

DIFFERENTIATING BIPOLAR SPECTRUM DISORDERS: THE DIAGNOSTIC UTILITY
OF THE BIS/BAS SCALES

Mian-Li Ong

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Masters of Arts in the Department of Psychology in the College of Arts and Sciences.

Chapel Hill
2014

Approved by:

Eric A. Youngstrom

Deborah C. Jones

Andrea M. Hussong

©2014
Mian-Li Ong
ALL RIGHTS RESERVED

ABSTRACT

Mian-Li Ong: Differentiating Bipolar Spectrum Disorders: The Diagnostic Utility of the
BIS/BAS Scales
(Under the direction of Eric A. Youngstrom)

The present study examined discriminative validities of the Behavioral Inhibition System/Behavioral Approach System (BIS/BAS) scales in differentiating bipolar spectrum disorders (BSDs) from other disorders. Participants were youth recruited from a combination of community mental health center and university medical facility. Receiver Operating Characteristic (ROC) analyses tested the BIS/BAS scales in distinguishing between BSD-positive and BSD-negative youth. We calculated diagnostic likelihood ratios in keeping with recommendations from evidence-based medicine. Binary logistic regressions tested for incremental value in combining subscales. BAS subscales discriminated between participants with BSD-positive and BSD-negative diagnoses, with areas under the curve ranging from .54 to .64. The BIS/BAS scales achieved statistical significance in identifying cases with BSDs, but effect sizes for discriminative comparisons were too small to be clinically useful. Upgrading clinical training to: (a) include prevalence of BSDs and (b) teaching clinicians more evidence based assessment strategies is important to improve assessment and diagnosis of BSDs.

To Eric and Jen Youngstrom, thank you for your support and continued guidance. To my family, thank you for making me the person I am today. Thank you Yen-Ling Chen and Carisa Ruiz for proofreading. And finally, to Merisa, my dearest: I could not have done this without your unyielding support and love.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES.....	vii
LIST OF ABBREVIATIONS.....	viii
DIFFERENTIATING BIPOLAR SPECTRUM DISORDERS: THE DIAGNOSTIC UTILITY OF THE BIS/BAS SCALES	1
Introduction.....	1
The Behavioral Approach System	3
METHOD	7
Participants.....	7
Measures	8
Schedule of Affective Disorders and Schizophrenia for Children.....	8
Index Test.....	9
BIS/BAS Scales	9
Procedure	10
Data Analytic Plan	10
Receiver Operating Characteristics.....	10
Diagnostic Likelihood Ratios	11
RESULTS	13
Descriptive Statistics.....	13
Differentiating Diagnostic Categories	13
Calculating Diagnostic Likelihood Ratios	14
Logistic Regressions and Interactions	15
CONCLUSION.....	16

Strengths and Limitations	18
Clinical Implications.....	19
APPENDIX 1: DEMOGRAPHICS OF STUDY SAMPLE.....	20
APPENDIX 2: ROC/AUC ANALYSES OF DIAGNOSTIC DIFFERENTIATION USING BIS/BAS TOTAL AND BIS/BAS SUBSCALES.....	21
APPENDIX 3: ROC CURVE ON BAS TOTAL AND SUBSCALES ON BSD DIAGNOSIS ...	22
REFERENCES	23

LIST OF FIGURES

Figure 1 –Demographics of Study Samples.....	20
Figure 2 –ROC/AUC Analyses of Diagnostic Differentiation using BIS/BAS Total and BIS/BAS subscales	21
Figure 3– ROC Curve on BAS Total and Subscales on BSD Diagnosis.....	22

LIST OF ABBREVIATIONS

AUC	Area under curve
BSDs	Bipolar spectrum disorders
DLR	Diagnostic likelihood ratio
ROC	Receiver operating characteristic

DIFFERENTIATING BIPOLAR SPECTRUM DISORDERS: THE DIAGNOSTIC UTILITY OF THE BIS/BAS SCALES

Introduction

BSDs are characterized by abnormal and persistent dysregulation of mood and energy (Goodwin & Jamison, 2007) and affect approximately 1.8% of adolescents in the U.S. (Van Meter, Moreira, & Youngstrom, 2011). BSD diagnosis notoriously predicts a host of adverse outcomes in youth, including high rates of prospective suicide attempts (Goldstein et al., 2012), high rates of alcohol and substance use (Stewart et al., 2012) and increased rates of incarceration (Pliszka, Sherman, Barrow, & Irick, 2000). However, despite overwhelming evidence delineating its severity, diagnosing BSDs continues to be a challenging and time-consuming process for clinicians (Miller, Johnson, Kwapil, & Carver, 2011), with individuals often waiting from 5 to 15 years for formal diagnosis of BSDs to be made (Marchand, Wirth, & Simon, 2006). There is a crucial need for accurate screening tools that can be administered quickly to individuals at risk for BSDs. This study examined the discriminative validity of the BIS/BAS scales in differentiating BSDs in youth from other disorders presenting to two separate clinical infrastructures.

Why do clinicians find BSDs so difficult to diagnose in youth? First, BSDs in youth have high comorbidity with other mood disorders. Notable examples include attention-deficit hyperactivity disorder (ADHD)(Kim & Miklowitz, 2002) and anxiety disorders (Merikangas et al., 2007). For these disorders, the primary issue is due to overlapping diagnostic criteria. Examples include accelerated speech and irritability in ADHD and separation anxiety in anxiety

disorders. In turn, these symptoms become less helpful in differentiating between these diagnostic groups (Chan, Stringaris, & Ford, 2011; Geller et al., 2002). Second, youth with BSDs are being misdiagnosed with major depression, with one study finding that 40% of BSD-positive patients have previously received an incorrect diagnosis of major depression (Chilakamarri, Filkowski, & Ghaemi, 2011). In line with findings of adolescents spending more time in depressed moods than hypomanic or manic episodes (Axelson et al., 2011; Judd et al., 2002), patients do not perceive their symptoms of elevated mood to be problematic; in fact, they might even consider them to be adaptive behavior. In turn, they are likely to seek clinical help only during states of pronounced depression (ten Have, Vollebergh, Bijl, & Nolen, 2002), which subsequently results an increased likelihood of a major depression diagnosis. Treatment of BSDs with antidepressants is not efficacious; in fact, it may exacerbate hypomania, mania or cycling (Chilakamarri et al., 2011). Lastly, current standards of practice rely on unstructured interviews and impressionistic interpretations of assessment tools, with research suggesting that agreement between practitioners about diagnosis of individual cases is typically only slightly better than chance (Jenkins, Youngstrom, Washburn, & Youngstrom, 2011). Assessment techniques routinely taught and used in mental health disciplines (Cashel, 2002) possess little research supporting their use in decision-making about individual cases.

Moreover, researchers have not been able to agree on the best method to diagnose BSDs reliably. This creates more confusion for clinicians, exacerbating the problem of early and accurate identification of BSDs (Rettew, Lynch, Achenbach, Dumenci, & Ivanova, 2009). First, some experts focus on the importance of symptoms such as grandiosity and elation (Geller et al., 2002), while others identify separate manic episodes as the primary distinguishing feature of BSDs (Carlson, 1990). Second, while the general consensus has been to trust the judgment of adult informants for externalizing symptoms (Carlson & Youngstrom, 2011; Meyers &

Youngstrom, 2008), and to trust the child for internalizing symptoms, experts disagree on the utility of teacher report (Carlson & Blader, 2011; Carlson & Klein, 2014) or parent report (Youngstrom, Findling, & Calabrese, 2003) as “top-shelf” measures in accurately describing symptoms of BSDs. Early and accurate identification of BSDs helps clinicians provide effective targeted treatment and reduces taxpayer dollars spent on less efficacious interventions.

There are few assessments today that have the discriminative validity needed for high-stakes clinical decisions (Watkins & Glutting, 2000; Wood, Nezworski, & Stejskal, 1996). Yet, while these well-established measures greatly improve diagnostic certainty (Geller et al., 2001; Kaufman et al., 1997), they are cumbersome, time-consuming and expensive (Ebesutani, Bernstein, Chorpita, & Weisz, 2012). A case in point would be the Parent-General Behavior Inventory (P-GBI), a 73-item parent-report screen for depressive, hypomanic and manic symptoms in youth aged 5-17. The P-GBI has good reason for its fame: it reliably discriminates BSD from comorbid diagnoses such as ADHD and depressive disorders (Youngstrom, Findling, Danielson, & Calabrese, 2001). While these characteristics of the P-GBI increase the possibility of capturing the intended construct of mood disorder—particularly the element of episodic change in presentation—it is longer (10 pages) and requires an increased minimum level of reading (12th grade). There is a strong need for a diagnostic instrument for clinicians that (i) is empirically supported, (ii) has strong validity, (iii) maps well to both clinical conceptualizations as well as more fundamental underlying processes (e.g., Research Domain Criteria; RDoC) (Sanislow et al., 2010), and (iv) is efficient to administer.

The Behavioral Approach System and BSDs

The Behavioral Inhibition System/Behavioral Approach System (BIS/BAS) scales (Carver & White, 1994) offer a credible alternative to the incumbent measures that exist today. The Behavioral Approach System (BAS) is hypothesized to be associated with happiness and

elation, and is critical in regulating behavior (Depue & Iacono, 1989; Gray, 1981, 1982, 1994). It provides drive towards cues of reward and approach behaviors, such as fun seeking or appetitive aggression. Of note, the BAS has played a prominent role in theoretical methods of developmental psychopathology, especially BSDs (Alloy & Abramson, 2010; Alloy, Bender, et al., 2009; Urosevic, Collins, Muetzel, Lim, & Luciana, 2012). The traits that correlate with the BAS map well onto the symptoms of BSDs in DSM-5 (American Psychiatric Association, 2013).

However, while the BIS/BAS scales boast a considerable amount of research relating to BSDs, they do not cover every symptom of mania. This peculiarity makes the scales a tricky instrument to use in differential diagnosis of BSDs. On one hand, high BAS levels correlate with a host of symptomatology associated with mania and hypomania, such as (a) high levels of energy (Alloy, Abramson, et al., 2009), (b) goal-oriented behavior (Nusslock, Abramson, Harmon-Jones, Alloy, & Hogan, 2007), (c) positive affect, and (d) irritability and anger (Harmon-Jones et al., 2002). High self-reported BAS scores also predict later emotional instability and progression to BSDs (Alloy et al., 2012). Findings from Gruber et al. (2013) have also suggested that parent-reported BAS scores were associated with increased symptoms of mania, suggesting a transdiagnostic link between reward dysregulation and manic mood severity in adolescence. As such, the BAS may be a risk factor in the developmental trajectory of BSDs from adolescence to adulthood. On the other hand, despite the plethora of research supporting the links between BAS and BSDs, the BAS appears to be an incomplete construct for differentiating BSD. Indeed, the BIS/BAS scales, a well-known measure of BAS (Carver & White, 1994), omits key features that pertain to mania, such as sleep or grandiosity. This conundrum becomes pertinent when one considers that the BAS happens to be one of the dimensions identified in the RDoC initiative (Cuthbert & Insel, 2010; Insel et al., 2010; Sanislow et al., 2010). Assessing the utility of the BAS scales as a diagnostic tool might provide insight into its role in other disorders,

inasmuch as the RDoC dimensions are thought to be transdiagnostic and are seen to be intermediate functions that are not themselves clinical symptoms (Cuthbert & Insel, 2010). More importantly, if validated as a diagnostic aid, the BAS scales could potentially identify BSDs before full syndromal onset, saving millions of dollars in the treatment of pediatric BSDs (Kim & Miklowitz, 2002; Miklowitz & Chang, 2008). Taken together, a close investigation of the BIS/BAS scales for clinical use is warranted.

The clinical utility of the BIS/BAS scales in distinguishing BSDs from other disorders remains unclear. There are a variety of potential advantages to the BAS scales. Firstly, the BAS scales are brief and easily tabulated: they only consist of 20 items on a Likert-type scale, and the questionnaire only takes roughly five minutes to complete; conversely, the P-GBI consists of 73 items and takes roughly 20-30 minutes to complete. Second, the BAS scales are theoretically relevant to domains of functioning associated with BSDs, such as manic symptoms, positive affect and irritability (Harmon-Jones et al., 2002), which dovetail with the drive to obtain a homogeneous endophenotype for BSD presentation (Hasler, Drevets, Gould, Gottesman, & Manji, 2006). Third, the BAS is conceptually “clean” due to it being developed as a research instrument. It has been found to map onto the left front-mid cortical region of the brain (Nusslock et al., 2012), which has been associated with conversion from bipolar II to bipolar I (Alloy, Bender, et al., 2009). Again, the BAS has been identified as a key domain of RDoC, which makes an investigation of its clinical application highly pertinent (Cuthbert & Insel, 2010).

The purpose of this study was to examine the discriminative and diagnostic validity of the BIS/BAS scales and the associated subscales in differentiating youth with BSDs (BP-I, BP-II, cyclothymic disorder, BSD-NOS) from those with any other disorder within a treatment-seeking sample. We hypothesized that BSD-positive youth would score significantly higher on the

BIS/BAS scales and subscales than those with a BSD-negative diagnosis. A second goal was to develop multi-level DLRs to facilitate assessment and diagnostic decision-making about individual cases (Straus, Glasziou, Richardson, & Haynes, 2011). A third goal was to use binary logistic regressions and interactions to determine any incremental value in combining subscales, after controlling for age and gender, both of which are known to impact BSD outcome (Merikangas et al., 2011). To the best of our knowledge, this study was the first to examine the discriminative validity of the BIS/BAS scales for delineating BSDs from other disorders in youth. It was also the first study to develop DLRs for interpretation of BIS/BAS scores in this age range, which makes it much more feasible for clinicians to adopt an evidence-based assessment framework for interpreting the BIS/BAS (Youngstrom, 2013a).

METHOD

Participants

Table 1 presents demographical information, child global assessment functioning, mean number of diagnoses and descriptive information of the BAS-Total scale and its subscales. Participants ($n = 819$, 152 BSDs, ages 5 to 17 years and 11 months) were youth in a consecutive case series recruited from a combination of community mental health center and university medical facility (Youngstrom, Youngstrom, & Starr, 2005) as part of a larger project (R01 MH066647, PI: E. A. Youngstrom). The study design was prospective: Data collection was planned before the index test and reference standard were performed (Bossuyt et al., 2003). Participants were recruited from two different clinical infrastructures: one being a community mental health center with four urban sites (Youngstrom, Youngstrom, et al., 2005). In those sites, the only inclusion criteria was that the patient needed to be between ages 5 and 18, and both the caregiver and child needed to be able to communicate in English at a conversant level in order to complete the diagnostic interview and questionnaires. The other clinical infrastructure was an outpatient academic medical center with many pharmacotherapy studies (Findling et al., 2001). In this site, recruitment was based on the presentation of bipolar-like symptoms and willingness to participate in treatment protocols. The sample was enriched by referrals of children whose parents had a diagnosed BSD and were currently undergoing treatment or research at an affiliated adult mood disorders clinic. Inclusion criteria for the study were (1) youth between ages 5-18 and (2) presenting for an outpatient evaluation for which the participants provided written assent and the primary caregiver provided written consent for participation. Exclusion criteria included (1) inability to communicate in English in order to complete the both the

diagnostic interview and questionnaires; (2) having a pervasive developmental disorder as determined by psychiatric history or psychiatric interview or having an Autism Screening Questionnaire score of 15 or higher (Berument, Rutter, Lord, Pickles, & Bailey, 1999); and (3) suspected moderate or severe mental retardation documented by either educational history, standardized cognitive ability scores of below 70 or a Peabody Picture Vocabulary Test-Third Edition (Dunn & Dunn, 1997). Study procedures were approved by the Institutional Review Board at Case Western Reserve University (Cleveland, OH) and Applewood Centers.

Measures

Reference Standard: Semistructured Diagnostic Interview Using the Schedule of Affective Disorders and Schizophrenia for Children (KSADS).

All participating families completed the Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime (KSADS-PL) (Kaufman et al., 1997), combined with the mood disorders module from the Washington University KSADS (WASH-U-KSADS) (Geller et al., 2002). The KSADS is the most widely used semi-structured diagnostic procedure for investigations of pediatric BSDs (Nottelmann, 2001). Bipolar I (BP-I), Bipolar II (BP-II), cyclothymic disorder, and Bipolar Not Otherwise Specified (BP-NOS) diagnoses were made in accordance with DSM-IV diagnostic criteria, including a strong emphasis that mood symptoms needed to represent a clear change in functioning and follow an episodic presentation.

Research assistants ($N = 34$) received extensive training prior to administering the KSADS. Training consisted of rating interviews along with an experienced rater while observing five K-SADS interviews, and leading at least five interviews that were re-scored by a reliable rater and passing with $\kappa \geq .85$ at item level (Findling et al., 2001; Youngstrom, Meyers, et al., 2005). Once trained, the same interviewer would evaluate the answers given by both informants,

and any discrepancies were resolved using best clinical judgment. The operational definition of BSD included any child diagnosed with BP-I, BP-II, cyclothymic disorder or BP-NOS.

Index Test

BIS/BAS Scales

The BIS/BAS Scales consist of 20 items using a Likert-type scale (1 = *strongly disagree*, 4 = *strongly agree*), comprising three BAS subscales (Reward Responsiveness, Fun Seeking and Drive) and one BIS subscale. The BAS Reward Responsiveness subscale has five items designed to assess positive response to reward stimuli. The BAS Drive subscale has four items indicative of persistence in pursuit of reward. The BAS Fun Seeking subscale contains four items indicating willingness to approach novel and rewarding stimuli. Finally, the BAS-Total sums the three BAS subscales to assess the sensitivity of the BAS. In this study, the primary caregiver provided all BIS/BAS data about the child. Although self-report data on the BIS/BAS scales were available in the research grant that this study is part of, we chose to use the caregiver report as it has consistently shown greater validity for discriminating pediatric bipolar disorder from other conditions than self or teacher report (Geller, Warner, Williams, & Zimmerman, 1998; Kahana, Youngstrom, Findling, & Calabrese, 2003; Youngstrom, 2007; Youngstrom, Jenkins, Doss, & Youngstrom, 2012; Youngstrom, Meyers, et al., 2005), making it the preferred format for the present study. Also, self-report was only available for the subset of adolescents 11 years and older. Internal consistencies (*Cronbach's alphas* = .66–.76) and test–retest reliabilities ($r_s = .59$ –.69) for the subscales have been satisfactory (Carver & White, 1994). For the present study, internal consistency statistics ranged from satisfactory to good (Cronbach's alphas: *BAS-Fun Seeking* = .66; *BAS-Reward Responsiveness* = .77; *BAS-Drive* = .83). The consistency statistics are affected by the low number of items in each subscale, and it is reassuring to note that internal consistency was as high or higher in the present sample compared to previous reports.

Procedure

The parent or guardian (primary caregiver) provided written consent for the participation of their child, and all youth provided written assent. Both youth and their primary caregiver completed the K-SADS and the WASH-U-KSADS. While the participants were being interviewed, the primary caregiver also completed the BIS/BAS questionnaires. Participants and parents did not have access to each other's responses on the rating scales. The interviewer would share psychiatric and treatment history with the supervising clinical psychologist to confirm diagnoses.

Data Analytic Plan

Preliminary analyses included descriptive statistics, missing value analyses, chi-square analyses and checking of assumptions. Logistic regressions included interaction terms for age and BAS subscale. While the primary analyses used nonparametric methods, it is still important to examine score distributions for evidence of “degeneracy” (Youngstrom, 2013b; Zhou, Obuchowski, & McClish, 2011). Degeneracy refers to distributions that either: (a) have a bimodal score distribution; or (b) have regions where score frequencies fail to progress monotonically (Zhou et al., 2011).

Receiver operating characteristics (ROC)

ROC curves depict the balance between the probability of a true positive test result for those who have the target condition (known as sensitivity) and the probability of a true negative test result for those who do not have the condition (known as specificity). In this study, ROC analyses quantified the ability of each scale to distinguish cases with BSDs from all other cases (McFall & Treat, 1999). Sensitivity and specificity statistics were then used to calculate the area under the curve (AUC), which is an effect size quantifying diagnostic accuracy of scores. An AUC of 1.0 would indicate that the test performed with perfect diagnostic accuracy, while an

AUC of .50 would indicate chance performance – that is, the resulting ROC curve would fall along a diagonal line (performing at 50% chance), also referred to as the chance diagonal (Obuchowski, 2003; Zhou et al., 2011). The Venkatraman (Venkatraman, 2000) test was computed in R (R Development Core Team, 2014) to compare the areas under the curve for each subscale to see if any performed significantly better than others in discriminating between youth with BSDs and without BSDs. DLRs quantified changes in odds of BSD diagnosis corresponding to test score ranges (e.g., low, indeterminate, high, etc.) (Straus et al., 2011). BAS scores were categorized into thirds (each containing approximately 33% of the scores) for easy classification. Finally, binary logistic regression analyses tested whether any combination of scales provided any incremental value than the best single scale does at identifying cases with BSDs. The BAS-total and all three subscale (Reward Responsiveness, Drive and Fun Seeking) variables were predictor variables, and BSD was the outcome variable. In step 1, race (white/not white), gender (male/female) and site (community academic center/outpatient clinic) was entered into the model as dummy variables, while age of child was entered as a continuous variable. In step 2, BAS-Reward Responsiveness, BAS-Fun Seeking and BAS-Drive entered the model. Finally, in step 3, multiple interactions entered the model (RR*FS; RR*Drive; Drive*FS; RR*FS*Drive).

Diagnostic likelihood ratios

DLRs capture more detailed diagnostic information for decision making about individual cases. Conceptually, the diagnostic likelihood ratio (DLR) is the change in the risk of diagnosis based on assessment results. It repackages the older concepts of diagnostic sensitivity and specificity. This method makes it easier for clinicians to use the information from test results to estimate posterior predictive values (Straus et al., 2011). Clinicians can combine DLRs with the prior probability of the diagnosis by means of a probability nomogram, online calculators or

Bayes' Theorem (Jaeschke, Guyatt, & Sackett, 1994) to obtain updated risk estimates for the disorder. We estimated likelihood ratios for multiple score ranges, dividing the sample into thirds (Jaeschke et al., 1994; Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). This (1) preserved more diagnostic information from the test results; (2) ensured that the scales behaved monotonically (Zhou et al., 2011); and (3) minimized ambiguity for high-stakes clinical decision making.

RESULTS

Descriptive Analyses

Table 1 presents group differences between BSD-positive youth and BSD-negative youth across (a) gender; (b) race; (c) age; (d) BAS-Total and subscale scores; and (e) child global assessment functioning score (CGAS). There were no pre-existing significant differences between gender breakdown in BSD-positive youth (45% female, $n = 69$) versus BSD-negative youth (39% female, $n = 258$), $\chi^2(1) = 2.33, p = .13$. We assessed whether there were pre-existing significant differences between racial breakdown in the BSD diagnostic group. As expected when comparing a community academic center with an outpatient clinic, nonwhites were significantly more likely to be found in the BSD-negative group (81% nonwhite, $n = 539$) than the BSD-positive group (63% nonwhite, $n = 96$), $\chi^2(1) = 22.14, p < .0005$. BSD-positive youth were significantly more likely to have higher CGAS ($M = 54.13, SD = 6.96$) than those with no BSD diagnosis ($M = 52.26, SD = 8.81$), $t(819) = 2.75, d = .24$. Hence, we included site in Block 1 of the logistic regression model. BSD-positive youth were significantly more likely to have higher parent BAS-Total scores than those with no BSD diagnosis ($M = 25.75, SD = 7.62$), $t(819) = 4.36, d = .38$. There were no pre-existing group differences in BAS-Drive scores, BAS-Reward Responsiveness scores and BAS-Fun Seeking scores.

Differentiating Diagnostic Categories

Table 2 presents findings from ROC analyses of the index test for the criterion groups (presence of K-SADS diagnosis). AUC values quantified the value of the BIS/BAS scales for

distinguishing between groups with BSD versus without BSD. The BAS-Total scale, the BAS-Drive and BAS-Fun Seeking subscales achieved statistical significance in distinguishing BSD-positive youth from BSD-negative youth ($p < .0005$); however, effect sizes were small, reflecting modest discrimination (BAS-Total: $AUC = .62$; BAS-Drive: $AUC = .64$; BAS-Fun Seeking: $AUC = .59$). BAS-Reward Responsiveness did not achieve statistical significance in distinguishing between BSD-positive and BSD-negative youth ($AUC = .54, p = .17$). Results from the Venkatraman (2000) procedure indicated that the BAS-Drive subscale was significantly better than the BAS-Total scale in its ability to discriminate BSD diagnosis in youth ($E = 8142, p < .05$) (Robin et al., 2011). In contrast, both BAS-Total and BAS-Drive scales identified BSD diagnosis significantly better than the BAS-Fun Seeking scale ($E = 13488, p < .0001$ for BAS-Total; $E = 14798, p < .01$ for BAS-Drive).

Calculating Diagnostic Likelihood Ratios

The scores were initially divided into quintiles (with the bottom ~20% of scores to be considered as very low, then the next ~20% as low, etc.). However, some categories were found to be too sparse, as the scales were not found to behave monotonically. As such, these categories were pooled, and scores on each subscale were divided into three categories (low, indeterminate, high). Table 2 reports the DLRs for the BAS-Total and the individual subscales. Overall, increases in odds of BSD diagnoses were fair when comparing BSD-positive participants to those with no BSD. On the BAS-Drive subscale, youth who are BSD-positive were more than twice as likely to score 11+ than those who did not receive a BSD diagnosis. Conversely, youth who are BSD-positive were half as likely to score 5 or below than those who were BSD-negative. On the BAS-Total scale, youth who are BSD-positive were modestly more likely to score 31 or above than those who did not receive a BSD diagnosis. Conversely, youth who are BSD-positive were half as likely to score 22 or below than those who were BSD-negative. On

the BAS-Fun Seeking subscale, youth who are BSD-positive were only slightly more likely to score 10 or above than those who did not receive a BSD diagnosis. Conversely, youth who were BSD-positive were half as likely to score 6 or below than those who were BSD-negative.

Logistic Regressions and Interactions

Logistic regressions involving the BAS-total and all three subscales did not predict BSD diagnosis after controlling for age and gender, $p > .05$. Similarly, interactions involving the BAS-Total and all three subscales did not predict BSD diagnosis after controlling for age and gender, $p > .05$.

DISCUSSION

The present study was the first to examine discriminative and diagnostic validities of parent rated BIS/BAS scales in differentiating youth with BSDs from those with other disorders. As hypothesized, BAS-Total scales, BAS-Fun Seeking and BAS-Drive provided statistically significant results when delineating BSDs from other disorders. Based on ROC analyses, the diagnostic efficiency of the BAS-total scale, BAS-Drive and BAS-Fun Seeking were fair to poor, with AUCs of .62, .64, and .59, respectively, for discriminating BSDs from other disorders. The Venkatraman procedure indicated that BAS-Total and BAS-Drive scales were significantly different from each other in discriminating between BSD-positive and BSD-negative youth. Both scales outperformed the BAS-Fun Seeking subscale in identifying BSDs. Taken together, due to poor effect sizes, the BAS scales do not meet the standards of high-stakes clinical decision-making (Youngstrom & De Los Reyes, in press), although results replicate prior findings of a statistically significant association with bipolar diagnoses (Alloy, Abramson, et al., 2009).

The second aim was to develop multilevel DLRs to facilitate clinical decision making about individual cases using the BIS/BAS scales. DLR values suggest that both the BAS-Total and Fun Seeking scales are not clinically helpful in differentiating BSDs from other disorders, with high scores on the BAS-Total increasing risk only modestly and high scores on the BAS-Fun Seeking increasing risk modestly. However, the BAS-Drive subscale is somewhat clinically helpful in differentiating a bipolar diagnosis from other disorders, with high scores on the BAS-Drive subscale increasing risk twofold. All three scales were modestly clinically helpful in ruling out BSDs, as low scores on all three scales decreased risk of BSDs by approximately twofold.

The third aim was to use binary logistic regressions and interactions to determine any incremental value in combining subscales. No subscale interacted with one another to predict BSD diagnosis, with $p > .05$, suggesting that BAS-Drive, BAS-Fun Seeking and BAS-Reward Responsiveness did not provide any additional utility in forecasting a BSD diagnosis. Clinicians need not interpret combinations of BAS subscales when assessing patients at risk for BSDs.

Strengths and Limitations of Study

Unique strengths to the current study were the analyses (ROC, DLRs and binary logistic regressions) relying on multiple methods for evaluating diagnostic efficiency. This offered clear indication of the clinical utility of the BIS/BAS scales in differentiating BSDs. While researchers traditionally have split groups into those with BSD and those without, and proceed to test differences in group means (i.e., already knowing the diagnosis), we note that clinicians work the opposite way – they are required to present the assessment to the patient before assessing the probability that the assessment has correctly classified the patient as BSD-positive (Youngstrom & De Los Reyes, in press). Clinical significance is often a higher bar than statistical significance, as effect sizes need to be much larger (for instance, a Cohen's d of .8 only converts into a mediocre 71% chance of predicting BSD using ROC). Requiring larger effect sizes serves as a useful signpost for judging the clinical importance of assessments. Secondly, ROC enables clinicians to identify tools that can help them to predict pediatric BSD at better-than-chance rates. Conceptually, ROC provided a method for researchers and clinicians to determine if the BIS/BAS ranked a randomly chosen case with a positive diagnosis of pediatric BSD higher than a randomly chosen case with a negative diagnosis of pediatric BSD. Similarly, clinicians can use DLRs from this study to clarify diagnostic decision-making in potentially ambiguous presentations of pediatric BSD. Because base rates of pediatric BSD are likely to be different in diverse demographic and workplace settings, the results from these tests can contribute to

accurate diagnosis by systematically assessing symptoms in a standardized fashion, and decrease the effects of cognitive biases and heuristics in situations where the clinician is uncertain about differential diagnosis of pediatric BSD (Youngstrom, Findling, Youngstrom, & Calabrese, 2005). This is because DLRs help to decrease false-positive diagnoses (Harrell, Califf, Pryor, Lee, & Rosati, 1982) in settings in which BSDs are likely to be uncommon, because their interpretation requires integrating the base rate or prior probability of BSD with the new information from the test. Taken together, the presentation of ROC and DLRs made it easier for clinicians to use Bayesian methods to integrate test results with other risk factors, generating posterior probabilities for the risk of PBD (Jenkins et al., 2011; Straus et al., 2011).

However, there are a couple of limitations of the study. First, the current sample included few Asian or Hispanics; thus results are not representative of these cultures. It will be important for future research to establish if these measures perform similarly with these diverse populations. Second, the BIS/BAS scales are focused on processes that are not intrinsically pathological. Better diagnostically discriminating measures (such as the P-GBI) include symptom-level measures such as mania and depressive symptoms, which in turn focus more on psychopathology as a result. Third, the BIS/BAS as a construct has been found to be correlated to other psychopathology, such as substance abuse, attention-deficit hyperactivity disorder and conduct disorder (Braddock et al., 2011; Markarian, Pickett, Deveson, & Kanona, 2013). As such, there might be elevated scores in the comparison group, contributing to the small effect sizes in the sample. However, because comorbidity is high between youth afflicted with BSDs and those disorders, we maintain that this analytical method maps most closely onto what a clinician would potentially encounter in a clinic.

Clinical Implications

While the present results provide support for the BIS/BAS scales as a promising research

dimension, the scales appear to have limited clinical utility. The BIS/BAS scales were only somewhat useful when individuals were at moderate risk for BSD and got a low score on the BIS/BAS. Second, it should also be noted that the BIS/BAS scales and subscales are in no way sufficient for finalizing a BSD diagnosis in isolation. They were not originally intended to be diagnostic instruments, and they do not systematically assess all the hallmark features associated with BSDs (such as cycling and duration of illness). The BIS/BAS cannot substitute for a thorough evaluation administered by a trained professional familiar with the diagnostic criteria for pediatric BSD. Third, using the BIS/BAS scales as a RDoC dimension (Insel et al., 2010) and as a possible endophenotype for BSD presentation might be helpful from a research perspective, but will not be as immediately helpful to the clinician as compared to current measures, such as the YMRS and P-GBI (Youngstrom, Gracious, Danielson, Findling, & Calabrese, 2003).

The results from this study should encourage clinical practitioners to: (1) track local base rates of diagnoses and common presenting problems, as they would be able to use them in conjunction with DLRs to obtain updated base rates; (2) select assessment tools that have demonstrated discriminative validity based on empirically-validated statistical tools such as the ROC and DLRs, and (3) to have DLRs available along with means of integrating disparate pieces of information, such as the probability nomogram (Jenkins et al., 2011; Youngstrom, 2013b). Future research should evaluate revised versions of the BIS/BAS that have been validated for use in diverse populations in BSDs, for the current BIS/BAS scales are multidimensional and lack configural invariance when assessing BIS and BAS in diverse samples (Demianczyk, Jenkins, Henson, & Conner, 2014). The BIS/BAS seems to have greatest value in the role of self or collateral report of transdiagnostic dimensions by complementing performance measures, and perhaps providing a narrower or cleaner measure of a construct than symptom based measures.

APPENDIX 1: DEMOGRAPHICS OF STUDY SAMPLE

	Any Bipolar Spectrum Disorder (<i>n</i> = 152)	No Bipolar Spectrum Disorder (<i>n</i> = 667)	<i>d</i>	χ^2
Race				22.14****
White	56	128	-	
Nonwhite	96	539	-	
Gender				2.33
Male	83	409	-	
Female	69	258	-	
Age --mean, (<i>SD</i>)	11.12 (3.68)	10.85 (3.37)	.077	-
Mean number of diagnoses-- mean, (<i>SD</i>)	4.10 (1.82)	3.63 (1.69)	.27	-
CGAS--mean, (<i>SD</i>)	54.13 (6.96)	52.26 (8.81)	.24*	-
Index Test--mean, (<i>SD</i>)				
BAS-Total	25.75 (7.62)	22.92 (7.11)	.38*	-
BAS-Drive	9.05 (3.41)	7.39 (3.27)	.50	-
BAS-Fun Seeking	8.42 (2.83)	7.62 (2.58)	.12	-
BAS-Reward Responsiveness	8.29 (3.18)	7.91 (3.04)	.30	-

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Nonwhites significantly more likely to be found in the BSD-negative group (80.8% nonwhite, $n = 539$) than the BSD-positive group (63.2% nonwhite, $n = 96$), $\chi^2(1) = 22.14$, $p < .0005$.

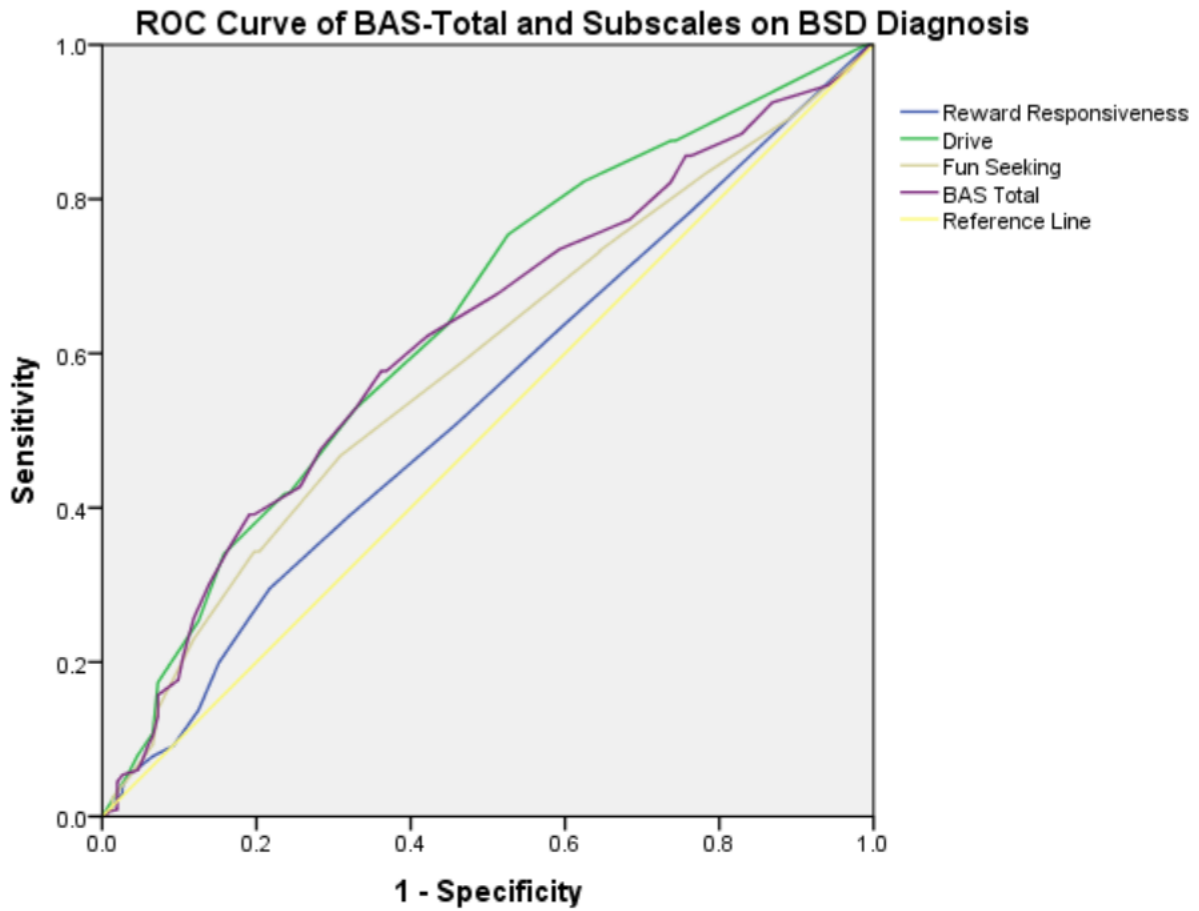
**APPENDIX 2: ROC/AUC ANALYSES OF DIAGNOSTIC DIFFERENTIATION USING
BIS/BAS TOTAL AND SUBSCALES**

Comparisons	AUC (SE)	95% CI	DLRs		
			<u>0-22</u> <i>(Low)</i>	<u>23-31</u> <i>Indeterminate</i>	<u>32+</u> <i>(High)</i>
BAS Total	.62 (.02)***	.57 - .67	0.51	0.77	1.53
BAS Drive	.64 (.03)***	.59 - .69	<u>0-5</u> <i>(Low)</i> 0.49	<u>6-11</u> <i>Indeterminate</i> 0.87	<u>12+</u> <i>(High)</i> 2.12
BAS Fun Seeking	.59 (.02)***	.54 - .64	<u>0-6</u> <i>(Low)</i> 0.52	<u>7-10</u> <i>Indeterminate</i> 1.04	<u>11+</u> <i>(High)</i> 1.33
BAS Reward Responsiveness	.54	.49 - .59	-	-	-

Note: * = $p < .05$, ** = $p < .01$, *** $p < .0005$.

Venkatraman test for two correlated ROC curves are as follows: (a) BAS-Total vs BAS-Drive: $E = 8142$, $p < .05$; (b) BAS-Total vs BAS-Fun Seeking: $E = 13488$, $p < .0001$; (c) BAS-Drive vs BAS-Fun Seeking: $E = 14798$, $p < .01$. Venkatraman test was conducted with bootstrapping of 2000 replications (default).

APPENDIX 3: ROC CURVE OF BAS-TOTAL AND SUBSCALES ON BSD DIAGNOSIS



There was no scale among the BAS-Total scale and the subscales that differentially predicted BSD diagnosis in youth.

REFERENCES

- Alloy, L. B., & Abramson, L. Y. (2010). The Role of the Behavioral Approach System (BAS) in Bipolar Spectrum Disorders. *Current Directions in Psychological Science*, 19(3), 189-194. doi: 10.1177/0963721410370292
- Alloy, L. B., Abramson, L. Y., Walshaw, P. D., Gerstein, R. K., Keyser, J. D., Whitehouse, W. G., . . . Harmon-Jones, E. (2009). Behavioral approach system (BAS)-relevant cognitive styles and bipolar spectrum disorders: Concurrent and prospective associations. *Journal of Abnormal Psychology*, 118(3), 459-471. doi: 10.1037/a0016604
- Alloy, L. B., Bender, R. E., Wagner, C. A., Whitehouse, W. G., Abramson, L. Y., Hogan, M. E., . . . Harmon-Jones, E. (2009). Bipolar spectrum-substance use co-occurrence: Behavioral approach system (BAS) sensitivity and impulsiveness as shared personality vulnerabilities. *Journal of Personality & Social Psychology*, 97(3), 549-565. doi: 10.1037/a0016061
- Alloy, L. B., Bender, R. E., Whitehouse, W. G., Wagner, C. A., Liu, R. T., Grant, D. A., . . . Abramson, L. Y. (2012). High Behavioral Approach System (BAS) sensitivity, reward responsiveness, and goal-striving predict first onset of bipolar spectrum disorders: A prospective behavioral high-risk design. *Journal of Abnormal Psychology*, 121(2), 339-351. doi: 10.1037/a0025877
- American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders : DSM-5. from <http://dsm.psychiatryonline.org/book.aspx?bookid=556>
- Axelson, D. A., Birmaher, B., Strober, M. A., Goldstein, B. I., Ha, W., Gill, M. K., . . . Keller, M. B. (2011). Course of subthreshold bipolar disorder in youth: Diagnostic progression from bipolar disorder not otherwise specified. *Journal of the American Academy of Child and Adolescent Psychiatry*, 50(10), 1001-1016 e1003. doi: 10.1016/j.jaac.2011.07.005
- Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry*, 175, 444-451.
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., . . . de Vet, H. C. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Croatian Medical Journal*, 44(5), 635-638.
- Braddock, K. H., Dillard, J. P., Voigt, D. C., Stephenson, M. T., Sopory, P., & Anderson, J. W. (2011). Impulsivity partially mediates the relationship between BIS/BAS and risky health behaviors. *Journal of Personality*, 79(4), 793-810. doi: 10.1111/j.1467-6494.2011.00699.x
- Carlson, G. A. (1990). Child and adolescent mania--diagnostic considerations. *Journal of Child Psychology and Psychiatry*, 31(3), 331-341.

- Carlson, G. A., & Blader, J. C. (2011). Diagnostic implications of informant disagreement for manic symptoms. *Journal of Child and Adolescent Psychopharmacology*, *21*(5), 399-405. doi: 10.1089/cap.2011.0007
- Carlson, G. A., & Klein, D. N. (2014). How to understand divergent views on bipolar disorder in youth. *Annual Review of Clinical Psychology*, *10*, 529-551. doi: 10.1146/annurev-clinpsy-032813-153702
- Carlson, G. A., & Youngstrom, E. A. (2011). Two opinions about one child--what's the clinician to do? *Journal of Child and Adolescent Psychopharmacology*, *21*(5), 385-387. doi: 10.1089/cap.2011.2159
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319-333.
- Cashel, M. L. (2002). Child and adolescent psychological assessment: Current clinical practices and the impact of managed care. *Professional Psychology: Research and Practice*, *33*(5), 446-453. doi: 10.1037//0735-7028.33.5.446
- Chan, J., Stringaris, A., & Ford, T. (2011). Bipolar disorder in children and adolescents recognised in the UK: A clinic-based study. *Child and Adolescent Mental Health*, *16*, 71-78. doi: 10.1111/j.1475-3588.2010.00566.x
- Chilakamarri, J. K., Filkowski, M. M., & Ghaemi, S. N. (2011). Misdiagnosis of bipolar disorder in children and adolescents: a comparison with ADHD and major depressive disorder. *Annals of Clinical Psychiatry*, *23*(1), 25-29.
- Cuthbert, B. N., & Insel, T. R. (2010). Toward new approaches to psychotic disorders: the NIMH Research Domain Criteria project. *Schizophrenia Bulletin*, *36*(6), 1061-1062. doi: 10.1093/schbul/sbq108
- Demianczyk, A. C., Jenkins, A. L., Henson, J. M., & Conner, B. T. (2014). Psychometric evaluation and revision of Carver and White's BIS/BAS scales in a diverse sample of young adults. *Journal of Personality Assessment*, *96*(5), 485-494. doi: 10.1080/00223891.2013.870570
- Depue, R. A., & Iacono, W. G. (1989). Neurobehavioral aspects of affective disorders. *Annual Review of Psychology*, *40*, 457-492. doi: 10.1146/annurev.ps.40.020189.002325
- Dunn, L. M., & Dunn, L. M. (1997). *Examiner's manual for the Peabody Picture Vocabulary Test*. (3rd ed.). Circle Pines, MN: American Guidance Service.
- Ebesutani, C., Bernstein, A., Chorpita, B. F., & Weisz, J. R. (2012). A transportable assessment protocol for prescribing youth psychosocial treatments in real-world settings: Reducing

- assessment burden via self-report scales. *Psychological Assessment*, 24(1), 141-155. doi: 10.1037/a0025176
- Findling, R. L., Gracious, B. L., McNamara, N. K., Youngstrom, E. A., Demeter, C. A., Branicky, L. A., & Calabrese, J. R. (2001). Rapid, continuous cycling and psychiatric comorbidity in pediatric bipolar I disorder. *Bipolar Disorders*, 3(4), 202-210.
- Geller, B., Warner, K., Williams, M., & Zimmerman, B. (1998). Prepubertal and young adolescent bipolarity versus ADHD: Assessment and validity using the WASH-U-KSADS, CBCL and TRF. *Journal of Affective Disorders*, 51(2), 93-100.
- Geller, B., Zimmerman, B., Williams, M., Bolhofner, K., Craney, J. L., DelBello, M. P., & Soutullo, C. (2001). Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(4), 450-455. doi: 10.1097/00004583-200104000-00014
- Geller, B., Zimmerman, B., Williams, M., Delbello, M. P., Bolhofner, K., Craney, J. L., . . . Nickelsburg, M. J. (2002). DSM-IV mania symptoms in a prepubertal and early adolescent bipolar disorder phenotype compared to attention-deficit hyperactive and normal controls. *Journal of Child and Adolescent Psychopharmacology*, 12(1), 11-25. doi: 10.1089/10445460252943533
- Goldstein, T. R., Ha, W., Axelson, D. A., Goldstein, B. I., Liao, F., Gill, M. K., . . . Birmaher, B. (2012). Predictors of prospectively examined suicide attempts among youth with bipolar disorder. *Archives of General Psychiatry*, 69(11), 1113-1122. doi: 10.1001/archgenpsychiatry.2012.650
- Goodwin, F. K., & Jamison, K. R. (2007). *Manic-depressive illness: Bipolar disorders and recurrent depression*. New York, NY: Oxford University Press.
- Gray, J. A. (1981). A critique of Eysenck's theory of personality. In H. J. Eysenck (Ed.), *A model for personality* (pp. 246-276.). Berlin: Springer-Verlag.
- Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press.
- Gray, J. A. (1994). A general model of the limbic system and basal ganglia: Applications to schizophrenia and compulsive behavior of the obsessive type. *Revue Neurologique*, 150(8-9), 605-613.
- Gruber, J., Gilbert, K. E., Youngstrom, E., Youngstrom, J. K., Feeny, N. C., & Findling, R. L. (2013). Reward dysregulation and mood symptoms in an adolescent outpatient sample. *Journal of Abnormal Child Psychology*, 41(7), 1053-1065. doi: 10.1007/s10802-013-9746-8
- Harmon-Jones, E., Abramson, L. Y., Sigelman, J., Bohlig, A., Hogan, M. E., & Harmon-Jones, C. (2002). Proneness to hypomania/mania symptoms or depression symptoms and

- asymmetrical frontal cortical responses to an anger-evoking event. *Journal of Personality & Social Psychology*, 82(4), 610-618.
- Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L., & Rosati, R. A. (1982). Evaluating the yield of medical tests. *JAMA*, 247(18), 2543-2546. doi: doi:10.1001/jama.1982.03320430047030
- Hasler, G., Drevets, W. C., Gould, T. D., Gottesman, II, & Manji, H. K. (2006). Toward constructing an endophenotype strategy for bipolar disorders. *Biol Psychiatry*, 60(2), 93-105. doi: 10.1016/j.biopsych.2005.11.006
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167(7), 748-751. doi: 10.1176/appi.ajp.2010.09091379
- Jaeschke, R., Guyatt, G. H., & Sackett, D. L. (1994). Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*, 271(9), 703-707.
- Jenkins, M. M., Youngstrom, E. A., Washburn, J. J., & Youngstrom, J. K. (2011). Evidence-based strategies improve assessment of pediatric bipolar disorder by community practitioners. *Professional Psychology: Research and Practice*, 42(2), 121-129. doi: 10.1037/a0022506
- Judd, L. L., Akiskal, H. S., Schettler, P. J., Endicott, J., Maser, J., Solomon, D. A., . . . Keller, M. B. (2002). The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Archives of General Psychiatry*, 59(6), 530-537.
- Kahana, S. Y., Youngstrom, E. A., Findling, R. L., & Calabrese, J. R. (2003). Employing parent, teacher, and youth self-report checklists in identifying pediatric bipolar spectrum disorders: An examination of diagnostic accuracy and clinical utility. *Journal of Child and Adolescent Psychopharmacology*, 13(4), 471-488. doi: 10.1089/104454603322724869
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., . . . Ryan, N. (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(7), 980-988. doi: 10.1097/00004583-199707000-00021
- Kim, E. Y., & Miklowitz, D. J. (2002). Childhood mania, attention deficit hyperactivity disorder and conduct disorder: A critical review of diagnostic dilemmas. *Bipolar Disorders*, 4(4), 215-225.

- Marchand, W. R., Wirth, L., & Simon, C. (2006). Delayed diagnosis of pediatric bipolar disorder in a community mental health setting. *Journal of Psychiatric Practice, 12*(2), 128-133.
- Markarian, S. A., Pickett, S. M., Deveson, D. F., & Kanona, B. B. (2013). A model of BIS/BAS sensitivity, emotion regulation difficulties, and depression, anxiety, and stress symptoms in relation to sleep quality. *Psychiatry Research, 210*(1), 281-286. doi: 10.1016/j.psychres.2013.06.004
- McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology, 50*, 215-241. doi: 10.1146/annurev.psych.50.1.215
- Merikangas, K. R., Akiskal, H. S., Angst, J., Greenberg, P. E., Hirschfeld, R. M., Petukhova, M., & Kessler, R. C. (2007). Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Archives of General Psychiatry, 64*(5), 543-552. doi: 10.1001/archpsyc.64.5.543
- Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., . . . Zarkov, Z. (2011). Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Archives of General Psychiatry, 68*(3), 241-251. doi: 10.1001/archgenpsychiatry.2011.12
- Meyers, O. I., & Youngstrom, E. A. (2008). A Parent General Behavior Inventory subscale to measure sleep disturbance in pediatric bipolar disorder. *Journal of Clinical Psychiatry, 69*(5), 840-843.
- Miklowitz, D. J., & Chang, K. D. (2008). Prevention of bipolar disorder in at-risk children: theoretical assumptions and empirical foundations. *Development and Psychopathology, 20*(3), 881-897. doi: 10.1017/S0954579408000424
- Miller, C. J., Johnson, S. L., Kwapil, T. R., & Carver, C. S. (2011). Three studies on self-report scales to detect bipolar disorder. *Journal of Affective Disorders, 128*(3), 199-210. doi: 10.1016/j.jad.2010.07.012
- Nottelmann, E., Biederman, J., Birmaher, B., Carlson, G. A., Chang, K. D., Fenton, W. S., . . . Weller, E. (2001). National Institute of Mental Health research roundtable on prepubertal bipolar disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(8), 871-878.
- Nusslock, R., Abramson, L. Y., Harmon-Jones, E., Alloy, L. B., & Hogan, M. E. (2007). A goal-striving life event and the onset of hypomanic and depressive episodes and symptoms: perspective from the behavioral approach system (BAS) dysregulation theory. *Journal of Abnormal Psychology, 116*(1), 105-115. doi: 10.1037/0021-843X.116.1.105
- Nusslock, R., Harmon-Jones, E., Alloy, L. B., Urosevic, S., Goldstein, K., & Abramson, L. Y. (2012). Elevated left mid-frontal cortical activity prospectively predicts conversion to

- bipolar I disorder. *Journal of Abnormal Psychology*, 121(3), 592-601. doi: 10.1037/a0028973
- Obuchowski, N. A. (2003). Receiver operating characteristic curves and their use in radiology. *Radiology*, 229, 3-8.
- Pliszka, S. R., Sherman, J. O., Barrow, M. V., & Irick, S. (2000). Affective disorder in juvenile offenders: A preliminary study. *The American Journal of Psychiatry*, 157(1), 130-132.
- R Development Core Team. (2014). A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research*, 18(3), 169-184. doi: 10.1002/mpr.289
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Muller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77. doi: 10.1186/1471-2105-12-77
- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-Based Medicine: How to practice and teach EBM*. (2nd ed.). New York: Churchill Livingstone.
- Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heinssen, R. K., . . . Cuthbert, B. N. (2010). Developing constructs for psychopathology research: research domain criteria. *Journal of Abnormal Psychology*, 119(4), 631-639. doi: 10.1037/a0020909
- Stewart, A. J., Theodore-Oklot, C., Hadley, W., Brown, L. K., Donenberg, G., DiClemente, R., & Project, S. S. G. (2012). Mania symptoms and HIV-risk behavior among adolescents in mental health treatment. *Journal of Clinical Child and Adolescent Psychology*, 41(6), 803-810. doi: 10.1080/15374416.2012.675569
- Straus, S. E., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2011). *Evidence-based medicine: How to practice and teach evidence-based medicine*. Edinburgh: Churchill Livingstone Elsevier.
- ten Have, M., Vollebergh, W., Bijl, R., & Nolen, W. A. (2002). Bipolar disorder in the general population in The Netherlands (prevalence, consequences and care utilisation): Results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Journal of Affective Disorders*, 68(2-3), 203-213.

- Urosevic, S., Collins, P., Muetzel, R., Lim, K., & Luciana, M. (2012). Longitudinal changes in behavioral approach system sensitivity and brain structures involved in reward processing during adolescence. *Developmental Psychology, 48*(5), 1488-1500. doi: 10.1037/a0027502
- Van Meter, A. R., Moreira, A. L., & Youngstrom, E. A. (2011). Meta-analysis of epidemiologic studies of pediatric bipolar disorder. *Journal of Clinical Psychiatry, 72*(9), 1250-1256. doi: 10.4088/JCP.10m06290
- Venkatraman, E. S. (2000). A permutation test to compare receiver operating characteristic curves. *Biometrics, 56*(4), 1134-1138. doi: 10.1111/j.0006-341X.2000.01134.x
- Watkins, M. W., & Glutting, J. J. (2000). Incremental validity of WISC-III profile elevation, scatter, and shape information for predicting reading and math achievement. *Psychological Assessment, 12*(4), 402-408.
- Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996). The Comprehensive System for the Roschach: A critical examination. *Psychological Science, 7*, 3-10.
- Youngstrom, E. A. (2007). Pediatric bipolar disorder. In E. Mash & R. Barkley (Eds.), *Assessment of child disorders* (4th ed., pp. 253-304). New York: Guilford Press.
- Youngstrom, E. A. (2013a). Future directions in psychological assessment: Combining evidence-based medicine innovations with psychology's historical strengths to enhance utility. *Journal of Clinical Child and Adolescent Psychology, 42*(1), 139-159. doi: 10.1080/15374416.2012.736358
- Youngstrom, E. A. (2013b). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology*. doi: 10.1093/jpepsy/jst062
- Youngstrom, E. A., & De Los Reyes, A. (in press). Moving towards cost-effectiveness in using psychophysiological measures in clinical assessment: Validity, decision-making, and adding value. *Journal of Clinical Child and Adolescent Psychology, 1*-28.
- Youngstrom, E. A., Findling, R. L., & Calabrese, J. R. (2003). Who are the comorbid adolescents? Agreement between psychiatric diagnosis, youth, parent, and teacher report. *Journal of Abnormal Child Psychology, 31*(3), 231-245.
- Youngstrom, E. A., Findling, R. L., Danielson, C. K., & Calabrese, J. R. (2001). Discriminative validity of parent report of hypomanic and depressive symptoms on the General Behavior Inventory. *Psychological Assessment, 13*(2), 267-276.
- Youngstrom, E. A., Findling, R. L., Youngstrom, J. K., & Calabrese, J. R. (2005). Toward an evidence-based assessment of pediatric bipolar disorder. *Journal of Clinical Child and Adolescent Psychology, 34*(3), 433-448. doi: 10.1207/s15374424jccp3403_4

- Youngstrom, E. A., Gracious, B. L., Danielson, C. K., Findling, R. L., & Calabrese, J. (2003). Toward an integration of parent and clinician report on the Young Mania Rating Scale. *Journal of Affective Disorders, 77*(2), 179-190.
- Youngstrom, E. A., Jenkins, M. M., Doss, A. J., & Youngstrom, J. K. (2012). Evidence-Based Assessment Strategies for Pediatric Bipolar Disorder. *Israel Journal of Psychiatry and Related Sciences, 49*(1), 15-27.
- Youngstrom, E. A., Meyers, O., Demeter, C., Youngstrom, J. K., Morello, L., Piiparinen, R., . . . Findling, R. L. (2005). Comparing diagnostic checklists for pediatric bipolar disorder in academic and community mental health settings. *Bipolar Disorders, 7*(6), 507-517. doi: 10.1111/j.1399-5618.2005.00269.x
- Youngstrom, E. A., Youngstrom, J. K., & Starr, M. (2005). Bipolar diagnoses in community mental health: Achenbach Child Behavior Checklist profiles and patterns of comorbidity. *Biological Psychiatry, 58*(7), 569-575. doi: 10.1016/j.biopsych.2005.04.004
- Zhou, X., Obuchowski, N. A., & McClish, D. K. (2011). *Statistical methods in diagnostic medicine*. Hoboken, NJ: John Wiley & Sons, Inc.