

**Comorbidity and Continuity of Depression and Conduct Problems
from Elementary School to Adolescence**

Heather McDonough-Caplan

The Ohio State University

Author note: Thank you to Dr. Theodore Beauchaine for his mentorship and Dr. Daniel Klein for his assistance with the present research findings. Research reported in this article was supported by Grant MH63699 to Theodore P. Beauchaine from the National Institute of Mental Health. Address correspondence and reprint requests to Theodore P. Beauchaine, Department of Psychology, The Ohio State University, 1835 Neil Avenue, Columbus, OH 43210; email: beauchaine.1@osu.edu.

No previous publications evaluating longitudinal symptom growth or development of heterotypic comorbidity have appeared using data from this sample. Previous publications have addressed patterns of psychopathology at entry into the study, links between maternal nicotine exposure and children's externalizing behaviors, and interactions between parenting practices and children's autonomic nervous system responding.

Abstract

Despite non-overlapping criterion sets, conduct disorder and depression co-occur at much higher rates than expected by chance. Contemporary model-based approaches to explaining *heterotypic comorbidity* use factor analysis and its variants to evaluate inter-relations among symptoms in large population-based and twin samples. These analyses invariably yield broadband internalizing and externalizing factors, which load on a higher-order general liability factor—findings that are robust across age and informant. Although model-based approaches elucidate structural aspects of comorbidity, they are variable-centered, and usually cross-sectional. Most therefore do not assess *developmental continuity* of comorbidity, or whether non-comorbid individuals are prospectively vulnerable to heterotypic comorbidity. We use an accelerated longitudinal design to evaluate growth in parent-reported conduct problems (CPs) and depression among children, ages 8-15 years, who were recruited at study entry into depressed only ($n=27$), CPs only ($n=28$), comorbid ($n=81$), and control ($n=70$) groups based on levels of symptoms. Consistent with normative developmental trends across this age range, steep growth in depression was exhibited by all groups, including those who reported only CPs at study entry. In contrast, growth in CPs was restricted to those who reported high symptoms at intake (with or without comorbid depression), compared with low and stable among depressed only and control participants. To our knowledge, this is the first study to demonstrate, using carefully ascertained “pure” versus comorbid groups who were followed naturalistically, that comorbid depression is likely to develop among those with pure CPs, but comorbid CPs are not likely to develop among those with pure depression.

Keywords: conduct disorder, depression, comorbidity, heterotypic, development

General Scientific Summary: This study indicates that children who experience conduct problems early in life are just as vulnerable as other children to developing depression in adolescence. In contrast, children who experience depression are not vulnerable to developing conduct problems in adolescence. These findings may be useful for helping us understand shared causes of different disorders, and for assigning children to appropriate treatments.

**Comorbidity and Continuity of Depression and Conduct Problems
from Elementary School to Adolescence**

Traditionally, most psychiatric disorders of childhood and adolescence have been viewed as either distinct diagnostic entities, as exemplified in the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013), or as variants along factor analytically-derived internalizing or externalizing dimensions (Achenbach & Edelbrock, 1983, 1991). When following the DSM-perspective, differential diagnosis is prioritized, which sometimes obscures common etiological mechanisms among internalizing disorders and among externalizing disorders (see e.g., Beauchaine, Zisner, & Sauder, 2017; Tackett et al., 2013). Even empirically-based taxonomies, such as the Child Behavior Checklist (Achenbach & Edelbrock), which acknowledge shared liability *within* the internalizing and externalizing spectra, can reify—however unwittingly—distinctions *between* internalizing and externalizing disorders. In recent years it has become clear that strong distinctions between and among internalizing and externalizing disorders are not always warranted (see e.g., Beauchaine & Constantino, 2017; Zisner & Beauchaine, 2016), and that *transdiagnostic* vulnerabilities to psychopathology extend both within *and* across the internalizing and externalizing spectra (e.g., Krueger, 1999; Lahey, Van Hulle, Singh, Waldman, & Rathouz, 2011; Lahey, Krueger, Rathouz, Waldman, & Zald, 2017; Tackett et al., 2013). As a result, both homotypic comorbidity (co-occurrence of *either* multiple internalizing *or* multiple externalizing disorders within individuals) and heterotypic comorbidity (co-occurrence of *both* internalizing *and* externalizing disorders within individuals) are observed at rates that far exceed those expected by chance (Angold, Costello, & Erkanli, 1999; Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Gilliom & Shaw, 2004; Keiley, Bates, Dodge, & Pettit, 2000; Kessler et al., 1994; Klein & Riso, 1993; Merikangas et al., 2010).

Although both homotypic and heterotypic comorbidity are common, etiological mechanisms are not fully understood (Beauchaine & Cicchetti, 2016). In this article, we explore longitudinal patterns of symptom development from ages 8-15 years among children who were recruited at study entry for “pure” externalizing behaviors, “pure” internalizing behaviors, comorbid internalizing and externalizing behaviors, and no psychiatric symptoms. As outlined below, comparing symptom development among contrasted groups provides certain advantages for addressing questions about shared versus specific etiologies of internalizing and externalizing behaviors. To our knowledge, this is the first study to date that examines longitudinal patterns of symptom development among contrasted groups recruited from the community and followed naturalistically. Other studies have evaluated comorbidity and continuity of internalizing and externalizing behaviors among treatment-seeking samples, which differ from more naturalistic samples in initial levels of functional impairment, symptom severity and duration, rates of comorbidity, and exposure to intervention (Goodman et al., 1997). Thus, longitudinal courses of their symptom profiles likely differ from those of untreated samples. Before describing study procedures, we first review contemporary structural accounts of comorbidity, and then consider how contrasted groups designs complement rather than compete with model-based approaches.

Hierarchical Latent Structure of Internalizing and Externalizing Behaviors

Model-based, structural approaches to characterizing psychopathology apply factor analysis and its variants to symptoms expressed among large population-based samples or twin registries. Such analyses almost invariably yield (1) a broadband internalizing factor that accounts for a preponderance of covariation among first-order syndromes including anxiety, depression, and withdrawal (this factor is sometimes parsed further into fear and distress); (2) a broadband externalizing factor that accounts for a preponderance of covariation among first-order

syndromes including impulsivity, delinquency, and overt aggression, and (3) a higher-order general liability factor that both the internalizing and externalizing factors load on, and accounts for considerable covariation between internalizing and externalizing syndromes (e.g., Lahey et al., 2012). In an era of non-replication (see Tackett et al., 2017), this bifactor latent structure is remarkably consistent, as evidenced by convergence across child, adolescent, and adult samples; male, female, and mixed samples; and self-, parent-, and teacher-reports of symptoms (e.g., Caspi et al., 2014; Lahey et al., 2011, 2014; Olino, Dougherty, Bufferd, Carlson, & Klein, 2014; Tackett et al., 2013). Furthermore, high scores on the general liability factor predict poor concurrent and prospective function across diverse settings for both children and adults (e.g., Caspi et al., 2014; Lahey et al., 2014; Martel et al., 2017).

Bifactor model-fitting—especially given such consistent results—offers several advantages for furthering our understanding of comorbidity. First, it suggests *transdiagnostic* etiological influences for internalizing and externalizing disorders. Only a decade ago, such interpretations enjoyed little support. Now, however, evidence of shared etiology is substantial, and includes (1) common genetic influences for internalizing and externalizing disorders (e.g., Cosgrove et al., 2011); (2) common genetic influences for general liability to psychopathology and negative emotionality—a personality trait that characterizes both internalizing and externalizing disorders (e.g., Tackett, Waldman, Van Hulle, & Lahey, 2011; Tackett et al., 2013); and (3) central nervous system accounts of internalizing and externalizing behavior that link both to common neural substrates of anhedonia/negative affectivity (Beauchaine & Constantino, 2017; Zisner & Beauchaine, 2016). Thus, bifactor models have yielded refinements in both theory and research, and ushered in new ways of thinking about comorbidity (Beauchaine & Zisner, 2017; Carver, Johnson, & Timpano, 2017; Forbes, Tackett, Markon, & Krueger, 2016).

Limits of Structural Models

A major strength of model-based approaches is that they provide—by reducing vast amounts of data—a single representation that best characterizes patterns of covariance among symptoms in the population. When derived from twin samples, structural models also inform us about heritability (e.g., Tackett et al., 2013). Thus, they are parsimonious, powerful statistically, and, as noted above, yield insights into etiology. Nevertheless, structural models can be insensitive to etiological heterogeneity, especially when such heterogeneity characterizes low base rate subgroups within the larger population. As reviewed by Plichta and Scheres (2014), for example, etiological mechanisms of impulsivity—a core vulnerability to externalizing disorders (Beauchaine et al., 2017; Mann et al., 2017)—may differ when the trait is manifested in normal personality versus psychopathology. In such situations, unique patterns of symptom covariance among psychopathological subgroups can be swamped by ordinary variation in symptoms (for further discussion see Shader et al., 2017). Similarly, etiologically meaningful subtypes of psychopathology may not be captured by an otherwise well-fitting structural model. For example, studies of externalizing behavior identify subsets of delinquent children and adults who display very little anxiety (see e.g., Beauchaine et al., 2017). This low base rate presentation responds poorly to treatment, and has been linked to risk for psychopathy (Frick, Ray, Thornton, & Kahn, 2014; Herpers, Scheepers, Bons, Buitelaar, & Rommelse, 2014). Thus, externalizing behaviors in the absence of internalizing symptoms—which are not well accounted for by population-based structural models—are diagnostically and prognostically meaningful.

Finally, most bifactor models conducted to date have been cross-sectional, and therefore do not evaluate continuity or change in patterns of symptoms—including homotypic continuity, heterotypic continuity, or emerging or diminishing comorbidity among individuals over time (for

exceptions see Caspi et al., 2014 and Forbes et al., 2016; Lahey et al., 2017). It is possible and perhaps even likely, for example, that some children with conduct problems never develop internalizing symptoms, and that some children with depression never develop conduct problems. Individual differences in longitudinal courses of symptoms often have implications for validating distinctions among heterogeneous subgroups, and for developing effective prevention and intervention programs (e.g., Robins & Guze, 1970). Such results would complement findings from model-based approaches, improving our understanding of etiological heterogeneity.

Comorbidity and Continuity in Longitudinal Samples

In addition to homotypic comorbidity, defined by multiple internalizing or multiple externalizing disorders within individuals (see above), homotypic *continuity*, defined by longitudinal progression of either internalizing or externalizing psychopathology across childhood and adolescence, is well documented (e.g., Beauchaine & McNulty, 2013; Cicchetti & Natsuaki, 2014; Cicchetti, Rogosch, & Toth, 1997; Hinshaw, 2015; Moffitt, 1993). Neither concurrent comorbidity nor homotypic continuity within internalizing or externalizing disorders is surprising given within-spectrum overlap of symptoms, personality characteristics, and both genetic and neural vulnerabilities (Beauchaine & Cicchetti, 2016; Tackett et al., 2011, 2013).

Notably, *heterotypic* continuity (sometimes called sequential comorbidity) between internalizing and externalizing disorders—a primary focus of this article—is also common, but is more difficult to explain given almost no overlap in symptoms. Nevertheless, both CPs and depression in childhood confer prospective vulnerability to one another in adolescence (e.g., Gilliom & Shaw, 2004; Kovacs, Paulauskas, Gatsonis, & Richards, 1988; Zoccolillo, 1992). In a recent meta-analysis of longitudinal studies comprising 17,712 children ages 12 years or below at initial assessment, Loth, Drabick, Leibenluft, and Hulvershorn (2014) found that childhood

externalizing disorders predicted adult depression in most studies, and in the sample as a whole. Among preschool children, externalizing behaviors also predict later growth in internalizing symptoms (Gilliom & Shaw, 2004). In contrast, others have demonstrated the opposite pattern, whereby internalizing disorders predict later CPs. Drabick, Gadow, and Sprafkin (2006), for example, reported that depressive symptoms predicted future conduct disorder among 6-10-year-old boys with ADHD. Early-onset internalizing and externalizing disorders predicting later onset of one another is of course expected if both arise from a common etiology (see above).

Strengths and Limitations of Existing Longitudinal Studies

To date, almost all studies that have evaluated heterotypic continuity of CPs and depression have done so using treatment-seeking samples with high levels of pre-existing comorbidity, rather than recruiting carefully ascertained comorbid and non-comorbid groups. On the one hand, comorbidity is characteristic of those who seek intervention, and may represent the bulk of those affected by internalizing and externalizing psychopathology (see above). Thus, results are likely generalizable to a large proportion of treatment-seeking people. Yet such generalizability comes at a cost. In attempting to disentangle heterotypic trajectories in symptoms among highly comorbid samples, lagged correlations (or similar statistical partialling approaches) must be used to establish prediction of CPs by depression versus prediction of depression by CPs. For example, one might determine whether depression at age 8 years predicts CPs at age 12, over-and-above CPs at age 8. Although this approach has intuitive appeal, statistical partialling is problematic whenever two disorders share a common etiology. If depression and CPs arise either partly or fully from transdiagnostic vulnerability (see Zisner & Beauchaine, 2016; Tackett et al., 2013), that very vulnerability is “partialled out” of the prediction equation. As a result, statistical partialling may obscure longitudinal relations between disorders when “controlling for” or

“partialling out” etiologically related conditions (Beauchaine, Hinshaw, & Pang, 2010). This situation is depicted in Figure 1. As reviewed by Miller and Chapman (2001), when two assumedly different disorders share symptoms, psychological vulnerabilities, and neural substrates, statistical partialling creates mathematical entities that may misrepresent/distort etiological relations between variables.

Following from this discussion, we sought to complement knowledge gleaned from structural models of comorbidity and previous longitudinal studies by evaluating—using an accelerated longitudinal design—patterns of heterotypic comorbidity and continuity in depression and CPs from ages 8-15 years among children who were recruited at study entry into depressed-only, CPs-only, heterotypically comorbid, and non-psychiatric control groups. We hypothesized that (1) those recruited for non-comorbid CPs would show continued growth in CPs, (2) those recruited for non-comorbid depression would show continued growth in depression, and (3) those recruited for pre-existing comorbidity would show continued growth in both internalizing and externalizing symptoms over time. Confirmation of these hypotheses would suggest that non-comorbid CPs and depression have potentially important prognostic value that has not been captured definitively by existing research. We included a non-psychiatric control group to ensure that any changes in symptoms could be indexed to normative growth in CPs and depression.

Method

Participants

Participants were recruited using advertisements placed in local newspapers, community publications, and city buses. To recruit the clinical groups, these advertisements described characteristics of depression and CPs, and asked interested parents to call if they felt that their child fit one or both descriptions. Separate advertisements seeking “well adjusted” children were

used to recruit non-psychiatric control participants. We received 445 inquiries from parents who completed (1) a 30-min computerized, structured phone interview including *DSM-IV-TR* (American Psychiatric Association, 2000) major depressive disorder (MDD), dysthymia (DYS), ADHD, oppositional defiant disorder (ODD), and CD subscales from the Child Symptom Inventory (CSI-4; Gadow & Sprafkin, 1997); and (2) the anxious/depressed, aggression, and attention problems subscales from the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983, 1991). Child-Symptom Inventory-4 items are assessed on 4-point scales (0 = *never*, 1 = *sometimes*, 2 = *often*, 3 = *very often*), with scores of 2 and 3 considered positive for each diagnostic criterion. Scores can also be summed dimensionally. In this study, we use dimensional scores for our primary analyses (described below), but we also report tentative diagnoses. Details regarding reliability and validity of these measures appear below. Responses were scored immediately by computer.

Based on these interviews, 212 children were assigned into the following four groups: conduct problems (CPs; $n=28$), depression (DEP; $n=27$), comorbid (CMB; $n=81$) and control (CTR; $n=70$). To be included in the CPs group, children were required to meet *DSM-IV* criteria for ODD and/or CD on the CSI, and/or score at or above the 98th percentile ($T \geq 70$) on the CBCL aggression subscale. They could not meet criteria for depression or dysthymia, and were required to score $T \leq 60$ on the CBCL anxious/depressed subscale. To be included in the DEP group, children were required to meet *DSM-IV* criteria for major depression or dysthymia on the CSI, and/or score at or above the 98th percentile ($T \geq 70$) on the CBCL anxious/depressed subscale. They could not meet criteria for CD or ODD, and were required to score $T \leq 60$ on the CBCL aggression subscale. Children in the CMB group were required to meet criteria for both the CPs and DEP groups, with no CBCL exclusions. Finally, children in the CTR group had to be free of

psychiatric diagnoses on all CSI scales, and score $T \leq 60$ on all CBCL subscales. Additional exclusion criteria included symptoms of psychosis, autism, or intellectual disability, as assessed during the structured phone interview. Of the 212 qualifying participants, 6 dropped out early and did not attend any lab sessions (described below). The final sample was therefore comprised of 206 children, including 134 boys and 72 girls. Racial composition was 61.7% Caucasian, 12.1% African American, 10.2% Hispanic, 6.3% Asian American, (2.4% Pacific Islander, 1.5% American Indian. The remaining 5.8% either entered a race other than those listed above, or declined to respond. Annual income averaged \$50,550, which falls slightly below median US household income, but nearly \$30,000 below median Seattle household income. Of note, \$72,000 is the low income threshold for a family of four in King County, WA (US Department of Housing and Urban Development, 2017).

Procedure

All procedures were approved by the local institutional review board (approval number 01-358), and researchers obtained parental consent and child assent prior to each yearly lab session (see below). Following the structured phone interview in which parents completed the CSI-4 and CBCL, those who met inclusion criteria were invited to the lab for their first of three annual assessments (Year 1). As described in detail elsewhere (e.g., Kopp & Beauchaine, 2007), these assessments included collection of additional questionnaires and participation in a broader lab protocol that is not relevant to the current paper (see Beauchaine, Hong, & Marsh, 2008; Shannon, Beauchaine, Brenner, Neuhaus, & Gatzke-Kopp, 2007; Vasilev et al., 2009). Similar assessments were conducted at Years 2 and 3, when CSI-4, CBCL, and laboratory tasks were again administered. Parents were compensated \$75 for participation at each annual assessment.

Measures

Child Symptom Inventory for *DSM-IV* (CSI-4; Gadow & Sprafkin, 1997). As described above, the CSI-4 assesses a wide range of *DSM-IV* psychiatric disorders of childhood. Parents completed the CD, ODD, ADHD, MDD, and dysthymia modules at each time point. Sensitivities and specificities vis-à-vis clinical diagnoses for the CSI scales used in this study range from .73 to .83, suggesting adequate validity (Gadow & Sprafkin, 1997). Reliabilities, as assessed via Cronbach's α coefficients calculated at each yearly assessment ranged from .85-.90.

Child Behavior Checklist (CBCL; Achenbach, & Edelbrock, 1991). At Year 1, parents completed the anxious/depressed, aggression, and attention problems subscales of the CBCL for group assignment purposes (see above). Cronbach's α s ranged from .90-.93.

Analyses

Among the 206 children who were assigned to groups, 28 (13.6%) dropped out before Year 2, and 20 (9.7%) dropped out before Year 3. Participants who dropped out scored higher on all CSI and CBCL scales. Rather than omitting these participants, which would have introduced bias into outcomes and analyses, we imputed their data in SPSS 24¹. Analyses were conducted across 30 imputations, according to established guidelines (see Graham, 2009).

We assessed growth in depressive symptoms by constructing multilevel models in Hierarchical Linear Modeling software, version 6.08 (HLM; Raudenbush, Bryk, & Congdon, 2004). Within-participant change in parent-reported depressive symptoms was modelled across age rather than timepoint at Level 1. This enabled us to evaluate symptom trajectories across the entire 8-year age span. As with all accelerated longitudinal designs, the sample size was

¹Debate has existed for some time in the literature regarding use of multiple imputation (MI) versus full-information maximum likelihood (FIML) estimation to replace missing data (e.g., Shin, Davison, & Long, 2017). Structural equation modeling programs often use FIML, whereas many statistics packages use MI (SAS provides both options). Although FIML sometimes outperforms MI when the number of imputations (m) is small, the procedures are equivalent when m is large (Graham, Olchowski, & Gilreath, 2007). Accordingly, we used 30 imputations (as few as 5 are typical). With large numbers of imputations, Monte Carlo studies show comparable parameter estimates for MI and FIML under a wide range of missing data conditions (Dong & Peng, 2013).

somewhat smaller at age extremes ($n = 51$ at age 8, $n = 81$ at age 9, $n = 123$ at age 10, $n = 111$ at age 11, $n = 117$ at age 12, $n = 73$ at age 13, $n = 47$ at age 14, and $n = 13$ at age 15). Age at Year 1 was entered at Level 2 to control for individual differences at entry into the accelerated longitudinal design. Group differences in growth of depression were assessed by entering orthogonal contrast codes at Level 2². Contrast 1 (C1) compared the CTR group to all clinical groups, Contrast 2 (C2) compared the CMB group to the DEP and CPs groups, and Contrast 3 (C3) compared the DEP and CPs groups. Sex effects were evaluated as a possible covariate at Level 2 given well-documented sex differences in growth of depression across this age range (e.g., Nolen-Hoeksema, 1987; Kendler & Gardner, 2014). The full model for DSM depressive symptoms was as follows:

$$\text{Level 1: depression/dysthymia symptoms}_{S_{ti}} = \pi_{0i} + \pi_{1i} * (\text{age}_{ti}) + e_{ti}$$

$$\text{Level 2: } \pi_{0i} = \beta_{00} + \beta_{01} * (\text{sex}_i) + \beta_{02} * (\text{C1}_i) + \beta_{03} * (\text{C2}_i) + \beta_{04} * (\text{C3}) + \beta_{05} * (\text{age}_{\text{Time1}}) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} * (\text{sex}_i) + \beta_{12} * (\text{C1}_i) + \beta_{13} * (\text{C2}_i) + \beta_{14} * (\text{C3}_i) + \beta_{15} * (\text{age}_{\text{Time1}}) + r_{1i}$$

Note that we specify depression/dysthymia symptoms together since *DSM-IV* dysthymia symptoms and *DSM-IV* major depression symptoms are largely overlapping. We therefore combined modules. Next, to evaluate growth in CPs, we ran a parallel model with DSM conduct disorder (CD) symptoms as the outcome. For both sets of analyses, we restricted outcomes to symptoms of DSM disorders (depression, CD), since almost all of the existing literature on heterotypic comorbidity and continuity has done so. This enabled us to compare results to previously reported findings. Finally, we evaluated growth in ODD symptoms, as requested by an anonymous reviewer.

Results

²Orthogonal contrast codes were constructed to ensure that all contrasts were independent, and to control for size differences between groups. Orthogonal coding provides numerous statistical advantages over other approaches, including lower probability of Type I error (see e.g., Pedhazur, 1997).

Descriptive Statistics

Descriptive statistics at study entry are reported by group in Table 1. As described in detail elsewhere (e.g., Vasilev et al., 2009), recruitment was effective in yielding groups with significant differences of large effect size across all measures of psychopathology, all F s (3,203) ≥ 23.8 , all p s $<.001$, all $\eta^2 \geq .23$.

Growth in (1) depressive/dysthymic symptoms and (2) CD and ODD symptoms across Years 1-3 are reported in Tables 2 and 3, respectively. Table 2 includes mean symptom levels and percentages of those who met *DSM-IV* criteria for MDD and/or dysthymia at each assessment, and Table 3 includes mean levels of and percentages of those who met *DSM-IV* criteria for CD and at each assessment. Significant symptoms of both depression and CD persisted across time, as reported in detail below.

Results from all of the multilevel models evaluating growth (slopes) in depressive symptoms, CD symptoms, and ODD symptoms are summarized in Table 4. Uncentered intercept effects are also reported. Given the complexity of the three models, we discuss specific findings for depressive symptoms as outcomes (top third of Table 4), CD symptoms as outcomes (middle third of Table 4), and ODD symptoms as outcomes (bottom third of Table 4) separately below.

Growth in Depression

In a preliminary HLM model without contrast codes, age predicted Level-1 slopes in parent-reported depressive symptoms, $b = 0.59$, $t(204) = 6.22$, $p < .001$. Thus, symptoms of depression increased over time sample-wide. Given the age range of the sample, normative growth in depression should be expected (e.g., Garber, Keiley, & Martin, 2002). Interestingly, however, when entered as a Level 2 fixed effect, sex did not predict growth in depressive symptoms, $b = -0.035$, $t(200) = -1.302$, $p = .195$. This is not entirely surprising given that so many participants

were recruited specifically for having high depression scores—regardless of sex.

Level 2 orthogonal contrasts comparing growth in depression across groups were evaluated next. The contrast comparing the non-psychiatric control group to all three clinical groups (C1) indicated steeper growth in depressive symptoms among *controls* than among the three clinical groups, $b = -0.0003$, $t(200) = -2.580$, $p = .011$. Growth in depression for all groups is depicted in Figure 1. As this finding indicates, children who were free from depression at study entry showed marked growth in symptoms from middle school to adolescence. Again, such findings are expected in this age range (e.g., Garber et al., 2002). Nevertheless, controls' overall level of depression was lower across the assessment period, as indicated by a significant mean-centered intercept for C1, $b = 0.003$, $t(201) = 10.44$, $p < .001$ (this intercept is not reported in Table 4, which lists uncentered intercepts).

The next contrast, which compared the CMB group to the CP and DEP groups (C2), approached significance, $b = 0.0004$, $t(200) = 1.900$, $p = .058$. This indicates steeper growth in depressive symptoms among the CP and DEP groups than among the CMB group. As Figure 1 illustrates, this finding is almost certainly attributable to a ceiling effect, since CMB participants were the most symptomatic across the entire age range.

The final contrast (C3), which compared the CP and DEP groups, was not significant, $b = -0.001$, $t(200) = -1.186$, $p = .238$. Thus, there was no difference in growth in depression between children with CPs only at study entry vs. those with depression/dysthymia only. Notably, however, both groups exhibited steep increases in depressive symptoms from ages 8-15 years.

Finally, to confirm that CPs only were associated with development of depression, we evaluated the slope effect for the CPs only group. This analysis indicated significant growth in depression from ages 8 to 15 among those recruited for CPs only, $b = 0.80$, $t(26) = 3.62$, $p = .001$

(this is not reported in Table 4). Thus, CPs alone were associated with growth in depression that was similar to that observed in the depression only group.

Growth in Conduct Problems

Next, we assessed group differences in parent-reported CD symptom growth from ages 8-15 years. As with depressive symptoms, age predicted slopes in CD symptoms at Level 1 in a preliminary model without contrast codes, $b = 0.39$, $t(205) = 7.55$, $p < 0.001$. Thus, CD symptoms increased over time sample-wide. As with depression, this is not entirely surprising given normative developmental increases in CD symptoms across this age range—particularly adolescent-onset symptoms (e.g., Moffitt, 1993). When entered as a Level 2 fixed effect, sex did not predict growth in CD symptoms, $b = -0.196$, $t(200) = -1.192$, $p = .235$. This is not unexpected given that many participants were recruited based on CD symptoms, regardless of sex.

The contrast comparing non-psychiatric controls to all clinical groups (C1) indicated steeper growth in CD symptoms by the control group than the three clinical groups, $b = -0.002$, $t(200) = -2.932$, $p = .004$. Growth in CD symptoms for all groups is depicted in Figure 2. Despite sample-wide growth, CD symptoms were lower overall among non-psychiatric controls than the clinical groups, as confirmed by a significant mean-centered intercept of C1, $b = -0.002$, $t(201) = -2.71$, $p = .008$ (not reported in Table 4). The next contrast, which compared the CMB group to the CP and DEP groups (C2), was also significant, $b = 0.003$, $t(200) = 2.030$, $p = .043$. Participants in the CP and DEP groups showed steeper growth in CD symptoms than those in the CMB group (see Figure 2).

The final contrast (C3), which compared the CPs and DEP groups, was not significant, $b = 0.007$, $t(200) = 1.296$, $p = .197$. Thus, the CPs and DEP groups showed similar growth *rates* in CD symptoms. However, the CPs group exhibited much higher levels of CD symptoms across

the entire age range than the DEP group, which exhibited very few symptoms (see Figure 2).

This difference in overall levels of CD symptoms was confirmed by a significant mean-centered intercept for C3, $b = -0.15$, $t(200) = -2.84$, $p = .006$ (not reported in Table 4)³.

To assess whether depression alone was associated with development of CPs, we evaluated the slope effect for the depression only group (not reported in Table 4). *This analysis indicated no growth in CPs from ages 8 to 15 years among those recruited for depression only*, $b = -0.74$, $t(25) = -0.86$, $p = .40$. Thus, depression alone was not associated with growth in CPs. We also evaluated the slope effect for controls only (not reported in Table 4), who also exhibited no growth in CD symptoms, $b = -0.19$, $t(68) = -0.83$, $p = .41$. We discuss these findings in greater detail below.

Growth in Oppositionality

Finally, we evaluated growth in ODD symptoms across groups. Doing so was important for at least two reasons. First, ODD often emerges before CD in longitudinal progressions of externalizing behavior (see Beauchaine & McNulty, 2013; Beauchaine et al., 2017). Thus, ODD symptoms may portend vulnerability to CD, and may be a more sensitive indicator of emerging externalizing problems than CD symptoms. Second, ODD is associated with high levels of irritability and negative emotionality, and predicts development of depression (see Burke, Hipwell, & Loeber, 2010; Burke & Loeber, 2010; Herzhoff & Tackett, 2015). Thus, ODD portends vulnerability to *both* CD and depression, and is marked by the very personality trait—persistent negative emotionality—that accounts for common heritable variance in internalizing and externalizing behaviors (Tackett et al., 2013) and derives from well-articulated neural

³At the request of the action editor, we also evaluated growth in aggressive vs. non-aggressive symptoms of CD separately. In these analyses, the CSI-4 CD subscale was partitioned into aggressive (*DSM-IV* symptoms 1-7) and non-aggressive (*DSM-IV* symptoms 8-15) criteria. No significant group differences in slopes emerged from either analysis. Thus, aggressive and non-aggressive symptoms of CD demonstrated additive rather than independent effects.

substrates (Zisner & Beauchaine, 2016).

As with depressive and CD symptoms, age predicted slopes in ODD symptoms at Level 1 in a preliminary model without contrasts codes, $b = 2.88$, $t(204) = 3.11$, $p = .002$. Thus, ODD symptoms increased over time sample-wide. The Level 2 sex effect was not significant, $b = 0.018$, $t(200) = 0.069$, $p = .946$. Once again, this is not surprising since so many participants were recruited based on ODD symptoms, regardless of sex.

The contrast comparing control participants to all clinical groups (C1) indicated steeper growth in ODD symptoms among non-psychiatric controls than the three clinical groups, $b = -0.005$, $t(200) = -4.282$, $p < .001$. Growth in ODD symptoms for all groups is depicted in Figure 3. Importantly, despite their steeper slope, controls exhibited overall lower ODD symptoms than the clinical groups, as indicated by a significant mean-centered intercept for C1, $b = 0.03$, $t(201) = 12.50$, $p < .001$ (not reported in Table 4). The next contrast (C2), which compared the CMB group to the CP and DEP groups, was not significant, $b = 0.004$, $t(200) = 1.923$, $p = .056$.

Similarly, the final contrast (C3), which compared the CP and DEP groups, was not significant, $b = 0.013$, $t(200) = 1.661$, $p = .098$. Thus, both groups showed similar growth rates in ODD symptoms. Notably, however, the mean-centered intercept for C3 was significant, indicating greater ODD symptoms for the CPs group than the DEP group from ages 8-15, $b = -0.26$, $t(200) = -3.04$, $p = .003$ (not reported in Table 4).

We ran one additional, unplanned contrast to determine whether growth in ODD symptoms differed for the DEP only versus CTR groups. The slope parameter was not significant, $b = -0.007$, $t(203) = -1.65$, $p = .10$. Thus, growth in ODD symptoms was indistinguishable between depressed participants and non-psychiatric controls.

Discussion

We sought to disentangle—to the extent possible with parent-report data—whether early depression confers vulnerability to later CD, early CPs confer vulnerability to later depression, or both. Although our findings do not answer these questions fully, they do help to clarify the existing literature. Below we consider findings specific to depression, CD, and ODD symptoms in turn, before presenting a more general discussion.

First, similar growth rates in depression were observed for all groups, including those recruited for CPs only and no psychiatric symptoms at study entry. As noted above, growth in depression is expected in this age range (Garber et al., 2002; Nolen-Hoeksema, 1987; Kendler & Gardner, 2014), so these findings aren't entirely surprising—except perhaps for the CPs-only group. Several authors have proposed that CPs in the absence of internalizing symptoms mark a more virulent, treatment-resistant phenotype that portends vulnerability to future psychopathy (e.g., Frick et al., 2014; Herpers et al., 2014). This perspective is often rooted in trait-based assumptions that would not predict longitudinal increases in depression among children recruited for CD-only. Our findings suggest that CPs without depression do not identify a subgroup of children who are invulnerable to heterotypic comorbidity and continuity. In all likelihood, assessment of additional constructs—most notably callous-unemotional traits—would be required to identify such individuals (e.g., Golmaryami & Frick, 2017). Our findings indicate that children with “pure” CPs are just as vulnerable to depression as other high-risk groups.

It is also notable that (a) children recruited for high levels of depression (98th percentile vis-à-vis national norms) continued to show worsening symptoms over time, and (b) CMB participants showed the highest levels of symptoms across the entire age range. The latter finding in particular suggests that heterotypic comorbidity is unlikely to relent from ages 8-15. It is troubling that depressed adolescents with comorbid CD are especially vulnerable to suicidal

behaviors (Nock, Hwang, Sampson, & Kessler, 2010). Although we did not measure suicidal behavior in this study, our findings of continuity in (and perhaps amplification of) heterotypic comorbidity suggest that early diagnosis may be useful for prevention.

Longitudinal trajectories in CD symptoms were more differentiated. Although sample-wide increases in symptoms were observed, these increases were driven entirely by those in the CPs and CMB groups, who, despite being recruited for high levels of CD symptoms, developed worsening externalizing behavior over time. In contrast, those in the DEP-only and non-psychiatric control groups exhibited no growth in CD symptoms from ages 8-15. This finding indicates that pure depression does not confer vulnerability to CPs, thereby clarifying a longstanding controversy in the literature. Importantly, this question could not have been resolved by applying statistical control techniques to highly comorbid samples (Beauchaine et al., 2010; Miller & Chapman, 2001). Such studies suggest that CD portends vulnerability to later depression and that depression portends vulnerability to later CPs (e.g., Drabick et al., 2006; Gilliom & Shaw, 2004; Loth et al., 2014). By using a contrasted groups design, we demonstrated that pure CPs follow a heterotypic developmental pathway, whereas pure depression does not. We note, however, that both depression and CD are dimensional constructs in nature, so we had to impose relatively strict yet arbitrary cutoffs to create these “pure” groups.

Oppositional symptoms demonstrated an “intermediate” growth pattern. Similar to depressive symptoms, increases in ODD were observed across the sample, despite recruitment based on high (CPs, CMB) versus low (DEP, CTR) symptoms at study entry. Yet similar to CD symptoms, lower levels of ODD were observed in the control and DEP groups across the entire age range. Furthermore, growth in ODD symptoms was similar among DEP and non-psychiatric control participants. This indicates age-normative increases in ODD symptoms rather than

pathological growth trajectories among those with pure depression. Thus, those with pure depression exhibited normal growth patterns in *both* ODD and CD.

In considering findings collectively, heterotypically comorbid participants appear to be most impaired—as expected given our recruitment strategy. Heterotypically comorbid participants showed (1) persistently high rates of CD symptoms across elementary school to mid-adolescence, (2) growth in depression across the same time span, and (3) the highest rates of depressive symptoms at every age evaluated. These children and adolescents therefore appear to be worse off than any of their peers—at least in terms of symptoms. This is consistent with the existing literature, which indicates that individuals with comorbid depression and CD show more severe symptoms of both disorders than individuals with either disorder alone (e.g. Ezpeleta, Domenech, & Angold, 2006; Marmorstein & Iacono, 2003).

It is important to note, however, that symptoms are not the only metric for evaluating impairment. For example, comorbid internalizing symptoms predict better responses to behavioral treatments for CD (Beauchaine, Webster-Stratton, & Reid, 2005), as well as less physical aggression, better peer ratings, and fewer police contacts (Walker et al., 1991). Moreover, comorbid internalizing symptoms confer partial protection from structural compromises in several brain structures among those with CD (Sauder, Beauchaine, Gatzke-Kopp, Shannon, & Aylward, 2012).

Even control participants—who were recruited for clear absence of symptoms—showed steep increases in depression from ages 8-15. In fact, by the Year 3 assessment, 29% met criteria for dysthymia or depression based on parent reports. Although it might be tempting to explain these findings away as a result of biased parent reporting or some other artifact, sensitivity and specificity of the CSI-4 mood disorder scales are strong, and the observed rate of depression,

although higher than that reported in the National Comorbidity Survey Replication Adolescent Supplement (Merikangas et al., 2010), is very close to that found in the National Longitudinal Study of Adolescent Health (Rushton et al., 2002). Thus, although growth in depression among non-psychiatric controls is concerning it is not atypical.

Although CD symptoms expressed by the CPs group seemed to decline at age 14 (see Figure 3), we purposefully do not interpret this for several reasons. First, sample size in any accelerated longitudinal study is smallest at age extremes. Furthermore, previous research shows that (1) early-onset CPs are unlikely to abate by mid-adolescence (e.g., Moffitt, 1993), and (2) parents know less about their children's externalizing behaviors in adolescence than in middle school (see De Los Reyes & Kazdin, 2005). Unfortunately, this sample is no longer being followed so we cannot address this issue with more data.

Limitations

A primary limitation of this study is sole reliance on parent-report data using symptom checklists. As noted above, both the CSI and CBCL demonstrate strong reliability and validity. Nevertheless, structured clinical interviews would have strengthened our findings. Moreover, although parents become slightly poorer informants as their children age (see above; De Los Reyes & Kazdin, 2005; Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985), this would cause under-reporting of symptoms over time, and therefore cannot explain increases in depression seen in all groups. We also relied on a single informant—a further limitation.

In addition, with three time points, we could not assess non-linear trends in emerging CD or depression given our objective of evaluating within-person growth trajectories. Although most growth trajectories appear to be near-linear (see Figs. 2-4), quadratic analyses may have been especially useful to assess patterns of growth in both CD and depression. Furthermore, since we

used an accelerated longitudinal design, there were relatively few participants at age extremes. We should be careful, for example, in making inferences about non-linear growth based on apparent leveling off of depression among the CD group (see Figure 2).

Conclusion

Depression and CD co-occur significantly more often than expected by chance, which may suggest shared etiology. In attempts to elucidate etiology and devise more effective treatments, several research groups have sought to determine which disorder precedes the other. To date, most analyses have evaluated lagged associations across time points in attempts to establish temporal precedence of CD versus depression via statistical control. As we note above however, statistical partialling has several limitations when disentangling comorbidity and continuity in psychopathology. In this study, we recruited relatively ‘pure’ groups who were vulnerable to CD, depression, or both based on high levels of symptoms. Our findings indicate that CPs alone are associated with normative increases in depression throughout childhood and adolescence. In contrast, depression alone is not associated with development of CPs. Notably, however, all groups showed steep growth in depressive symptoms from ages 8-15 years. We hope future research elucidates specific mechanisms of shared vulnerability to CPs and depression—both biological and psychological (e.g., Carver et al., 2017; Tackett et al., 2013; Zisner & Beauchaine, 2016)—as we seek to develop more effective prevention and intervention programs.

References

- Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the CBCL and Revised Child Behavior Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., & Edelbrock, C. S. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont Press.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Press.
- Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry*, *40*, 57-87.
- Beauchaine, T. P., & Cicchetti, D. (2016). A new generation of comorbidity research in the era of neuroscience and the Research Domain Criteria. *Development and Psychopathology*, *28*, 891-894. doi:10.1017/S0954579416000602
- Beauchaine, T. P., & Constantino, J. N. (2017). Redefining the endophenotype concept to accommodate transdiagnostic vulnerabilities and etiological complexities. *Biomarkers in Medicine*, *11*, 769-780. doi:10.221/bmm-2017-0002
- Beauchaine, T. P., Hinshaw, S. P., & Pang, K. L. (2010). Comorbidity of ADHD and early-onset conduct disorder: Biological, environmental, and developmental mechanisms. *Clinical Psychology: Science and Practice*, *17*, 327-336. doi:10.1111/j.1468-2850.2010.01224.x
- Beauchaine, T. P., Hong, J., & Marsh, P. (2008). Sex differences in autonomic correlates of conduct problems and aggression. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*, 788-796. doi:10.1097/CHI.0b013e318172ef4b
- Beauchaine, T. P., & McNulty, T. (2013). Comorbidities and continuities as ontogenic processes: Toward a developmental spectrum model of externalizing behavior. *Development and Psychopathology*, *25*, 1505-1528. doi:10.1017/S0954579413000746
- Beauchaine, T. P., Webster-Stratton, C., & Reid, M. J. (2005). Mediators, moderators, and predictors of one-year outcomes among children treated for early-onset conduct problems: A latent growth curve analysis. *Journal of Consulting and Clinical Psychology*, *73*, 371-388. doi:10.1037/0022-006X.73.3.371
- Beauchaine, T. P., & Zisner, A. (2017). Motivation, emotion regulation, and the latent structure of psychopathology: An integrative and convergent historical perspective. *International Journal of Psychophysiology*, *119*, 108-118. doi:10.1016/j.ijpsycho.2016.12.014
- Beauchaine, T. P., Zisner, A., & Sauder, C. L. (2017). Trait impulsivity and the externalizing spectrum. *Annual Review of Clinical Psychology*, *13*, 343-368. doi:10.1146/annurev-clinpsy-021815-093253.
- Burke, J. D., Hipwell, A. E., & Loeber, R. (2010). Dimensions of oppositional defiant disorder as predictors of depression and conduct disorder in preadolescent girls. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*, 484-492. doi:10.1016/j.jaac.2010.01.016
- Burke, J. D., & Loeber, R. (2010). Oppositional defiant disorder and the explanation of the comorbidity between behavioral disorders and depression. *Clinical Psychology Science and Practice*, *17*, 319-326. doi:10.1111/j.1468-2850.2010.01223.x
- Carver, C. S., Johnson, S. L., & Timpano, K. R. (2017). Toward a functional view of the p factor in psychopathology. *Clinical Psychological Science*, *5*, 880-889. doi:10.1177/2167702617710037

- Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., ... Moffitt, T. E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2*, 119-137. doi:10.1177/2167702613497473
- Cicchetti, D., & Natsuaki, M. N. (2014). Multilevel developmental perspectives toward understanding internalizing psychopathology: Current research and future directions. *Development and Psychopathology, 26*, 1189-1190. doi:10.1017/S0954579414000959
- Cicchetti, D., Rogosch, F. A., & Toth, S. L. (1997). Ontogenesis, depressotypic organization, and the depressive spectrum. In S. S. Luthar, J. Burack, D. Cicchetti, & J. Weisz (Eds.), *Developmental psychopathology: Perspectives on adjustment, risk, and disorder* (pp. 273-313). New York, NY: Cambridge University Press.
- Cosgrove, V. E., Rhee, S. H., Gelhorn, H. L., Boeldt, D., Corley, R. C., Ehringer, M. A., ... Hewitt, J. K. (2011). Structure and etiology of co-occurring internalizing and externalizing disorders in adolescents. *Journal of Abnormal Child Psychology, 39*, 109-123. doi:10.1007/s10802-010-9444-8
- Costello, E. J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry, 60*, 837-844. doi:10.1001/archpsyc.60.8.837
- De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483-509. doi:10.1037/0033-2909.131.4.483
- Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus, 2*, 222. doi:10.1186/2193-1801-2-222
- Drabick D. A., Gadow K. D., & Sprafkin J. (2006). Co-occurrence of conduct disorder and depression in a clinic-based sample of boys with ADHD. *Journal of Child Psychology and Psychiatry, 47*, 766-774. doi:10.1111/j.1469-7610.2006.01625.x
- Edelbrock, C., Costello, A. J., Dulcan, M. K., Kalas, R., & Conover, N. C. (1985). Age differences in the reliability of the psychiatric interview of the child. *Child Development, 56*, 265-275. doi:10.2307/1130193
- Ezpeleta, L., Dome`nech, J. M., & Angold, A. (2006). A comparison of pure and comorbid CD/ODD and depression. *Journal of Child Psychology and Psychiatry, 47*, 704-712. doi:10.1111/j.1469-7610.2005.01558.x
- Forbes, M. K., Tackett, J. L., Markon, K. E., & Krueger, R. F. (2016). Beyond comorbidity: Toward a dimensional and hierarchical approach to understanding psychopathology across the life span. *Development and Psychopathology, 28*, 971-986. doi:10.1017/S0954579416000651
- Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin, 140*, 1-57. doi:10.1037/a0033076
- Gadow, K. D., & Sprafkin, J. (1997). *CSI 4 norms manual*. Stony Brook, NY: Checkmate Plus.
- Garber, J., Keiley, M. K., & Martin, N. C. (2002). Developmental trajectories of adolescents' depressive symptoms: Predictors of change. *Journal of Consulting and Clinical Psychology, 70*, 79-95. doi:10.1037/0022-006X.70.1.79
- Gilliom, M., & Shaw, D. S. (2004). Codevelopment of externalizing and internalizing problems in early childhood. *Development and Psychopathology, 16*, 313-333.

doi:10.1017/S0954579404044530

- Golmaryami, F. N., & Frick, P. J. (2017). Callous-unemotional traits and the development of externalizing spectrum disorders. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (pp. 360-374). NY: Oxford.
- Goodman, S. H., Lahey, B. B., Fielding, B., Dulcan, M., Narrow, W., & Regier, D. (1997). Representativeness of clinical samples of youths with mental disorders: A preliminary population-based study. *Journal of Abnormal Psychology, 106*, 3-14. doi:10.1037/0021-843X.106.1.3
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549-576. doi:10.1146/annurev.psych.58.110405.085530
- Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*, 206-213. doi:10.1007/s11121-007-0070-9
- Herpers, P., Scheepers, F., Bons, D., Buitelaar, J., & Rommelse, N. (2014). The cognitive and neural correlates of psychopathy and especially callous-unemotional traits in youths: A systematic review of the evidence. *Development and Psychopathology, 26*, 245-273. doi:10.1017/S0954579413000527
- Herzhoff, K., & Tackett, J. L. (2015). Subfactors of oppositional defiant disorder: Converging evidence from structural and latent class analyses. *Journal of Child Psychology and Psychiatry, 57*, 18-29. doi:10.1111/jcpp.12423
- Hinshaw, S. P. (2015). Developmental psychopathology, ontogenic process models, gene-environment interplay, and brain development: An emerging synthesis. *Journal of Abnormal Psychology, 124*, 771-775. doi:10.1037/abn0000110.
- Keiley, M. K., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2000). A crossdomain growth analysis: Externalizing and internalizing behaviors during 8 years of childhood. *Journal of Abnormal Child Psychology, 28*, 161-179. doi:10.1023/A:1005122814723
- Kendler, K. S. & Gardner, C. O. (2014). Sex differences in the pathways to major depression: A study of opposite-sex twin pairs. *American Journal of Psychiatry, 171*, 426-435. doi:10.1176/appi.ajp.2013.13101375
- Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., ...Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States. *Archives of General Psychiatry, 51*, 8-19. doi:10.1001/archpsyc.1994.03950010008002
- Klein, D. N., & Riso, L. P. (1993). Psychiatric disorders: Problems of boundaries and comorbidity. In C. G. Costello (Ed.), *Basic issues in psychopathology* (pp. 19-66). NY: Guilford.
- Kopp, L. M. & Beauchaine, T. P. (2007). Patterns of psychopathology in the families of children with conduct problems, depression, and both psychiatric conditions. *Journal of Abnormal Child Psychology, 35*, 301-312. doi:10.1007/s10802-006-9091-2.
- Kovacs, M., Paulauskas, S., Gatsonis, C., & Richards, C. (1988). Depressive disorders in childhood III. A longitudinal study of comorbidity with and risk for conduct disorders. *Journal of Affective Disorders, 15*, 205-217. doi:10.1016/0165-0327(88)90018-3
- Krueger, R. F. (1999). The structure of common mental disorders. *Archives of General Psychiatry, 56*, 921-926. doi:10.1001/archpsyc.58.6.597
- Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology, 121*, 971-977. doi:10.1037/a0028355

- Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2017). A hierarchical causal taxonomy of psychopathology across the life span. *Psychological Bulletin*, *143*, 142-186. doi:10.1037/bul0000069
- Lahey, B. B., Rathouz, P. J., Keenan, K., Stepp, S. D., Loeber, R., & Hipwell, A. E. (2014). Criterion validity of the general factor of psychopathology in a prospective study of girls. *Journal of Child Psychology and Psychiatry*, *56*, 415-422. doi:10.1111/jcpp.12300
- Lahey, B. B., Van Hulle, C. A., Singh, A. L., Waldman, I. D., & Rathouz, P. J. (2011). Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. *Archives of General Psychiatry*, *68*, 181-189. doi:10.1001/archgenpsychiatry.2010.192
- Loth, A. K., Drabick, D. A., Leibenluft, E., & Hulvershorn, L. A. (2014). Do childhood externalizing disorders predict adult depression? A meta-analysis. *Journal of Abnormal Child Psychology*, *42*, 1103-1113. doi:10.1007/s10802-014-9867-8
- Mann, F. D., Engelhardt, L., Briley, D. A., Grotzinger, A. D., Patterson, M. W., Tackett, J. L., ...Harden, P. (2017). Sensation seeking and impulsive traits as personality endophenotypes for antisocial behavior: Evidence from two independent samples. *Personality and Individual Differences*, *105*, 30-39. doi:10.1016/j.paid.2016.09.018
- Marmorstein, N. R., & Iacono, W. G. (2003). Major depression and conduct disorder in a twin sample: Gender, functioning, and risk for future psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 225-233. doi:10.1097/00004583-200302000-00017
- Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., ...Bressan, R. A. (2017). A general psychopathology factor (p factor) in children: Structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal Psychology*, *126*, 137-148. doi:10.1037/abn0000205
- Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., ...Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication—Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*, 980-989. doi:10.1016/j.jaac.2010.05.017
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*, 40-48. doi:10.1037/0021-843X.110.1.40
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674-701. doi:10.1037/0033-295X.100.4.674
- Nock, M. K., Hwang, I., Sampson, N. A., & Kessler, R. C. (2010). Mental disorders, comorbidity and suicidal behavior: Results from the National Comorbidity Survey Replication. *Molecular Psychiatry*, *15*, 868-876. doi:10.1038/mp.2009.29
- Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin*, *101*, 259-282. doi:10.1037//0033-2909.101.2.259
- Olino, T.M., Dougherty, L.R., Bufferd, S.J., Carlson, G.A., & Klein, D.N. (2014). Testing models of psychopathology in preschool-aged children using a structured interview-based assessment. *Journal of Abnormal Child Psychology*, *42*, 1201-1211. doi:10.1007/s10802-014-9865-x.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). San Diego, CA: Harcourt Brace.

- Plichta, M. M., & Scheres, A. (2014). Ventral-striatal responsiveness during reward anticipation in ADHD and its relation to trait impulsivity in the healthy population: A meta-analytic review of the fMRI literature. *Neuroscience and Biobehavioral Reviews*, *38*, 125–134. doi:10.1016/j.neubi.2014.05.002
- Raudenbush, S., Bryk, A., & Congdon, R. (2004). *HLM 6: Hierarchical linear and non-linear modeling*. Lincolnwood, IL: Scientific Software International.
- Robins, E., & Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *American Journal of Psychiatry*, *126*, 983-987. doi:10.1176/ajp.126.7.983
- Rushton, J. L., Forcier, M., & Schectman, R. M. (2002). Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*, 199-205. doi:10.1097/00004583-200202000-00014
- Sauder, C., Beauchaine, T. P., Gatzke-Kopp, L. M., Shannon, K. E., & Aylward, E. (2012). Neuroanatomical correlates of heterotypic comorbidity in externalizing male adolescents. *Journal of Clinical Child and Adolescent Psychology*, *41*, 346-352. doi:10.1080/15374416.2012.658612
- Shader, T. M., Gatzke-Kopp, L. M., Crowell, S. E., Reid, M. J., Thayer, J. F., Vasey, M. W., ...Beauchaine, T. P. (2017). Quantifying respiratory sinus arrhythmia: Effects of misspecifying breathing frequencies across development. *Development and Psychopathology*. Epublshed ahead of print. doi:10.1017/S0954579417000669
- Shannon, K. E., Beauchaine, T. P., Brenner, S. L., Neuhaus, E., & Gatzke-Kopp, L. (2007). Familial and temperamental predictors of resilience in children at risk for conduct disorder and depression. *Development and Psychopathology*, *19*, 701-727. doi:10.1017/S0954579407000351
- Shin, T., Davison, M. L., & Long, J. D. (2017). Maximum likelihood versus multiple imputation for missing data in small longitudinal samples with nonnormality. *Psychological Methods*, *22*, 426-449. doi:10.1037/met0000094
- Tackett, J. L., Lahey, B. B., van Hulle, C., Waldman, I., Krueger, R. F., & Rathouz, P. J. (2013). Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *Journal of Abnormal Psychology*, *122*, 1142-1153. doi:10.1037/a0034151
- Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., ...Shrout, P. E. (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science*, *12*, 742-756. doi:10.1177/1745691617690042
- Tackett, J. L., Waldman, I., Van Hulle, C. A., & Lahey, B. B. (2011). Shared genetic influences on negative emotionality and major depression/conduct disorder comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*, 818-827. doi:10.1016/j.jaac.2011.05.007
- US Department of Housing and Urban Development (2017). *Income limits*. retrieved from <https://www.huduser.gov/portal/datasets/il.html> on 11/7/2017.
- Vasilev, C. A., Crowell, S. E., Beauchaine, T. P., Mead, H. K., & Gatzke-Kopp, L. M. (2009). Correspondence between physiological and self-report measures of emotion dysregulation: A longitudinal investigation of youth with and without psychopathology. *Journal of Child Psychology and Psychiatry*, *50*, 1357-1364. doi:10.1111/j.1469-7610.2009.02172.x
- Walker, J. L., Lahey, B. B., Russo, M. F., Frick, P. J., Christ, M. A. G., McBurnett, K., ...Green, S. M. (1991). Anxiety, inhibition, and conduct disorder in children: I. Relations to social

impairment. *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 187-191. doi:10.1097/00004583-199103000-00004

Zisner, A., & Beauchaine, T. P. (2016). Neural substrates of trait impulsivity, anhedonia, and irritability: Mechanisms of heterotypic comorbidity between externalizing disorders and unipolar depression. *Development and Psychopathology*, 28, 1179-1210. doi:10.1017/S0954579416000754.

Zoccolillo, M. (1992). Co-occurrence of conduct disorder and its adult outcomes with depressive and anxiety disorders: A review. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31, 547–556. doi:10.1097/00004583-199205000-00024

Table 1*Psychopathology Measures at Time 1 by Group*

	Control (<i>n</i> =70)		Depressed (<i>n</i> =27)		CD (<i>n</i> =28)		Comorbid (<i>n</i> =81)	
Percent male	47.1		66.7		78.6		75.3	
Variable	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age at Time 1	9.83	1.52	10.22	1.48	9.54	1.60	9.95	1.46
Child Symptom Inventory								
MDD symptom severity	0.70	1.01	6.56	3.43	2.82	2.82	10.22	4.92
MDD symptom count	0.03	0.17	1.56	1.53	0.21	0.79	2.72	2.16
DYS symptom severity	0.93	1.11	7.00	2.27	2.93	1.90	9.74	4.04
DYS symptom count	0.03	0.17	2.26	1.23	0.36	0.49	2.98	1.94
CD symptom severity	0.67	1.01	1.44	1.48	7.07	5.07	7.40	4.69
CD symptom count	0.03	0.17	0.15	0.36	1.86	1.90	1.80	1.91
ODD symptom severity	4.00	2.96	7.37	3.44	16.32	4.06	17.04	4.64
ODD symptom count	0.27	0.66	1.56	1.67	5.79	1.83	5.95	2.14
Child Behavior Checklist								
anxious/depressed (<i>T</i>)	52.80	3.61	74.37	7.47	66.75	8.61	81.56	8.34
aggression (<i>T</i>)	51.16	2.38	57.04	5.81	77.64	8.90	79.65	9.16

Notes. MDD=major depressive disorder; DYS=dysthymia; CD=conduct disorder; ODD=oppositional defiant disorder.

Table 2

Symptom Severity and Percentages of Participants Meeting Diagnostic Thresholds for Major Depression and/or Dysthymia by Group Over Time*

Group	Time 1			Time 2			Time 3		
	Severity		% meeting criteria	Severity		% meeting criteria	Severity		% meeting criteria
	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	
Control	0.70	1.01	0.0%	9.96	1.54	23.5%	10.66	2.60	29.0%
Depressed	6.56	3.43	85.2%	12.81	3.25	68.0%	13.13	4.01	73.9%
Conduct	2.82	2.82	0.0%	12.81	3.59	76.9%	13.64	3.84	61.9%
Comorbid	10.22	4.92	79.0%	14.59	4.49	82.6%	16.13	4.99	73.7%

*MDD and dysthymia are combined because symptoms are largely overlapping across disorders.

Table 3*Symptom Severity and Percentages of Participants Meeting Diagnostic Thresholds for Conduct Disorder by Group Over Time*

Group	Time 1			Time 2			Time 3		
	Severity		% meeting criteria	Severity		% meeting criteria	Severity		% meeting criteria
	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	
Control	0.67	1.07	0.0%	0.43	0.94	4.4%	0.82	1.29	4.8%
Depressed	1.44	1.48	0.0%	0.97	1.45	20.0%	1.33	1.78	12.5%
Conduct	7.07	5.07	42.9%	4.12	3.55	42.9%	5.27	4.07	68.2%
Comorbid	7.40	4.69	38.3%	3.96	3.04	56.5%	4.93	3.73	64.9%

Table 4. Summary of Fixed Slope and Uncentered Intercept Effects from Multilevel Models

Variable	coefficient	standard error	t-statistic	p-value
Predicting MDD symptom severity				
For Intercept 1				
Intercept 2	0.105	0.936	0.112	.911
sex	0.263	0.280	0.936	.351
C1	0.006	0.001	4.540	<.001
C2	-0.007	0.002	-3.090	.003
C3	0.013	0.009	1.454	.147
age at study entry	-0.317	0.096	-3.291	.002
For Age slope				
Intercept 2	0.417	0.090	4.619	<.001
sex	-0.035	0.027	-1.302	.195
C1	<-0.001	<0.001	-2.580	.011
C2	<0.001	<0.001	1.900	.058
C3	-0.001	<0.001	-1.186	.238
age at study entry	-0.003	0.009	-0.353	.724
Predicting CD symptom severity				
For Intercept 1				
Intercept 2	0.205	5.936	0.035	.973
sex	1.316	1.713	0.769	.443
C1	0.044	0.007	6.514	<.001
C2	-0.155	0.016	-2.755	.007
C3	-0.155	0.055	-2.835	.006
age at study entry	0.869	0.618	1.407	.161
For Age slope				
Intercept 2	-0.301	0.529	-0.568	.571
sex	-0.196	0.164	-1.192	.235
C1	-0.002	<0.001	-2.932	.004
C2	0.003	0.001	2.030	.043
C3	0.007	0.004	1.296	.197
age at study entry	-0.020	0.052	-0.228	.735
Predicting ODD symptom severity				
For Intercept 1				
Intercept 2	5.751	8.939	0.643	.520
sex	-1.873	2.740	-0.684	.495
C1	0.091	0.013	6.806	<.001
C2	-0.068	0.023	-2.986	.004
C3	-0.258	0.087	-2.945	.004
age at study entry	-1.236	0.918	-1.347	.180
For Age slope				
Intercept 2	2.830	0.841	3.366	.001
sex	0.018	0.257	0.069	.946
C1	-0.005	0.001	-4.282	<.001
C2	0.004	0.002	1.923	.056
C3	0.013	0.008	1.661	.098
age at study entry	-0.088	0.082	-1.066	.288

Table notes. Degrees of freedom were 200 for all *t*-tests. C1=Contrast 1 (all clinical groups vs. healthy controls); C2=Contrast 2 (comorbid group vs. depressed and conduct problems groups); C3=Contrast 3 (depressed vs. conduct problems groups).

Figure 1. Effects of using statistical partialling techniques (e.g., ANCOVA, hierarchical regression, lag correlational designs) to evaluate heterotypic continuity of traits with shared etiologies. The left panel depicts the concurrent correlation (comorbidity) between depression and conduct problems (CPs) at age 8. The right panel depicts prospective prediction of CPs at age 12 by depression at age 8, controlling for CPs at age 8. Statistical partialling removes shared vulnerability (A in both the left and right panels), which (1) creates statistical entities (depression without vulnerability to CPs, CPs without vulnerability to depression) that misrepresent/distort etiological relations between disorders, and (2) under-estimate longitudinal associations by artificially attenuating common variance among predictors and outcomes (C, right panel). Despite recognition of these issues for decades (see Pedhazur, 1997), partialling procedures remain common in the psychopathology literature. Contrasted groups designs do not rely on statistical partialling and therefore offer advantages over research conducted with only comorbid samples.

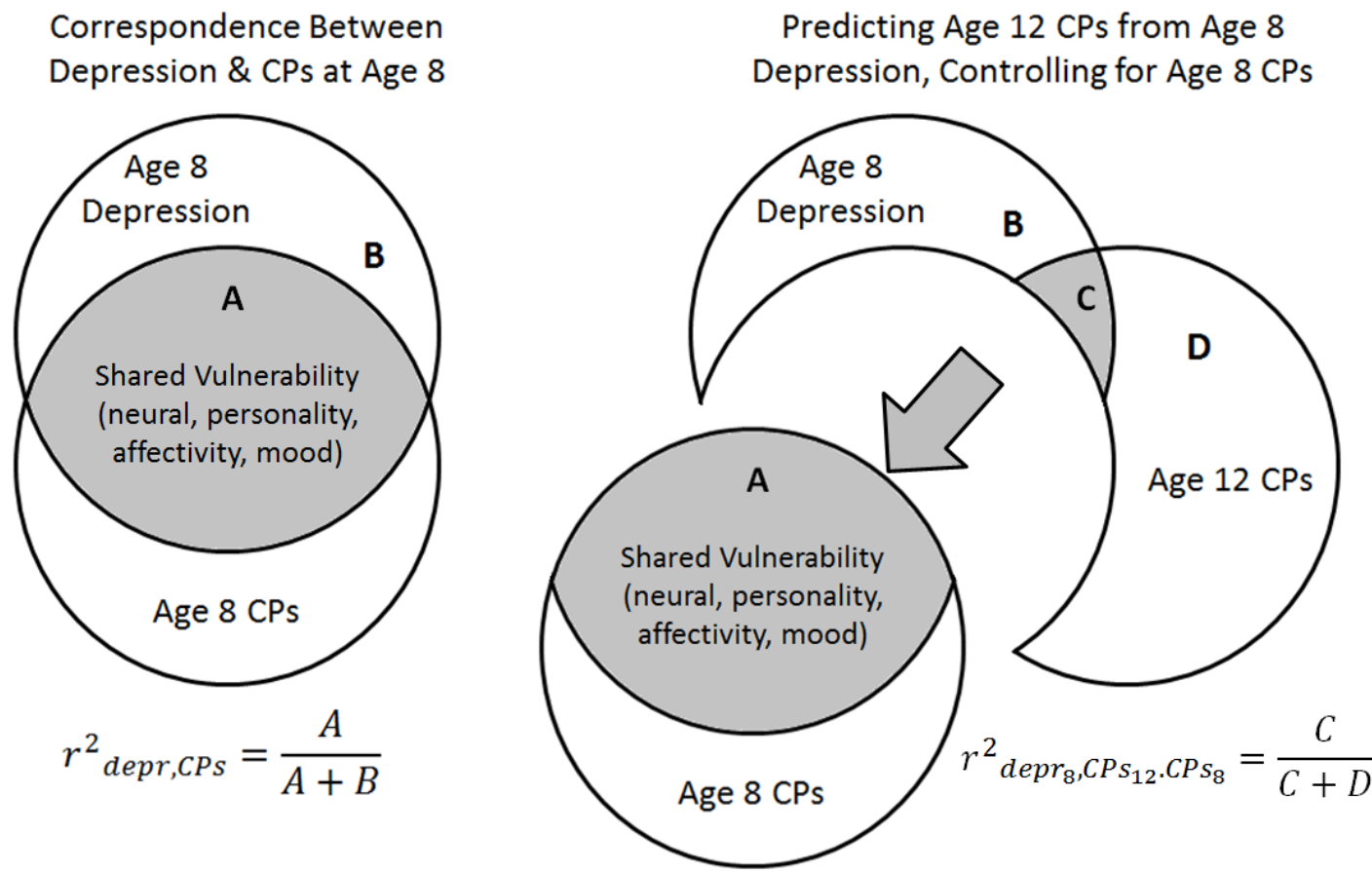


Figure 2. Growth in depressive symptoms by group from ages 8-14 years. Age 15 data are omitted because only 13 participants were distributed across the four groups, so means are likely unreliable.

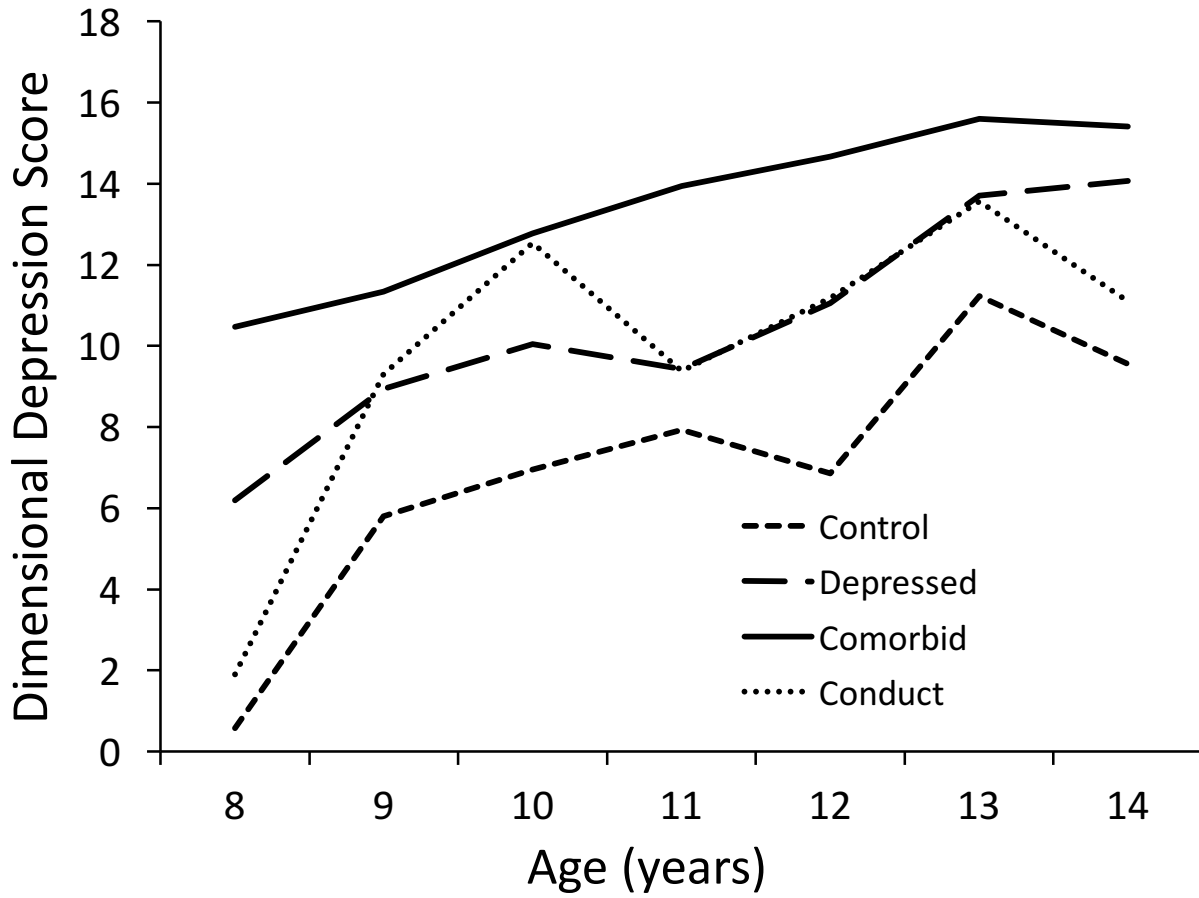


Figure 3. Growth in conduct disorder symptoms by group from ages 8-14 years. Age 15 data are omitted because only 13 participants were distributed across the four groups, so means are likely unreliable.

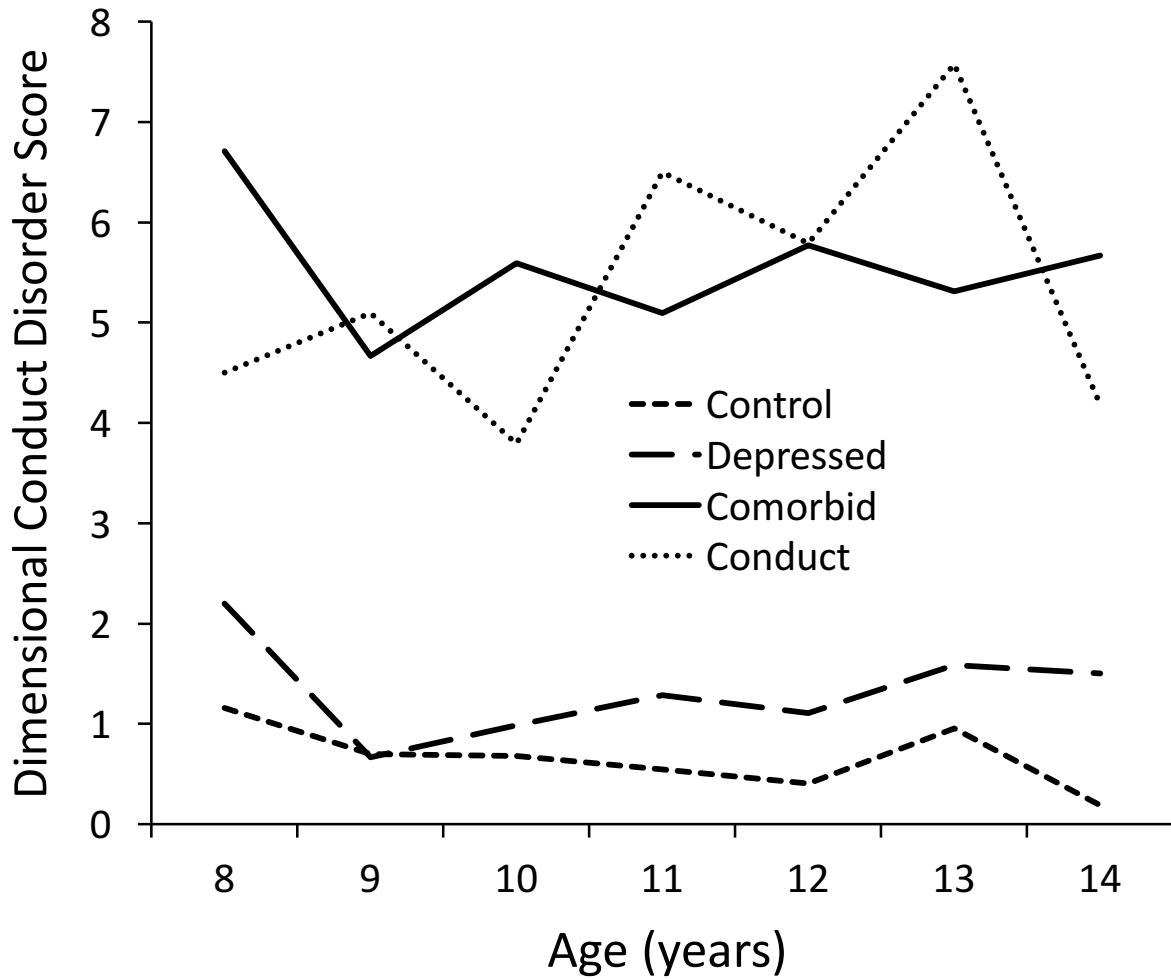


Figure 4. Growth in oppositional defiant disorder symptoms by group from ages 8-14 years. Age 15 data are omitted because only 13 participants were distributed across the four groups, so means are likely unreliable.

