Utility of the CBCL DSM-Oriented Scales in Assessing Emotional Disorders in Youth with Autism

Caroline I. Magyar[a] and **Vincent Pandolfi**[b]

[a]Caroline I. Magyar, Ph.D., BCBA, Associate Professor, University of Rochester School of Medicine, 601 Elmwood Ave, Box 671, Rochester, NY 14642, United States

[b]Vincent Pandolfi, Associate Professor, Rochester Institute of Technology, Psychology Department, 18 Lomb Memorial Drive, Rochester, N.Y. 14623-5604, United States

Abstract

Background—Youth with autism spectrum disorder (ASD) are at risk for one or more emotional disorders (ED) including depressive and anxiety conditions. DSM-5 diagnostic guidelines indicate that co-occurring ED must be specified when present (APA, 2013). While ED may be evaluated for during initial diagnostic assessment, routine monitoring and screening is needed to identify emerging ED in later childhood and adolescence, a period of high risk.

Method—Confirmatory factor analysis, convergent and divergent validity analyses, criterion-related validity, and diagnostic accuracy analyses of the CBCL's Affective Problems and Anxiety Problems DSM Oriented Scales was completed on 93 well-characterized youth, ages 6 to 18 years with ASD (6:1 M:F), with and without intellectual impairment. These youth were from predominately white, middle-class backgrounds.

Results—Each scale measured a single construct reliably (depressive and anxiety disorders), neither scale measured symptoms of ASD, and youth with a depressive disorder had other ED co-morbidities.

Conclusions—Findings demonstrate the DSM Oriented Affective and Anxiety Problem Scales can be used to screen for depression and anxiety in youth with ASD. Replication is needed with various subgroups representing gender, age, developmental level, autism, and mental health severity differences, and with groups across a broader set of demographics.

Keywords

Autism; Assessment; Anxiety; Depression; CBCL; Emotional Disorders; DSM

Depression and anxiety are two emotional disorders (ED) reported to affect youth with autism spectrum disorder (ASD) at a high rate (American Psychiatric Association, 2013). The most commonly diagnosed anxiety disorders include Separation Anxiety Disorder, Specific Phobia, Social Anxiety Disorder (e.g., see van Steensel, Bogels, & Perrin, 2011 for a review), Generalized Anxiety Disorder, and related Obsessive Compulsive Disorder. Major Depression is reported more often than bi-polar disorder (e.g., Joshi et al, 2013; Weissman & Bates, 2010). Prevalence estimates vary, but the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (DSM-5) reports about 70% of individuals with ASD are likely to have one and 40% two or more mental health conditions (APA, 2013, pp. 58) indicating many present with complex symptom profiles.

DSM-5 diagnostic guidelines for ASD indicate that co-occurring ED must be specified when present (APA, 2013). This implies that practitioners should be assessing for ED conditions in individuals with ASD. While this may occur upon initial diagnostic assessment routine monitoring for ED through formal screening is important for early identification (see Magyar & Pandolfi, 2012; Pandolfi & Magyar, 2016; Pandolfi, Magyar, & Dill, 2012; and Pandolfi, Magyar, & Norris, 2014). The use of third party report when used as part of a multimethod/informant approach is critical for a couple of reasons (e.g., see Tarbox, La Cava, & Hoang for a discussion on considerations in ASD assessment). One, self-report data obtained through routine clinical interview methods may not be reliable for many youth with ASD due to language and cognitive impairments associated with their ASD. Two, ED symptom presentation may be moderated by neurodevelopmental impairments (e.g., intellectual disability; see Stewart, Barnard, Pearson, Hasan, & O'Brien, 2006), and the social communication impairments and restricted, repetitive behaviors, interests or activities associated with ASD. Such might include restricted facial expression, deficits in nonverbal communicative behavior such as gestures, use of idiosyncratic phrases to describe personal phenomena, and hypo/hyper sensitivity to environmental stimuli that can result in disruptive behaviors. These issues challenge practitioners to accurately identify ED-specific problems that may co-occur with ASD. For example, an adolescent with ASD who exhibits a flat and restricted facial expression that is not directed toward the examiner, and whose communication offers little in the way of detail and is not integrated with emotional or empathic gestures may be thought of as depressed, when in fact these behaviors are symptoms of his/her ASD.

While more work is needed to identify ED risk factors and correlates for youth with ASD and to determine the extent to which these might differ from those identified for youth in general (see Pandolfi & Magyar, 2016 for a discussion), older children and adolescents with ASD appear to be at particular risk for developing ED. For those affected by ED, the emergence of symptoms is common throughout this period of development (Gotham, Brunwasser, & Lord, 2015). Thus, any changes in baseline behavior and functioning during this time period should be noted and youth should be screened to determine if the observed changes represent developmental changes, symptoms of an emerging ED, or both. Screening results can provide the data required to inform the need for further assessment and the selection of methods and measures when a differential diagnostic assessment is appropriate.

The Child Behavior Checklist 6–18 (CBCL; Achenbach & Rescorla (2001) is one widely used screening measure for emotional and behavioral problems (EBP) and was recently evaluated for its potential use in assessing youth with ASD (Pandolfi, Magyar, & Dill, 2012; Pandolfi, Magyar, & Norris, 2014). This norm-referenced caregiver completed questionnaire screens for a wide range of EBP. The CBCL contains two empirically-derived (i.e., through factor analysis) broadband scales representing internalizing and externalizing problems, eight empirically-derived syndrome scales representing different patterns of co-occurring EBP, and six DSM-Oriented scales derived through expert consensus. The syndrome scales reflect sets of co-occurring problems and each scale does not map onto a DSM diagnostic category. Each DSM Oriented scale reflects a broad emotional or behavioral problem that corresponds to a broad DSM diagnostic category (Achenbach & Rescorla, 2001). This organizational structure makes the CBCL an attractive measure to use in screening and diagnostic assessment protocols.

In studies of youth with ASD, the CBCL's syndrome scales demonstrated good reliability, validity, and diagnostic accuracy (see Pandolfi, Magyar, & Dill, 2012; Pandolfi, Magyar, & Norris, 2014). Yet, for those youth with more complex clinical presentations (i.e., developmental changes in their ASD symptom profile, presence of one or more ED with/ without behavior problems), the interpretation of syndrome scale scores alone might be insufficient to assist the practitioner in making decisions about targeted screening for specific categories of ED. This is particularly so for youth who screen positive on many of the syndrome scales. For example, clinically significant scores on *Anxious/Depressed*, *Withdrawn/Depressed*, and *Somatic Complaints* might suggest significant negative affect; however, it is difficult to determine the relative contribution of a depressive or anxiety disorder to the youth's functional impairment and distress. Thus, elevations on one or more scales may not sufficiently narrow the range of diagnostic possibilities and suggest the need for more targeted screening. Practitioners, therefore, are likely to benefit from targeted screening data that can assist in decision making about the need for more comprehensive diagnostic assessment.

According to the test authors, the CBCL's DSM-Oriented *Affective Problems* and *Anxiety Problems* scales align with disorders listed in the DSM's sections containing the depressive and anxiety disorders, respectively. The specific disorders screened by *Affective* and *Anxiety Problems* are among the EDs most often observed in youth with ASD. The current study sought to evaluate the reliability and validity of the CBCL's DSM Oriented *Affective Problems* and *Anxiety Problems* scales for assessment of depression and anxiety in youth with ASD.

Method

Participants

Parent and youth participant dyads provided the archival data analyzed in this study (*N*=93). Participant data were collected from a study that explored genotype-phenotype relationships in youth with ASD. Participants were recruited from central and western New York State. The study was approved by the host university's institutional review board and written informed consent was obtained from all participants.

Data from all 93 participants were used for the present study's confirmatory factor analyses (CFA) and they ranged in age from six to 18 years (M= 11 years, 7 months, SD= 3 years, 5 months). Data from a subset of these participants were used for the remaining analyses (N= 77; described below).

Expert clinical consensus regarding ASD diagnostic status was informed by each child participant's developmental history, and the administration of the Autism Diagnostic Interview-Revised (ADI-R; Rutter, LeCouteur, & Lord, 2003) and the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 2002). The evaluation team included a licensed psychologist and several trained nondoctoral evaluators. All had considerable experience evaluating youth with autism spectrum disorder and all personnel were trained to reliably administer the ADOS and ADI-R within a research context.

Characteristics for the participants are presented in Table 1.

Males outnumbered females by a 6 to 1 ratio. The participants were generally white, non-Hispanic, from middle to upper socioeconomic status backgrounds, and had parents that were well educated. Cognitive data indicated that 12.4% of participants earned IQ scores less than 70. The participants were administered either the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003), Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997), Stanford-Binet-Fifth Edition (Roid, 2003), or the Differential Ability Scales (DAS; Elliott, 1990). Most participants fell into the low range of adaptive functioning as indicated by the Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchietti, 1984).

Because K-SADS (Kaufman et al., 1997) data could not be collected on 16 participants (failed to complete entire interview), a total of 77 participants provided all of the psychiatric data that were needed for the study's other analyses (mean age= 11 years, 11 months; SD= 3 years, 3 months). Table 2 presents additional information on the characteristics of these participants.

Most of the participants were male, had parents that were generally well educated, and were white and non-Hispanic. Most had IQ scores above 70 and Vineland Adaptive Behavior Scale scores of 70 or less.

The psychiatric assessments included a parent-completed medical history form, administration of the Child Behavior Checklist 6–18 and administration of the K-SADS screener with the parent as informant. Direct observation and a semi-structured interview were completed with the youth participant. A multi-method/multi-informant approach to diagnostic determination was made, which has been recommended previously (e.g., see Pandolfi & Magyar, 2012). This included data from the parent responses to the psychiatric questions on the medical history form asking about the presence of emotional/mental and behavioral disorders in their child (yes/no/not sure), the profile of scores from the CBCL Syndrome Scales (not the DSM Oriented Scales), the results of the K-SADS screening with the parent as respondent, results of the evaluator's semi-structured child interview data that inquired about a range of emotions and their correlates (e.g., Do you ever feel sad? What kinds of things make you feel sad?), and the evaluator„s observational data on the youth

participant's mental health status, The assessments were completed by the evaluation team (described above) and all diagnostic decisions were made by and/or reviewed by the licensed clinical psychologist from the team.

With respect to DSM diagnosis 19.5% were diagnosed with a depressive disorder (Major Depressive Disorder or Dysthymic Disorder). All of these youth exhibited at least one additional co-occurring disorder with anxiety disorder (73.3%), ADHD (53.3%), ODD (20.0%), and tic disorder (20.0%) being the most common, and 53.3% of those with a depressive disorder were diagnosed with two or more additional disorders. A total of 45.5% of participants were diagnosed with an anxiety disorder. For children with an anxiety disorder 71.4% were diagnosed with a co-occurring disorder and 28.6% with two or more DSM disorders. Of those diagnosed with a co-occurring disorder, 37.1% were diagnosed with ADHD, 31.4% with a depressive disorder, and 20% with a tic disorder.

Measures

CBCL 6–18—The CBCL is a Likert-type norm-referenced caregiver completed rating scale that describes a child's functioning during the previous six months (see Achenbach & Rescorla, 2001). All items are scored on a three point Likert scale (0= "Not True," 1= "Somewhat or Sometimes True," or 2= "Very True or Often True). All CBCL scales have a T-score mean of 50 and standard deviation of 10 and different norms are provided for each gender across the 6–11 and 12–18 year age ranges. The six norm-referenced DSM-Oriented Scales include *Affective Problems*, *Anxiety Problems*, *Somatic Problems*, *Attention Deficit/ Hyperactivity Problems*, *Oppositional Defiant Problems*, and *Conduct Problems*. These scales were created based on expert consensus and were developed to assist practitioners in the differential diagnostic process. The *Affective Problems* scale assesses for symptoms of Major Depressive Disorder and Dysthymic Disorder and the *Anxiety Problems* scale assesses for symptoms of Separation Anxiety Disorder, Specific Phobia, and Generalized Anxiety Disorder. There is substantial psychometric support for the various CBCL scales (see Achenbach & Rescorla, 2001; Berub & Achenbach, 2010).

K-SADS—The K-SADS is a semi-structured diagnostic interview based on the DSM-III-R (APA, 1987) and DSM-IV (APA, 1994). It provides severity ratings for children and adolescents across 20 diagnostic areas (Ambrosini, 2000; Kaufman, et al., 1996). The measure contains a screening interview and supplemental diagnostic interviews. Reliability and validity data are generally favorable (see Ambrosini, 2000; Kaufman et al, 1997) and the K-SADS has been used in published studies of youth with ASD (see Gjevik et al., 2011; Leyfer et al., 2006).

The K-SADS data analyzed were collected from items belonging to the depression and anxiety screening interviews. K-SADS interview items are scored in the following way: 0 = "No Information Available," 1 = "Not present," 2 = "Sub-threshold," and 3 = "Threshold." The K-SADS depression items were consistent with the kinds of problems assessed by the DSM Oriented *Affective Problems* scale. The anxiety interview items analyzed in this study were those developed to assess for *Separation Anxiety Disorder*, *Agoraphobia/Specific Phobia*, and *Overanxious Disorder/Generalized Anxiety Disorder*, which is consistent with

the kinds of problems assessed by the DSM Oriented *Anxiety Problems* scale. We created two K-SADS scores from these screening interviews: a K-SADS Depression score and a K-SADS Anxiety score. The Depression score reflected the sum of item scores on the depression interview scale. The Anxiety score reflected the sum of item scores across the *Separation Anxiety*, *Agoraphobia/Specific Phobia*, and *Overanxious Disorder/Generalized Anxiety Disorder* interview scales.

Autism Diagnostic Interview-Revised (ADI-R; Rutter, LeCouteur, & Lord, 2003) —This diagnostic interview is used to assess individuals suspected of ASD. Caregivers familiar with the individual's developmental history and present behavior serve as respondents. The ADI-R assesses for the presence of ASD across three subdomains: Language/Communication, Reciprocal Social Interaction, and Restrictive, Repetitive, and Stereotyped Behaviors and Interests. A current behavior algorithm allows for an assessment of the individual's current symptom profile. This algorithm was used for the present analyses because scores reflect functioning during a timeframe that is consistent with the CBCL. The current algorithm of the ADI-R was preferred to the ADOS because different ADOS modules were used with the participants based on their functional levels. Without an assurance of psychometric equivalence of scores across modules, aggregating data across the different modules might have biased the analyses. Use of the ADI-R current algorithm also allowed for multiple methods of data collection with the same informant (interview, paper and pencil CBCL scale) which helped decrease the likelihood of method variance contributing to the correlations between the DSM Oriented Scales and the ADI-R.

Vineland Adaptive Behavior Scales (Vineland; Sparrow et al., 1984)—This standardized norm-referenced measure of adaptive behavior is appropriate for use with individuals aged birth to 90 years. The Survey/Interview Form was used to describe the participants' level of functioning. The Adaptive Behavior Composite indicates overall level of functioning and it also provides scores for specific adaptive behavior domains: Communication, Daily Living Skills, Socialization, and Motor Skills (ages birth to 6 years, only). The Adaptive Behavior Composite and four subdomains have a mean of 100 and standard deviation of 15. The technical manuals provide evidence of its reliability and validity.

Cognitive Measures—Different intelligence tests were used based on participant age and developmental levels. These included the WISC-IV, WAIS-III, Stanford-Binet-Fifth Edition and the DAS. All of these tests are norm-referenced with a mean of 100 and standard deviation of 15.

Data Analysis

Several statistical analyses were used to help understand a wide range of measurement properties for the DSM Oriented *Affective Problems* and *Anxiety Problems* scales. Because data analyses were performed on archival data, a priori power analyses were not possible. Statistical significance and effect sizes (proportions of variance, standardized mean differences) were obtained. Confirmatory factor analysis was used on dichotomized item-level data to assess the extent to which each scale measured a single construct using LISREL

8.80 (Joreskog & Sorbem, 2006). In addition, scale reliability for *Affective Problems* and *Anxiety Problems* was also determined using CFA parameters (factor loadings and error variances). The remaining analyses were conducted using SPSS (SPSS, Inc., 2012) and the *Affective* and *Anxiety Problems* scale scores reflected the sum of their respective item scores (scored 0, 1, or 2). Convergent and discriminant correlations were computed among *Affective Problems* and *Anxiety Problems* scales, the remaining DSM Oriented Scales, the KSADS Depression and Anxiety scales, and the current algorithm of the ADI-R representing levels of ASD symptoms within the three months preceding the evaluation. Criterion-related validity was assessed by independent *t*-tests which compared the *Affective* and *Anxiety Problems* scores obtained by those with a diagnosis of a depressive or anxiety disorder against those without the diagnosis. Finally, diagnostic accuracy was evaluated for each scale using ROC analyses.

Results

Confirmatory Factor Analysis

The extent to which *Affective Problems* and *Anxiety Problems* each measured a single construct was evaluated by CFA. Two statistics evaluated the adequacy of two single-factor models (presumably *Affective Problems* and *Anxiety Problems*): (a) the Root Mean Square Error of Approximation (*RMSEA*; Steiger & Lind, 1980), with values < .10 indicating support for the model (MacCallum, Browne, & Sugawara, 1996), and (b) the Comparative Fit Index (*CFI*; Bentler, 1990), with values close to .95 indicating adequate fit (Hu & Bentler, 1999). Initial CFA results suggested two correlated errors for *Affective Problems* (i.e., *overtired without good reason* with *underactive, slow moving, or lacks energy and trouble sleeping: describe*, and *sleeps less than most kids*). The modeling of correlated errors is appropriate in instances when items within a measurement scale contain similar wording, have overlapping item content or seem to ask for redundant information (see Brown, 2006). The CFA on this *Affective Problems* model indicated good fit (*RMSEA*= .06, *CFI*= .98). With the exception of one item (i.e., *sleeps less than most kids*) all factor loadings were statistically significant ($a < .05$) The median loading of .59 indicated that, on average, 35% of an item's variance was accounted for the latent factor (affective problems).CFA results for *Anxiety Problems* (*RMSEA*= .03, *CFI*= 1.00) indicated that this scale measured one construct. All factor loadings on *Anxiety Problems* were statistically significant ($a < .05$) and the median loading of .63 indicated that, on average, 40% of an item's variance was accounted for by the latent factor (anxiety problems). CFA parameters were used to compute scale reliabilities and results reflect the proportion of true score variance measured by the scale (see Brown, 2006; Raykov, 1997; 2001). Results for *Affective Problems* (.86) and *Anxiety Problems* (.81) indicated acceptable scale reliabilities for screening measures. Thus, CFA results indicated that each scale measured a single construct with good reliability.

Correlations With Other Variables

Affective Problems and *Anxiety Problems* scores were correlated with several other measures (*N*= 77), whose reliability was evaluated with Guttman's Lambda-2 (Guttman, 1945). Lambda-2 is preferred to coefficient alpha because of its less restrictive assumptions (see Green & Yang, 2009; Sijtsma, 2009). The K-SADS Depression (.79), K-SADS Anxiety

(.80), and ADI-R (.72) all demonstrated acceptable scale reliability. Correlational analyses were then performed using the bootstrap method in AMOS 16 (Arbuckle, 2007)[1] because the Pearson r normality assumption did not hold. Here, 10,000 samples of size N=77 were drawn with replacement and we derived bias-corrected confidence intervals for correlations between the DSM Oriented *Affective Problems* and *Anxiety Problems* scales and several other measures. Results are presented in Table 3.

Affective Problems showed a statistically significant and moderately strong correlation with the interview-based K-SADS Depression measure (r= .59, p < .001), and *Anxiety Problems* showed a similar relationship with the K-SADS Anxiety measure (r= .66, p < .001). Each scale also showed statistically significant but more modest correlations with all remaining K-SADS and DSM Oriented scales (see Table 3). *Affective Problems* and *Anxiety Problems* correlated .57 with one another (p < .001). Table 3 also shows that CBCL *Affective Problems* correlated significantly with K-SADS Anxiety (.52, <.001) and CBCL *Anxiety Problems* correlated significantly with K-SADS Depression (.43, p<.001). Neither scale correlated significantly with the ADI-R current algorithm. Finally, the K-SADS Depression and Anxiety measures correlated .43 *(p < .001)* and had a 95% bias-corrected confidence interval of .22 to .64 (not presented in the table).

Criterion-Related Validity

Independent t-tests were used to see if those diagnosed with a depressive or anxiety disorder differed from those without these disorders with respect to scores on *Affective Problems* and *Anxiety Problems*. Because of skewed data distributions, the analyses were run using bootstrapped regression analyses in AMOS 16: 10,000 bootstrap samples were used[2]. However, we found that the bootstrap results were not appreciably different than results of standard t-tests that were run on the sample data, so we report the results of the standard t-test. Four t-tests were conducted and the Dunn-Bonferroni correction was used to maintain the experimentwise error rate at a_{EW} = .05, so each individual t-test was evaluated for statistical significance at a_{DB} = .0125. Hedges g was used to estimate effect size (Hedges & Olkin, 1985).

The first set of comparisons involved those with a depressive disorder (n = 15) vs. those without (n = 62). On the *Affective Problems* scale, those with a depressive disorder (M = 10.47, SD = 3.46, CI_{95}= 8.56, 12.38) scored higher than those without (M = 4.77, SD = 3.77, CI_{95}= 3.81, 5.73) and this difference was statistically significant (t = 5.33, df = 75, p < .001) and rather large (g = 1.54). On the *Anxiety Problems* scale, those with a depressive disorder (M = 6.67, SD = 3.22, CI_{95}= 4.89, 8.45) scored higher than those without (M =

---

[1]Bootstrapping can be used when the normality assumption does not hold and raises the possibility of biased sample correlations. Here, 10,000 samples of size N=77 were drawn with replacement. The Pearson correlation was computed for each sample which results in an estimated sampling distribution of the statistic. The procedure allows for two things: a comparison between the sample statistic and a bootstrap-derived statistic, and the creation of bias-corrected confidence intervals. In this study, we report the sample correlations which did not differ appreciably from the bootstrap correlations (maximum difference was .008, modal difference was .002). We report the 95% bias-corrected confidence intervals, the results of which were consistent with the results of significance tests conducted on the sample correlations.

[2]The use of dummy coding resulted in regression coefficients that were equivalent in value to the mean between-group difference of interest.

3.55, $SD = 2.49$, $CI_{95} = 2.92$, 4.18) and this difference was also statistically significant ($t = 4.10$, $df = 75$, $p < .001$) and large ($g = 1.18$).

The second set of comparisons involved those with an anxiety disorder ($n = 35$) vs. those without ($n = 42$). On the *Anxiety Problems* scale, those with an anxiety disorder ($M = 5.26$, $SD = 2.84$, $CI_{95} = 4.29$, 6.23) scored higher than those without ($M = 3.24$, $SD = 2.66$, $CI_{95} = 2.41$, 4.07) and this difference was statistically significant ($t = 3.22$, $df = 75$, $p = .002$) and moderately large ($g = 0.76$). On the *Affective Problems* scale, those with an anxiety disorder ($M = 6.86$, $SD = 4.64$, $CI_{95} = 5.27$, 8.45) scored higher than those without ($M = 5.07$, $SD = 3.93$, $CI_{95} = 3.85$, 6.29) but this difference was not statistically significant ($t = 1.83$, $df = 75$, $p = .071$; $g = 0.42$).

Diagnostic Accuracy

ROC analyses were used to determine how well the *Affective Problems* and *Anxiety Problems* scales identify those with and without a depressive and anxiety disorder, respectively. Table 4 displays the results.

Affective Problems evidenced good overall diagnostic accuracy. With a raw score cut-off of 6.5, sensitivity (.93) was excellent and specificity was good (.74). *Anxiety Problems* demonstrated good overall accuracy; however, sensitivity was better than specificity. Increasing the cut-off score from 1.5 (sensitivity = .91, specificity = .31) to 2.5 decreased sensitivity (.83) and did not appreciably improve specificity (.52).

Discussion & Implications

General Findings

This study examined the reliability and validity of the CBCL's DSM Oriented *Affective Problems* and *Anxiety Problems* scales in a well-characterized sample of youth with ASD. The sample was notable in that the majority of youth presented with several co-occurring problems that included ED-specific problems in addition to ASD and other developmental impairments such as ID. Results demonstrated strong psychometric support for the scales. Each scale reliably measured a single construct and convergent and discriminant correlations provided evidence that both scales measured what they were designed to measure: depression and anxiety, respectively. Importantly, findings also indicated that neither scale correlated with the ADI-R current behavior algorithm indicating that the problems evaluated by the *Affective* and *Anxiety Problems* scales are not ASD-specific.

Criterion-related validity analyses showed the extent to which *Affective Problems* and *Anxiety Problems* discriminated between groups of youth with and without depressive and anxiety disorders. Results showed that the group of youth with a depressive disorder scored significantly higher on *Affective Problems* than those youth without a depressive disorder and they also scored high on the *Anxiety Problems* scale. The latter finding can probably be explained by the fact that all youth with a diagnosis of a depressive disorder also had one or more co-occurring psychiatric disorders, with 73% having a co-occurring anxiety disorder. Results also showed that the group of youth with an anxiety disorder scored significantly higher on the *Anxiety Problems* scale than youth without an anxiety disorder.

Diagnostic accuracy analyses showed that the *Affective Problems* scale was excellent at identifying depressive disorder in this sample of youth with ASD and showed good specificity. The *Anxiety Problems* scale demonstrated good ability to detect the presence of an anxiety disorder, but specificity was low. The reason for this finding is unclear; however, we offer some possible reasons. First, the six items comprising *Anxiety Problems* reflect a narrow sampling of anxiety disorder symptoms such as fears and worries, and the scale does not include items assessing behavioral avoidance. Second, one item, *dependent* does not appear anxiety-specific and many youth with ASD are dependent on others to help meet every day needs (e.g., activities of daily living, academic/vocational tasks, etc.). Indeed, most participants in this sample had Vineland Adaptive Behavior Scale scores in the low range. It is also possible that those youth who presented with more complex clinical presentations may have had other diagnosable problems in addition to anxiety that were deemed more impairing and in need of treatment. Some might argue that anxiety is part of ASD itself and might account for lack of specificity. While this psychometric study cannot resolve that issue the results suggest that the CBCL might help future research that aims to address that important conceptual problem.

Limitations and Future Directions

In addition to attempts to further explore diagnostic accuracy issues, replication of this study is needed using more diverse ASD samples including samples with more diverse demographics (I.e., race, ethnicity, socio-economic, geographic) and larger numbers of female participants. While the characteristics of the youth participants in this study represent a significant proportion of the larger ASD population, replication is needed with samples representing different developmental, adaptive, and psychopathology severity levels. Data analysis on more narrow age ranges would also be helpful such as on the two normative age ranges of the CBCL (6–11, and 12–18 years). Although the sample that provided the data for the present study might be typical of the complex clinical presentations observed in everyday settings, obtaining data from relatively less complex cases would also add to our understanding of the CBCL's measurement properties in youth with ASD. An example would be participants with more narrowly defined problems such as those with a single diagnosis of a depressive or anxiety disorder. This would help us to better understand issues related to diagnostic accuracy and how additional psychiatric and developmental problems (e.g., ID severity) might moderate CBCL scores and interpretation.

Future studies might also incorporate other measures and methods of assessment when evaluating the psychometric properties of the CBCL. Such work might correlate CBCL data with data obtained from anxiety- and depression-specific measures, rather than from more broad based measures such as the K-SADS. In this study, clinical judgment regarding psychiatric diagnosis was informed by all data obtained from developmental and psychological assessment measures including the CBCL's *empirically-derived scales*, interview data from parent and youth participant dyads, and observational data of youth participants. We note here that the DSM Oriented Scales were not used to inform diagnostic decision-making. However, future such studies might have evaluators be completely blind to any CBCL data when rendering a psychiatric diagnosis in order to avoid any issues of circularity.

Conclusions

Findings from this initial investigation extends work into the relative utility of the CBCL in the assessment of ED in youth with ASD. The CBCL's caregiver-completed paper and pencil measure is a cost-effective and cost-efficient way to screen for a broad array of EBP in youth with ASD. It may be particularly suitable for schools that apply a multi-tiered system of support and intervention for emotional and behavioral problems (e.g., see Magyar & Pandolfi, 2012 and Pandolfi & Magyar, 2016 for a discussion) and in clinic settings. The CBCL's eight syndrome scales cut across DSM diagnostic categories and alert the practitioner to possible ED and ED co-morbidities. The DSM Oriented Scale data can be used for targeted screening and to develop hypotheses about possible ED. For example, one might better understand elevations on the empirically-derived *Internalizing Domain* and *Anxious/Depressed* and *Withdrawn/Depressed* syndrome scales by analyzing results from the *Affective* and *Anxiety Problems* scales. These data can be evaluated within the context of the youth's ASD and other developmental features and verified through the differential diagnostic process using a multi-method/informant assessment approach to determine final diagnostic outcome. This helps ensure diagnostic accuracy and increases the likelihood of appropriate treatment recommendations.

Acknowledgments

References

Achenbach, TM., Rescorla, LA. Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2001.

Ambrosini PJ. Historical development and present status of the Schedule for Affective Disorders and Schizophrenia for school-age children (K-SADS). Journal of the American Academy of Child and Adolescent Psychiatry. 2000; 39(1):49–58. [PubMed: 10638067]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3. Washington, D.C.: Author; 1987.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, D.C.: Author; 1994.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, D.C.: Author; 2000.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5. Washington, D.C.: Author; 2013.

Arbuckle, JL. Amos (Version 16.0) [Computer Program]. Chicago: IBM SPSS; 2007.

Bentler PM. Comparative fit indexes in structural models. Psychological Bulletin. 1990; 107:238–246. [PubMed: 2320703]

Berubé, RL., Achenbach, TM. Bibliography of published studies using the Achenbach System of Empirically Based Assessment: 2006 Edition. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2010. Available online at www.ASEBA.org

Brown, TA. Confirmatory factor analysis for applied research. New York: Guilford Press; 2006.

Elliott, CD. Differential Ability Scales. Pearson; San Antonio, TX: 1990.

Gjevik E, Eldevik S, Fjaeran-Granum T, Sponheim E. Kiddie-SADS reveals high rates of DSM-IV disorders in children and adolescents with autism spectrum disorders. Journal of Autism and Developmental Disorders. 2011; 41:761–769. DOI: 10.1007/s10803-010-1095-7 [PubMed: 20824493]

Gotham K, Brunwasser SM, Lord C. Depressive and anxiety symptom trajectories from school age through young adulthood in samples with autism spectrum disorder and developmental delay. Journal of the American Academy of Child and Adolescent Psychiatry. 2015; 54:369–376.e3. [PubMed: 25901773]

Green SB, Yang Y. Commentary on coefficient alpha: A cautionary tale. Psychometrika. 2009; 74(1): 121–135.

Guttman L. A basis for analyzing test-retest reliability. Psychometrika. 1945; 10:255–282. [PubMed: 21007983]

Hedges, LV., Olkin, I. Statistical methods for meta-analysis. San Diego, CA: Academic Press; 1985.

Hollingshead, AA. Four-factor index of social status. Yale University; New Haven, CT: 1975. Unpublished manuscript

Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999; 6:1–55.

IBM SPSS Statistics 21. Chicago, IL: SPSS, Inc; 2012.

Jöreskog, K., Sörbom, D. LISREL 8.80. Lincolnwood, IL: Scientific Software International, Inc; 2006.

Joshi G, Biederman J, Petty C, Goldin R, Furtak SL, Wozniak J. Examining the comorbidity of bipolar disorder and autism spectrum disorder: A large controlled analysis if phenotypic and familial correlates in a referred population of youth with bipolar I disorder with and without autism spectrum disorder. Journal of Clinical Psychiatry. 2013; 74:578–586. DOI: 10.4088/JCP.12m07392 [PubMed: 23842009]

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Ryan, N. Kiddie-Sads-Present and Lifetime Version. 1996. Version 1.0 of October, 1996http://www.wpic.pitt.edu\ksads

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADSPL): Initial reliability and validity data. Journal of the American Academy of Child and Adolescent Psychiatry. 1997; 36(7):980–988. [PubMed: 9204677]

Kraijer, DW. Pervasive Developmental Disorder in Mental Retardation Scale Manual. 2. ITS B.V., Leiden: Psychologische Instrumenten Tests En Services; 2006.

Leyfer OT, Folstein SE, Bacalman S, Davis NO, Dinh E, Morgan J, Tager-Flusberg H, Lainhart JE. Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. Journal of Autism and Developmental Disorders. 2006; 36:849–861. [PubMed: 16845581]

Lord, C., Rutter, M., DiLavore, P., Risi, S. Autism Diagnostic Observation Schedule: Manual. Los Angeles: Western Psychological Services; 2002.

MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychological Methods. 1996; 1:130–149.

Magyar CI, Pandolfi V. Considerations for establishing a multi-tiered problem-solving model for students with autism and emotional and behavioral disorders. Psychology in the Schools. 2012; 49(10):975–987.

Pandolfi, V., Magyar, CI. Assessment of Co-Occurring Emotional and Behavioral Disorders in Youth with ASD using the Child Behavior Checklist 6–18. In: Patel, VB.Preedy, VR., Martin, CR., editors. The Comprehensive Guide to Autism. Springer-Verlag; Berlin Heidelberg: 2012.

Pandolfi, V., Magyar, CI. Psychopathology. In: Matson, JM., editor. Comorbid Conditions Among Children with Autism Spectrum Disorders. New York: Springer; 2016.

Pandolfi V, Magyar CI, Dill CA. An initial psychometric evaluation of the Child Behavior Checklist 6–18 in a sample of youth with autism spectrum disorders. Research in Autism Spectrum Disorders. 2012; 6:96–112. [PubMed: 22059091]

Pandolfi V, Magyar CI, Norris M. Validity study of the CBCL 6–18 for the assessment of emotional problems in youth with ASD. Journal of Mental Health Research in Intellectual Disabilities. 2014; 7(4):306–322. [PubMed: 25419257]

Raykov T. Estimation of composite reliability for congeneric measures. Applied Psychological Measurement. 1997; 21(2):173–184.

Raykov T. Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. British Journal of Mathematical and Statistical Psychology. 2001; 54:315–323. [PubMed: 11817096]

Roid, GH. Stanford-Binet Intelligence Scales. 5. Itasca, IL: Riverside; 2003.

Rutter, M., Bailey, A., Lord, C. Social Communication Questionnaire. Los Angeles: Western Psychological Services; 2003.

Rutter, M., LeCouteur, AL., Lord, C. Autism Diagnostic Interview- Revised. Los Angeles: Western Psychological Services; 2003.

Sijtsma K. On the use, the misuse, and very limited usefulness of Cronbach's alpha. Psychometrika. 2009; 74(1):107–120. [PubMed: 20037639]

Sparrow, SS., Balla, DA., Cicchetti, DV. Vineland Adaptive Behavior Scales- Interview Edition. Circle Pines, MN: American Guidance Service; 1984.

Steiger, JH., Lind, JM. Statistically based tests for the number of common factors. Paper presented at the meeting of the Psychometric Society; Iowa City, IA. 1980.

Stewart ME, Barnard L, Pearson J, Hasan R, O'Brien G. Presentation of depression in autism and Asperger's syndrome: a review. Autism. 2006; 10:103–116. [PubMed: 16522713]

Tarbox, J., La Cava, S., Hoang, K. Types of Assessment. In: Matson, JL., editor. Handbook of Assessment and Diagnosis of Autism Spectrum Disorder. New York: Springer; 2016. utism and Child Psychopathology Series

van Steensel, Bogels, Perrin. Anxiety disorders in children and adolescents with autistic spectrum disorders: A meta-analysis. Clinical Child and Family Psychological Review. 2011; 14:302–317. DOI: 10.1007/s10567-011-0097-0

Wechsler, D. Wechsler Adult Intelligence Scale- Third Edition. San Antonio, TX: The Psychological Corporation; 1997.

Wechsler, D. Wechsler Intelligence Scale for Children. Fourth. San Antonio, TX: Harcourt Assessment Inc; 2003.

Weissman AS, Bates ME. Increased clinical and neurocognitive impairment in children with autism spectrum disorders and comorbid bipolar disorder. Research in Autism Spectrum Disorders. 2010; 4:670–680. DOI: 10.1016/j.rasd.2010.01.005

Highlights

- Emotional disorders are common in youth with autism spectrum disorder

- Over 40% of individuals with autism spectrum disorder have more than two mental health conditions

- Routine screening during later childhood and adolescence is needed for early identification

- The Child Behavior Checklist's Syndrome Scales can assist with broad screening for a range of emotional and behavior disorders

- The Child Behavior Checklist's DSM-Oriented Scales can assist with the targeted assessment of depression and anxiety

Table 1

Participant Characteristics for Confirmatory Factor Analyses

	Number	Percent
Gender Male	80	86.0
Socioeconomic Status[a]		
Major Business/Professional	35	38.0
Medium Business/Minor Professional	39	42.4
Skilled Craftsman/Clerical/Sales	12	13.0
Machine Operator/Semi-skilled	4	5.3
Unskilled Laborer/Menial Services	2	2.2
Race[b]		
White	89	95.7
Asian	2	2.2
Black	2	2.2
Ethnicity: Non-Hispanic	93	100.0
Full Scale IQ <70[c]	11	12.4
Vineland Adaptive Behavior Scales Classification[d]		
Moderately High (115–129)	1	1.1
Adequate (86–114)	8	8.6
Moderately Low (71–85)	22	23.7
Low (70)	62	66.7
Mild (55–70)	42	45.2
Moderate (40–54)	11	11.8
Severe (25–39)	8	8.6
Profound (24)	1	1.1

Note. N=93 unless otherwise indicated.

[a]Hollingshead (1975) scale ($n = 92$), sum of percentages exceed 100 due to rounding error.

[b]Sum of percentages exceed 100 due to rounding error.

[c]$n = 89$.

[d]Vineland (Sparrow et al., 1984; n=93), sum of percentages exceed 100 due to rounding error.

Table 2

Characteristics of Participants Evaluated for a Co-occurring Psychiatric Disorder

	Number	Percent
Gender Male	67	87.0
Socioeconomic Status[a]		
Major Business/Professional	28	36.8
Medium Business/Minor Professional	31	40.8
Skilled Craftsman/Clerical/Sales	11	14.5
Machine Operator/Semi-skilled	4	5.3
Unskilled Laborer/Menial Services	2	2.6
Race		
White	74	96.1
Asian	2	2.6
Black	3	1.3
Ethnicity: Non-Hispanic	77	100.0
Full Scale IQ <70	9	11.7
Vineland Adaptive Behavior Scales Classification[b]		
Adequate (86–114)	8	10.4
Moderately Low (71–85)	17	22.1
Low (70)	52	67.5
Mild (55–70)	38	49.4
Moderate (40–54)	10	13.0
Severe (25–39)	4	5.2
Depressive Disorder[c]	15	19.5
With ADHD	8	53.3
With Adjustment Disorder	0	0
With Anxiety Disorder	11	73.3
With Conduct Disorder	1	0.7
With Oppositional Defiant Disorder	3	20.0
With Tic Disorder	3	20.0
With Other Disorder	1	0.7
Anxiety Disorder	35	45.5
With ADHD	13	37.1
With Adjustment Disorder	1	0.3
With Conduct Disorder	1	0.3
With Depressive Disorder	11	31.4
With Oppositional Defiant Disorder	2	5.7
With Tic Disorder	7	20.0
With Other Disorder	4	11.4

Note. *N*=77 unless otherwise indicated.

[a] *n*=76.

[b]Percentages may exceed 100 due to rounding error.

[c]Major Depressive Disorder or Dysthymic Disorder.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Correlations Between Affective Problems, Anxiety Problems, and Other Measures

Measure	Affective Problems			Anxiety Problems		
	r^a	p	95% CI^b	r	p	95%CI
K-SADS						
Depression	.59	<.001	.41, .73	.43	<.001	.41, .73
Anxiety	.52	<.001	.33, .66	.66	<.001	.51, .76
DSM Oriented Scales						
ADHD	.35	.002	.14, .53	.29	.018	.05, .49
ODD	.46	<.001	.26, .62	.35	.002	.13, .53
Conduct Problems	.43	.002	.17, .64	.37	.003	.15, .55
Somatic Problems	.44	<.001	.25, .60	.30	.005	.09, .51
Affective Problems	----	----	----	.57	<.001	.39, .70
Anxiety Problems	.57	<.001	.39, .70	----	----	----
ADI-R	.10	.427	−.16, .33	.16	.167	−.07, .40

Note.

^{a}Pearson sample correlation.

^{b}Bias-corrected 95% confidence interval derived from 10,000 bootstrap samples with replacement.

Table 4

Diagnostic Accuracy Analyses (N = 77)

ROC Statistic[a]	DSM Oriented Scale		
	Affective Problems	*Anxiety Problems*	
AUC[b]	.88**	.71*	
Standard Error	.04	.06	
p	<.001	.001	
95%CI[c]	.80, .96	.60, .83	
Sensitivity[d]	.93	.83	.91
Specificity	.74	.52	.31
Cut Score	6.5	2.5	1.5

[a]ROC analyses evaluated how well *Affective Problems* identified those with and without a depressive disorder and how well *Anxiety Problems* identified those with and without an anxiety disorder.

[b]Area under the curve.

[c]95 percent confidence interval for AUC.

[d]Two cut scores obtained from the sample and their associated sensitivity and specificity data are provided for *Anxiety Problems*.