NIH Public Access
Author Manuscript
Psychol Assess. Author manuscript; available in PMC 2014 July 16.

Published in final edited form as:
Psychol Assess. 2014 March ; 26(1): 195–206. doi:10.1037/a0035138.

Differentiating Bipolar Disorder from Unipolar Depression and ADHD: The Utility of the General Behavior Inventory

Laura L. Pendergast,
Department of Psychological, Organizational, and Leadership Studies in Education, Temple University

Eric A. Youngstrom,
Department of Psychology, University of North Carolina – Chapel Hill

Kristen G. Merkitch,
Department of Psychology, Temple University

Katie A. Moore,
Department of Psychology, Temple University

Chelsea Black,
Department of Psychology, Temple University

Lyn Abramson, and
Department of Psychology, University of Wisconsin-Madison

Lauren B. Alloy
Department of Psychology, Temple University

Abstract

Adolescence and early adulthood are the peak ages for the onset of unipolar and bipolar mood disorders. Moreover, for most individuals with attention deficit hyperactivity disorder (ADHD), symptoms and impairment begin in childhood but persist well into adolescence and adulthood (e.g., Barkley, 2010). Thus, adolescence and early adulthood represent a developmental window wherein individuals can be affected by mood disorders, ADHD, or both. Because treatment protocols for unipolar depression (UPD), bipolar disorder (BD), and ADHD are quite different, it is crucial that assessment instruments used among adolescents and young adults differentiate between these disorders. The primary objectives of this study were to evaluate the predictive and diagnostic validity of General Behavior Inventory (GBI; Depue, Krauss, Spoont, & Arbisi, 1989) scores in discriminating BD from UPD and ADHD. Participants were drawn from adolescent (n=361) and young adult (n=614) samples. Based on findings from logistic regression and receiver operating characteristics analyses, the diagnostic efficiency of the GBI scales range from *fair* (discriminating UPD from BD) to *good* (discriminating BD participants from nonclinical controls). Multilevel diagnostic likelihood ratios are also provided to facilitate individual decision making.

Correspondence concerning this article should be sent to Laura L. Pendergast, laura.pendergast@temple.edu.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Bipolar disorder (BD) is an affective condition that affects approximately 4% of the U.S. population (Merikangas et al., 2007) and is often associated with adverse outcomes, including increased use of health services, difficulty with employment and interpersonal relationships, and high rates of suicide attempts (Dennehey et al., 2011; Judd & Akiskal, 2003; Merikangas et al., 2007; Robins & Regier, 1991; Sanchez-Moreno et al., 2009). Pharmacological and psychosocial treatments can considerably reduce BD symptoms and prevent relapse (see Fountoulakis & Vieta, 2008 for a review). However, diagnostic challenges emerge when differentiating BD from other disorders that have a high degree of shared symptomatology, particularly unipolar depression (UPD) and attention deficit hyperactivity disorder (ADHD; Galanter & Leibenluft, 2008; Geller, Zimerman, Williams, Bolhofner, & Craney, 2001; Sala, Axelson, & Birmaher, 2009). Actuarial assessment instruments are thought to increase the objectivity and reliability of predictions, and have been shown to improve diagnostic certainty and identification of appropriate treatment (Dawes, Faust, & Meehl, 1989). The General Behavior Inventory (GBI; Depue, Krauss, Spoont, & Arbisi, 1989) is commonly used to assess BD symptoms and has demonstrated robust psychometric properties relative to other instruments (Miller, Johnson, & Eisner, 2009). However, the utility of the GBI in distinguishing BD from other symptomatically similar disorders remains unclear. The purpose of this study is to examine the diagnostic value of GBI scores in differentiating BD from UPD, as well as ADHD.

Overview of Bipolar Disorder

BD is characterized by intense and fluctuating states of depression and (hypo)mania—persistent and abnormal periods of elevated, expansive, or irritable mood. There are several BD subtypes delineated largely by the severity and duration of (hypo)manic symptoms. Bipolar I is defined by a history of at least one manic episode that causes significant impairment or hospitalization (a history of depressive episodes may or may not be present), whereas bipolar II is defined by a history of at least one hypomanic and major depressive episode. Hypomania criteria are comparable to those of mania but to a lesser degree: rather than significant impairment, hypomania is discerned by a significant change in functioning. Cyclothymia is a chronic disorder defined by alternating periods of potentially brief hypomanic and depressive symptoms, but with the mood *dysregulation* persisting for two or more years (one year in adolescents). Bipolar disorder not otherwise specified (BDNOS; American Psychiatric Association, 2001) is characterized by (hypo)manic symptoms with or without depressive symptoms that are insufficient in severity, duration, or persistence to meet the full criteria for mania, hypomania, cyclothymia, or depression (American Psychiatric Association, DSM-IV-TR, 2000; cf. “other specified bipolar and related disorder” in DSM-5, American Psychiatric Association, 2013). Although bipolar spectrum disorders may emerge at any time, research suggests that adolescence may be a particular age of risk for first onset of bipolar disorder (see Alloy, Abramson, Walshaw, Keyser, & Gerstein, 2006 for a review and synthesis). One retrospective study found that the peak age of onset for BD symptoms was between 15 and 19 years (Lish, Dime-Meenan, Whybrow, Price, & Hirschfeld, 1994). Individuals with BD may develop symptoms in childhood, which often leads to misdiagnosis of more common pediatric disorders, such as ADHD, that have substantial symptomatic overlap (Maniscalco & Hamrin, 2008).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Diagnosis of Bipolar Disorder

Several studies have demonstrated that individuals with bipolar disorder may wait between 5–15 years before a formal diagnosis of BD is made (e.g., Ghaemi, Boiman, & Goodwin, 2000; Ghaemi, Sachs, Chiou, Pandurangi, & Goodwin, 1999; Hirschfeld, Lewis, & Vornik, 2003). Differentiating BD from disorders with similar symptoms, such as UPD and ADHD, is particularly difficult (Galanter & Leibenluft, 2008; Geller, Zimerman, Williams, Bolhofner, & Craney, 2001; Sala, Axelson, & Birmaher, 2009).

BD and UPD are challenging to differentiate (Akiskal, 1995; Bowden, 2005; Ghaemi et al., 2000). Many individuals with BD are initially appropriately diagnosed with major depression because the onset of a depressive episode precedes the onset of a (hypo)manic episode (Ghaemi et al., 1999), particularly for individuals with early-onset BD (Bowden, 2001; Lish et al., 1994). However, individuals with BD often do not perceive symptoms of elevated mood as problematic and may even consider them to be adaptive, decreasing the likelihood of spontaneous reporting to clinicians (Akiskal, 1983; Dunner & Tay, 1993). Moreover, current depression may inhibit recall of (hypo)manic symptoms (Judd et al., 2002). Because patients are less likely to report (hypo)manic symptoms, UPD is commonly misdiagnosed in individuals with BD. Differential diagnosis between BD and UPD has substantial treatment implications. Indeed, one of the primary approaches to treatment for UPD—antidepressant monotherapy—is a frequently used but controversial treatment for bipolar disorder, because its efficacy in treating bipolar depression remains unclear (Pacchiarotti et al., 2010; Sidor &MacQueen, 2011; Undurraga et al., 2012; see Bauer, Ritter, Grunze, & Pfennig, 2012 for a review). Some research has shown that antidepressant monotherapy is ineffective in ameliorating BD depression and in preventing (hypo)mania, and it may increase the risk of serious side effects (Pacchiarotti et al., 2010). Considerations about the safety and efficacy of antidepressant medication use among individuals with BD cannot be given appropriate consideration when the diagnosis itself is wrong (Baldessarini, Vieta, Calabrese, Tohen, & Bowden, 2010; Pacchiarotti et al., 2010). These serious treatment implications necessitate the initiation of assessment techniques aimed at accurate and early diagnosis of BD.

Due to overlapping symptomatology, it is also difficult to differentiate BD from ADHD, particularly among children and adolescents. Core symptoms of pediatric BD, such as irritability, distractibility, and accelerated speech, are also very common among individuals with ADHD, and may therefore be of limited utility in differentiating the two groups (Chan et al., 2011; Geller et al., 2002). Both BD and ADHD are often associated with aggression, anxiety, hyperactivity, and mood and sleep disturbances (Asherson, 2005; Cahill, Green, Jairam, & Malhi, 2007; see Skirrow, 2009 for a review). It is particularly difficult to differentiate pediatric BD from ADHD when cross-sectional, as opposed to longitudinal, assessment techniques are used.

Another diagnostic complication is the high comorbidity rate between BD and ADHD in younger samples (e.g. Cahill et al., 2007). Rates of comorbidity of ADHD in samples ascertained for bipolar disorder range as high as 98% (see meta-analysis by Kowatch, Youngstrom, Danielyan, & Findling, 2005). Although rates of BD in samples ascertained for

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

ADHD are often substantially lower, there still is notable co-occurrence that may be attributable to the high base rate of ADHD in most pediatric clinical settings (Galanter & Leibenluft, 2008; Youngstrom, Arnold, & Frazier, 2010). Because ADHD includes similar symptoms and is much more common in most clinical settings, it complicates the accurate detection of bipolar disorder.

For most affected individuals, symptoms of ADHD persist well into adolescence and adulthood (see Barkley, 2006 for review). This exacerbates the challenge for clinicians working with patients in emerging adulthood: ADHD continues to manifest longer than described in most training programs, and with less known about evidence-based assessment strategies (Barkley, 2010; Wender, 1998). The combination of bipolar disorder often manifesting earlier than thought and ADHD continuing longer through development than previously thought means that these two competing diagnostic formulations overlap during adolescence and emerging adulthood, creating a clear need for assessment methods that can help identify when either or both conditions are occurring.

Correct differential diagnosis between ADHD and BD is crucial for both research and treatment purposes. If an individual with BD is misdiagnosed as having only ADHD, the best case scenario is merely a delay in initiating treatments that would help stabilize acute mood disturbance (Baldessarini, Tondo, Hennen, & Viguera, 2002). However, untreated BD is associated with high rates of substance misuse, vocational and social disruption, and increased likelihood of risky behavior, accidents, and incarceration (Lopez, Mathers, Ezzati, Jamison, & Murray, 2006; Stewart, et al., 2012), as well as progression to more severe BD (e.g., Alloy et al., 2012b). Further, individuals with BD often present with suicidal ideation whereas those with ADHD do not (Cahill et al., 2007; Geller et al., 2002). Those misdiagnosed with ADHD thus experience a delay in receiving efficacious treatment with mood stabilizers, which have been shown to reduce serious symptoms, including suicidality. Conversely, misdiagnosing someone who has ADHD with BD leads to a dramatically different prescription, with atypical antipsychotics or anti-epileptic agents as the front line treatment, and combination treatment using more than one drug simultaneously being common (Yatham, et al., 2005). These medications have little or no evidence of efficacy for treating ADHD, yet they will still expose the patient to all the potential side effects, which can be major (Correll, 2008).

Assessment of Bipolar Disorder

Poor assessment techniques and instruments are partially responsible for the delay in formal diagnosis of BD (Lish et al., 1994; Mantere, Suiminen, Leppamaki, Arvilommi, & Isometsa, 2004; Miller, Johnson, Kwapil, & Carver, 2010). Inasmuch as a BD diagnosis is founded on the frequency, length, and severity of mood and energy disturbance, diagnostic complications may arise when practitioners fail to assess symptoms in a specific and sequential manner. This is often the case in clinical practice as many practitioners use unstructured diagnostic methods, increasing the likelihood of inaccurate diagnosis of BD (Rettew, Lynch, Achenbach, Dumenci, & Ivanova, 2009; Zimmerman & Mattia, 1999; Brickman, LoPicollo, & Johnson, 2002). Actuarial assessment techniques, such as combining behavior rating scales with statistical prediction rules (Swets, Dawes, &

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Monahan, 2000), may improve diagnosis of BD because such techniques tend to be more consistent, more accurate, and substantially less prone to cognitive biases (Jenkins, Youngstrom, Washburn, & Youngstrom, 2011).

Structured clinical interviews are considered to be preferred tools in the assessment of bipolar disorder and could potentially alleviate diagnostic uncertainty. However, structured interviews lack brevity and cost-efficiency – qualities that are in high demand in modern day clinical practice (Ebesutani, Bernstein, Chorpita, & Weisz, 2012; Starfield et al., 1994). Various behavior checklists have been evaluated as potential actuarial measures for BD (see Johnson, Miller, & Eisner, 2008; Youngstrom, Freeman, & Jenkins, 2009, for review). The Child Behavior Checklist (CBCL; Achenbach, 1991) differentiates BD from ADHD and UPD in youths (Mick, Biederman, Pandina, & Faraone, 2003; Youngstrom, Findling, Calabrese, et al., 2004), but it has not been evaluated in young adults for discriminating BD. The Mood Disorder Questionnaire (MDQ; Hirschfeld et al., 2000) can distinguish between BD and UPD in adult clinical populations (Miller, Johnson, & Eisner, 2009), but it has not been evaluated for discriminating BD from ADHD in this age range.

Among the numerous assessments for BD available, the General Behavior Inventory (GBI) has the greatest combined sensitivity (.78) and specificity (.98), and overall most robust psychometric properties (Depue, Krauss, Spoont, & Arbisi, 1989; Klein, Dickstein, Taylor, & Harding, 1989; Miller et al., 2010), as well as small standard errors of measurement (Danielson, Youngstrom, Findling, & Calabrese, 2003). The GBI was specifically created to encompass the major symptoms of BD, including both manic and depressive features (Miller, Johnson, & Eisner, 2009). The Depression scale of the GBI discriminates adolescents with a mood disorder from both those with a non-mood disorder and those with no disorder (Reichart et al., 2003). The GBI also has demonstrated efficacy at case-finding in nonclinical populations (Alloy et al., 2008; 2012b; Angst & Cassano, 2005, Depue et al., 1989). Additionally, the GBI Hypomanic/Biphasic scale can identify bipolar versus unipolar cases with minimal false positives (Depue et al., 1989). One study also found that the GBI differentiated youths with BD from those with ADHD, oppositional defiant disorder, and conduct disorder (Danielson, Youngstrom, Findling, & Calabrese, 2003).

Although substantial research supports the GBI as a useful instrument for identifying bipolar disorder among both adolescents and adults, studies examining the extent to which scores on the GBI differentiate BD from ADHD and UPD are sparse, and no studies to date have examined the ability of the GBI to differentiate between BD, UPD and ADHD in young adults. The purpose of this study is to examine the predictive and diagnostic validity of GBI scores in differentiating individuals with BD from nonclinical controls and those with ADHD or UPD in late adolescence and emerging adulthood – the age range where it is clinically crucial to differentiate between persisting ADHD versus emerging BD, while also accurately discerning UPD. We hypothesized that individuals with BD would score significantly higher than nonclinical controls and individuals with ADHD or UPD on the GBI Hypomanic/Biphasic subscale (consistent with findings in adolescents; Danielson et al., 2003). Further, we hypothesized that the BD and UPD groups would score significantly higher on the GBI Depression subscale than the ADHD or nonclinical controls, with the two mood disorder groups showing similar elevations to each other on GBI Depression. In

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

particular, we believed that GBI subscale scores would differentiate individuals with BD from those with ADHD because the items focus on mood and energy, and they embed the concept of change in functioning and phasic, rather than chronic, presentation. Another objective was to develop multi-level diagnostic likelihood ratios (DLRs) to facilitate assessment and decision-making about individual cases (Straus, Glasziou, Richardson, & Haynes, 2011). To our knowledge, this paper is the first to examine the GBI's discriminative validity for separating BD and ADHD in young adults. It is also the first to develop DLRs for interpretation of GBI scores in this age range – making it much more feasible for clinicians to adopt an evidence based assessment (EBA) framework for interpreting the GBI (Youngstrom, 2013).

Method

Participants

Participants from two samples were included in this study. Participants in Sample 1 were 359 adolescents ages 14–19 who volunteered to participate in the Teen Emotion and Motivation (TEAM) project (Alloy et al., 2012a). Participants in Study 2 were 18–24 year old students (*n*=614) who participated in the Longitudinal Investigation of Bipolar Spectrum disorders (LIBS) project (Alloy et al., 2008; 2012b). The TEAM and LIBS projects were conducted at the same university by the same primary investigators, were very similar methodologically, and employed samples that were comparable in regard to race, sex, and SES (see Table 1). Details about sample recruitment and exclusionary criteria are reported elsewhere (Alloy et al., 2008; Alloy et al., 2012a,b). Both studies over-sampled individuals at risk for bipolar spectrum disorders. Notably, although data from participants with a history of (hypo)mania that started prior to enrolling in Project TEAM were excluded from the main study (see Alloy et al., 2012a), data from TEAM participants with a history of BD were included in this study.

Measures

Diagnostic Reference Standard: Schedule for Affective Disorders and Schizophrenia – Lifetime Version—The Schedule for Affective Disorders and Schizophrenia – Lifetime Version (SADS-L) is a widely used semi-structured diagnostic interview that has long been regarded as a preferred tool in clinical assessment (e.g., Gallagher, 1987). An expanded version of the SADS-L (Exp-SADS-L) was administered to participants in projects TEAM and LIBS that included extended coverage of mood symptoms. The Exp-SADS-L adds questions and probes to better capture symptoms of a variety of constructs and disorders including depression, hypomania, mania, cyclothymia, eating disorders, ADHD, and acute stress disorder. The Exp-SADS-L was administered by highly trained post-doctoral fellows, doctoral students, and post-baccalaureate research assistants (see Alloy et al., 2008, 2012a, 2012b, for additional details about interviewer training and exp-SADS-L modifications). All interviewers had at least a bachelor's degree and completed over 200 hours of training (i.e., didactics, case vignettes, audiotaped assessments, role playing, and exams), and were blind to participants' scores on the GBI. Additionally, diagnoses were regularly reviewed by senior diagnosticians and an expert psychiatric diagnostic consultant. In regard to interrater reliability, diagnostic kappa values

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

exceed .96 for mood disorders and .93 for ADHD. Interrater reliability was assessed regularly throughout the project. A subset of the cases (approximately 10%) were independently by three senior diagnosticians, and kappa coefficients were calculated based on ratings from the three senior diagnostician as well as the original interviewer.

General Behavior Inventory (GBI)—The GBI (Depue, 1987) taps depressive and hypomanic/manic symptoms in adults, and it has also been validated in children and adolescents (Danielson et al.; Youngstrom, Meyers, et al., 2005; Youngstrom, Findling et al., 2004). The GBI includes 73 items on which respondents use a 4-point Likert-type scale to indicate the frequency with which they experience a particular phenomenon (e.g., *"Have people said that you looked sad or lonely?"*). Higher scores reflect increased pathology. The Depression scale sums 46 of the items, and the Hypomanic/Biphasic scale sums 28 items; internal consistencies for both scales consistently exceed .90 in prior samples. In this study, alpha coefficients for the (hypo)manic and depressive factors were .93 and .96, respectively, in the adolescent (TEAM) sample and .95 and .98 in the young adult (LIBS) sample.

Procedures

Eligible participants were contacted and invited to visit the lab. LIBS project participants completed the GBI prior to coming to the lab, whereas project TEAM participants completed the GBI during their lab visit. All participants received monetary compensation for participation. Written parental consent and youth assent was obtained for participants under age 18, and participants over 18 independently provided written informed consent. All study procedures were approved by the Institutional Review Board at Temple University.

Data Analyses

Analyses were also conducted on each sample separately, and the findings were comparable (results are available by request). Logistic regressions included interaction terms for sample, and all interactions were non-significant, indicating that the predictions did not change significantly across samples. After establishing that there were no significant differences in GBI performance across samples (all p values $\geq .18$), the data were pooled to maximize the precision of DLRs and diagnostic efficiency estimates (Kraemer, 1992). All participants were grouped into four categories based on exp-SADS-L *DSM-IV* diagnoses: (a) those with BSD (of individuals with BSD, 3% had bipolar I, 50% had bipolar II, 19% bipolar NOS, and 28% had cyclothymia), (b) those with unipolar depression diagnoses (current or past major depressive disorder, dysthymia, depressive disorder NOS, or subthreshold major depressive disorder) but no history of hypomania or mania, (c) those with current or past ADHD (any subtype), and (d) nonclinical controls who did not meet criteria for BD, UPD, or ADHD. Notably, nearly all participants with a past unipolar depression (UPD) or ADHD diagnosis reported some residual symptoms. These categories are hierarchical and allow comorbidity. For example, cases included in the BD group for analyses might also have depression, ADHD, or other diagnoses (e.g., anxiety disorders, eating disorders, etc.); cases with UPD could have ADHD or other comorbid diagnoses (except bipolar); and the ADHD group could include any comorbidities. This hierarchical arrangement is consistent with prior investigations and maximizes the clinical generalizability of findings relative to more

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

"distilled" designs (Youngstrom et al., 2006). In this study, 11% of participants with BD also had ADHD, and 8% of participants with UPD also had ADHD.

A series of analyses evaluated the discriminative validity of GBI subscale scores in differentiating individuals with BSD from those with UPD, ADHD, and persons without the aforementioned diagnoses. First, logistic regression analyses evaluated the extent to which the two GBI subscale scores could differentiate between (a) individuals with BD versus no diagnosis (nonclinical controls), (b) those with any mood disorder versus those without (clinical and nonclinical controls), (c) individuals with BD versus those without (clinical and nonclinical controls), (d) individuals with BD versus those with any other diagnosis (i.e., either UPD or ADHD – clinical controls), (e) individuals with BD versus those with UPD, and (f) individuals with BD versus ADHD. The first comparison (BD vs. nonclinical controls) provides a "best case scenario" for GBI performance (Youngstrom, Meyers, Youngstrom, Calabrese, & Findling, 2006). The second comparison (mood disorder vs. non-mood disorder) provides the extent to which GBI scores can differentiate individuals with mood disorders from those without them – an important first step in assessing BD symptoms. The third and fourth comparisons (BD vs. all controls and BD vs. clinical controls) assessed the utility of GBI scores in differentiating between individuals with BD and all controls (i.e., clinical and nonclinical) as well as BD vs. clinical controls, respectively. Participants included in the BD vs. any diagnosis comparison group are most similar to those typically seen in a clinical setting. The final two comparisons (BD vs. UPD and BD vs. ADHD) examine the utility of GBI scores in contributing to the differential diagnosis of BD and two disorders that are notoriously difficult to differentiate from BD. Logistic regressions also tested whether there was incremental value in combining the GBI scales, versus interpreting the more discriminating scale by itself. Subsequently, receiver operating characteristics (ROC) analyses quantified the relative value of GBI subscale scores in making the aforementioned distinctions (Swets, et al., 2000), and the Hanley & McNeil (1983) procedure tested for significant differences in ROC performance. Finally, DLRs quantified the change in odds of BD diagnosis relative to ranges of test score (e.g., Low, Moderate, High, etc.; Straus, et al., 2011).

Results

Differentiating Diagnostic Categories

Table 2 presents descriptive statistics, findings from logistic regression analyses, and findings from ROC analyses. In all scenarios, the GBI provided statistically significant differentiation of individuals with BD and those from the respective comparison groups. Specifically, scores from the GBI Hypomanic/Biphasic subscale significantly ($p < .001$) and uniquely contributed to all diagnostic comparisons, with Nagelkerke R^2 estimates ranging from .13 (BD vs. UPD) to .33 (BD vs. nonclinical control). GBI Depression subscale scores significantly and uniquely contributed only when comparing individuals with BD to those with no diagnosis (nonclinical controls) and when comparing those with any mood disorder (UPD or BD) to all other participants (all controls; ADHD or no diagnosis).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Diagnostic Efficiency Statistics

ROC analyses examined the value of the two GBI subscale scores for distinguishing between diagnostic groups. ROC curves depict the balance between the probability of a true positive test result (sensitivity) and the probability of a true negative test result (specificity). The Area under the Curve (AUC) reflects the diagnostic accuracy of scores, whereby an AUC of 1 would indicate perfect diagnostic accuracy, and an AUC of .50 indicates chance levels of diagnostic performance. AUC values for both GBI subscales were significantly better than chance across all four comparison groups (see Table 2). Calculations were conducted to evaluate whether the Hypomanic/Biphasic or Depressive subscale was more effective in differentiating individuals with the target disorder from those without it (Hanley & McNeil, 1983). For most comparison groups, the Hypomanic/Biphasic and Depressive subscales were roughly equal in their ability to differentiate those with the target disorder from those without it. However, the Hypomanic/Biphasic subscale was significantly better than the Depressive subscale in differentiating individuals with BD from those with UPD. ROC curves for the GBI depressive and hypomanic subscales are displayed in Figure 1.

Diagnostic Likelihood Ratios

Diagnostic likelihood ratios (DLRs) capture more detailed diagnostic information for decision-making about individual cases. DLRs repackage the older concepts of diagnostic sensitivity and specificity, making it easier to use the information from test results to estimate posterior predictive values (Straus, et al., 2011). Using a multi-level approach – estimating the DLR for several intervals of score ranges – preserves more information from the test results (Guyatt & Rennie, 2002). Conceptually, the DLR is the change in the odds of the diagnosis based on the assessment results. The DLR can be combined with the prior probability of the diagnosis by means of a probability nomogram (Jaeschke, Guyatt, & Sackett, 1994) or via applets or online calculators. DLRs of 1.0 mean that the test result is not changing the odds of the diagnosis at all; numbers less than 1.0 indicate that the test result decreases the odds of the diagnosis. Numbers greater than 1.0 indicate increased odds (and probability) of the diagnosis. Suggested rules of thumb are that DLRs between .5 and 2.0 rarely have much impact on the decision-making process, whereas DLR's in the 3 to 7 range are helpful, and values greater than 10 (or smaller than .10) are potentially decisive pieces of information (Straus, et al., 2011).

We emphasize DLRs for scores on the hypomanic/biphasic subscale of the GBI because prior analyses revealed that Hypomanic/Biphasic subscale scores differentiated BSD from other diagnostic groups more effectively than did scores on the depression subscale.

When estimating multi-level DLRs, the goal is to preserve information but not create unnecessary complexity or split to the point that estimates do not behavior monotonically (Zhou, Obuchowski, & McLish, 2002). To calculate DLRs, the scores were initially divided into sextiles (the bottom ~17% of scores were considered Very Low, then next ~17% Low, then Somewhat Low, and so on). However, several categories were found to be redundant (i.e., the confidence intervals for the DLRs overlapped almost completely). Redundant categories were combined, and subsequently, scores on each subscale were divided into three categories: Low, Moderate, and High.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3 reports the DLRs for the GBI Hypomanic/Biphasic subscale. DLRs also were calculated for a two-step diagnostic process whereby the Depressive subscale was used to determine the presence or absence of a mood disorder, and the Hypomanic/Biphasic subscale was used to determine the presence or absence of a BSD; see Table 4.

Overall, increases in odds of BD diagnoses were particularly evident when comparing individuals with BD to those with no diagnosis (nonclinical controls) and to individuals with ADHD. Most notably, when comparing participants with BD to those with ADHD, those with GBI Hypomanic/Biphasic scores of 20 or higher were nearly 5 times more likely to receive bipolar diagnoses using the Exp-SADS-L interview. Conversely, those with very low scores (0 or 1) were 5 times less likely to have a bipolar disorder based on the semi-structured diagnostic interview (DLR = 0.18).

Discussion

The primary goal of the present study was to evaluate the predictive and diagnostic validity of a promising instrument, the General Behavior Inventory, when attempting to discriminate bipolar spectrum disorders from unipolar depression or ADHD in two samples of adolescents and emerging adults. The GBI has performed well in other age groups, but emerging adulthood is a crucial developmental stage to demonstrate the ability of an instrument to discriminate between bipolar disorder, depression, and ADHD. On the one hand, adolescence and early adulthood are the peak ages of risk for the onset of mood disorder (Merikangas & Pato, 2009), and on the other hand, ADHD is more likely to persist into this age range than had previously been appreciated (Barkley, 2010). Differentiating these conditions changes the treatment prescription, so it would be valuable to have effective assessment tools validated for this age group. As hypothesized, the GBI provided statistically significant and clinically meaningful discrimination of cases with bipolar disorder versus the other diagnostic comparison groups. Specifically, the GBI Depression score separated the two groups with mood disorder (BPD and UPD) from the comparison groups without mood disorder (ADHD and Nonclinical controls); and the Hypomanic/Biphasic score discriminated those with BPD from all other groups. Based on ROC analyses, the diagnostic efficiency of the GBI scales ranged from “fair,” with AUCs ~.64 for discriminating UPD from bipolar disorders, to “good” (AUC > .80) at discriminating mood disorders from those with no diagnoses. Logistic regressions indicated that the Hypomanic/Biphasic scale was sufficient to capture the diagnostic information. The exception was attempting to discriminate “any mood disorder” from “all others;” then both the Depression and Hypomanic/Biphasic scales added incremental information, suggesting that a sequential interpretation process might be helpful: First, use the depression scale to determine the presence or absence of mood disorder, and then examine the hypomanic/biphasic scale to differentiate whether the mood disorder was bipolar versus a unipolar depression.

The second aim was to develop multilevel diagnostic likelihood ratios (DLRs) to facilitate clinical decision-making about individual cases using the GBI. The DLR values in Tables 3 and 4 show that GBI scores can contribute helpful information in differentiating bipolar disorders from all comparison groups, with very low or very high scores changing the odds by as much as sixfold. There are several published examples of using DLRs and a

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

probability nomogram or calculator in the assessment of bipolar disorder as well as ADHD (Frazier & Youngstrom, 2006; Jenkins, et al., 2011). Comparing these values to other instruments used to assess bipolar disorder, the DLRs are comparable to those found using parent report on the GBI to detect bipolar spectrum disorder in children and adolescents (Youngstrom, Findling, Calabrese, et al., 2004) and larger than adolescent self-report on the GBI demonstrated in the same sample (Youngstrom, Findling, Calabrese, et al., 2004; Youngstrom, Frazier, Findling, & Calabrese, 2008).

Most of the participants in the present samples were older, had a better reading level, and may have had more maturity and insight into their behavior than did the adolescents involved in prior investigations of the GBI. Intriguingly, recent evidence suggests that people with ADHD develop increased self-awareness over time, leading to more valid self-report in adolescence and young adulthood (Barkley, Knouse, & Murphy, 2011). It is less clear that improvements in insight would also apply to bipolar disorder, and compromised insight is an associated feature of hypomania or mania (Pini, Dell'Osso, & Amador, 2001; Youngstrom, Findling, & Calabrese, 2004). However, the GBI includes item content assessing changes in energy as well as changes in mood, which may be less prone to state-dependent biases in recall (Angst, et al., 2011). A second, more methodological explanation could be that the predictor and criterion measure used in the present samples relied on more similar information than the criterion used in the pediatric studies (which included parents in the diagnostic interviews). The exp-SADS-L that determined diagnoses interviewed the same person that completed the GBI, combining their self-perceptions with clinical judgment about the content and clinical observations of their mental status and behavior during the interview. Consequently, the present samples had more shared source variance between the predictor and criterion (Campbell & Fiske, 1959). However, the exp-SADS-L format better approximates the methods that would be viable in clinical practice (e.g., it is unlikely that parents would routinely be involved in diagnostic interviews about young adults).

Because it is possible to convert effect sizes such as Cohen's *d* into an estimated area under the curve from a ROC analysis (Hasselbad & Hedges, 1995), results also indicate that the diagnostic efficiency of the GBI outperforms what would be accomplished by most neurocognitive tasks when trying to separate bipolar from ADHD, with the possible exception of verbal memory tasks (Walshaw, Alloy, & Sabb, 2010; pooled Cohen's *d* for 11 different different tests ranged from .31 to .96, median $d = .40$ and AUC=.61, when comparing cases with bipolar to heatlhy controls). The "utility" of a test takes into account not just its validity, but also the costs and benefits associated with its use (Swets, et al., 2000). The GBI has several key advantages from the perspective of utility: It is in the public domain, and may be used at no charge. It requires no special training to administer or score (although a free Excel worksheet is available to calculate the subscale scores and combine prior probabilities with the DLR). The low costs of training and administration, and easy accessibility to the instrument, give the GBI an edge in terms of utility even when neurocognitive or neuroimaging tasks develop to the point that they demonstrate equal diagnostic efficiency. The main drawbacks to the GBI are its length and reading level. The detailed item content requires an 11th or 12th grade reading level, which will not be suitable

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

for use in some clinical settings. Strengths of the present study include that (a) it used a semi-structured diagnostic interview administered by highly trained raters to establish the criterion diagnoses, (b) it tested diagnostic efficiency in two samples, (c) it pooled the samples to increase the precision of the diagnostic likelihood ratios after establishing that the GBI performed similarly in both samples, (d) it used diagnostically heterogeneous samples that included high rates of diagnoses that are challenging to distinguish from bipolar disorder, and (e) it addressed a key gap in the literature of differentiating between bipolar disorder versus depression or unremitted ADHD at an age range where little prior work has been done on assessment despite being a developmental epoch where these disorders are highly likely to overlap. Additionally, the use of semi-structured versus fully structured interview, along with the inclusion of diagnostically heterogeneous comparison groups, increases the generalizability of results (Youngstrom, et al., 2006).

Limitations of the study include reliance on a community sample. Results may not generalize to other clinical settings. Changes in the severity of the mood disorder will affect the sensitivity of the GBI or any other test (Zhou et al., 2002). All other factors being equal, it is easier to detect more severe presentations. The inclusion of cyclothymic disorder and bipolar NOS in the bipolar group resulted in a more conservative estimate of the sensitivity of the GBI. Similarly, clinical settings will vary widely in terms of the rates of “confounding” diagnoses that are likely to generate false positive results on a test (Youngstrom, et al., 2006). Samples with high rates of ADHD, severe depression, and psychosis will make it more difficult for the GBI, or any tool, to tease apart bipolar disorder from the competing diagnoses. Although the present paper extends the investigation of the GBI into early adulthood, it does not address the validity of the GBI in late life, which remains another gap in the assessment literature pertaining to bipolar disorder. The modification and use of the SADS-L as a reference standard for the assessment of ADHD is another limitation of this study. Although DSM-based, clinical interviews, in general, are considered to be preferred tools in assessment of adult ADHD (Adler & Cohen, 2004), there is little evidence supporting the use of this particular clinical interview, the SADS-L, as a diagnostic measure of ADHD.

Overall, the present results provide support for the diagnostic efficiency of the GBI as a method of discriminating between bipolar and other diagnoses in emerging adulthood. Given the prevalence of bipolar disorder in outpatient and community settings, low GBI scores may be decisive in excluding a bipolar diagnosis. High GBI scores would raise the posterior probability of a bipolar disorder into a moderate range, warranting additional systematic assessment (Straus, et al., 2011; Youngstrom, Jenkins, Jensen-Doss, & Youngstrom, 2012). Future research should evaluate the incremental information that family history (Tsuchiya, Byrne, & Mortensen, 2003), neurocognitive testing (Walshaw, et al., 2010), and other sources of information could contribute in tandem with the GBI in evidence based assessment of bipolar disorder.

Acknowledgments

This research was supported by NIMH grants MH52617 and MH77908 to Lauren B. Alloy.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

References

Alloy LB, Abramson LY, Walshaw PD, Cogswell A, Sylvia LG, Hughes ME, Iacoviello BM, Whitehouse WG, Urosevic S, Nusslock R, Hogan ME. Behavioral Approach System (BAS) and Behavioral Inhibition System (BIS) sensitivities and bipolar spectrum disorders: Prospective prediction of bipolar mood episoes. Bipolar Disorders. 2008; 10:310–322. [PubMed: 18271911]

Alloy LB, Bender RE, Whitehouse WG, Wagner CA, Liu RT, Grant DA, Abramson LY. High behavioral approach system (BAS) sensitivity, reward responsiveness, and goal-striving predict first onset bipolar spectrum disorders: A prospective behavioral high-risk design. Journal of Abnormal Psychology. 2012a; 121

Alloy LB, Urosevic S, Abramson LY, Jager-Hyman S, Nusslock R, Whitehouse WG, Hogan ME. Progression along the bipolar spectrum: A longitudinal study of predictors of conversion from bipolar spectrum conditions to bipolar I and II disorders. Journal of Abnormal Psychology. 2012b; 121:16–27. [PubMed: 21668080]

Angst J, Azorin JM, Bowden CL, Perugi G, Vieta E, Gamma A, Young AH. Prevalence and characteristics of undiagnosed bipolar disorders in patients with a major depressive episode: the BRIDGE study. Archives of General Psychiatry. 2011; 68:791–798. [PubMed: 21810644]

Barkley RA. Differential diagnosis of adults with ADHD: the role of executive function and self-regulation. J Clin Psychiatry. 2010; 71:e17. [PubMed: 20667287]

Barkley RA, Knouse LE, Murphy KR. Correspondence and disparity in the self- and other ratings of current and childhood ADHD symptoms and impairment in adults with ADHD. Psychological Assessment. 2011; 23:437–446. [PubMed: 21381833]

Campbell DT, Fiske DW. Convergent and discriminant validation by multitrait-multimethod matrix. Psychological Bulletin. 1959; 56:81–105. [PubMed: 13634291]

Carlson GA. The bottom line. Journal of Child and Adolescent Psychopharmacology. 2003; 13:115–118. [PubMed: 12886908]

Correll CU. Assessing and maximizing the safety and tolerability of antipsychotics used in the treatment of children and adolescents. Journal of Clinical Psychiatry. 2008; 69(Suppl 4):26–36. [PubMed: 18533766]

Danielson CK, Youngstrom EA, Findling RL, Calabrese JR. Discriminative validity of the General Behavior Inventory using youth report. Journal of Abnormal Child Psychology. 2003; 31:29–39. [PubMed: 12597697]

DelBello MP, Soutullo CA, Hendricks W, Niemeier RT, McElroy SL, Strakowski SM. Prior stimulant treatment in adolescents with bipolar disorder: association with age at onset. Bipolar Disorders. 2001; 3:53–57. [PubMed: 11333062]

Ebesutani C, Bernstein A, Chorpita BF, Weisz JR. A transportable assessment protocol for prescribing youth psychosocial treatments in real-world settings: reducing assessment burden via self-report scales. Psychological Assessment. 2012; 24:141–155. [PubMed: 21859220]

Frazier TW, Youngstrom EA. Evidence-Based Assessment of Attention-Deficit/Hyperactivity Disorder: Using Multiple Sources of Information. Journal of the American Academy of Child & Adolescent Psychiatry. 2006; 45:614–620. [PubMed: 16670656]

Galanter CA, Leibenluft E. Frontiers between attention deficit hyperactivity disorder and bipolar disorder. Child and Adolescent Psychiatric Clinics of North America. 2008; 17:325–346. [PubMed: 18295149]

Gallagher, D. Assessment of depression by interview methods and psychiatric rating scales. In: Poon, L., editor. Handbook of clinical memory assessment of older adults. Washington, DC: American Psychological Association; 1986.

Geller B, Zimerman B, Williams M, Bolhofner K, Craney JL. Bipolar disorder at prospective follow-up of adults who had prepubertal major depressive disorder. The American Journal of Psychiatry. 2001; 158:125–127. [PubMed: 11136645]

Guyatt, GH.; Rennie, D., editors. Users' Guides to the Medical Literature. Chicago: AMA Press; 2002.

Hanley JA, McNeil BJ. A method of comparing the areas under receiver operator characteristic curves derived from the same cases. Radiology. 1983; 148:839–843. [PubMed: 6878708]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Hasselbad V, Hedges LV. Meta-analysis of screening and diagnostic tests. Psychological Bulletin. 1995; 117:167–178. [PubMed: 7870860]

Hirschfeld RM, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. Journal of Clinical Psychiatry. 2003; 64:161–174. [PubMed: 12633125]

Jaeschke R, Guyatt GH, Sackett DL. 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? Journal of the American Medical Association. 1994; 271:703–707. [PubMed: 8309035]

Jenkins MM, Youngstrom EA, Washburn JJ, Youngstrom JK. Evidence-based strategies improve assessment of pediatric bipolar disorder by community practitioners. Professional Psychology: Research and Practice. 2011; 42:121–129.

Johnson, SL.; Miller, CJ.; Eisner, L. Bipolar Disorder. In: Hunsley, J.; Mash, EJ., editors. A Guide to Assessments That Work. New York: Oxford University Press; 2008. p. 121-137.

Kowatch RA, Youngstrom EA, Danielyan A, Findling RL. Review and meta-analysis of the phenomenology and clinical characteristics of mania in children and adolescents. Bipolar Disorders. 2005; 7:483–496. [PubMed: 16403174]

Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet. 2006; 367:1747–1757. [PubMed: 16731270]

Merikangas KR, Pato M. Recent developments in the epidemiology of bipolar disorder in adults and children: Magnitude, correlates, and future directions. Clinical Psychology: Science and Practice. 2009; 16:121–133.

Mick E, Biederman J, Pandina G, Faraone SV. A preliminary meta-analysis of the child behavior checklist in pediatric bipolar disorder. Biological Psychiatry. 2003; 53:1021–1027. [PubMed: 12788247]

Pini S, Dell'Osso L, Amador XF. Insight into illness in schizophrenia, schizoaffective disorder, and mood disorders with psychotic features. American Journal of Psychiatry. 2001; 158:122–125. [PubMed: 11136644]

Rettew DC, Lynch AD, Achenbach TM, Dumenci L, Ivanova MY. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. International Journal of Methods in Psychiatric Research. 2009; 18:169–184. [PubMed: 19701924]

Sala R, Axelson D, Birmaher B. Phenomenology, longitudinal course, and outcome of children and adolescents with bipolar spectrum disorders. Child Adolesc Psychiatr Clin N Am. 2009; 18:273–289. vii. [PubMed: 19264264]

Scheffer RE, Kowatch RA, Carmody T, Rush AJ. Randomized, placebo-controlled trial of mixed amphetamine salts for symptoms of comorbid ADHD in pediatric bipolar disorder after mood stabilization with divalproex sodium. American Journal of Psychiatry. 2005; 162:58–64. [PubMed: 15625202]

Stewart AJ, Theodore-Oklota C, Hadley W, Brown LK, Donenberg G, Diclemente R. Project Style Study, G. Mania Symptoms and HIV-Risk Behavior Among Adolescents in Mental Health Treatment. Journal of Clinical Child & Adolescent Psychology. 2012

Straus, SE.; Glasziou, P.; Richardson, WS.; Haynes, RB. Evidence-based medicine: How to practice and teach EBM. 4th ed.. New York: Churchill Livingstone; 2011.

Swets JA, Dawes RM, Monahan J. Psychological science can improve diagnostic decisions. Psychological Science in the Public Interest. 2000; 1:1–26.

Tsuchiya KJ, Byrne M, Mortensen PB. Risk factors in relation to an emergence of bipolar disorder: A systematic review. Bipolar Disorders. 2003; 5:231–242. [PubMed: 12895201]

Walshaw PD, Alloy LB, Sabb FW. Executive function in pediatric bipolar disorder and attention-deficit hyperactivity disorder: in search of distinct phenotypic profiles. Neuropsychology Review. 2010; 20:103–120. [PubMed: 20165924]

Wender PH. Attention-deficit hyperactivity disorder in adults. Psychiatric Clinics of North America. 1998; 21:761–774. [PubMed: 9890120]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Yatham LN, Kennedy SH, O'Donovan C, Parikh S, MacQueen G, McIntyre R, Gorman CP. Canadian Network for Mood and Anxiety Treatments (CANMAT) guidelines for the management of patients with bipolar disorder: consensus and controversies. Bipolar Disorders. 2005; 7(Suppl 3): 5–69. [PubMed: 15952957]

Youngstrom EA. Future directions in psychological assessment: Combining Evidence-Based Medicine innovations with psychology's historical strengths to enhance utility. Journal of Clinical Child & Adolescent Psychology. 2013; 42:139–159. [PubMed: 23153181]

Youngstrom EA, Arnold LE, Frazier TW. Bipolar and ADHD Comorbidity: Both Artifact and Outgrowth of Shared Mechanisms. Clinical Psychology: Science & Practice. 2010; 17:350–359. [PubMed: 21278822]

Youngstrom EA, Findling RL, Calabrese JR. Effects of adolescent manic symptoms on agreement between youth, parent, and teacher ratings of behavior problems. Journal of Affective Disorders. 2004; 82:S5–S16. [PubMed: 15571790]

Youngstrom EA, Findling RL, Calabrese JR, Gracious BL, Demeter C, DelPorto Bedoya D, Price M. Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5 to 17 years. Journal of the American Academy of Child & Adolescent Psychiatry. 2004; 43:847–858. [PubMed: 15213586]

Youngstrom EA, Frazier TW, Findling RL, Calabrese JR. Developing a ten item short form of the Parent General Behavior Inventory to assess for juvenile mania and hypomania. Journal of Clinical Psychiatry. 2008; 69:831–839. [PubMed: 18452343]

Youngstrom EA, Freeman AJ, Jenkins MM. The assessment of children and adolescents with bipolar disorder. Child and Adolescent Psychiatric Clinics of North America. 2009; 18:353–390. [PubMed: 19264268]

Youngstrom EA, Jenkins MM, Jensen-Doss A, Youngstrom JK. Evidence-Based Assessment Strategies for Pediatric Bipolar Disorder. Israel Journal of Psychiatry & Related Sciences. 2012; 49:15–27. [PubMed: 22652926]

Youngstrom EA, Meyers OI, Youngstrom JK, Calabrese JR, Findling RL. Comparing the effects of sampling designs on the diagnostic accuracy of eight promising screening algorithms for pediatric bipolar disorder. Biological Psychiatry. 2006; 60:1013–1019. [PubMed: 17056395]

Zhou, X-H.; Obuchowski, NA.; McClish, DK. Statistical methods in diagnostic medicine. New York: Wiley; 2002.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

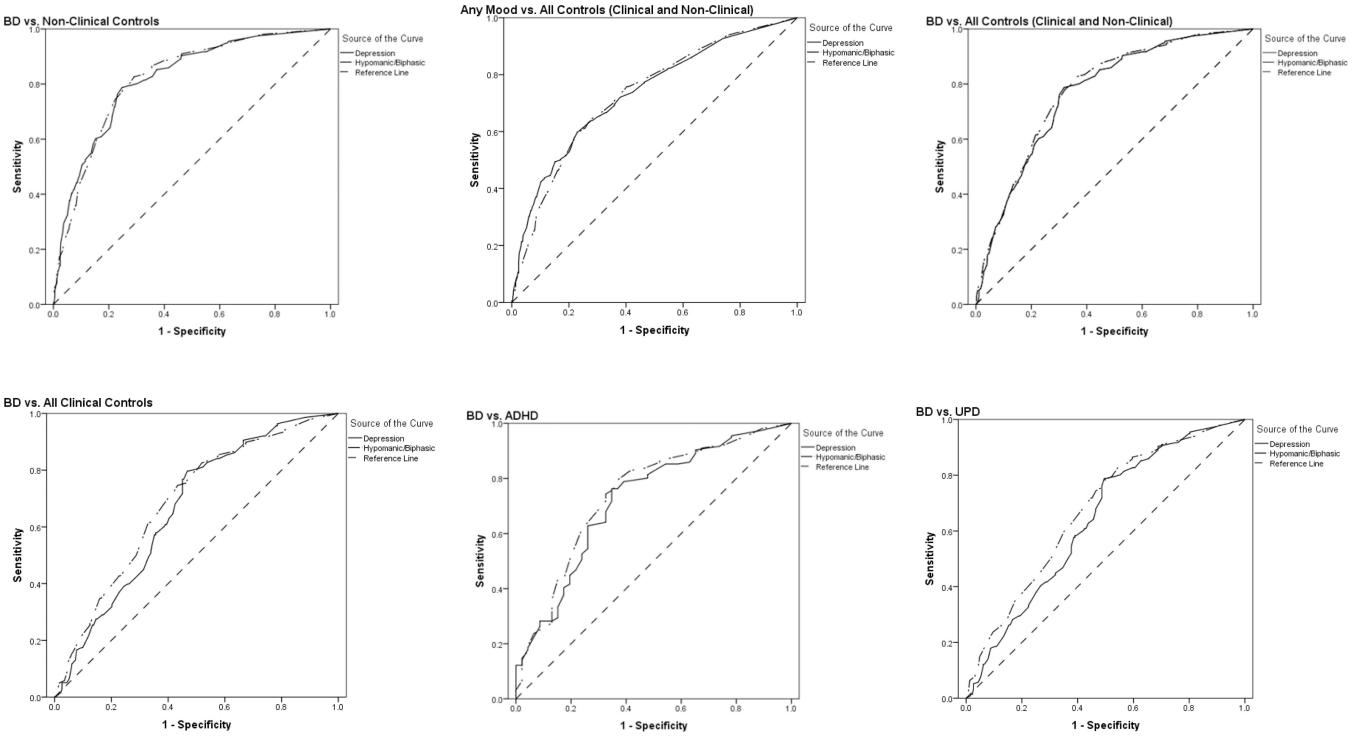

Figure 1.
ROC Curves for Analyses of GBI Subscales.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Demographics of Samples for Logistic Regression and ROC Analyses

	Adolescent Subsample Ages 14–19	Young Adult Subsample Ages 18–24
Race		
Black	85	150
White	208	320
Asian	35	40
Multi-Racial	6	80
Native Amer.	4	0
Other	21	24
Sex		
Male	109	409
Female	250	204
Mean Age (SD)	18.35 (1.52)	19.63 (1.76)
Diagnosis		
ADHD	33	30
BD	37	119
UPD	68	160
No Dx	233	327
Total	359	614

Note. A total of 17 participants met criteria for both BD and ADHD.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 2

Logistic Regression and Receiver Operating Characteristics/Area Under the Curve Analyses of Diagnostic Differentiation using GBI Depression and Hypomanic/Biphasic Subscale Scores and Descriptive Statistics

	Logistic Regression Weights			ROC/AUC Findings				AUC[D]
Comparison	**Depression**	**Hypomanic /Biphasic**	**Constant**	**Depression**		**Hypomanc/ Biphasic**		**z**
				AUC	**SE**	**AUC**	**SE**	
BD (156) vs. Nonclinical Controls (563)	.04**	.11***	−2.94***	.82***	.02	.82***	.02	0
Any Mood (384) vs. All Controls[1] (589)	.05***	.05*	−1.41***	.73***	.02	.73***	.02	0
BD (156) vs. All[1] (817)	.01	.12***	−3.11***	.77***	.02	.78***	.02	−.84
BD (156) vs. Clinical Controls[2] (254)	−.01	.12***	−1.65***	.66***	.03	.69***	.03	−1.22
BD (156) vs. UPD (228)	−.02	.11***	−1.47***	.64***	.03	.67***	.02	−1.97*
BD (156) vs. ADHD (46)	.02	.12**	−.29	.72***	.04	.74***	.04	−.65

	Descriptive Statistics by Dx				
	Depression		**Hypomanic/ Biphasic**		***r***
	M	***SD***	***M***	***SD***	
Nonclinical Controls	7.24	9.43	5.66	9.43	.88
BD	20.76	12.01	14.22	6.65	.76
ADHD	14.45	12.80	10.36	7.83	.85
UPD	14.81	13.09	9.64	7.22	.85
All Participants	11.20	11.96	7.60	7.32	.88

Note.

* denotes $p < .05$,

** denotes $p < .01$,

*** denotes $p < .001$.

Sample sizes are noted in parentheses.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

[1] refers to all non-target cases (with and without diagnoses).

[2] refers to diagnoses of either UPD or ADHD. Also, comorbid-ities were allowed in BD, UPD, and ADHD groups. AUC^D refers to use of the Hanley and McNeil (1983) method for testing for statistically significant differences between ROC curves. *r* refers to the correlation between factors.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Likelihood Ratios for Scores on GBI Hypomanic/Biphasic Subscale

Combined Sample			
	Range		
Comparisons	**Low 0–5**	**Moderate 6–19**	**High 20+**
Screening			
BD (*n* = 156) vs. Nonclinical Controls (*n* = 565)			
DLR	.09	4.21	6.49
CI	.06–.15	2.89–6.14	3.57–11.80
Any Mood (*n* = 384) vs. All Controls (*n* = 589)			
DLR	.24	2.90	3.87
CI	.18–.34	2.20–3.78	2.25–6.66
Differential Diagnosis			
BD (*n* = 156) vs. All Controls (*n* = 817)			
LR	.12	3.24	5.17
CI	.08–.20	2.26–4.25	3.07–8.67
BD (*n* = 156) vs. Clinical Controls (*n* = 252)			
DLR	.24	1.91	3.54
CI	.14–.41	1.26–2.90	1.86–6.74
BD (*n* = 156) vs. UPD (*n* = 228)			
DLR	.30	1.80	3.16
CI	.16–.57	1.18–2.75	1.65–6.02
BD (*n* = 156) vs. ADHD (*n* = 63)			
DLR	.18	2.18	5.24
CI	.09–.37	1.12–4.25	1.20–22.83

Note. Because there were no significant differences in diagnostic efficiency of the GBI when cross-validating in the second sample, both samples were pooled to provide more precise and stable estimates (Kraemer, 1992). DLR = Diagnostic Likelihood Ratio; the multiplicative change in the odds of the target diagnosis given the range in which the observed test score fell. CI = 95% confidence interval around the diagnostic likelihood ratio.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Likelihood Ratios for Two-Step Classification Process

Combined Sample			
Range for Depressive Subscale			
Comparisons	Low 0–2	Moderate 3–14	High 15+
Step 1. Any Mood (*n* = 384) vs. All Controls (*n* = 589)			
DLR	.14	.78	4.75
CI	.08–.25	.53–1.14	3.31–6.82
Range for Hypomanic/Biphasic Subscale			
Comparisons	Low 0–5	Moderate 6–19	High 20+
Step 2. BD (*n* = 156) vs. UPD (*n* = 228)			
DLR	.30	1.80	3.16
CI	.16–.57	1.18–2.75	1.65–6.02

Note. DLR = Diagnostic Likelihood Ratio; the multiplicative change in the odds of the target diagnosis given the range in which the observed test score fell. CI = 95% confidence interval around the diagnostic likelihood ratio.