Running head: NEGATIVE AFFECT SYMPTOM DIFFERENTIATION IN YOUTH

Symptom differentiation of anxiety and depression across youth development and clinic-referred/non-referred samples: An examination of competing factor structures of the Child Behavior Checklist DSM-Oriented Scales

Maggi Price¹
Charmaine Higa-McMillan¹
Chad Ebesutani²
Kelsie Okamura³
Brad Nakamura³
Bruce Chorpita⁴
John Weisz⁵,⁶

Psychology Department
University of Hawaii, Hilo, Hilo, Hawaii¹

Psychology Department
Yonsei University, Seoul, South Korea²

Psychology Department
University of Hawaii, Manoa, Manoa, Hawaii³

Psychology Department
University of California, Los Angeles, Los Angeles, California⁴

Psychology Department
Harvard University, Cambridge, Massachusetts⁵

Judge Baker Children’s Center
Harvard Medical School, Cambridge, Massachusetts⁶

Send correspondence to Charmaine Higa-McMillan; 200 W. Kawili St, Hilo, HI 96720; Phone: 808-933-3166; E-mail: higac@hawaii.edu
Abstract

This study examined the psychometric properties of the *DSM*-oriented scales of the Child Behavior Checklist (CBCL/6-18; Achenbach, Dumenci, & Rescorla, 2003) using confirmatory factor analysis to compare the six-factor structure of the *DSM*-oriented scales to competing models consistent with developmental theories of symptom differentiation. We tested these models on both clinic-referred (*N* = 757) and school-based, non-clinic-referred (*N* = 713) samples of youth in order to assess the generalizability of the factorial structures. Although previous research has supported the fit of the six-factor *DSM*-oriented structure in a normative sample of youth ages 7 to 18 (Achenbach & Rescorla, 2001), tripartite model research indicates that anxiety and depressive symptomatology is less differentiated among children compared to adolescents (Jacques & Mash, 2004). We thus examined the relative fit of a six- and five-factor model (collapsing anxiety and depression) with younger (ages 7-10) and older (ages 11-18) youth subsamples. Results revealed that the six-factor model fit the best in all samples except among younger non-clinical children. Results extended the generalizability of the rationally-derived six-factor structure of the *DSM*-oriented scales to clinic-referred youth, and also provided further support to the notion that younger children in non-clinical samples exhibit less differentiated symptoms of anxiety and depression.

Key words: Symptom differentiation, Negative affect, Anxiety, Depression, Development
Although symptoms of anxiety and depression tend to co-exist throughout the lifespan, empirical evidence indicates that rates of comorbidity are highest among youth (Rohde, Lewinsohn, & Seeley, 1991). High rates of comorbidity have led some researchers to conclude that anxiety and depression actually represent a unitary construct in youth (e.g., Achenbach, Connors, Quay, Verhulst, & Howell, 1989; Brady & Kendall, 1992; Jacques & Mash, 2004). Research also indicates that anxiety and depression symptom differentiation in youth may differ as a function of psychopathological disturbance and/or age (cf. Brady & Kendall, 1992). Such research has led some to question the validity of the current conceptualizations of anxiety and depressive disorders in the Diagnostic and Statistical Manual of Mental Disorders (APA, 2000). The difficulty discriminating between anxiety and depression coupled with the high rates of comorbidity in both adult and child populations suggest that depression and anxiety are strongly related to one another and therefore may not be best conceived as separate and discrete disorders (e.g., Lahey et al., 2008; Watson, 2005; Higa-McMillan, Smith, Chorpita, & Hayashi, 2008). In fact, treatments have also more recently begun to be developed and tested to treat this broad class of emotional problems and disturbances related to anxiety and depression, such as the Unified Protocol for Treatment of Emotional Disorders in Youth (Ehrenreich, Goldstein, Wright, & Barlow, 2009). Proposed changes for the upcoming fifth addition of the DSM reflect this growing concern. For example, although a Mixed Anxiety/Depression disorder currently exists in the Appendix of the DSM-IV-TR, the Mood Disorders workgroup has proposed that it be included in the DSM-V as a separate diagnostic category (Fawcett, 2009). While research in this area is growing in the adult literature, it is critical that research in developmental psychopathology examine the implications of development and severity of symptoms on the structural similarities and differences between anxiety and depression.
Comorbidity and Symptom Severity

The high rate of comorbid anxious and depressive symptomology in youth has been well established in the research literature (e.g., Kovacs, Gatsonis, Paulauskas, & Richards, 1990; Masi, Mucci, Favilla, Romano, & Poli, 1999; van Lang Ferdinand, Ormel, & Verhulst, 2006). In a review of research studies that examined anxiety and depression comorbidity in youth, Brady and Kendall (1992) found that comorbidity rates of the studies assessed ranged from 15.9 to 61.9% depending on the sample studied. For example, lower comorbidity rates were found in non-clinical samples (Anderson, Williams, McGee, & Silva, 1987; Costello et al., 1988) and higher rates were found among youth inpatients (Carey, Finch, & Imm, 1989) and adolescent outpatients (Kovacs et al., 1990). In contrast, research that has attempted to address the unitary versus two-factor structure of depression and anxiety in youth has concluded that higher comorbidity rates are found among less disturbed samples and that symptom differentiation may be a function of psychopathology severity (Gurley, Cohen, Pine, & Brook, 1996). Among a large community sample of youth ($N = 976$, ages 9-20) and their caregivers who completed structured clinical interviews, Gurley et al. (1996) found that anxiety and depression could not be discriminated among those without a diagnosis but could be discerned in their diagnosed subsample. Results also indicated that the overlap between anxiety and depressive symptoms decreased (i.e., symptoms became more differentiated) as emotional psychopathology increased. Interestingly, comorbidity was slightly more common among younger youth and a unitary depression/anxiety factor best fit those who were below the most stringent diagnostic threshold for all disorders, regardless of age. The authors concluded that the ability to discriminate between depression and anxiety increases as the severity of emotional psychopathology increases. While studies using adult participants have replicated these findings (Clark & Watson, 1991;
Hiller, Zaudig, & von Bose, 1989; Katon & Roy-Byrne, 1991; Zinbarg et al., 1994), few, if any, studies have done so with youth samples. Thus, the present study seeks to address this research need.

**Tripartite Model and Development**

Research on the tripartite model has also shed light on the overlap and differentiation of anxiety and depression symptoms in youth. The tripartite model purports that anxiety and depression can be understood along three dimensions (Clark & Watson, 1991). The first dimension, Negative Affect (NA), accounts for general emotional distress and includes states such as fear, anger, guilt and sadness (Watson & Clark, 1984; Watson & Tellegen, 1985). Individuals with anxiety and/or depression are hypothesized to have moderate to high levels of NA. The second dimension, known as physiological hyperarousal (PH), is theorized to be specific to anxiety (i.e., high levels of PH should be found among individuals with anxiety, and PH and depression should have low or nonsignificant correlations). Lastly, positive affect (PA), characterized by energy, engagement, and high levels of interest, should be significantly negatively correlated with depression and relatively unrelated to anxiety symptoms.

Some tripartite model research has found that the tripartite model is applicable to both younger children and adolescents. For example, results from Turner and Barrett (2003) indicate that the tripartite model differentiates between anxiety and depression in nonreferred third-, sixth-, and ninth-grade children. Results from this study provided little evidence for increasing differentiation between depression and anxiety. Among a sample of child and adolescent psychiatric inpatients, Joiner, Catanzaro and Laurent (1996) found that the tripartite model was well supported. Additionally, Chorpita, Albano, and Barlow (1998) assessed a clinical sample of children with anxiety disorder and/or mood disorder diagnoses and found that a model in which
fear, anxiety, and depression were conceptualized as distinct but related factors fit the data best. Interestingly, Chorpita, Plummer, and Moffitt (2000) found that among a child clinical sample, nearly all aspects of the tripartite model were supported (e.g., NA was significantly related to anxiety, PA was significantly related to depression), except the PH dimension, which did not correlate significantly with Generalized Anxiety Disorder symptomology. Similarly, in a large school-based sample, Chorpita (2000) found that PH was only related to Panic Disorder but not to other anxiety disorders. Hence, it appears that the constructs of NA and PA are well supported in the child psychopathology literature; however there is evidence that suggests that the PH factor may not be applicable to youth.

Despite the growing empirical support for the applicability of the tripartite model to younger youth samples, there is stronger support for its applicability in samples of older children and adolescents. For example, Jacques and Mash (2004) found that among a large community sample of elementary school children and high school youth (N = 472), the tripartite model was better supported in the older sample. Using confirmatory factor analytic methods to assess items from a battery of youth anxiety and depression measures, Cole, Truglio, and Peeke (1997) found that a unified construct model (combining depression and anxiety) fit data from a non-clinical sample of thirds graders (N = 280) better than the tripartite model as well as a dual model of anxiety and depression (i.e., anxiety and depression specified as separate factors). On the contrary, these researchers found that the tripartite model fit the sixth grade sample (N = 211) data better compared to the unified one-factor model. Furthermore, anxiety and depression symptoms were less correlated in the older sample compared to the younger sample. Similarly, Lonigan, Hooe, David, and Kistner (1999) found that the tripartite model fit data from older children from a non-clinical sample (N = 213, ages 12-17) better than their younger counterparts.
Results from these studies provide support for the notion that there is an increasing differentiation between anxiety and depression across development (De Bolle & De Fruyt, 2010).

Alternative theories used to explain the development of anxiety and depression in youth are also worth noting. Cannon and Weems (2006) and Weems, Saltzman, Reiss, and Carrion (2003) have found support for the notion that primary anxiety symptoms may precede the development of depression. For instance, in an investigation of the applicability of the tripartite model in a large community sample of youth, Cannon and Weems (2006) found that in the younger subsample (ages 6 to 11) there was a distinct anxiety-only group but a less distinct depression-only group. In the older subsample (ages 12 to 17), there were more instances of depression-only and less youth who solely had anxiety problems. The authors suggest that the comorbidity of anxiety and depression in youth may not be a result of heightened susceptibility to a common component (such as negative affect). Rather, comorbidity may be better explained by the distress and impairment caused by anxiety which leads to increased depressive symptomology. Similarly, results from a study of children with a history of traumatic experience(s) suggest that the emotional numbing of depression may develop as a result of the chronic hyperarousal associated with anxiety in youth (Weems et al., 2003).

In sum, previous studies indicate that the applicability of the tripartite model, specifically, the differentiation between depression and anxiety, may differ as a function of age in both clinical and nonreferred samples. However, other research indicates that the tripartite model is appropriate for both younger children and adolescents. Additionally, some epidemiological research suggests that anxiety and depression symptoms may be less differentiated among non-clinical samples (Gurley et al., 1996). This phenomenon may be partially explained by the higher
comorbidity rates of anxiety and depression in younger versus older youth. Given the inconsistent research evidence from previous studies, we sought to assess the applicability of three related factor structures in order to shed light on the two-factor versus unitary conceptualizations of depression and anxiety symptomology in youth.

**Present Study**

In order to examine anxiety and depression symptom differentiation in youth, we assessed competing factor structures based on the items from the Child Behavior Checklist (CBCL) DSM-Oriented Scales (Achenbach et al., 2003) in a large sample of referred and non-referred youth as well as subsamples of older (ages 11-18) and younger (ages 7-10) youth. This age cut-off was chosen because the formal operations stage begins around age 11 (Piaget, 1983). At this time, youth begin to have abstract thoughts and are capable of meta-cognitions and thus worry and self-consciousness. Moreover, the definition of "adolescence" has recently been modified so as to include children as young as 11 years of age; this is because children at this age have begun to experience more adolescent-typical problems (e.g., substance abuse, eating disorders and depression; Saluja et al., 2004). Given that past research has assessed the Tripartite Model and symptom differentiation of anxiety and depression through child self-report, use of parent report also represents an important step in understanding these constructs as reported by parents. The CBCL is one of the most commonly used parent-report measures of youth behavioral and emotional problems and competencies. The rationally-derived CBCL DSM-Oriented scales were created by a panel of child mental health professionals with the aim of producing a CBCL model (and set of DSM-oriented scales) congruent with DSM nosology (Achenbach et al., 2003). Each scale represents one or more DSM diagnoses. The scales and their associated diagnoses include: (1) Affective Problems (2) Anxiety Problems, (3) Attention
Deficit/Hyperactive Problems, (4) Conduct Problems, (5) Oppositional Defiant Problems, and (6) Somatic Problems. The construction of the DSM-oriented scales thus provided us a set of items upon which to examine the differentiation of anxiety and depressive problems across sample types.

Specifically, we examined the six-factor CBCL DSM-Oriented structure as well as the applicability of two competing, yet related, DSM-Oriented models. We tested a five-factor DSM-Oriented model by collapsing the Affective Problems and Anxiety Problems scales into a broader “Negative Affect” factor. We also tested a two-factor model of the DSM-Oriented scales, which consisted of the Internalizing Problems and the Externalizing Problems scale (Achenbach, 1991). The Internalizing Problems scale includes the Affective, Anxiety, and Somatic Problems scales and the Externalizing Problems scale includes the Oppositional Defiant, Conduct, and Attention Deficit/Hyperactive Problems scales.

Based on previous psychometric research on the CBCL DSM-Oriented scales, we hypothesized that the six-factor model would yield good fit for the entire sample (N = 1470) as well as all subsamples (e.g., younger, older, clinical, and non-clinical youth). However, given the inconsistent results regarding the relationship between age, psychopathology severity, and depression and anxiety symptom differentiation in youth, we predicted that the six- and five-factor structures would provide variable fit for the subsamples. In other words, we expected that the model encompassing a one-factor construct combining depression and anxiety (i.e., the five-factor structure), to differentially fit younger and older participants as well as clinic-referred and non-referred youth, compared to the six-factor structure. Lastly, we were interested in potential differences in reporting styles and item interpretations of the CBCL items across sex. Thus, we examined the data for differential item functioning across boys and girls. We expected that
measurement invariance tests would identify differences (with respect to differential item functioning) across genders as previous studies have found that symptom presentation (e.g., anxiety, depression) often differs between boys and girls (Essau, Conradt, & Petermann, 2000; Lewinsohn, Rohde, & Seeley, 1998).

**Method**

**Participants**

The present sample \((N = 1,470)\) consisted of clinic-referred \((n = 757)\) and school-based \((n = 713)\) youth whose caregivers completed the CBCL/6-18.

*Clinic-referred sample.* The clinical sample was comprised of youth referred to a University clinic located in Honolulu, Hawaii and a clinic in a children’s hospital in Boston, Massachusetts for mental health evaluations. The mean age of the clinic-referred sample was 13.07 years \((SD = 3.04; \text{range} = 7-18)\). This sample consisted of 528 males \((69.7\%)\) and 224 females \((29.6\%);\) youth gender data were missing for five \((0.7\%)\) participants. Major ethnic groups of the youths included Multiethnic \((n = 397; 52.4\%);\) Asian American \((n = 85; 11.2\%);\) White \((n = 84; 11.1\%);\) Hawaiian/Pacific Islander \((n = 42; 5.5\%);\) other \((n = 18; 2.4\%);\) African American \((n = 12; 1.6\%);\) and Latino/Hispanic \((n = 9; 1.2\%);\) Ethnicity data for this sample were missing for 110 \((14.5\%)\) participants. Of the 757 youth in the clinical sample, 372 had at least one diagnosis \((58.2\% \text{ had one diagnosis and } 18.1\% \text{ had two or more diagnoses});\) Of youth with a diagnosis, 168 had an anxiety disorder \((\text{Separation Anxiety Disorder } n = 24; 3.1\%; \text{ Obsessive Compulsive Disorder } n = 20; 2.5\%; \text{ Social Anxiety Disorder } n = 52; 6.9\%; \text{ Specific Phobia } n = 41; 5.4\%; \text{ Generalized Anxiety Disorder } n = 31; 4.1\%); 42 \((5.5\%)\) were diagnosed with Major Depressive Disorder, 119 \((15.7\%)\) were diagnosed with Oppositional Defiant Disorder and 130
(17.2%) youth received a Conduct Disorder diagnosis. Eighteen (2.4%) participants were diagnosed with Pervasive Developmental Disorder and 28 (3.7%) received an “other” diagnosis.

School-based sample. The non-clinical (school-based) sample was comprised of children and adolescents recruited from private and public schools in grades 3-12 on O’ahu, Hawaii. The mean age of this sample was 12.86 years ($SD = 2.8$; range = 7-18) and consisted of 295 males (41.4%) and 416 females (58.3%); youth gender data were missing for two (0.3%) participants. Major ethnic groups of the youths included Asian American ($n = 380; 53.3$%), Multiethnic ($n = 288; 40.4$%), Hawaiian/Pacific Islander ($n = 20; 2.8$%), White ($n = 19; 2.7$%), and Latino/Hispanic ($n = 2; 0.3$%). Ethnicity data for this sample was missing for 24 (3.4%) participants.

Six separate t-tests were conducted in order to investigate CBCL $DSM$-oriented scale score differences between the clinic-referred and school-based samples. All analyses were performed using a 99.7% confidence interval (alpha of .003). Results indicated that the clinic-referred group had significantly higher total scores on all $DSM$-Oriented scales compared to the school-based, non-referred group ($p < .003$).

Procedure

For both samples, participants and their caregivers underwent standardized Institutional Review Board-approved notice of privacy and consent procedures. Upon receiving signed consent, the primary caregivers of the school-based youths were mailed a blank CBCL (with a University-addressed, stamped envelope) to fill out and mail back to the University. For the clinic-referred sample, the caregivers were asked to fill out a CBCL regarding their children (following consent provided at the initial meeting with the youths and their primary caretakers). Clinic-referred youth underwent diagnostic intake evaluations (including completing
questionnaires and structured diagnostic interviews) and their caregivers also completed a packet of intake questionnaires and participated in structured diagnostic interviews. Diagnostic determinations were made by Ph.D. level clinical child psychologists and doctoral students in clinical psychology.

**Measure**

*Child Behavior Checklist for Ages 6–18 (CBCL/6–18; Achenbach & Rescorla, 2001)* The CBCL is a parent-report instrument that includes 113 items measured on a 3-point Likert scale rated as Not True (0), Somewhat or Sometimes True (1), or Very True or Often True (2). The *DSM*-Oriented Scales are comprised of 55 of the 113 items. The six *DSM*-Oriented scales include: (1) Affective problems (2) Anxiety Problems, (3) Attention Deficit/Hyperactive Problems, (4) Conduct Problems, (5) Oppositional Defiant Problems, and (6) Somatic Problems. We transformed the narrow-band six factor structure into a more broad-band five-factor structure by collapsing the Affective Problems and Anxiety Problems into a single Negative Affect factor. A third broad-band factor structure, consisting of two factors, was also tested. The two-factor structure is composed of an Internalizing Problems scale (Affective Problems, Anxiety Problems and Somatic Problems scales are collapsed) and an Externalizing Problems scale (Attention Deficit/Hyperactive Problems, Conduct Problems, and Oppositional Defiant Problems scales are collapsed; see Figure 1).

The *DSM*-Oriented Scales have evidenced strong test-retest reliability, cross-informant agreement and internal consistency (Achenbach et al., 2003). Supportive scale reliability and convergent and discriminative validity have been demonstrated in a large sample of clinic-referred children (Nakamura, Ebensutani, Bernstein, & Chorpita, 2009) and support for the factor structure has been found in a community sample of youth (Achenbach et al., 2003). The clinical
utility of the CBCL DSM-Oriented scales has also been explored by Ferdinand (2008) and Ebesutani et al. (2010). Ferdinand (2008) found that when using parent/child impairment ratings, the Anxiety Problems and Affective Problems scales could predict relevant diagnoses (e.g., scores on the Anxiety Problems scale could predict SAD, GAD, and SPEC). Among a large sample of clinic-referred children, five of the six CBCL DSM-Oriented scales (excluding the Somatization Problems scale, which was not tested due to absence of Somatization Disorder diagnoses in the sample) corresponded significantly with the individual diagnoses targeted by the scale (Ebesutani et al., 2010). Furthermore, each of these scales was able to discriminate youth within the diagnostic group targeted by the scale from youth without such diagnoses.

Data Analytic Strategy

Confirmatory factor analysis. To examine competing factor structures of the CBCL DSM-Oriented Scales, we conducted confirmatory factor analysis (CFA) using Mplus version 4.21 (Muthén & Muthén, 2006). Due to the categorical (i.e., ordinal) nature of our data, we employed the robust weighted least-squares with mean and variance adjustment (WLSMV) estimator as it has been found to handle ordinal data well (Brown, 2006; Muthén, du Toit, & Spisic, 1997; Flora & Curran, 2004). The fit of the six-, five-, and two-factor DSM-Oriented structures were examined for the entire sample (clinical and non-clinical samples combined) and the clinical and non-clinical samples separately. We also examined the fit of these factor structures among different subsamples including (a) younger (ages 7-10) and (b) older (ages 11-18) youth subsamples (both clinical and non-clinical participants combined in each subgroup) as well as more specific subgroups crossing age by sample type: (c) younger clinical, (d) older clinical, (e) younger non-clinical, and (f) older non-clinical. The fit of the DSM-Oriented Scale structures were evaluated via various fit indices including Root Mean Square Error of
Approximation (RMSEA; Steiger, 1990), Comparative Fit Index (CFI, Bentler, 1990), and the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973). Values of .90 and .95 and above represent good and excellent model fit, respectively, for the CFI and TLI indices (Bentler, 1990; Hu & Bentler, 1999). For the RMSEA index, values of .08 or lower indicate good fit, and RMSEA values of .05 or lower indicate excellent fit (Browne & Cudeck, 1993). To examine the relative fit of the two-, five- and six-factor structures, we conducted chi-square ($\chi^2$) difference tests for each subsample.\footnote{Chi-square difference tests were calculated using the “difftest” command in Mplus. Rather than calculating the $\chi^2$ difference tests by hand, the “difftest” command had to be used due to the use of a limited information estimator (i.e., WLSMV). When using the WLSMV estimator, degrees of freedom are estimated and $\chi^2$ values are not distributed as standard $\chi^2$. Therefore, degrees of freedom are estimated and differences in degrees of freedom are not the same as those produced by estimators such as the ML estimator. When using the DIFFTEST command, the $p$-value is the only parameter that should be interpreted. This is because, in the process of Mplus executing the DIFFTEST procedures, "chi-square and degrees of freedom are adjusted to obtain a correct $p$-value" (Muthén, 2007). Details about the “difftest” command and the procedure used to estimate degrees of freedom can be found in the Mplus’ Technical Appendices (at www.statmodel.com/download/webnotes/webnote10.pdf; Asparouhov & Muthén, 2006) and in the Mplus User's Guide (Muthén & Muthén, 2004).}

Multi-group confirmatory factor analysis. We then conducted multi-group confirmatory factor analysis (MG-CFA) using Mplus version 4.21 to determine whether the best fitting CBCL DSM-Oriented structure was associated with measurement invariance across sex (i.e., male and female youth) and sample-type (i.e., clinical and non-clinical participants). MG-CFA is commonly considered to be the most robust technique for the examination of measurement invariance and offers researchers the ability to examine all aspects of population heterogeneity and measurement invariance (Brown, 2006). As recommended by Brown (2006), we employed a “step-up” approach for evaluating model restrictions within multiple-groups CFA. The model-building approach allowed us to identify and examine any aspects of poor fit associated with each additional model restriction (Brown, 2006). Hence, after testing the CFA model in each group separately (i.e., single-sample CFA tests), we conducted simultaneous tests (for all samples of interest) of equal form (i.e., equal factor structure, also referred to as configural...
invariance). Due to the categorical nature of our data, we did not model item intercepts (as would be done with continuous data) but rather, we modeled item thresholds (Lubke & Muthén, 2004). The number of thresholds is equal to the number of categories for each item minus one (Brown, 2006). Thus we modeled two thresholds for each item due to the 3-point Likert-scale underlying the CBCL. We also set all scaling factors to 1.0, as is required when conducting MG-CFA with categorical data (cf. Mplus FAQ).

We tested the various forms of measurement invariance in the following ways. First, we assessed the fit of equal form across groups using the same fit index cut-offs mentioned above. Support for equal form across groups indicates configural invariance and justifies proceeding to the test of equal factor loadings (i.e., metric invariance, also referred to as weak factorial invariance) and then to the test of equal item thresholds. We freed and constrained both factor loadings and item thresholds in tandem so that meaningful comparisons of factor distributions across groups could be made due to the categorical nature of our data (Muthén & Asparouhov, 2002). We determined whether these constraints significantly degraded fit using the $\chi^2$ difference test as well as by examining the change in CFI values between the ‘equal form and equal thresholds’ test and the nested ‘equal factor loading’ test. If the $\Delta$CFI is greater than .01 or the $\chi^2$ difference test results are significant, the model fit of the equal thresholds and equal factor loading model is considered significantly degraded compared to the equal form model (Cheung & Rensvold, 1998), suggesting a lack of full invariance of the tested parameters (i.e., that the parameters generally vary) across groups. Multiple tests (i.e., $\Delta$CFI and $\chi^2$ difference testing) were used to provide a more conservative and robust evaluation of measurement invariance. Partial measurement invariance was examined if these results indicated that full measurement invariance was not supported. Partial measurement invariance may be present and supported
when the lack of overall measurement invariance is due to only *some* of the indicators being non-invariant across groups (Byrne, Shavelson, & Muthén, 1989). To test for partial measurement invariance, the most non-invariant parameters are freed one at a time, and tests are performed to examine for invariance under such conditions.

**Results**

**Confirmatory Factor Analysis**

The 55 items comprising the six-, five- and two-factor CBCL *DSM*-Oriented Scales were subject to CFA for the full sample (non-clinical and clinical samples; \(N = 1,470\)). The resulting fit statistics and chi-square difference test results appear in Table 1. The two-factor model (comprised of the Externalizing Problems factor and Internalizing Problems factor) demonstrated poor to moderate fit in nearly all subsamples—except for the non-clinical (all ages) sample and the non-clinical older youth sample, which evidenced good fit (i.e., RMSEA < .08, CFI > .90, and TLI > .95).

The five-factor model, which is composed of the Negative Affect (combined Affective Problems and Anxiety Problems), Attention Deficit Problems, Somatic Problems, Conduct Problems, and Oppositional Defiant Problems scales, provided good fit for five of the nine samples examined, including all non-clinical samples (all ages non-clinical, young non-clinical, and old non-clinical) as well as the younger and older subsamples (both clinical and non-clinical participants). The five-factor model provided moderate fit for the entire sample and all clinical samples (all ages clinical, young clinical, and old clinical). Results from the chi-square difference tests also indicated that the six- and five-factor models provided significantly better fit in all samples compared to the two-factor model as evidenced by \(p\) values of less than .0001 for
each test. These results suggest that the five-factor represents the structure of the data better than the two-factor model.

The original *DSM*-oriented scales are composed of six factors (similar to the five factor model tested in the present paper but splitting Negative Affect into Affective Problems and Anxiety Problems). The resulting fit statistics represent good fit for this six-factor model in the entire sample as well as adequate to good fit for all subsamples. Although the CFI values for both the clinical and older clinical youth samples do not represent good fit by conventional standards (i.e., CFI < .90), both TLI and RMSEA values indicate good fit for the six-factor structure of the *DSM*-Oriented Scales in these two samples. Chi-square difference tests were also conducted to compare the fit of the six- and five-factor models for each subsample. Results from these tests indicate that the six-factor model provides significantly better fit for all samples ($p$'s < .01) except for the non-clinical young subsample [$\chi^2_{diff}(4) = 9.377, p = .052$]. The non-significant results indicate that the six-factor model does not provide a significantly better fit relative to the five-factor model for the non-clinical young sample—thereby supporting the more parsimonious, good-fitting five-factor model among the non-clinical young subsample. In other words, results suggest that the five-factor structure, which represents a unitary negative affect factor (i.e., depression and anxiety are undifferentiated), provides better fit for the younger youth sample characterized by less severe psychopathology (i.e., non-clinic-referred). Conversely, the six factor structure, which includes separate factors for Affective and Anxiety Problems, provides optimal fit for the younger clinic-referred sample (i.e., the younger subsample with more severe psychopathology).

**Multi-group confirmatory factor analysis**
**Measurement invariance across gender.** As tested and reported above, single-sample solutions indicated that the six-factor DSM-Oriented model provided good fit for all subsamples assessed—and also provided significantly better fit relative to the five-factor model for all subsamples, except the non-clinical young sample (whose data are best represented by a five-factor structure). Therefore, we proceeded to test measurement invariance of the six-factor structure across gender for the entire sample excluding the non-clinical young participants ($n = 1238$). To determine whether model fit (of the six-factor model) differed across gender, we first conducted a test of equal form by placing equality constraints on the general factorial structure across boys and girls. Strong model fit indices (i.e., $CFI = .960$, $TLI = .968$, $RMSEA = .064$) resulting from the test of equal form indicated that configural invariance of the six-factor model was supported across gender among this subsample. We then assessed the next level of measurement invariance by simultaneously constraining factor loadings and item thresholds to be equal across groups. Results from this test indicated that these constraints significantly degraded model fit, as evidenced by significant chi-square difference test results [$\chi^2_{\text{diff}}(45) = 265.289\ p<.0001$] and a $\Delta CFICFI$ value greater than .01 ($\Delta CFICFI = .039$) between the restricted and unrestricted solutions. In other words, results indicated that the CBCL items had different relationships to their respective DSM-Oriented factor across males and females.

We then proceeded to test partial metric invariance by freeing constraints on non-invariant items. We determined which items to free by calculating the difference in factor loadings for each item across groups. For example, the Somatic Problems scale item 56G, “vomiting, throwing up” was associated with the largest difference between boys and girls in factor loadings. For the boys-only group, the item factor loading was .70 and for girls it was .95, thus we eliminated the constraints previously imposed on this item's factor loading and item
threshold to allow this item to freely load on the Somatic Problems factor across group. We then tested for partial invariance by constraining all other factor loading and threshold parameters to be equivalent across samples. Results indicated that partial metric invariance was not supported as evidenced by a significant $\chi^2_{\text{diff}}$ test result and a change in $\Delta$CFI greater than .01 ($\chi^2_{\text{diff}}(44) = 257.558$ $p < .0001$, $\Delta$CFI = .039). This procedure was repeated four more times (i.e., four items associated with large factor loading differences were added to the model, one at a time). Chi-square difference test results and $\Delta$CFI values indicated that the model continued to be significantly degraded compared to the less constrained model, with each new addition. After freeing the 5 items associated with the largest factor loading differences between groups, we concluded that partial invariance was not supported (i.e., after 5 parameters were freed, $\chi^2_{\text{diff}}(43) = 222.360$ $p < .0001$, $\Delta$CFI = .036). We decided to discontinue partial invariance testing at this juncture as Vandenberg and Lance (2000) suggest that if there a large number of indicators are found to be noninvariant, partial invariance tests should not be continued and full invariance should be assumed. These results thus suggest that the item-factor relationships are not the same across boys and girls, and that the DSM-oriented scales are associated with differential item functioning with respect to gender among this subsample.

*Measurement invariance across sample-type.* The single-sample CFA results indicate that the five-factor model fit the young non-clinical data best and the six-factor model fit the young clinical sample data best. Therefore, we did not assess measurement invariance of any given factor structure across sample-type in the young subsample but did so for the older subsample (since the six-factor structure fit the best among the older clinical and older non-clinical youth). Results from the test of equal form across these groups indicate that configural invariance was present across clinical and non-clinical older youth (CFI = .916, TLI = .954, RMSEA = .065).
We then tested the next level of measurement invariance by constraining factor loadings and item thresholds to be equal across groups. Results from this analysis revealed that measurement invariance was not supported at this level (i.e., $\chi^2_{\text{diff}}(47) = 917.011 \ p < .0001 \text{ and } \Delta\text{CFI} = .095$). In other words, results indicated that the CBCL items had different relationships to their respective DSM-Oriented factor across older clinical and older non-clinical samples. We tested partial invariance by following the steps outlined above and concluded that partial invariance was not supported for the six-factor model across clinical and non-clinical participants in the older youth sample (i.e., after freeing the 5 most noninvariant parameters; $\chi^2_{\text{diff}}(44) = 834.298 \ p < .0001 \text{ and } \Delta\text{CFI} = .083$). These results suggest that older-youth participants answered many CBCL items associated with the DSM-Oriented scales in systematically different ways based on sample type (i.e., clinic referred or school-based).

**Discussion**

This study examined symptom differentiation of anxiety and depression across development and sample types by examining various competing factor structures using the DSM-Oriented scales of the CBCL. All three models tested provided at least adequate fit, suggesting that the CBCL DSM-Oriented Scales (and hierarchical combinations of these scales) are structurally consistent with DSM nosology. Given that past research of these scales have been limited to samples of youth who had moderate to high CBCL Total Problems Scale scores (Achenbach et al., 2003) and given the mixed results regarding the factorial validity of the scales (Hartman et al., 2001), this study provides more conclusive supportive evidence for the structural validity of the CBCL DSM-Oriented scales.

As hypothesized, we found that the six factor model (i.e., Affective Problems, Anxiety Problems, Somatic Problems, Attentional Problems, Conduct Problems, and Oppositional
Defiant Problems) provided good fit for all subsamples. Interestingly, the five-factor model (with Anxiety Problems and Affective Problems collapsed into a single Negative Affect factor) was supported above the six-factor model among the young, non-clinical sample. These findings are consistent with previous research on symptom differentiation in clinical versus non-clinical samples (Gurley et al., 1996) and younger versus older youth (Cole et al., 1997; Jacques & Mash, 2004; Lonigan et al., 1999), suggesting that anxiety and depression are less differentiated in non-clinical youth and younger children. In addition to providing further support for the unique lack of differentiation of anxiety and depression among non-referred children, the present study also extended these findings to parent reports of youth emotional and behavioral problems.

Regarding this lack of differentiation among younger youth, it is possible that depression is a result of the distress and impairment caused by anxiety as Weems and colleagues suggest (Cannon & Weems, 2006; Weems et al., 2003) and thus the difference between anxiety and depression is less distinct in younger children. However, given that anxiety and depression were differentiated from each other in the younger clinical sample, an alternative explanation is that at a subclinical level, anxiety and depression are more alike than different in younger children. This conclusion appears to support the notion that there is an underlying trait of negative affect that accounts for this class of symptoms. Developmental theories and research have also shed light on possible mechanisms by which anxiety and depression differentiate in youth. For example, some research suggests that advances in cognitive development (i.e., self-concept development and self-esteem; Burnett, 1995; Ohannessian, Lerner, Lerner, & von Eye, 1999) and/or contextual factors in adolescence (e.g., peer relationships; Boivin, Hymel, & Bukowski, 1995; Boivin, Poulin, & Vitaro, 1994; Kupersmidt & Patterson, 1991) may play a role in the onset of depressive symptoms in older youth. Additionally, some research indicates that as emotions
become more specific and outcome-dependent with development, children’s increasingly sophisticated attributions result in previously undifferentiated feelings states (e.g., anxiety, depression; Weiner & Graham, 1988). Based on this theory, De Bolle, Decuyper, De Clercq and De Fruyt (2010) suggest that younger children who experience problems of depression or anxiety are likely to report a general malaise. Later in development, specific symptoms may be attributed to specific environmental entities or may be interpreted as signs of imminent danger which may result in specific fears or phobias and generalized anxiety, respectively. On the other hand, these signals may be interpreted as a failure to effectively cope with the situation, which will eventually affect one’s self-esteem and thus increase their risk for depression. In sum, these theories and supporting research suggest that cognitive development plays a primary role in negative affect symptom differentiation.

Beyond symptom differentiation among clinical and non-clinical samples and younger versus older youth, the partial invariance test results indicate that the parents of boys answered many CBCL items in systematically different ways than the parents of girls. These findings are not surprising given that research on gender differences across developmental psychopathology consistently find that the nature of symptom expression differs between girls and boys (Essau, Conradt, & Petermann, 2000; Lewinsohn, Rohde, & Seeley, 1998). Specifically, past research has often failed to find gender differences in depressive and anxiety symptomatology among pre-pubescent male and female youth (e.g., Cohen et al., 1993; Nolen-Hoeksema, Girgus, & Seligman, 1991). However, compared to male adolescents, higher rates of anxiety and depressive disorders are consistently found among female adolescents (Brooks-Gunn & Petersen, 1991; Essau, Conradt, & Petermann, 2000). Similar results were found for older clinical and older non-clinical youth (i.e., responses differed based on sample type). Although these samples may
in fact be responding differently, Brown (2006) notes that some parameters may differ by chance, especially if the sample size is large. Therefore, future research in this area should explore this matter more closely as it is of relevance to the generalizability and interpretation of the DSM-oriented scales.

Limitations to this study should be noted. First, this study only included parent report of child symptoms. Although more research is needed on the parent version of the ASEBA scales (i.e., the CBCL), the literature on cross-informant agreement suggests that on average, parents and children agree on symptom presence to a moderate degree (Achenbach, McConaughy, & Howell, 1987; De Los Reyes & Kazdin, 2005). Nevertheless, understanding parents' perspectives on the nature of their child's symptoms and the resulting structure of psychopathology is necessary as multi-informant assessment is routinely recommended (Barbosa, Tannock, & Manassis, 2002; Jensen et al., 1999). Furthermore, the possibility that parents are unable to detect slight differences in anxiety and affect among younger children with less severe psychopathology cannot be ruled out. Thus, longitudinal studies are needed to fully understand the complex interactions between the severity of symptoms, sequence of symptoms, developmental stage, and contextual factors in symptom differentiation between anxiety and depression across development. Second, whereas the clinical sample in this study was composed of clinic-referred youth from two different cities, the non-referred sample was only composed of youth from Hawaii. It is possible that youth in Hawaii differ in meaningful ways from youth from Massachusetts. Thus assumptions regarding the differences between the clinical and non-clinical samples should be made with caution. Lastly, the present sample size was not large enough to conduct analyses that would elucidate the developmental trajectory of negative affect differentiation. It is recommended that a larger sample size be utilized in future studies in order
to identify the particular age at which the differentiation between depression and anxiety begin to emerge.

Nonetheless, the present study provided additional needed empirical support for the relatively recently constructed CBCL *DSM*-oriented scales. Despite being rationally-derived based on the consensus of raters, the present confirmatory factor analytic findings empirically supported the 6-factor structure among younger and older clinic-referred samples. As noted above, the present study also further supports the lack of symptom differentiation of anxiety and depression among younger youth with less severe psychopathology, aiding to resolve the discrepant findings in the literature and also highlighting the need for further research to better understand why this particular subsample of young children tend to be associated with their own unique (unitary) representation of anxiety and depression.
References


Table 1. Fit Indices of the CBCL DSM-Oriented Scales

<table>
<thead>
<tr>
<th>Fit Indices of the CBCL DSM-Oriented Scales</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$\chi^2_{\text{diff}}$</th>
<th>$\Delta df$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-Factor Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All ($N = 1470$)</td>
<td>0.846</td>
<td>0.950</td>
<td>0.079</td>
<td>2580.33</td>
<td>251</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical ($n = 757$)</td>
<td>0.771</td>
<td>0.897</td>
<td>0.098</td>
<td>1650.59</td>
<td>200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical old ($n = 538$)</td>
<td>0.807</td>
<td>0.907</td>
<td>0.098</td>
<td>1016.42</td>
<td>166</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical young ($n = 219$)</td>
<td>0.832</td>
<td>0.884</td>
<td>0.095</td>
<td>287.23</td>
<td>97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-clinical ($n = 713$)</td>
<td>0.923</td>
<td>0.950</td>
<td>0.055</td>
<td>425.56</td>
<td>135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-clinical old ($n = 488$)</td>
<td>0.935</td>
<td>0.957</td>
<td>0.055</td>
<td>277.31</td>
<td>112</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-clinical young ($n = 225$)</td>
<td>0.892</td>
<td>0.922</td>
<td>0.076</td>
<td>130.85</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young ($n = 444$)</td>
<td>0.897</td>
<td>0.940</td>
<td>0.077</td>
<td>438.13</td>
<td>120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old ($n = 1026$)</td>
<td>0.871</td>
<td>0.955</td>
<td>0.080</td>
<td>1615.35</td>
<td>212</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five-Factor Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All ($N = 1470$)</td>
<td>0.895</td>
<td>0.968</td>
<td>0.064</td>
<td>1895.04</td>
<td>270</td>
<td>482.23</td>
<td>6</td>
<td>.000</td>
</tr>
<tr>
<td>Clinical ($n = 757$)</td>
<td>0.816</td>
<td>0.923</td>
<td>0.085</td>
<td>1374.72</td>
<td>214</td>
<td>220.10</td>
<td>5</td>
<td>.000</td>
</tr>
<tr>
<td>Clinical old ($n = 538$)</td>
<td>0.840</td>
<td>0.927</td>
<td>0.086</td>
<td>881.23</td>
<td>176</td>
<td>142.09</td>
<td>5</td>
<td>.000</td>
</tr>
<tr>
<td>Clinical young ($n = 219$)</td>
<td>0.890</td>
<td>0.925</td>
<td>0.085</td>
<td>222.66</td>
<td>98</td>
<td>102.71</td>
<td>6</td>
<td>.000</td>
</tr>
<tr>
<td>Non-clinical ($n = 713$)</td>
<td>0.948</td>
<td>0.967</td>
<td>0.045</td>
<td>331.17</td>
<td>137</td>
<td>165.28</td>
<td>6</td>
<td>.000</td>
</tr>
<tr>
<td>Non-clinical old ($n = 488$)</td>
<td>0.956</td>
<td>0.970</td>
<td>0.046</td>
<td>224.58</td>
<td>111</td>
<td>115.48</td>
<td>6</td>
<td>.000</td>
</tr>
<tr>
<td>Non-clinical young ($n = 225$)</td>
<td>0.934</td>
<td>0.953</td>
<td>0.059</td>
<td>103.34</td>
<td>58</td>
<td>75.57</td>
<td>5</td>
<td>.000</td>
</tr>
<tr>
<td>Young ($n = 444$)</td>
<td>0.944</td>
<td>0.968</td>
<td>0.056</td>
<td>297.25</td>
<td>123</td>
<td>148.24</td>
<td>5</td>
<td>.000</td>
</tr>
<tr>
<td>Old ($n = 1026$)</td>
<td>0.905</td>
<td>0.969</td>
<td>0.067</td>
<td>1257.32</td>
<td>226</td>
<td>319.41</td>
<td>6</td>
<td>.000</td>
</tr>
<tr>
<td>Six-Factor Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All ($N = 1470$)</td>
<td>0.906</td>
<td>0.972</td>
<td>0.059</td>
<td>1688.51</td>
<td>273</td>
<td>170.23</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>Clinical ($n = 757$)</td>
<td>0.850</td>
<td>0.938</td>
<td>0.076</td>
<td>1163.40</td>
<td>217</td>
<td>177.60</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>Clinical old ($n = 538$)</td>
<td>0.864</td>
<td>0.939</td>
<td>0.079</td>
<td>777.90</td>
<td>178</td>
<td>881.23</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>Group</td>
<td>RMSEA</td>
<td>CFI</td>
<td>TLI</td>
<td>χ²</td>
<td>df</td>
<td>p</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>----</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical young (n = 219)</td>
<td>0.914</td>
<td>0.941</td>
<td>0.067</td>
<td>195.40</td>
<td>98</td>
<td>302.47</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>Non-clinical (n = 713)</td>
<td>0.949</td>
<td>0.968</td>
<td>0.044</td>
<td>326.98</td>
<td>137</td>
<td>20.80</td>
<td>4</td>
<td>.000</td>
</tr>
<tr>
<td>Non-clinical old (n = 488)</td>
<td>0.957</td>
<td>0.971</td>
<td>0.045</td>
<td>221.47</td>
<td>111</td>
<td>17.37</td>
<td>4</td>
<td>.002</td>
</tr>
<tr>
<td>Non-clinical young (n = 225)</td>
<td>0.934</td>
<td>0.954</td>
<td>0.059</td>
<td>102.90</td>
<td>58</td>
<td>9.38</td>
<td>4</td>
<td>.052</td>
</tr>
<tr>
<td>Young (n = 444)</td>
<td>0.951</td>
<td>0.972</td>
<td>0.053</td>
<td>274.63</td>
<td>123</td>
<td>50.86</td>
<td>3</td>
<td>.000</td>
</tr>
<tr>
<td>Old (n = 1026)</td>
<td>0.915</td>
<td>0.973</td>
<td>0.063</td>
<td>1145.34</td>
<td>228</td>
<td>110.33</td>
<td>3</td>
<td>.000</td>
</tr>
</tbody>
</table>

*Note.* RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis Index.
Figure 1. The CBCL DSM-oriented 6-factor model and competing 5- and 2-factor models

6 Factor Model

5 Factor Model

Affective Problems

Negative Affect

Anxiety Problems

Somatic Problems

Somatic Problems

2 Factor Model

Internalizing Problems

Attention Deficit/Hyperactivity Problems

Attention Deficit/Hyperactivity Problems

Externalizing Problems

Conduct Problems

Conduct Problems

Oppositional Defiant Problems

Oppositional Defiant Problems