Above caution on any 2 scales <i>or</i> above warning on any 1 scale | | Above warning on any 1 scale | |
| | | LA | EA | LA | EA | LA | EA |
| Any disorder | Sensitivity | 0.74 | 0.79 | 0.48 | 0.50 | 0.31 | 0.25 |
| | Specificity | 0.52 | 0.56 | 0.74 | 0.81 | 0.90 | 0.94 |
| | PPV | 0.63 | 0.61 | 0.67 | 0.70 | 0.77 | 0.78 |
| | NPV | 0.64 | 0.75 | 0.56 | 0.65 | 0.54 | 0.59 |
| | AUC (95% CI) | 0.63 (0.57 – 0.68) | 0.67 (0.54 - 0.81) | 0.61 (0.55 – 0.66) | 0.66 (0.52 – 0.80) | 0.61 (0.55 – 0.66) | 0.60 (0.45 – 0.74) |
| Serious mental illness | Sensitivity | 0.90 | 0.88 | 0.69 | 0.71 | 0.51 | 0.35 |
| | Specificity | 0.45 | 0.51 | 0.70 | 0.81 | 0.86 | 0.93 |
| | PPV | 0.29 | 0.42 | 0.37 | 0.60 | 0.48 | 0.67 |
| | NPV | 0.95 | 0.92 | 0.90 | 0.88 | 0.88 | 0.78 |
| | AUC (95% CI) | 0.68 (0.62 – 0.74) | 0.70 (0.56 – 0.84) | 0.69 (0.62 – 0.76) | 0.76 (0.62 – 0.90) | 0.69 (0.62 – 0.76) | 0.64 (0.47 – 0.81) |

Note: “LA” denotes Latinos, “EA” denotes European Americans. For all comparisons, $n = 398$ Latinos, $n = 60$ European Americans.

Criterion is diagnostic categories as measured by the K-SADS.

Table 5

Utility of MAYSI-2 in Detecting Lifetime Psychopathology Among Latino Boys and Girls

| Criterion | | MAYSI-2 Scoring | | | | | |
|------------------------|--------------|------------------------------|-----------------------|---|-----------------------|------------------------------|-----------------------|
| | | Above caution on any 1 Scale | | Above caution on any 2 scales or above warning on any 1 scale | | Above warning on any 1 Scale | |
| | | Boys | Girls | Boys | Girls | Boys | Girls |
| Any disorder | Sensitivity | 0.67 | 0.88 | 0.39 | 0.68 | 0.22 | 0.51 |
| | Specificity | 0.61 | 0.35 | 0.80 | 0.62 | 0.91 | 0.88 |
| | PPV | 0.66 | 0.58 | 0.70 | 0.64 | 0.74 | 0.81 |
| | NPV | 0.62 | 0.74 | 0.53 | 0.66 | 0.50 | 0.64 |
| | AUC (95% CI) | 0.64 (0.57 – 0.71) | 0.62 (0.52 – 0.71) | 0.59 (0.52 – 0.66) | 0.65 (0.55 – 0.74) | 0.57 (0.50 – 0.64) | 0.69 (0.60 – 0.78) |
| Serious mental illness | Sensitivity | 0.90 | 0.89 | 0.61 | 0.79 | 0.42 | 0.62 |
| | Specificity | 0.52 | 0.29 | 0.75 | 0.57 | 0.89 | 0.81 |
| | PPV | 0.25 | 0.33 | 0.32 | 0.43 | 0.40 | 0.56 |
| | NPV | 0.97 | 0.87 | 0.91 | 0.87 | 0.89 | 0.84 |
| | AUC (95% CI) | 0.71 (0.64 – 0.79) | 0.59 (0.49 – 0.69) | 0.67 (0.58 – 0.76) | 0.68 (0.58 – 0.78) | 0.65 (0.55 – 0.75) | 0.71 (0.61 – 0.82) |

Note. For all comparisons, $n = 268$ boys, $n = 130$ girls. Base rates for boys: 53.7% any disorder, 15.3% serious mental illness. Base rates for girls: 50.0% any disorder, 28.5% serious mental illness. Criterion is diagnostic categories as measured by the K-SADS.

Serious mental illness includes all mood, anxiety, and psychotic disorders.

Table 6

Utility of MAYSI-2 Scales in Detecting Lifetime Psychopathology Among Latino Boys and Girls

| | | | MAYSI-2 Scale | | | | | | | | | | | |
|------------------------|---------------------|-------------|---------------|------|-----------------|-------------|-------------------|-------------|--------------------|------|-------------------|------|---------------------|------|
| | | | Alcohol-Drug | | Angry-Irritable | | Depressed-Anxious | | Somatic Complaints | | Suicidal Ideation | | Thought Disturbance | |
| | Criterion | | C | W | C | W | C | W | C | W | C | W | C | W |
| Boys | Mood Disorder | Sensitivity | 0.26 | 0.04 | 0.65 | 0.30 | 0.48 | 0.04 | 0.70 | 0.09 | 0.44 | 0.44 | 0.44 | 0.35 |
| | | Specificity | 0.83 | 0.97 | 0.80 | 0.98 | 0.80 | 0.96 | 0.71 | 0.98 | 0.89 | 0.93 | 0.63 | 0.87 |
| | | PPV | 0.13 | 0.13 | 0.23 | 0.54 | 0.19 | 0.09 | 0.18 | 0.33 | 0.28 | 0.37 | 0.10 | 0.20 |
| | | NPV | 0.92 | 0.92 | 0.96 | 0.94 | 0.94 | 0.91 | 0.96 | 0.92 | 0.94 | 0.95 | 0.92 | 0.93 |
| | Anxiety Disorder | Sensitivity | 0.12 | 0.00 | 0.29 | 0.00 | 0.35 | 0.18 | 0.77 | 0.06 | 0.29 | 0.18 | 0.47 | 0.29 |
| | | Specificity | 0.82 | 0.97 | 0.76 | 0.95 | 0.79 | 0.97 | 0.71 | 0.98 | 0.88 | 0.90 | 0.63 | 0.86 |
| | | PPV | 0.04 | 0.00 | 0.08 | 0.00 | 0.10 | 0.27 | 0.15 | 0.17 | 0.14 | 0.11 | 0.08 | 0.13 |
| | | NPV | 0.93 | 0.94 | 0.94 | 0.93 | 0.95 | 0.95 | 0.98 | 0.94 | 0.95 | 0.94 | 0.95 | 0.95 |
| | Behavioral Disorder | Sensitivity | 0.36 | 0.07 | 0.45 | 0.12 | 0.35 | 0.05 | 0.41 | 0.05 | 0.24 | 0.17 | 0.48 | 0.28 |
| | | Specificity | 0.88 | 0.98 | 0.81 | 0.97 | 0.81 | 0.96 | 0.70 | 0.99 | 0.90 | 0.92 | 0.65 | 0.89 |
| | | PPV | 0.45 | 0.50 | 0.40 | 0.54 | 0.34 | 0.27 | 0.28 | 0.50 | 0.39 | 0.37 | 0.28 | 0.40 |
| | | NPV | 0.83 | 0.79 | 0.84 | 0.80 | 0.82 | 0.79 | 0.81 | 0.79 | 0.81 | 0.80 | 0.82 | 0.82 |
| Substance Use Disorder | Sensitivity | 0.37 | 0.07 | 0.31 | 0.07 | 0.26 | 0.04 | 0.35 | 0.03 | 0.18 | 0.14 | 0.51 | 0.25 | |
| | Specificity | 0.93 | 0.99 | 0.80 | 0.97 | 0.80 | 0.96 | 0.69 | 0.98 | 0.89 | 0.92 | 0.70 | 0.91 | |
| | PPV | 0.75 | 0.88 | 0.46 | 0.54 | 0.42 | 0.36 | 0.39 | 0.50 | 0.47 | 0.48 | 0.49 | 0.60 | |
| | NPV | 0.72 | 0.66 | 0.68 | 0.65 | 0.66 | 0.64 | 0.66 | 0.65 | 0.66 | 0.66 | 0.72 | 0.69 | |
| Girls | Mood Disorder | Sensitivity | 0.22 | 0.04 | 0.52 | 0.15 | 0.74 | 0.19 | 0.67 | 0.15 | 0.52 | 0.44 | -- | -- |
| | | Specificity | 0.76 | 0.93 | 0.69 | 0.90 | 0.65 | 0.95 | 0.50 | 0.96 | 0.82 | 0.90 | -- | -- |
| | | PPV | 0.19 | 0.13 | 0.30 | 0.29 | 0.36 | 0.50 | 0.26 | 0.50 | 0.42 | 0.55 | -- | -- |
| | | NPV | 0.79 | 0.79 | 0.85 | 0.80 | 0.91 | 0.82 | 0.85 | 0.81 | 0.87 | 0.86 | -- | -- |
| | Anxiety Disorder | Sensitivity | 0.18 | 0.12 | 0.47 | 0.12 | 0.77 | 0.12 | 0.88 | 0.12 | 0.47 | 0.41 | -- | -- |
| | | Specificity | 0.75 | 0.95 | 0.66 | 0.89 | 0.62 | 0.93 | 0.51 | 0.95 | 0.78 | 0.87 | -- | -- |
| | | PPV | 0.10 | 0.25 | 0.17 | 0.14 | 0.23 | 0.20 | 0.21 | 0.25 | 0.24 | 0.32 | -- | -- |
| | | NPV | 0.86 | 0.88 | 0.89 | 0.87 | 0.95 | 0.88 | 0.97 | 0.88 | 0.91 | 0.91 | -- | -- |
| | Behavioral Disorder | Sensitivity | 0.57 | 0.26 | 0.65 | 0.26 | 0.52 | 0.13 | 0.35 | 0.09 | 0.30 | 0.17 | -- | -- |
| | | Specificity | 0.83 | 0.98 | 0.71 | 0.93 | 0.59 | 0.94 | 0.42 | 0.94 | 0.76 | 0.83 | -- | -- |
| | | PPV | 0.42 | 0.75 | 0.33 | 0.43 | 0.21 | 0.30 | 0.11 | 0.25 | 0.21 | 0.18 | -- | -- |
| | | NPV | 0.90 | 0.86 | 0.91 | 0.85 | 0.85 | 0.83 | 0.75 | 0.83 | 0.84 | 0.82 | -- | -- |
| Substance Use Disorder | Sensitivity | 0.60 | 0.22 | 0.46 | 0.22 | 0.35 | 0.05 | 0.51 | 0.14 | 0.24 | 0.19 | -- | -- | |
| | Specificity | 0.90 | 1.00 | 0.69 | 0.94 | 0.54 | 0.91 | 0.45 | 0.97 | 0.74 | 0.84 | -- | -- | |
| | PPV | 0.71 | 1.00 | 0.37 | 0.57 | 0.23 | 0.20 | 0.27 | 0.63 | 0.27 | 0.32 | -- | -- | |

| | | | | | | | | | | | | |
|-----|-------------|-------------|------|------|------|------|------|------|------|------|----|----|
| NPV | 0.85 | 0.76 | 0.76 | 0.75 | 0.68 | 0.71 | 0.70 | 0.74 | 0.71 | 0.72 | -- | -- |
|-----|-------------|-------------|------|------|------|------|------|------|------|------|----|----|

Note. “C” denotes caution cutoff; “W” denotes warning cutoff. Figures in bold indicate conceptual overlap between the MAYSI-2 and diagnoses as measured by the K-SADS. For all comparisons, $n = 268$ boys, $n = 130$ girls.