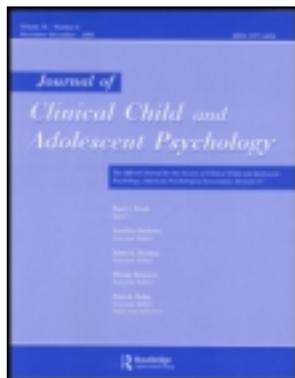


This article was downloaded by: [UNC University North Carolina]

On: 04 July 2011, At: 19:44

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Clinical Child & Adolescent Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/hcap20>

Internal Consistency and Associated Characteristics of Informant Discrepancies in Clinic Referred Youths Age 11 to 17 Years

Andres De Los Reyes^a, Eric A. Youngstrom^b, Shairy C. Pabón^a, Jennifer K. Youngstrom^b,
Norah C. Feeny^c & Robert L. Findling^d

^a Department of Psychology, University of Maryland at College Park

^b Department of Psychology, University of North Carolina at Chapel Hill

^c Department of Psychology, Case Western Reserve University

^d Department of Psychiatry, University Hospitals Case Medical Center

Available online: 11 Jan 2011

To cite this article: Andres De Los Reyes, Eric A. Youngstrom, Shairy C. Pabón, Jennifer K. Youngstrom, Norah C. Feeny & Robert L. Findling (2011): Internal Consistency and Associated Characteristics of Informant Discrepancies in Clinic Referred Youths Age 11 to 17 Years, *Journal of Clinical Child & Adolescent Psychology*, 40:1, 36-53

To link to this article: <http://dx.doi.org/10.1080/15374416.2011.533402>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan, sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Internal Consistency and Associated Characteristics of Informant Discrepancies in Clinic Referred Youths Age 11 to 17 Years

Andres De Los Reyes

Department of Psychology, University of Maryland at College Park

Eric A. Youngstrom

Department of Psychology, University of North Carolina at Chapel Hill

Shairy C. Pabón

Department of Psychology, University of Maryland at College Park

Jennifer K. Youngstrom

Department of Psychology, University of North Carolina at Chapel Hill

Norah C. Feeny

Department of Psychology, Case Western Reserve University

Robert L. Findling

Department of Psychiatry, University Hospitals Case Medical Center

In this study, we examined the internal consistency of informant discrepancies in reports of youth behavior and emotional problems and their unique relations with youth, caregiver, and family characteristics. In a heterogeneous multisite clinic sample of 420 youths (ages 11–17 years), high internal consistency estimates were observed across measures of informant discrepancies. Further, latent profile analyses identified systematic patterns of discrepancies, characterized by their magnitude and direction (i.e., which informant reported greater youth problems). In addition, informant discrepancies systematically and uniquely related to informants' own perspectives of youth mood problems, and these relations remained significant after taking into account multiple informants' reports of informant characteristics widely known to relate to informant discrepancies. These findings call into question the prevailing view of informant discrepancies as indicative of unreliability and/or bias on the part of informants' reports of youths' behavior.

This work was supported, in part, by NIH R01 5MH066647 (PI: Eric Youngstrom). Special thanks to Heather Marcinick, the members of the ABACAB research team and personnel at Applewood Centers, and the families who participated in this study.

Dr. Findling receives or has received research support, acted as a consultant, and/or served on a speaker's bureau for Abbott, Adrenex, AstraZeneca, Biovail, Bristol-Myers Squibb, GlaxoSmithKline, Johnson & Johnson, KemPharm, Lilly, Lundbeck, Neuropharm, Novartis, Noven, Organon, Otsuka, Pfizer, Sanofi-Aventis, Sepracore, Shire, Solvay, Supernus Pharmaceuticals, Validus, and Wyeth.

Correspondence should be addressed to Andres De Los Reyes, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland at College Park, Biology/Psychology Building, Room 3123H, College Park, MD 20742. E-mail: adelosreyes@psyc.umd.edu

The absence of definitive measures for assessing clinical conditions (e.g., anxiety, depression, disruptive behavior) in children and adolescents (hereafter referred to collectively as “youths”) makes it crucial to assess clinical functioning from the perspectives of multiple informants. These informants include the youth, parents, teachers, clinicians, laboratory observers, and official records. Multiple informants’ reports often disagree with each other. Indeed, these disagreements occur in reports on the number or severity of symptoms and whether a youth meets diagnostic criteria for a disorder (De Los Reyes & Kazdin, 2005), making informants’ reporting disagreements (hereafter referred to as “informant discrepancies”) some of the most consistent effects observed in clinical science (Achenbach, 2006; Achenbach, McConaughy, & Howell, 1987). These informant discrepancies influence how one draws research conclusions in that (a) multiple informants are often used in a single study to assess the same clinical condition and (b) relying on any one informant within the study often significantly changes the conclusions one might draw from the study’s findings (e.g., De Los Reyes & Kazdin, 2005, 2006a, 2008, 2009; Koenig, De Los Reyes, Cicchetti, Scahill, & Klin, 2009; Weisz, Jensen-Doss, & Hawley, 2005; Youngstrom, Findling, & Calabrese, 2003).

Whereas prior work has largely focused on documenting how informant discrepancies influence the methodology behind interpreting research conclusions, an emerging body of work suggests that these discrepancies have substantive implications for understanding the development and expression of youth psychopathology. For instance, greater parent–youth reporting discrepancies on reports of numerous constructs (e.g., youth’s behavior and emotional problems, negative parenting, parental monitoring) predict such varied outcomes as increases in youth aggressive and oppositional behaviors, youth mood and anxiety behaviors, and poor parent involvement in the youth’s therapy (De Los Reyes, Goodman, Kliewer, & Reid-Quinones, 2010; Ferdinand, van der Ende, & Verhulst, 2004; K. Guion, Mrug, & Windle, 2009; Israel, Thomsen, Langeveld, & Stormark, 2007). Further, informant discrepancies and in particular discrepancies between parent and teacher reports reflect variation in the circumstances in which youths express disruptive behavior symptoms (De Los Reyes, Henry, Tolan, & Wakschlag, 2009). Stated another way, informant discrepancies signal the development of poor youth outcomes as well as contextual variability in whether youths express specific behaviors.

In sum, research on informant discrepancies informs the science behind the development, assessment, and treatment of youth psychopathology. However, two pivotal issues warrant further attention. First, little is known of the basic psychometric properties of

discrepancies. In particular, it is unclear whether discrepancies measurements taken within clinical assessments pass internal consistency thresholds recommended for clinical assessments (e.g., alpha at or above .80; Nunnally & Bernstein, 1994). The ability of informant discrepancies to inform interpretations of research conclusions and contribute to our understanding of developmental psychopathology rests on assessments of these discrepancies yielding consistent scores across multiple domains. Otherwise, discrepancies may not reflect coherent psychological phenomena worthy of study, such as stable patterns of differences between informants’ perceptions of the behaviors being rated. Thus, the first specific aim of the present study was to examine the internal consistency of discrepancies across informants’ multiple reports of youths.

Second, research has identified a number of associated characteristics of informant discrepancies, including problem type being assessed, the informant pair examined, informants’ mood and anxiety symptoms, and parent and family stress (De Los Reyes & Kazdin, 2005). However, we have a poor understanding of why these discrepancies exist. A key reason why this gap remains is because traditional interpretations have often focused on single explanatory factors for why these discrepancies exist. In fact, previous research has focused on informants’ levels of psychosocial dysfunction, such as depressive and anxiety symptoms and levels of stress and conflict, as the key factors that explain discrepancies (for a review, see De Los Reyes & Kazdin, 2005). This is because interpretations of discrepancies have largely been based on the idea that an informant’s level of emotional distress leads to that informant attending to, encoding, and remembering negative youth behaviors (as opposed to positive or neutral behaviors), and predominantly using these remembrances to provide reports (Richters, 1992; Youngstrom, Izard, & Ackerman, 1999). However, prior work calls into question the utility of this explanation. Indeed, the same form of psychosocial dysfunction inconsistently relates to discrepancies and often multiple forms of such dysfunction relate to discrepancies when studied in combination (cf. Briggs-Gowan, Carter, & Schwab-Stone, 1996; De Los Reyes et al., 2008; van der Oord, Prins, Oosterlann, & Emmelkamp, 2006; Youngstrom et al., 1999). Further, informants for whom studying levels of psychosocial dysfunction has been of interest (e.g., parents) often share a close relationship with the youth being assessed. As a result, often domains of dysfunction also relate to each other and pose risk for the development of youth psychopathology (e.g., depressed mood, family discord, severity of youth problem behavior; Deater-Deckard, 1998; Granic & Patterson, 2006; Kazdin & Wassell, 2000). In addition, clinical conditions assessed in youths vary in their associated features (e.g.,

harsh and inconsistent parenting with conduct disorder vs. overcontrolled parenting with anxiety disorders; see Kazdin & De Los Reyes, 2007; Wood, McLeod, Sigman, Hwang, & Chu, 2003). In sum, attributing informant discrepancies to any one domain of psychosocial dysfunction is problematic because (a) many dysfunction domains relate to discrepancies, (b) no one domain consistently relates to discrepancies, (c) psychosocial dysfunction in some informants (parents) are also associated features of the youths' dysfunction being assessed, and (d) clinic populations vary widely in the associated features of the disorders for which youths meet diagnostic criteria.

In light of these limitations in prior work, the second specific aim of this study was to examine whether informant discrepancies are uniquely related to various domains of psychosocial dysfunction including informants' mood symptoms, family relations, and youth problem type and severity. To address this second specific aim, we took advantage of studying a heterogeneous multisite clinic sample for which prior research suggests the presence of (a) low to moderate levels of informant agreement, albeit with some exceptions; (b) mood symptoms, family conflict, and stress; and (c) both internalizing and externalizing behavior concerns in youths being assessed in the sample (see Du Rocher Schudlich, Youngstrom, Calabrese, & Findling, 2008; Youngstrom et al., 2003; Youngstrom, Findling, & Calabrese, 2004; Youngstrom, Findling, Calabrese, Gracious, et al., 2004; Youngstrom et al., 2006). To our knowledge, this study offers the widest lens to date through which to examine the interplay of various psychosocial dysfunction domains, and whether one domain more so than others relates to discrepancies.

In a heterogeneous clinic sample of youths, we extended the literature in three ways. First, we used latent profile analysis (LPA) to examine patterns of caregiver–youth reporting discrepancies across multiple domains of youth psychological concerns (Bartholomew, Steele, Moustaki, & Galbraith, 2002). Prior work suggests that two characteristics that describe the structure of discrepancies are how large the disagreements are and which of the two informants report greater levels of the behavior being assessed (Barker, Bornstein, Putnick, Hendricks, & Suwalsky, 2007). As such, we expected discrepancies to vary by *magnitude* (how much disagreement) and *direction* (i.e., which report was greater).

Second, we examined the internal consistency of discrepancies between caregiver and youth reports and across multiple domains of youth problem behavior. We surmised that informant discrepancies represent systematic and stable differences between informants' perspectives (De Los Reyes & Kazdin, 2005). As such, we expected to find acceptable levels of internal consistency across multiple “items” or measures of informant

discrepancies. Further, prior work has demonstrated statistically that the larger the individual differences between two measures, the greater the internal consistency in measures of the discrepancies between them (Rogosa, Brandt, & Zimowski, 1982; Rogosa & Willett, 1983). Thus, we expected to observe a positive relation between internal consistency estimates and the magnitude of discrepancies (i.e., greater discrepancies would translate into larger estimates of internal consistency).

Last, we used the groups of caregiver–youth discrepancies patterns identified through LPA to compare discrepancies reporting patterns across domains of various associated characteristics. In particular, we examined the association between reporting discrepancies patterns and three domains. First, we examined caregiver and youth mood symptoms given that, as mentioned previously, researchers have posited that informant discrepancies are largely accounted for by aspects of informants' psychosocial functioning, such as depressed mood. Second, we studied family functioning and more specifically global indices of such functioning that capture multiple domains typifying the relations among family members (e.g., communication patterns, problem solving, general family dynamics). We examined family functioning in this way in keeping with prior work in both clinic and nonclinic samples suggesting that discrepancies relate to various dimensions of family functioning (e.g., caregiver acceptance and involvement, parent–child conflict; see De Los Reyes & Kazdin, 2006b; Treutler & Epkins, 2003). Third, we studied interviewer reports of youth clinical severity given prior work suggesting informant discrepancies are related to clinical impressions of youth behavior (Hawley & Weisz, 2003), and that mood symptoms and family functioning relate to both discrepancies *and* youth functioning (for a similar argument, see Youngstrom, Loeber, & Stouthamer-Loeber, 2000).

Prior work is inconsistent on identifications of associated characteristics of discrepancies (De Los Reyes & Kazdin, 2005). As mentioned previously, the study of these associated characteristics has typically revolved around the idea that when present, they relate to informants' reporting biases of youth behavior and emotional problems, or “overestimated” reports of negative behaviors. However, in clinic samples of youths often an independent evaluator (interviewer) or set of evaluators (clinical team) has collected information from multiple informants and as such, has determined that the youth being assessed expresses concerns that warrant treatment (Hawley & Weisz, 2003). Under these circumstances, the discrepancies between informants' reports likely reflect differences between informants in their perspectives of the youth's behavior (self-report vs. other report) and the contexts in which they observe the youth's behavior (e.g., home vs. nonhome contexts; see

Kraemer et al., 2003), more so than they reflect that one or both of the informants are providing biased reports. Thus, we expected to find that independent caregiver and youth reports of a specific set of youth clinical symptoms would uniquely relate to the discrepancies between their reports of youth problem behavior, broadly construed.

In the present study, the specific youth clinical symptoms of interest were youth mood problems, given that in the sample we examined nearly 60% of youths met primary criteria for a *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev. [*DSM-IV-TR*]; American Psychiatric Association, 2000) mood disorder. Thus, we expected both caregiver and youth reports of youths' mood problems to be significantly related to caregiver-youth discrepancies on general problem behaviors and to remain significant when taking into account caregiver mood symptoms, interviewer ratings of youth clinical severity and youth mood symptoms, and caregiver and interviewer ratings of family functioning. We also expected that these relations would remain significant when taking into account the youth demographic characteristics of age, gender, and ethnicity/race, which sometimes relate to discrepancies (De Los Reyes & Kazdin, 2005). Because we assessed domains via youth, caregiver, and interviewer reports, these tests effectively controlled for not only shared variance accounted for by domain but also shared variance accounted for by the informant completing the measure.

METHOD

Participants

The present study is a secondary analysis of data gathered to establish the prevalence and clinical features of bipolar disorder in youths seeking outpatient mental health services (NIH R01MH066647, PI: E. Youngstrom). The main performance site was the largest community mental health center in the state of Ohio providing services to children and families. A representative sample of families was invited to participate (via consecutive case series when sufficient interview slots were available and random selection otherwise). Participating families all completed the same assessment battery, which included standard rating scales as well as a set of interview modules and checklists specific to bipolar symptoms.

Eligible youths were between the ages of 4 and 18 years (inclusive), but only youth ages 11 and older provided self-report on questionnaires. The present analyses are limited to this older subset. Patients with a diagnosed or suspected pervasive developmental disorder, a psychiatric disorder due to a general medical condition, or evidence of mental retardation usually were

referred to a different agency prior to recruitment and were not enrolled. Participants were recruited for this protocol from the sites noted previously. Participants included 420 youths (age range = 11–17 years; $M = 13.55$, $SD = 1.8$) and their caregivers. Youths were 48% female ($n = 203$); 68% African American, non-Hispanic ($n = 285$); 26% White, non-Hispanic ($n = 109$); 1% Hispanic ($n = 6$); and 5% of other ethnicities ($n = 20$).

Caregivers in the sample included a variety of caregivers serving a caregiving role in the youth's life. Specifically, caregivers in the sample were primarily biological mothers (73.2%), with the rest biological fathers (5.7%), adoptive mothers (5.5%), grandmothers (6.9%), adoptive fathers (1%), or other caregivers (e.g., grandfathers, aunts/uncles, foster parents; 7.5%) (Two participants did not provide this information).

Measures

DSM-IV-TR diagnoses. Youths in the sample were interviewed for the presence of *DSM-IV-TR* diagnoses using the Kiddie Schedule for Affective Disorders and Schizophrenia-Plus (KSADS-PL-Plus). This interview is a combination of the KSADS-PL (Kaufman et al., 1997) and the mood items from the Washington University in St. Louis KSADS (Geller et al., 2001), ascertaining *DSM-IV-TR* disorders with more extensive coverage of associated features of mood disorders, as recommended by the National Institute of Mental Health (Nottelmann et al., 2001). Caregivers and youths were interviewed separately, and in the presence of discrepant interviews, these discrepancies were resolved by additional interviewing and synthesizing with clinical judgment (Youngstrom et al., 2005). This was not an automatic algorithm but a systematic effort to gather additional information until a conclusive interpretation was possible.

Youths varied greatly in their primary diagnoses. Five characteristics are worthy of mention. First, 38% of youths met criteria for a unipolar depressive disorder. Second, 34% met criteria for Attention Deficit Hyperactivity Disorder (ADHD) or a disruptive behavior disorder without a comorbid mood disorder. Third, 5% of youths met criteria for bipolar I disorder per strict *DSM-IV-TR* criteria, and another 15% met criteria for bipolar II, cyclothymic disorder, or bipolar not otherwise specified. Fourth, 8% met primary diagnostic criteria for another *DSM-IV-TR* condition. Fifth, youths met criteria for between 0 and 8 axis I diagnoses ($Mdn = 3$), and 52% met criteria for ADHD as one of their diagnoses (this 52% includes those that met criteria for both ADHD and a mood disorder and those that met criteria for ADHD but did not meet criteria for a mood disorder). Thus, the sample was quite diagnostically

heterogeneous and comprised youths meeting primary diagnostic criteria for various internalizing and externalizing diagnostic conditions.

Youth behavior and emotional problems. Youth behavior and emotional problems were assessed via caregiver report and youth self-report. Specifically, caregivers and youths completed the Child Behavior Checklist (CBCL) and Youth Self-Report Form (YSR), respectively. The CBCL (Achenbach, 1991a) and YSR (Achenbach, 1991b) are commonly used to assess and report behavioral problems and social aptitude in youths. The CBCL is a caregiver-reported survey that rates behaviors on a 3-point scale of 0 (*not true of the youth*), 1 (*somewhat or sometimes true of the youth*), and 2 (*very true or often true of the youth*). Derived from the CBCL, the items on the YSR are nearly identical to the caregiver survey items formerly described (89 items being the same) and include slight rewordings to reflect the different perspectives of youth and caregiver. Both instruments are used to assess eight syndrome scales: anxious/depressed, withdrawn/depressed, somatic complaints, thought problems, social problems, attention problems, rule-breaking (called “delinquent behavior” on the 1991 version), and aggressive problems. We measured caregiver-youth discrepancies on common items.

Caregiver–youth reporting discrepancies. Caregivers’ and youths’ perceived youth behavior and emotional problems were assessed using raw score totals from the eight CBCL (caregiver) and YSR (youth) syndromes. We used raw score totals as opposed to *T* scores because CBCL and YSR *T* scores are calculated using item content that differs across caregiver and youth reports. To assess informant discrepancies it is important to hold item content constant, and thus we measured reporting discrepancies using identical item content across caregiver and youth raw score reports.

Discrepancies were measured using standardized difference scores (SDS), consistent with current recommendations and practices (e.g., De Los Reyes et al., 2008; De Los Reyes & Kazdin, 2004, 2006b; K. Guion et al., 2009; Owens, Goldfine, Evangelista, Hoza, & Kaiser, 2007). Specifically, SDS were created by first converting each youth’s ratings and their caregiver’s ratings of each CBCL/YSR syndrome into *z* scores, and then subtracting the youth’s *z* score for each subscale from the caregiver’s *z* score on that same subscale. This resulted in eight SDS (one for each CBCL/YSR syndrome), with negative scores representing instances in which the youth reported greater concerns on the syndrome, relative to the caregiver, and positive scores representing the caregiver reporting greater concerns on the syndrome,

relative to the youth. We assessed informant discrepancies using syndrome scores as opposed to at the item level (e.g., discrepancies calculated between each caregiver and youth report on each item) because, by construction, the reliability of informant discrepancies is heavily dictated by the reliability of the measures from which they are calculated. Because multi-item summary scores will demonstrate greater reliability than the individual items that comprise the summary scores (Nunnally & Bernstein, 1994), it appeared most prudent to assess discrepancies using summary scores.

The mathematical properties and associated characteristics of SDS, along with the rationale for choosing SDS over other discrepancies measures have been demonstrated, reported, and reviewed elsewhere (De Los Reyes & Kazdin, 2004, 2005; K. Guion et al., 2009; Owens et al., 2007). Briefly, we used SDS as opposed to, for instance, residual scores (regression-based measurements of unshared variance between two informants’ scores) or raw scores (unstandardized scores) because the two informants’ reports used to assess informant discrepancies often significantly differ in their variances (De Los Reyes & Kazdin, 2004). As a result, when bivariate correlations are calculated between the reports used to create the discrepancies scores and each of these three measures of discrepancies, residual scores and raw scores will correlate too highly with the report with larger variance, whereas the SDS will correlate equally with the two informants’ reports, regardless of inequality of variances (De Los Reyes & Kazdin, 2004). Thus, SDS do not exhibit the multicollinearity concerns of other scores.

One limitation of using SDS is that the scores lose information about differences in the rating variances across informants. This might have particularly important implications for the latent profile modeling findings we report next, as “dispersion” or within-individual variance is one of the three fundamental characteristics of multivariate profiles (along with the mean or “level” of problems and the “shape” or ranking of the problems in the profile; R. M. Guion, 1998). Thus, next we report comparisons of the latent profile solutions derived based on the eight SDS discrepancies scores across the CBCL/YSR syndromes and the eight raw discrepancies scores based on the same syndromes and item content.

Youth mood symptoms. An important refinement in this study was gathering separate ratings of manic symptoms in addition to other mood symptoms. Indeed, previously youth depressive or manic symptoms have been examined in relation to discrepancies and both separately relate to informant discrepancies (e.g., De Los Reyes et al., 2008; Youngstrom, Findling, & Calabrese, 2004). Youth depressive and manic symptoms were assessed via three different sources: youth self-report,

caregiver report, and clinical interviewer ratings based on observations during the interview as well as semi-structured interview of both the youth and caregiver.

Three different sources of information quantified each youth's level of depressive symptoms. These included youth self-report on the Depression scale of the General Behavior Inventory (A-GBID; Danielson, Youngstrom, Findling, & Calabrese, 2003; Depue et al., 1981), and caregiver report on the Depression scale of the caregiver version of the GBI (C-GBID; Youngstrom, Findling, Danielson, & Calabrese, 2001). Both of these scales contain 46 items rating symptoms of depression on a 0-to-3 scale. Both the adolescent and caregiver versions show exceptional internal consistency, with alphas exceeding .96 in both published samples (Youngstrom, Findling, Calabrese, Gracious, et al., 2004). In the present sample, alphas for youth and caregiver report were .97 and .96, respectively. The validity of the GBI as a measure of youth mood symptoms has been demonstrated by multiple samples and groups (Danielson et al., 2003; Depue, Krauss, Spont, & Arbisi, 1989; Depue et al., 1981; Findling et al., 2002; Klein, Depue, & Slater, 1986; Klein, Dickstein, Taylor, & Harding, 1989; Mallon, Klein, Bornstein, & Slater, 1986; Nusslock, Abramson, Harmon-Jones, Alloy, & Hogan, 2007; Youngstrom, Findling, Calabrese, Gracious, et al., 2004; Youngstrom, et al., 2005). In addition, the clinical interviewer rated the youth's severity of depressive symptoms on the Child Depression Rating Scale—Revised edition (CDRS; Poznanski, Miller, Salguero, & Kelsh, 1984). It has demonstrated good validity and psychometric properties (Overholser, Brinkman, Lehnert, & Ricciardi, 1995; Poznanski et al., 1984). It is the most widely used clinical rating scale for depression severity in youths, and scores produced an alpha of .90 in the present sample.

Each youth's level of manic symptoms was quantified via the Mood Disorder Questionnaire caregiver (MDQ-C) and adolescent (MDQ-A) self-report versions—adaptations of a screening instrument for bipolar disorder in adults (Hirschfeld, 2001) and youths. Both include items for each of the *DSM-IV-TR* symptoms of mania, along with an item asking if many of the symptoms co-occurred at the same time, and another item asking if there was impairment associated with the symptoms. Items are scored as being present or absent, with higher scores indicating greater manic symptoms. The MDQ has shown validity in both adult and pediatric populations (Hirschfeld, 2001; Wagner, et al., 2006; Youngstrom, et al., 2005). Internal consistency was .75 for adolescent self-report and .82 for caregiver report in the present sample. In addition, the clinical interviewer rated the severity of manic symptoms using the Young Mania Rating Scale (YMRS; Young, Biggs, Ziegler, & Meyer, 1978). The YMRS is

a clinical rating scale containing 11 items using a 0-to-4 scale, with some item scores doubled to weight for their clinical importance and low base rate (Young et al., 1978). The YMRS has been used extensively with juvenile bipolar spectrum disorders, demonstrating good reliability and good ability to discriminate bipolar spectrum disorder from ADHD (Fristad, Weller, & Weller, 1992, 1995; Youngstrom, Gracious, Danielson, Findling, & Calabrese, 2003). Alpha in the present sample was .86.

Caregiver mood symptoms. Caregivers also provided ratings of their own depressive and manic symptoms. Initially, they did this by completing GBIs about themselves as well as about their child. However, this proved burdensome, and so the protocol was modified in the 2nd year of data collection so that caregivers completed the MDQ (C-MDQC; Hirschfeld et al., 2000) and the Beck Depression Inventory (C-BDIC; Beck & Steer, 1987) about themselves. As a result, the C-MDQC and C-BDIC scores are not available for the first 100 participants. The C-MDQC had an alpha of .81 and the C-BDIC had alpha of .88 in the present sample.

Family psychosocial functioning. The short form of the Family Assessment Device (FAD) measured healthy and unhealthy family dynamics (Epstein, Baldwin, & Bishop, 1983). Using a 4-point Likert-type scale, caregivers rated the applicability of each statement to their own families in a 27-item questionnaire. Scores ranged from 1 (*strongly agree*) to 4 (*strongly disagree*), where low total scores indicated fewer problems and more adaptive family functioning. The FAD includes domains assessing General Family Functioning (e.g., *We feel accepted for what we are*), Problem Solving (e.g., *After our family tries to solve a problem, we usually discuss whether it worked or not*), and Communication (e.g., *When someone is upset the others know why*), yielding a total score across these domains. Psychometric properties are described in the literature (Miller, Kabacoff, Keitner, Epstein, & Bishop, 1986), and the FAD has demonstrated associations with youth behavior problems and both caregiver and youth psychiatric diagnoses (Du Rocher Schudlich et al., 2008). The FAD Total Score obtained an alpha of .91 in the present sample, and acceptable alpha estimates for its component family domains of General Functioning (.85), Problem Solving (.78), and Communication (.70).

The clinical rater also provided a global rating of family functioning at the end of the interview via the Global Family Environment Scale (GFES). The GFES is a global rating of the quality of family environment, scaled from 1 to 100 (high scores indicate better functioning). It has demonstrated good interrater reliability

(>.80) and criterion validity with disruptive behavior disorders (Rey et al., 1997).

Youth level of functioning. At the end of the interview, the rater quantified the youth's level of functioning using the Clinical Global Impressions (CGI) scale, ranging from 1 (*normal, not at all ill*) through 4 (*moderately ill*) to 7 (*among the most extremely ill patients*) (National Institute of Mental Health, 1985). The CGI has demonstrated good interrater reliability and sensitivity to treatment effects in clinical trials (e.g., Findling et al., 2007).

Youth demographic characteristics. Demographics were assessed through a packet and contact sheet completed by the caregiver and included youth age, gender, and ethnicity/race.

Procedure

The Institutional Review Boards for Human Investigation of the University Hospitals of Cleveland, Case Western Reserve University, the University of North Carolina at Chapel Hill, and Applewood Centers approved protocol procedures. All families sought mental health services at an outpatient clinic. They were invited to participate in a full day research interview that offered more thorough assessment of youth functioning and family history. As part of the invitation process (preconsent), participants were told that the goal of the study was to improve identification and treatment of children's emotional and behavioral problems. Caregivers provided written informed consent and youths provided written assent. Highly trained raters conducted

a semistructured diagnostic interview (KSADS-PL-Plus; Geller et al., 2001; Kaufman et al., 1997), sequentially interviewing the caregiver and the youth, and then using clinical judgment and reinterviewing to clarify diagnostic inconsistencies. While the caregiver was completing the interview, the youth filled out the rating scales, and vice versa. At the conclusion of their assessment, caregiver and youth were debriefed of the purpose of the overall study (assessment and diagnosis of youth's emotional/behavior symptoms, with an additional focus on pediatric bipolar disorder) and given some feedback on the youth's emotional/behavioral symptoms. In addition, if a family authorized the release of assessment outcomes to a mental health provider, a one-page summary form was sent to the provider that detailed clinically elevated symptoms endorsed on the KSADS-PL-Plus.

RESULTS

Descriptive Statistics and Preliminary Analyses

Means and standard deviations of all dependent variables are presented in Table 1. Means and standard deviations of the reports used to assess caregiver-youth discrepancies and construct latent profile models are presented in Table 2. Skewness and kurtosis were all close enough to normal (± 1.0) to fall within the range where planned analytic methods would be robust.

We conducted paired *t* tests of differences between caregiver CBCL and youth YSR reports, and Levene's tests of differences in variances between caregiver and youth reports, both using Holm's stepdown Bonferroni correction. Paired *t* tests revealed that the caregiver

TABLE 1
Means and Standard Deviations of Dependent Variable Measures for the Total Sample

Measure	<i>N</i>	<i>M</i> (<i>SD</i>)
Mood Disorder Questionnaire, Total Score (Caregiver/Youth)	418	4.98 (3.37)
General Behavior Inventory, Depression Score (Caregiver/Youth)	413	30.84 (23.51)
Mood Disorder Questionnaire, Total Score (Youth/Self)	420	5.29 ^a (3.05)
General Behavior Inventory, Depression Score (Youth/Self)	418	40.27 (28.27)
Youth Mania Rating Scale, Summary Score (Interviewer/Youth)	393	7.65 (9.20)
Children's Depression Rating Scale, Total Score (Interviewer/Youth)	395	34.57 (14.17)
Mood Disorder Questionnaire, Total Score (Caregiver/Self)	329	3.14 (3.00)
Beck Depression Inventory, Total Score (Caregiver/Self)	328	9.24 (8.31)
Global Family Environment Scale, Total Score (Interviewer/Family)	412	66.97 (12.06)
Family Assessment Device, Total Score (Caregiver/Family)	417	2.09 (.44)
Clinical Global Impression, Severity Score (Interviewer/Youth)	411	3.97 (.94)

Note. *N* = 420. For every measure, parenthetical notation at the end of each measure name refers, first, to who completed the measure, and second to the target of the measure (e.g., Caregiver/Youth = caregiver completed the measure about the youth's behavior).

^aPaired *t* tests (*n* = 329) revealed that youths self-reported significantly greater mood concerns on the Mood Disorder Questionnaire (*M* = 5.26) than caregivers self-reported on their version of the Mood Disorder Questionnaire (*M* = 3.14), *p* < .001.

TABLE 2
Raw Score Means (*SD*), *T* Score Means (*SD*), and Correlations Among Raw Score Subscales of Caregiver (CBCL) and Youth (YSR) Reports of Youth Psychopathology for the Total Sample

	<i>Variable</i>	<i>Caregiver Raw</i>	<i>Caregiver T^a</i>	<i>Youth Raw</i>	<i>Youth T</i>	<i>Cross-Informant Correlation</i>
1	Anxious-Depressed	7.03 (4.81)	62.23 (9.91)	6.20 (5.27)	57.88 (9.25)	.29*
2	Withdrawn/Depressed	5.88 (3.68)	66.36 (11.07)	5.00 (3.38)	59.79 (9.43)	.25*
3	Somatic Complaints	3.56 (3.37)	61.10 (9.19)	4.80 (3.89)	59.14 (9.01)	.37*
4	Social Problems	6.13 (4.27)	64.14 (9.71)	5.11 (3.82)	59.01 (8.76)	.29*
5	Thought Problems	5.57 (4.16)	64.96 (9.45)	5.85 (4.65)	58.98 (8.98)	.19*
6	Attention Problems	9.48 (4.40)	68.15 (11.51)	7.38 (3.98)	60.71 (10.33)	.22*
7	Rule-Breaking Problems	8.53 (5.04)	66.37(8.27)	6.49 (4.52)	59.09 (7.89)	.41*
8	Aggressive Behavior	16.16 (8.02)	71.05 (11.92)	10.79 (6.58)	60.99 (10.00)	.30*

Note: *N* = 420. The cross-informant correlations estimates are based on scores derived from the items present across the eight CBCL/YSR syndromes. CBCL = Child Behavior Checklist; YSR = Youth Self-Report.

^a*T* scores based on caregiver report are based on 419 participants.

**p* < .001.

reported significantly higher levels of youth behavior and emotional problems than the youth self-reported on the following syndromes (all *ps* < .01): anxious/depressed, withdrawn/depressed, social problems, attention problems, rule-breaking problems, and aggressive behavior. Conversely, the youth self-reported significantly higher levels than the caregiver reported on the somatic complaints syndrome (*p* < .001). We observed nonsignificant differences between thought problems syndrome reports. Levene’s tests revealed significantly greater variances for the YSR on the anxious/depressed, somatic complaints, and thought problems syndromes, and greater variances for the CBCL on the withdrawn/depressed, social problems, attention problems, rule-breaking problems, and aggressive behavior syndromes (all *ps* < .05). Thus, the differences in variances support our use of SDS to construct LPA models, because by construction there would be no differences in the means and variances of caregiver report and youth self-report scores used to construct the discrepancies used in the model (De Los Reyes & Kazdin, 2004).

To examine the association between caregiver and youth reports, Pearson product-moment correlations were conducted between caregiver and youth raw score reports on the CBCL and YSR syndromes, respectively (Table 2). Consistent with prior work (Achenbach, 2006), correlations between caregiver and youth reports were statistically significant but low to moderate in magnitude (*rs* ranging .19–.41; see Cohen, 1988, for benchmarks).

Latent Profile Modeling of Caregiver-Youth Discrepancies

Latent profile solution. Before testing the internal consistency of informant discrepancies and their relations with youth, caregiver, and family characteristics,

we were interested in statistically modeling reporting discrepancies patterns. To address this aim, we modeled caregiver–youth reporting discrepancies by conducting an exploratory LPA on the eight SDS computed across the CBCL/YSR syndromes (Bartholomew et al., 2002). Like cluster analysis, LPA attempts to identify groups of cases based on similar patterns of indicator variables. Like confirmatory factor analysis, LPA computes tests of relative model fit, yielding indices such as the Bayesian Information Criterion (BIC) to compare whether a given model is a more or less parsimonious solution to the data than competing solutions, with lower scores indicating greater parsimony (Raftery, 1986, 1995). LPA focuses on continuous indicators to identify case groupings; these procedures are a generalization of latent class analysis, which uses categorical or ordinal variables to identify groupings (McCutcheon, 1987). Latent profile analysis identifies groups within which there is local independence of indicators (i.e., indicators are statistically independent within levels of each group). Thus, LPA is a “person-centered” approach to data analysis that identifies case profiles exhibiting similar data patterns across indicators. Probabilities provided by a solution may be used to assess the confidence with which cases are assigned (McCutcheon, 1987).

We expected that the LPA would identify the following profiles of caregiver–youth rating discrepancies: (a) dyads within which the youth reported far greater behavior and emotional concerns than the caregiver, (b) dyads within which the youth reported slightly more behavior and emotional concerns than the caregiver, (c) dyads within which the caregiver reported slightly more behavior and emotional concerns than the youth, and (d) dyads within which the caregiver reported far greater behavior and emotional concerns than the youth. As is customary in LPA (Bartholomew et al., 2002; McCutcheon, 1987), we tested profile solutions

from one to five (i.e., one more than the expected four-profile solution), evaluating the fit and interpretability of each.

Evaluating model fit involved two metrics. First, we compared the BIC indices of one- to five-profile solutions to each other, with lower BIC indices indicating superior model fit (Bartholomew et al., 2002). Along these lines, the four-profile solution fit the data best, $LL = -4564.26$, $BIC = 9533.22$. The BIC of this solution was lower relative to those of the three- and five-profile solutions, $BICs = 9622.39$ and 9549.76 , respectively. The BIC of the four-profile solution differed from each of the three- and five-profile BICs by more than 10, indicating “very strong” evidence in support of this solution relative to competing solutions (Raftery, 1995).

The second metric we used to evaluate model fit was specific to the characteristics of the four-profile solution itself. Specifically, a key metric by which researchers assess the suitability of model fit in latent classification modeling is the mean participant probability of assignment within each group, with a common metric being mean values of assignment probability above .70 (see Nagin, 2006). The probabilities of latent profile assignment are the four assignment probabilities (one for each profile in the solution) accorded to each participant dyad in the sample. The highest of these four assignment probabilities dictated a participant dyad’s profile assignment. It is important to note that the higher the assignment probability, the greater the confidence that a dyad was “rightfully” assigned to the particular profile group to which they were assigned.

In Table 3 we report frequencies of latent profile groups for the four-profile solution as well as the mean probabilities of latent profile assignment and within-group mean SDS across the eight CBCL/YSR syndromes. Across the groups the mean assignment probability was above .90, and the mean assignment

probabilities within each group were at or above .90. This suggests superb model fit in that latent profile assignments were made with a great degree of confidence that dyads were assigned to profile groups within which they were (a) maximally similar in patterns of reporting discrepancies to other dyads assigned to their group and (b) maximally different in these reporting discrepancies patterns from dyads assigned to the other three groups.

Consistent with our hypotheses, the four-profile solution yielded the following profiles of caregiver–youth reporting discrepancies of the youth’s behavior and emotional concerns ($N = 420$; see Table 3): (a) dyads with a youth who consistently reported greatly higher levels of their own behavior and emotional concerns, relative to the caregiver (Youth Reports Much Greater on Average, $n = 56$), (b) dyads with a youth who consistently reported slightly higher levels of their own behavior and emotional concerns, relative to the caregiver (Youth Reports Slightly Greater on Average, $n = 147$), (c) dyads with a caregiver who consistently reported slightly higher levels of their youth’s behavior and emotional concerns, relative to the youth (Caregiver Reports Slightly Greater on Average, $n = 173$), and (d) dyads with a caregiver who consistently reported greatly higher levels of their youth’s behavior and emotional concerns, relative to the youth (Caregiver Reports Much Greater on Average, $n = 44$). As seen in Table 3, each of the profiles identified in the four-profile solution yielded consistently different “kinds” of caregiver–youth reporting discrepancies (i.e., relatively greater caregiver or youth reports of different magnitudes). Of note, the LPA profile reflected similar patterns of SDS, regardless of the CBCL/YSR syndrome. In other words, youths who reported greater behavior and emotional concerns relative to the caregiver when providing anxious/depressed syndrome reports also tended to engage in

TABLE 3
Latent Profiles of Caregiver–Youth Discrepancies on the CBCL and YSR, Respectively

Latent Profile	N	Latent Variables					
		Profile Prevalence	M Assignment Probability	M SDS: CBCL/YSR Domains	Child Gender (% Boys)	Child Age (M, SD)	Ethnicity (% African American)
Youth >> Caregiver	56	13.3%	.94	–1.50	42.9%	13.82 (2.04)	51.8%
Youth > Caregiver	147	35%	.90	–.44	47.6%	13.84 (1.78)	66.7%
Caregiver > Youth	173	41.2%	.91	.46	54.3%	13.30 (1.82)	74%
Caregiver >> Youth	44	10.5%	.91	1.56	65.9%	13.16 (1.74)	68.2%
Total	420	100%	.91				

Note: $N = 420$. The mean assignment probability is based on the value used to assign each individual caregiver–youth dyad to a latent profile; higher values indicate greater confidence that the caregiver–youth dyad was assigned to the correct class. Tests of analysis of variance suggested that there were no significant differences among the profiles in mean assignment probabilities. Youth >> Caregiver = Youth Reports Much Greater on Average; Youth > Caregiver = Youth Reports Slightly Greater on Average; Caregiver > Youth = Caregiver Reports Slightly Greater on Average; Caregiver >> Youth = Caregiver Reports Much Greater on Average; CBCL = Child Behavior Checklist; YSR = Youth Self-Report; SDS = Standardized difference scores based on caregiver (CBCL) and youth (YSR) reports.

the same pattern of reporting when providing aggressive behavior syndrome reports. The same was true for caregiver–youth dyads in which the caregiver reported greater problems relative to the youth.

Relations among latent profile groups and demographic and study characteristics. In Table 3 we report tests of relations between the latent profile groups and the child demographic characteristics of gender, age, and ethnicity. We observed nonsignificant relations between profile group composition and child gender, age, and ethnicity after correcting for multiple comparisons using Holm’s stepdown Bonferroni correction. Nonsignificant differences between profile group composition and ethnicity were found regardless of whether ethnicity was coded dichotomously (African American vs. not) or assessed across the multiple ethnic groups observed in the sample (Asian or Pacific Islander; African American, non-Hispanic; White, non-Hispanic; Hispanic; and “other” ethnicity). Despite this, we controlled for these child demographic characteristics in tests of our main hypotheses.

Along with tests of demographic relations with latent profile composition we also examined demographic relations with assessment site (coded as community [$n = 300$] vs. university [$n = 120$] clinic). Applying Holm’s stepdown Bonferroni correction we observed nonsignificant relations between assessment site and child age and child gender and significant relations between assessment site and ethnicity (coded as African American vs. not), $\kappa = .73$, $p < .001$. This level of kappa is considered “substantial” (Landis & Koch, 1977) and indicates that the ethnicity covariate we used in our regression tests of our main hypothesis was essentially statistically redundant with assessment site.

In addition, we noted previously that the first 100 caregiver participants did not complete self-reports of their mood symptoms on the BDI and MDQ. We examined this study characteristic in relation to the latent profile groups and observed a nonsignificant relation between latent profile groups and whether the caregiver completed self-reports on the BDI and MDQ versus the GBI.

One concern with this sample is that caregivers differed in terms of their relationship to the child being rated (e.g., biological mothers and fathers, adoptive mothers and fathers, stepmothers and stepfathers, grandparents), and prior work has identified differences in correlations among different pairs of informants (parent–youth, parent–teacher, teacher–youth; Achenbach, 2006). However, whether or not the caregiver was the biological mother of the youth was not significantly related to classification in informant discrepancies latent profile groups, $\chi^2(3) = 2.78$, $\Phi = .08$, $p > .40$. Further, nonsignificant relations between latent profile

classification and caregiver type were observed when specifically comparing biological mothers in the sample ($n = 306$) to biological fathers ($n = 24$), $\chi^2(3) = 1.54$, $\Phi = .07$, $p > .65$. This lack of significant relations between level of discrepancies between caregiver and child reports and caregiver type is consistent with prior work (De Los Reyes et al., 2008).

Comparability of four-class solution with solution based on raw difference scores. To ensure that our latent profile solution was not identified simply as a function of the method of discrepancies assessment used, we compared the four-profile solution identified using SDS indicators and the four-profile solution identified using raw difference score indicators (i.e., non-standardized scores; De Los Reyes & Kazdin, 2004). The structure and composition of the four-profile solution based on SDS indicators was essentially identical to the four-profile solution obtained when based on raw difference score indicators, Cramer’s $V = .96$, $\kappa = .96$, $p < .001$.

Internal Consistency of Caregiver–Youth Discrepancies

We tested the internal consistency of caregiver–youth reporting discrepancies by examining the eight caregiver–youth reporting discrepancies scores taken across the CBCL/YSR syndromes. Specifically, we estimated Cronbach’s alpha using the eight discrepancy scores for each dyad as “items” measuring a general tendency to disagree. We also repeated the analysis comparing the internal consistency for the extreme groups compared to the moderate groups in terms of latent profiles of agreement. That is, we conducted tests of internal consistency within groups of dyads in which either caregiver or youth reported much greater levels of youth psychological concerns (relative to each other), versus groups of dyads within which there was only slightly greater reporting by caregiver or youth (relative to each other). We conducted tests between these two groups as opposed to the four groups separately because LPA seeks to create profiles for cases exhibiting similar patterns of scores across indicators. As such, within any one profile the variance between cases is reduced so as to maximize the variance between different profiles. The internal consistency of informant discrepancies hinges on there being discrepancies between informants within a case, as well as variability in the extent of the discrepancies between cases of pairs of informants (Rogosa et al., 1982; Rogosa & Willett, 1983). Thus, profile groups that shared similar magnitudes of informant discrepancies but different directions in these discrepancies were grouped together to test whether the

magnitudes of informant discrepancies related to the internal consistency levels of discrepancies.

Based on the total sample, the internal consistency estimate for the eight SDS for the CBCL/YSR syndromes was $\alpha = .89$. This level of consistency is quite high given the low number of items on the scale and well within the range considered acceptable estimates of internal consistency for clinical instruments (Nunnally & Bernstein, 1994). Further, consistent with the idea that greater individual differences between caregiver and youth reports should translate into greater internal consistency estimates for informant discrepancies (Rogosa et al., 1982; Rogosa & Willett, 1983), the internal consistency estimate for the dyads assigned to the Youth Reports Much Greater on Average and Caregiver Reports Much Greater on Average profile groups was far higher ($\alpha = .96$, $n = 100$), relative to the internal consistency estimate for the dyads assigned to the Youth Reports Slightly Greater on Average and Caregiver Reports Slightly Greater on Average profile groups ($\alpha = .71$, $n = 320$). A Feldt (1969) test comparing the difference between these two alpha coefficients revealed a significant difference between them, $F(95, 219) = 7.25$, $p < .001$.

Caregiver–Youth Discrepancies and Youth, Caregiver, and Family Characteristics

We tested the relation between patterns of caregiver–youth reporting discrepancies and youth, caregiver, and family characteristics through a series of hierarchical multiple regression analyses. Specifically, in each analysis youth age, gender (coded “0” for males, $n = 217$; “1” for females, $n = 203$), and ethnicity (coded “0” for African American, $n = 285$; “1” for all other ethnicities, $n = 135$) were entered in the first step as independent variables, and the latent profile pattern was entered in the second step. The youth, caregiver, and family characteristics described previously were entered as the dependent variables in separate regression analyses (see Table 1). Finally, for those characteristics that uniquely related to discrepancies via these analyses, we examined whether these characteristics were still related with discrepancies when taking into account the variance explained by the other informant discrepancies examined. For instance, if we found that youth self-reports of mood problems related to discrepancies, we would conduct a follow-up test of this relation to examine whether these effects would remain significant when controlling for youth demographics, as well as interviewer reports of the youth, caregiver and interviewer reports of family functioning, and caregivers’ self-reported mood.

To enhance interpretability of the results, we recoded the latent profile assignments of caregiver–youth

reporting discrepancies so that scores in the positive direction reflected the youth self-reporting more problems than the caregiver reported in the youth (i.e., 1 = Caregiver Reports Much Greater on Average, 2 = Caregiver Reports Slightly Greater on Average, 3 = Youth Reports Slightly Greater on Average, 4 = Youth Reports Much Greater on Average).

Due to missing data on the dependent variables, sample sizes varied for each test. Specifically, the missing data were almost entirely composed of caregiver report about own mood, as the MDQ and BDI were added in a protocol change during the 2nd year of data collection. However, as mentioned previously whether caregivers completed self-reports based on the MDQ and BDI was not significantly related to the latent profile groups of caregiver–youth reporting discrepancies. Sample sizes for each of the regression analyses reported for separate characteristics are based on the sample sizes for the dependent variables reported in Table 1.

Youth characteristics. As shown in Table 4, results for analyses of the relation between caregiver–youth discrepancies and youth characteristics were consistent for youth self-reports and caregiver reports and not for interviewer reports. Variables entered in the first step contributed significant variance to the regression model for caregiver and youth reports of youth depressive symptoms. However, the first step was not significant for caregiver and youth reports of youth manic symptoms. In the second step, caregiver–youth discrepancies were significantly and positively related to caregiver and youth reports of youth depressive symptoms ($\beta = -.22$, part $r = -.22$; $\beta = .47$, part $r = .46$, respectively; both $ps < .001$). In the second step, caregiver–youth reporting discrepancies were significantly and positively related to caregiver and youth reports of youth manic symptoms ($\beta = -.19$, part $r = -.19$; $\beta = .29$, part $r = .29$, respectively; both $ps < .001$).

For interviewer reports, variables entered in the first step contributed significant variance to the model for the interviewer reports of youth depressive and manic symptoms. However, the first step was not significant for the interviewer report of youth clinical severity. In the second step, caregiver–youth discrepancies were not related to interviewer reports of youth depressive symptoms ($\beta = 0$, part $r = 0$, *ns*), interviewer reports of youth manic symptoms ($\beta = -.07$, part $r = -.07$, *ns*), or the interviewer reports of youth clinical severity ($\beta = -.06$, part $r = -.06$, *ns*).

Caregiver characteristics. Results for analyses of the relation between caregiver–youth discrepancies and caregiver characteristics were consistently nonsignificant for both caregiver self-reports. For both the caregiver

TABLE 4

Dependent Variables for Separate Multiple Regression Tests of the Relation Between Caregiver-Youth Reporting Discrepancies and Youth, Caregiver, and Family Characteristics, Controlling for Youth Age, Youth Gender, and Youth Ethnicity/Race

Measure	First Step R	First Step R ² Δ	First Step FΔ (df)	Second Step R ² Δ	Second Step FΔ (df)
Caregiver re: Mania	.13	.02	2.58 (3, 414)	.03	15.91*** (1, 413)
Caregiver re: Depression	.25	.06	9.21*** (3, 409)	.05	22.44*** (1, 408)
Adolescent re: Mania	.12	.01	2.26 (3, 416)	.08	38.33*** (1, 415)
Adolescent re: Depression	.25	.06	9.02*** (3, 414)	.21	119.31*** (1, 413)
Interviewer re: Mania	.16	.02	3.58* (3, 389)	0	2.01 (1, 388)
Interviewer re: Depression	.32	.10	15.17*** (3, 391)	0	0 (1, 390)
Caregiver: Own Mania	.07	0	.64 (3, 325)	0	1.33 (1, 324)
Caregiver: Own Depression	.04	0	.21 (3, 324)	0	1.02 (1, 323)
Interviewer re: Family	.17	.03	3.93** (3, 408)	.01	3.35 (1, 407)
Caregiver re: Family	.15	.02	3.28* (3, 413)	0	2.13 (1, 412)
Interviewer Severity	.07	0	.63 (3, 407)	0	1.47 (1, 406)

Note: Caregiver re: Mania = Caregiver report of youth manic symptoms on Mood Disorder Questionnaire (MDQ); Caregiver re: Depression = Caregiver report of youth depressive symptoms on General Behavior Inventory (GBI); Adolescent re: Mania = Youth self-report of manic symptoms on MDQ; Adolescent re: Depression = Youth self-report of depressive symptoms on GBI; Interviewer re: Mania = Interviewer report of youth manic symptoms on Young Mania Rating Scale; Interviewer re: Depression = Interviewer report of youth depressive symptoms on Child Depression Rating Scale-Revised; Caregiver: Own Mania = Caregiver self-report of manic symptoms on MDQ; Caregiver: Own Depression = Caregiver ratings of own depressive symptoms on Beck Depression Inventory; Interviewer re: Family = Interviewer report of family functioning on Global Family Environment Scale; Caregiver re: Family = caregiver report of family functioning on Family Assessment Device; Interviewer Severity = Interviewer report of youth clinical severity on Clinical Global Impressions scale.

* $p < .05$. ** $p < .01$. *** $p < .001$, two tailed. All other tests $p > .05$. All tests remained statistically significant (i.e., $p < .05$) after applying Holm's stepdown Bonferroni correction.

self-reports of depressive and manic symptoms, variables entered in the first step did not contribute significant variance to the model. In the second step, caregiver-youth discrepancies were not related to caregiver self-reports of depressive and manic symptoms ($\beta = -.06$, part $r = -.05$, ns ; $\beta = -.06$, part $r = -.06$, ns , respectively).

Family characteristics. Results for analyses of the relation between caregiver-youth discrepancies and family characteristics were nonsignificant for caregiver and interviewer reports. For the caregiver report, variables entered in the first step contributed significant variance to the model. However, in the second step caregiver-youth reporting discrepancies were not related to caregiver reports of family functioning ($\beta = -.07$, part $r = -.07$, ns). For the interviewer report, variables entered in the first step contributed significant variance to the model. However, in the second step, reporting discrepancies were not related to interviewer reports of family functioning ($\beta = -.09$, part $r = -.09$, ns).

Tests jointly considering youth, caregiver, and family characteristics. Caregiver reports and youth self-reports of youth characteristics were consistently related to caregiver-youth reporting discrepancies. Thus, we were interested in testing whether these relations would remain significant when taking into account not only youth demographic characteristics but also

caregiver and family characteristics and interviewer reports of youth characteristics. To test this we ran the same hierarchical regression analyses described previously except that in the first step, along with youth demographic characteristics we simultaneously entered the two caregiver and interviewer reports of family functioning (FAD and GFES), the two caregiver self-reports of mood problems (C-MDQC and C-BDIC), and the three interviewer reports of youth mood problems and clinical severity (YMRS, CDRS, and CGI). Results are presented in Table 5.

For the caregiver and youth reports of youth depressive and manic symptoms, variables entered in the first step contributed significant variance to the model. In the second step each of the caregiver reports and youth self-reports of youth depressive and manic symptoms remained significantly related to caregiver-youth discrepancies: caregiver-reported youth depressive symptoms ($\beta = -.15$, part $r = -.14$, $p < .01$), caregiver-reported youth manic symptoms ($\beta = -.12$, part $r = -.12$, $p < .05$), youth self-reported youth depressive symptoms ($\beta = .52$, part $r = .50$, $p < .001$), and youth self-reported youth manic symptoms ($\beta = .33$, part $r = .32$, $p < .001$).

Follow-up analyses: relations between informant characteristics and individual informants' reports. To supplement the main tests of our hypotheses, we calculated bivariate correlations between all of the informant characteristics indices listed in Table 4 and the broadband indices from the CBCL and the YSR: Total

TABLE 5
 Dependent Variables for Separate Multiple Regression Tests of the Relation Between Caregiver-Youth Reporting Discrepancies and Youth Characteristics, Controlling for Youth Age, Youth Gender, Youth Ethnicity/Race, Caregiver and Family Characteristics, and Interviewer Reports of Youth Mood and Severity

Measure	First Step R	First Step R ² Δ	First Step FΔ (df)	Second Step R ² Δ	Second Step FΔ (df)
Caregiver re: Mania	.54	.29	12.57*** (10, 300)	.01	6.26* (1, 299)
Caregiver re: Depression	.68	.46	25.97*** (10, 296)	.02	11.57** (1, 295)
Adolescent re: Mania	.34	.11	3.98*** (10, 300)	.10	39.78*** (1, 299)
Adolescent re: Depression	.43	.18	6.79*** (10, 298)	.25	130.17*** (1, 297)

Note: Caregiver re: Mania = Caregiver report of youth manic symptoms on Mood Disorder Questionnaire (MDQ); Caregiver re: Depression = Caregiver report of youth depressive symptoms on General Behavior Inventory (GBI); Adolescent re: Mania = Youth self-report of manic symptoms on MDQ; Adolescent re: Depression = Youth self-report of depressive symptoms on GBI.

p* < .05. *p* < .01. ****p* < .001, two tailed. All tests remained statistically significant (i.e., *p* < .05) after applying Holm's stepdown Bonferroni correction.

Externalizing, Total Internalizing, and Total Problem (Table 6). The two interviewers' reports of youth mood symptoms were significantly related with all but one of the caregiver-reported scales and related to only two of the child-reported scales. This is consistent with prior work suggesting that in clinical assessments of the youth, when caregiver and child disagree about the youths' problem behavior the interviewer tends to agree more with the caregiver than the youth (Hawley & Weisz, 2003). Further, caregiver self-reported mood symptoms were consistently and positively related to their reports of youth problem behavior, consistent with prior work suggesting that greater levels of caregivers' mood symptoms are related to caregivers reporting greater levels of problem behavior in the youth

(Richters, 1992). In addition, interviewer reports and not caregiver reports of family functioning were related to reports of youth problem behavior and only for youth self-report of externalizing behavior and total problem behavior, but not internalizing problems. This is consistent with prior work in community samples indicating that the relation between reports of youth problem behavior and family functioning tend to largely arise based on youth report (although in this research the youth was also the reporter of family functioning; see Treutler & Epkins, 2003). Moreover, CBCL and YSR reports were positively correlated with interviewer reports of youth clinical severity, consistent with prior work (Achenbach, 1991a, 1991b). Therefore, the patterns of the relations among informant characteristics

TABLE 6
 Bivariate Correlations Between Broadband CBCL and YSR Scales and Dependent Variables Used in Multiple Regression Tests of the Relation Between Caregiver-Youth Reporting Discrepancies and Youth, Caregiver, and Family Characteristics

Measure	CBCL Ext	CBCL Int	CBCL Tot	YSR Ext	YSR Int	YSR Tot
Caregiver re: Mania	.41	.27	.45	.20	.08	.19
Caregiver re: Depression	.23	.57	.43	.11	.24	.22
Adolescent re: Mania	.09	.14	.13	.49	.45	.55
Adolescent re: Depression	.02	.23	.11	.51	.72	.70
Interviewer re: Mania	.21	.19	.22	.14	.08	.14
Interviewer re: Depression	.03	.45	.20	.16	.35	.29
Caregiver: Own Mania	.20	.31	.31	.14	.16	.19
Caregiver: Own Depression	.21	.26	.26	.18	.15	.19
Interviewer re: Family	-.10	-.02	-.04	-.17	-.15	-.18
Caregiver re: Family	.13	.16	.10	.09	.01	.04
Interviewer Severity	.32	.31	.37	.21	.22	.25

Note: CBCL = Child Behavior Checklist; YSR = Youth Self-Report; Ext = Total Externalizing, Int = Total Internalizing; Tot = Total Problem; Caregiver re: Mania = Caregiver report of youth manic symptoms on Mood Disorder Questionnaire (MDQ); Caregiver re: Depression = Caregiver report of youth depressive symptoms on General Behavior Inventory (GBI); Adolescent re: Mania = Youth self-report of manic symptoms on MDQ; Adolescent re: Depression = Youth self-report of depressive symptoms on GBI; Interviewer re: Mania = Interviewer report of youth manic symptoms on Young Mania Rating Scale; Interviewer re: Depression = Interviewer report of youth depressive symptoms on Child Depression Rating Scale-Revised; Caregiver: Own Mania = Caregiver self-report of manic symptoms on MDQ; Caregiver: Own Depression = Caregiver ratings of own depressive symptoms on Beck Depression Inventory; Interviewer re: Family = Interviewer report of family functioning on Global Family Environment Scale; Caregiver re: Family = caregiver report of family functioning on Family Assessment Device; Interviewer Severity = Interviewer report of youth clinical severity on Clinical Global Impressions scale.

Bold indicates *p* < .001.

and individual caregiver and child reports were in line with prior research.

DISCUSSION

Summary of Main Findings

There were four main findings that extended the literature on informant discrepancies in clinical assessments. First, using LPAs we identified caregiver–youth dyads that varied in terms of the direction of disagreement and the magnitudes of these disagreements. Second, we identified high levels of internal consistency among eight measures of discrepancies, with these internal consistencies being particularly high for caregiver–youth dyads that exhibited large reporting discrepancies. Third, both caregiver reports and youth self-reports of youth’s mood problems were uniquely related to caregiver–youth reporting discrepancies, even when taking into account caregiver mood symptoms; family functioning; interviewer reports of the youth’s clinical severity and mood symptoms; and youth age, gender, and ethnicity/race. Of interest, caregiver mood symptoms and family functioning, and interviewer reports of the youth’s clinical severity and mood symptoms did not significantly relate to discrepancies.

Fourth, we identified a dose-response relation between caregiver–youth reporting discrepancies and caregiver reports and youth self-reports of the youth’s mood problems taken from measures that were completed independently of the measures used to assess caregiver–youth discrepancies. Specifically, as the “dose” of youths reporting greater problems relative to caregivers steadily increased across the profile groupings, youth self-reported mood problems steadily increased and caregiver reports of youth mood problems steadily decreased (Tables 4 and 5). This is an interesting finding in that it not only demonstrates a linear relation between caregiver–youth discrepancies and caregiver and youth reports of youth mood problems, it demonstrates this relation across groups of caregiver–youth dyads that systematically differed in the magnitude and direction of their reporting discrepancies. Thus, informant discrepancy measures exhibit stable psychometric properties and conceptually meaningful relations with the perspectives of the youth’s functioning of the informants providing these reports.

In light of our findings a key question arises: Why did informant discrepancies not relate to associated characteristics that have often been identified in prior work, such as parents’ mood symptoms and family functioning? One reason might be methodological. Prior work has been inconsistent on whether or which informants’ reports of mood symptoms and family functioning relate to informant discrepancies (parent vs. youth; De

Los Reyes et al., 2008; De Los Reyes & Kazdin, 2005; Kolko & Kazdin, 1993; Treutler & Ekins, 2003). At the bivariate level the individual caregiver and youth reports of youth problem behavior were related to caregiver, youth, and interviewer reports of youth, caregiver, and family characteristics, in line with prior work (Table 6). A large body of work suggests that the youth, caregiver, and family characteristics we examined in relation to informant discrepancies are not only related to discrepancies but to the development and maintenance of the youth psychopathology for which caregivers and youths provided discrepant reports. Stated another way, youth, caregiver, and family characteristics substantively covary with the target of the behavioral reports used to assess informant discrepancies (i.e., youth psychopathology). This is a phenomenological issue endemic to much of the prior work on informant discrepancies (e.g., De Los Reyes et al., 2008; De Los Reyes & Kazdin, 2005; K. Guion et al., 2009; Pelton, Steele, Chance, & Forehand, 2001). Most crucially, as far as we are aware this study comprises the most comprehensive examination to date of informant characteristics in relation to discrepancies, in that we (a) were well powered to detect effects identified in prior work ($N = 420$); (b) included 11 indices of informant characteristics taken from caregiver, youth, and interviewer reports; (c) examined eight indices of informant discrepancies via sophisticated latent analytic modeling; and (d) examined informant characteristics jointly to account for both shared and unique variance in relation to discrepancies.

In sum, methodological aspects of the study are an unlikely explanation of our findings. Rather, we argue that the most parsimonious interpretation of our findings is that caregivers and youths have unique, reliable, and valid perspectives on how and where they observe youth problem behavior. This could be why only caregiver and youth reports of youth mood problems related to caregiver–youth reporting discrepancies on measures completed independently of the youth mood reports. This could also be why informant characteristics widely postulated to contribute to informants’ reporting biases did not relate to discrepancies. Indeed, the unique perspectives of caregivers and youths as informants of youths’ behavior may not be indicative of faulty reporting or bias. Instead, informants’ unique perspectives may reflect the notion that caregivers observe youth behaviors that youths either do not observe or do not attend to when providing reports and vice versa (see Kraemer et al., 2003). Of importance, these disagreements between caregiver and youth report should not be automatically interpreted as “bias.” Indeed, interviewer ratings of youth mood symptoms were not significantly related to caregiver–youth discrepancies (Table 4).

Limitations

There are limitations to the present study. First, discrepancies were assessed using standardized difference scores. Prior work has raised concerns about the reliability of difference scores for assessing constructs such as discrepancies and general variation between scores (e.g., De Los Reyes & Kazdin, 2004; Nunnally & Bernstein, 1994; Rogosa et al., 1982; Rogosa & Willett, 1983). However, prior work has noted that when differences between measurements are high, difference scores demonstrate acceptable levels of reliability (Rogosa et al., 1982; Rogosa & Willett, 1983). Indeed, we observed this in the present study. Because the size of the discrepancies and their reliability may vary across samples, we encourage future research to use other strategies besides difference scores for assessing discrepancies, including direct assessments of informants' perceptions of discrepancies.

Second, sample characteristics could limit the generalizability of the findings. We studied a clinic sample that provided a useful test of our hypotheses. At the same time, our findings may not generalize to other samples and informant pairs. For example, in clinic samples in which informants are unrelated to each other and view the behavior being rated in different contexts (e.g., parents and teachers), informants may also differ in whether they both evidence mood concerns of their own and/or deficits in family functioning. As such, depressed mood and/or family functioning may *appear* to influence the degree of dyadic discrepancies, in large part because one informant (e.g., parent) consistently evidences characteristics that the other informant does not (e.g., teacher). However, in this case it would still be unclear whether depression actually played a role in one informant reporting behaviors that the other informant did not report or alternatively that one informant just happened to differ from the other informant on certain characteristics. In any event, it is important that future work extends our findings to other clinic and nonclinic samples for which informant discrepancies are a concern.

Third, we examined caregiver–youth discrepancies and yet caregiver type varied widely in the study. We demonstrated that variation in caregiver type in the sample had no bearing on the nature and structure of informant discrepancies. Yet, like most research on informant discrepancies, the sample was predominantly composed of biological mothers (see De Los Reyes & Kazdin, 2005). We encourage future work to replicate and extend our findings to the study of caregiver–youth discrepancies with larger samples that exhibit increased heterogeneity of types of caregivers in the sample.

Implications for Research, Policy, and Practice

Our findings have significant research and clinical implications for the assessment and treatment of youth psychopathology. First, the greatest levels of internal consistency we observed were within dyads that substantially disagreed in their reports. Although consistent with prior work on the psychometrics of difference scores, these findings contradict prevailing views in the clinical literature, where discrepancies are often viewed as representative of the unreliability of informants' reports. The view of discrepancies as indicating the unreliability of informants' reports might be partially attributable to our field's focus on single characteristics such as an informant's mood symptoms as explanatory factors for the presence of discrepancies. Focusing on a unique characteristic also increases the likelihood that when discrepancies arise, one is quicker to dismiss an informant's ratings as unreliable if they happen to possess this characteristic. An example may be if the parent and teacher ratings of a youth are discrepant and the parent also exhibits depressive symptoms. Often the assumption will be that if the parent is depressed, then they cannot possibly be a reliable informant of the youth's behavior problems.

Perhaps thinking about discrepancies as primarily due to single characteristics of the informants does a disservice to understanding discrepancies and using the outcomes of clinical assessments in assessment and treatment research. Instead, our findings suggest that future work should focus on characteristics that relate specifically to the assessed behaviors and the circumstances through which informants observe the assessed behaviors. In particular, two of these characteristics, the perspectives by which informants observe behavior (e.g., self vs. other) and the contexts within which informants observe behavior (e.g., home vs. school), have been implicated as key domains for explaining why informants disagree (Achenbach, 2006; De Los Reyes & Kazdin, 2005; Kraemer et al., 2003). Indeed, conceptualizing the study of informant discrepancies in this way is consistent with decades of basic research focused on understanding why different people often observe the same behaviors in different ways (e.g., Pronin, 2008). Thus, future investigations should focus squarely on identifying the circumstances in which informant discrepancies reveal meaningful information on how (perspectives) and where (context) informants observe children's behavior.

Second, if discrepant reports are internally consistent across measurements, they can be used to understand how informants view clinical problems similarly or differently. Because informants often disagree on which concerns warrant treatment (Hawley & Weisz, 2003), understanding these discrepancies prior to treatment

may function as a tool for engaging clients in treatment and establishing treatment goals that are perceived as concordant with the client's views. This may facilitate treatment planning, treatment adherence, and monitoring treatment progress. Specifically, researchers have posited that understanding informant discrepancies may aid in rapport building and improvements in treatment outcomes (Yeh & Weisz, 2001). Consistent with these ideas, three recent studies have found that informant discrepancies on pretreatment youth concerns predict posttreatment outcomes (fewer discrepancies predict better outcomes) and treatment process factors (fewer discrepancies predict fewer session cancels, lower likelihood of treatment dropout, and greater number of therapy visits; Brookman-Frazee, Haine, Gabayan, & Garland, 2008; Ferdinand, van der Ende, & Verhulst, 2006; Jensen Doss & Weisz, 2008).

Our findings suggest that researchers and practitioners should view informant discrepancies as opportunities to gather rich clinical information on the problems being assessed. Gathering information about how informants view the youth's problems differently may result in a greater understanding of how to intervene to change these problems. For example, if informants who are also participants in treatment disagree on the presence of some concerns (e.g., aggression) but agree on the presence of other concerns (e.g., parent-youth conflict), it may benefit rapport building and treatment adherence if the clinician targeted "agreement" concerns before attempting to target "discrepant" concerns. Thus, use of discrepancies to plan treatment may increase the likelihood that participants in treatment will be actively engaged in treatment. However, the systematic use of discrepancies to plan treatment has not been the subject of experimental research. Therefore, we encourage future research to address these issues in both laboratory and nonlaboratory (e.g., community mental health center) settings.

REFERENCES

- Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist 4-18 and 1991 Profile*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T. M. (1991b). *Manual for the Youth Self Report Form and 1991 Profile*. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T. M. (2006). As others see us: Clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science*, *15*, 94-98.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213-232.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Barker, E. T., Bornstein, M. H., Putnick, D. L., Hendricks, C., & Suwalsky, J. T. D. (2007). Adolescent-mother agreement about adolescent problem behaviors: Direction and predictors of disagreement. *Journal of Youth and Adolescence*, *36*, 950-962.
- Bartholomew, D. J., Steele, F., Moustaki, I., & Galbraith, J. I. (2002). *The analysis and interpretation of multivariate data for social scientists*. Philadelphia: Chapman & Hall/CRC.
- Beck, A. T., & Steer, R. A. (1987). *Beck Depression Inventory manual*. San Antonio, TX: Psychological Corporation.
- Briggs-Gowan, M. J., Carter, A. S., & Schwab-Stone, M. (1996). Discrepancies among mother, child, and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology*, *24*, 749-765.
- Brookman-Frazee, L., Haine, R. A., Gabayan, E. N., & Garland, A. F. (2008). Predicting frequency of treatment visits in community-based youth psychotherapy. *Psychological Services*, *5*, 126-138.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Danielson, C. K., Youngstrom, E. A., Findling, R. L., & Calabrese, J. R. (2003). Discriminative validity of the General Behavior Inventory using youth report. *Journal of Abnormal Child Psychology*, *31*, 29-39.
- Deater-Deckard, K. (1998). Parenting stress and child adjustment: Some old hypotheses and new questions. *Clinical Psychology: Science and Practice*, *5*, 314-332.
- De Los Reyes, A., Goodman, K. L., Kliever, W., & Reid-Quinones, K. R. (2008). Whose depression relates to discrepancies? Testing relations between informant characteristics and informant discrepancies from the perspectives of both informants. *Psychological Assessment*, *20*, 139-149.
- De Los Reyes, A., Goodman, K. L., Kliever, W., & Reid-Quinones, K. R. (2010). The longitudinal consistency of mother-child reporting discrepancies of parental monitoring and their ability to predict child delinquent behaviors two years later. *Journal of Youth and Adolescence*, *39*, 1417-1430.
- De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology*, *37*, 637-652.
- De Los Reyes, A., & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment*, *16*, 330-334.
- 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.
- De Los Reyes, A., & Kazdin, A. E. (2006a). Conceptualizing changes in behavior in intervention research: The range of possible changes model. *Psychological Review*, *113*, 554-583.
- De Los Reyes, A., & Kazdin, A. E. (2006b). Informant discrepancies in assessing child dysfunction relate to dysfunction within mother-child interactions. *Journal of Child and Family Studies*, *15*, 643-661.
- De Los Reyes, A., & Kazdin, A. E. (2008). When the evidence says, "Yes, no, and maybe so": Attending to and interpreting inconsistent findings among evidence-based interventions. *Current Directions in Psychological Science*, *17*, 47-51.
- De Los Reyes, A., & Kazdin, A. E. (2009). Identifying evidence-based interventions for children and adolescents using the range of possible changes model: A meta-analytic illustration. *Behavior Modification*, *33*, 583-617.
- Depue, R. A., Krauss, S., Spont, M. R., & Arbisi, P. (1989). General Behavior Inventory identification of unipolar and bipolar affective conditions in a nonclinical university population. *Journal of Abnormal Psychology*, *98*, 117-126.

- Depue, R. A., Slater, J. F., Wolfstetter-Kausch, H., Klein, D. N., Goplerud, E., & Farr, D. A. (1981). A behavioral paradigm for identifying persons at risk for bipolar depressive disorder: A conceptual framework and five validation studies. *Journal of Abnormal Psychology, 90*, 381–437.
- Du Rocher Schudlich, T. D., Youngstrom, E. A., Calabrese, J. R., & Findling, R. L. (2008). The role of family functioning in bipolar disorder in families. *Journal of Abnormal Child Psychology, 36*, 849–863.
- Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy, 9*, 171–180.
- Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. *Psychometrika, 34*, 363–373.
- Ferdinand, R. F., van der Ende, J., & Verhulst, F. C. (2004). Parent-adolescent disagreement regarding psychopathology in adolescents from the general population as a risk factor for adverse outcome. *Journal of Abnormal Psychology, 113*, 198–206.
- Ferdinand, R. F., van der Ende, J., & Verhulst, F. C. (2006). Prognostic value of parent-adolescent disagreement in a referred sample. *European Child & Adolescent Psychiatry, 15*, 156–162.
- Findling, R. L., Frazier, T. W., Youngstrom, E. A., McNamara, N. K., Stansbrey, R. J., Gracious, B. L., et al. (2007). Double-blind, placebo-controlled trial of divalproex monotherapy in the treatment of symptomatic youth at high risk for developing bipolar disorder. *Journal of Clinical Psychiatry, 68*, 781–788.
- Findling, R. L., Youngstrom, E. A., Danielson, C. K., DelPorto, D., Papish David, R., Townsend, L., et al. (2002). Clinical decision-making using the General Behavior Inventory in juvenile bipolarity. *Bipolar Disorders, 4*, 34–42.
- Fristad, M. A., Weller, E. B., & Weller, R. A. (1992). The Mania Rating Scale: Can it be used in children? A preliminary report. *Journal of the American Academy of Child & Adolescent Psychiatry, 31*, 252–257.
- Fristad, M. A., Weller, R. A., & Weller, E. B. (1995). The Mania Rating Scale (MRS): Further reliability and validity studies with children. *Annals of Clinical Psychiatry, 7*, 127–132.
- Geller, B., Zimmerman, B., Williams, M., Bolhofner, K., Craney, J. L., DelBello, M. P., et al. (2001). Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*, 450–455.
- Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review, 113*, 101–131.
- Guion, K., Mrug, S., & Windle, M. (2009). Predictive value of informant discrepancies in reports of parenting: Relations to early adolescents' adjustment. *Journal of Abnormal Child Psychology, 37*, 17–30.
- Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Hillsdale, NJ: Erlbaum.
- Hawley, K. M., & Weisz, J. R. (2003). Child, parent, and therapist (dis)agreement on target problems in outpatient therapy: The therapist's dilemma and its implications. *Journal of Consulting and Clinical Psychology, 71*, 62–70.
- Hirschfeld, R. M. (2001). Bipolar spectrum disorder: Improving its recognition and diagnosis. *Journal of SClinical Psychiatry, 62 Suppl 14*, 5–9.
- Hirschfeld, R. M. A., Williams, J. B. W., Spitzer, R. L., Calabrese, J. R., Flynn, L., Keck, P. E., et al. (2000). Development and validation of a screening instrument for bipolar spectrum disorder: The Mood Disorder Questionnaire. *American Journal of Psychiatry, 157*, 1873–1875.
- Israel, P., Thomsen, P. H., Langeveld, J. H., & Stormark, K. M. (2007). Parent-youth discrepancy in the assessment and treatment of youth in usual clinical care setting: Consequences to parent involvement. *European Child & Adolescent Psychiatry, 16*, 138–148.
- Jensen Doss, A., & Weisz, J. R. (2008). Diagnostic agreement predicts treatment process and outcomes in youth mental health clinics. *Journal of Consulting and Clinical Psychology, 76*, 711–722.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 980–988.
- Kazdin, A. E., & De Los Reyes, A. (2007). Conduct disorder. In R. J. Morris, & T. R. Kratochwill, (Eds.), *The practice of child therapy* (4th ed., pp. 207–247). Philadelphia: Taylor & Francis.
- Kazdin, A. E., & Wassell, G. (2000). Therapeutic changes in children, parents, and families resulting from treatment of children with conduct problems. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 414–420.
- Klein, D. N., Depue, R. A., & Slater, J. F. (1986). Inventory identification of cyclothymia. IX. Validation in offspring of bipolar I patients. *Archives of General Psychiatry, 43*, 441–445.
- Klein, D. N., Dickstein, S., Taylor, E. B., & Harding, K. (1989). Identifying chronic affective disorders in outpatients: Validation of the General Behavior Inventory. *Journal of Consulting and Clinical Psychology, 57*, 106–111.
- Koenig, K., De Los Reyes, A., Cicchetti, D., Scahill, L., & Klin, A. (2009). Group intervention to promote social skills in school-age children with pervasive developmental disorders: Reconsidering efficacy. *Journal of Autism and Developmental Disorders, 39*, 1163–1172.
- Kolko, D. J., & Kazdin, A. E. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent, and teacher reports. *Journal of Child Psychology & Psychiatry, 34*, 991–1006.
- Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry, 160*, 1566–1577.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.
- Mallon, J. C., Klein, D. N., Bornstein, R. F., & Slater, J. F. (1986). Discriminant validity of the General Behavior Inventory: an outpatient study. *Journal of Personality Assessment, 50*, 568–577.
- McCutcheon, A. L. (1987). *Latent class analysis*. Newbury Park, CA: Sage.
- Miller, I. W., Kabacoff, R. I., Keitner, G. I., Epstein, N. B., & Bishop, D. S. (1986). Family functioning in families of psychiatric patients. *Comprehensive Psychiatry, 27*, 302–312.
- Nagin, D. S. (2006). *Group-based modeling of development*. Cambridge, MA: Harvard University Press.
- National Institute of Mental Health. (1985). Rating scales and assessment instruments for use in pediatric psychopharmacology research. *Psychopharmacology Bulletin, 21*, 839–843.
- Nottelmann, E., Biederman, J., Birmaher, B., Carlson, G. A., Chang, K. D., Fenton, W. S., et al. (2001). National Institute of Mental Health research roundtable on prepubertal bipolar disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*, 871–878.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Nusslock, R., Abramson, L. Y., Harmon-Jones, E., Alloy, L. B., & Hogan, M. E. (2007). A goal-striving life event and the onset of hypomanic and depressive episodes and symptoms: Perspective from the behavioral approach system (BAS) dysregulation theory. *Journal of Abnormal Psychology, 116*, 105–115.

- Overholser, J. C., Brinkman, D. C., Lehnert, K. L., & Ricciardi, A. M. (1995). Children's Depression Rating Scale-Revised: Development of a short form. *Journal of Clinical Child Psychology, 24*, 443-452.
- Owens, J. S., Goldfine, M. E., Evangelista, N. M., Hoza, B., & Kaiser, N. M. (2007). A critical review of self-perceptions and the positive illusory bias in children with ADHD. *Clinical Child and Family Psychology Review, 10*, 335-351.
- Pelton, J., Steele, R. G., Chance, M. W., & Forehand, R. (2001). Discrepancy between mother and child perceptions of their relationship: II. Consequences for children considered within the context of maternal physical illness. *Journal of Family Violence, 16*, 17-35.
- Poznanski, E. O., Miller, E., Salguero, C., & Kelsh, R. C. (1984). Preliminary studies of the reliability and validity of the Children's Depression Rating Scale. *Journal of the American Academy of Child Psychiatry, 23*, 191-197.
- Pronin, E. (2008). How we see others and how we see ourselves. *Science, 320*, 1177-1180.
- Raftery, A. E. (1986). Choosing models for cross-classifications. *American Sociological Review, 51*, 145-146.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111-163.
- Rey, J. M., Singh, M., Hung, S.-f., Dossetor, D. R., Newman, L., Plapp, J. M., et al. (1997). A global scale to measure the quality of the family environment. *Archives of General Psychiatry, 54*, 817-822.
- Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin, 112*, 485-499.
- Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92*, 726-748.
- Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement, 20*, 335-343.
- Treutler, C. M., & Epkins, C. C. (2003). Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent-child relationships? *Journal of Abnormal Child Psychology, 31*, 13-27.
- van der Oord, S., Prins, P. J. M., Oosterlann, J., & Emmelkamp, P. M. G. (2006). The association between parenting stress, depressed mood and informant agreement in ADHD and ODD. *Behaviour Research and Therapy, 44*, 1585-1595.
- Wagner, K. D., Hirschfeld, R., Emslie, G. J., Findling, R. L., Gracious, B. L., & Reed, M. L. (2006). Validation of the Mood Disorder Questionnaire for bipolar disorder in adolescents. *Journal of Clinical Psychiatry, 67*, 827-830.
- Weisz, J. R., Jensen Doss, A., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A review and critique of the evidence base. *Annual Review of Psychology, 56*, 337-363.
- Wood, J. J., McLeod, B. D., Sigman, M., Hwang, W.-C., & Chu, B. C. (2003). Parenting and childhood anxiety: Theory, empirical findings, and future directions. *Journal of Child Psychology and Psychiatry, 44*, 134-151.
- Yeh, M., & Weisz, J. R. (2001). Why are we here at the clinic? Parent-child (dis)agreement on referral problems at outpatient treatment entry. *Journal of Consulting and Clinical Psychology, 69*, 1018-1025.
- Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: Reliability, validity, and sensitivity. *British Journal of Psychiatry, 133*, 429-435.
- Youngstrom, E. A., Findling, R. L., & Calabrese, J. R. (2003). Who are the comorbid adolescents? Agreement between psychiatric diagnosis, youth, parent, and teacher report. *Journal of Abnormal Child Psychology, 31*, 231-245.
- Youngstrom, E. A., Findling, R. L., & Calabrese, J. R. (2004). Effects of adolescent manic symptoms on agreement between youth, parent, and teacher ratings of behavior problems. *Journal of Affective Disorders, 82S*, S5-S16.
- Youngstrom, E. A., Findling, R. L., Calabrese, J. R., Gracious, B. L., Demeter, C., DelPorto Bedoya, D., et al. (2004). Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5 to 17 years. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*, 847-858.
- Youngstrom, E. A., Findling, R. L., Danielson, C. K., & Calabrese, J. R. (2001). Discriminative validity of parent report of hypomanic and depressive symptoms on the General Behavior Inventory. *Psychological Assessment, 1*, 267-276.
- Youngstrom, E. A., Gracious, B. L., Danielson, C. K., Findling, R. L., & Calabrese, J. R. (2003). Toward an integration of parent and clinician report on the Young Mania Rating Scale. *Journal of Affective Disorders, 77*, 179-190.
- Youngstrom, E. A., Izard, C., & Ackerman, B. (1999). Dysphoria-related bias in maternal ratings of children. *Journal of Consulting and Clinical Psychology, 67*, 905-916.
- Youngstrom, E. A., Loeber, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting and Clinical Psychology, 68*, 1038-1050.
- Youngstrom, E. A., Meyers, O. I., Demeter, C., Kogos Youngstrom, J., Morello, L., Piiparinen, R., et al. (2005). Comparing diagnostic checklists for pediatric bipolar disorder in academic and community mental health settings. *Bipolar Disorders, 7*, 507-517.
- Youngstrom, E. A., Meyers, O. I., Youngstrom, J. K., Calabrese, J. R., & Findling, R. L. (2006). Comparing the effects of sampling designs on the diagnostic accuracy of eight promising screening algorithms for pediatric bipolar disorder. *Biological Psychiatry, 60*, 1013-1019.