

Informant Discrepancies in Clinical Reports of Youths and Interviewers' Impressions of the Reliability of Informants

Andres De Los Reyes, Ph.D.,¹ Eric A. Youngstrom, Ph.D.,² Anna J. Swan, B.A.,¹
Jennifer K. Youngstrom, Ph.D.,² Norah C. Feeny, Ph.D.,³ and Robert L. Findling, M.D.⁴

Abstract

Objective: In this study the authors examined whether discrepancies between parent and youth reports of the youth's emotional and behavioral functioning are related to interviewers' reliability ratings of parents and youths.

Methods: In a consecutive case series analysis of 328 parents and youths aged 11–17 years, parents and youths provided reports of youth emotional and behavioral functioning and participated in structured clinical interviews. At the conclusion of the interviews, interviewers rated the reliability of informants. Interviewers rated youths' clinical severity and parents and youths provided information on youth demographics. Nominal logistic regressions tested patterns of discrepancies between parent and youth reports (i.e., which informant consistently reported greater degrees of youth emotional and behavioral concerns than the other) as predictors of interviewers' ratings of the reliability of parents and youths. All analyses controlled for variance explained by youth demographics and youth severity.

Results: When parents reported greater degrees of youth emotional and behavioral concerns than youths self-reported, interviewers were likely to rate the youth as an unreliable informant, and were unlikely to rate the youth as an unreliable informant when parents reported less concerns than youths self-reported. However, interviewers' ratings of parents' reliability did not relate to the discrepancies between reports, regardless of which informant reported greater degrees of youth concerns.

Conclusions: Prior research indicates that informant discrepancies potentially reveal important information of youths' emotional and behavioral concerns, such as the settings in which youths express these concerns. Yet, when parents and youths disagree in their clinical reports of the youth's functioning, this relates to whether a clinical interviewer views the youth as a reliable informant of their own functioning. To increase the cost-effectiveness and clinical utility of multi-informant clinical evaluations, practitioners and researchers should anticipate informant discrepancies and predict what they may represent before conducting clinical evaluations.

Introduction

PRACTITIONERS AND RESEARCHERS who follow best practices while conducting clinical assessments of children and adolescents (collectively referred to as “youths” unless otherwise specified) likely gather multiple informants' reports (Hunsley and Mash 2007). Practitioners and researchers might gather reports from those with whom the youth is well acquainted, such as parents and teachers; research and clinical personnel such as other healthcare professionals and trained laboratory observers; or the youth himself or herself. One challenge in keeping with these best practices is that multiple informants' reports often yield different conclusions, both in research findings and clinical evaluation outcomes (Achenbach et al. 1987; Perrin et al. 2000; Johnson and Wang 2008; De Los Reyes and Kazdin 2009). These differences among multiple informants' reports can collectively be

referred to as “informant discrepancies” and they are some of the most frequently encountered challenges facing practitioners and researchers who work with youths.

The traditional view is that informant discrepancies reflect bias or unreliability in some or all informants' reports (Richters 1992; Krosnick 1999; De Los Reyes and Kazdin 2005, 2006, 2008; Fisher et al. 2006). Yet, two lines of evidence indicate that informant discrepancies may yield information beyond that of unreliability or bias. First, informants agree more when they observe youths in the same setting (e.g., two teachers at school, both parents at home), when what is being assessed is relatively easy for informants to observe (e.g., aggression vs. worry), and when the patient is a younger child (ages 6–11 years) versus an adolescent (Achenbach et al. 1987). Thus, practitioners and researchers can often predict which informants' reports will disagree. Second, because informants often

¹Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland at College Park, College Park, Maryland.

²Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.

³Department of Psychology, Case Western Reserve University, Cleveland, Ohio.

⁴Department of Psychiatry, University Hospitals Case Medical Center, Cleveland, Ohio.

This work was supported, in part, by NIH R01 5MH066647 (PI: E. Youngstrom).

systematically vary in the settings within which they primarily observe youths, discrepancies among informants' reports often reveal meaningful information on where youths express emotional and behavioral problems (for a review, see De Los Reyes 2011). For instance, discrepancies between parent and youth reports about youths' emotional and behavioral functioning are stable both across assessed domains and over the course of controlled trials (De Los Reyes Alfano et al. 2010; De Los Reyes Youngstrom et al. 2011). Further, recent work indicates that discrepancies between parent and teacher reports—both of preschool children's oppositional behavior and older children's aggressive behavior—relate to differences in the specific settings within which children express these behaviors (e.g., home vs. school) (De Los Reyes et al. 2009; Hartley et al. 2011).

In light of work reviewed previously on the utility of informant discrepancies, it is important to note that research also indicates that practitioners hold particular views as to who are the "optimal informants" (e.g., parents, teachers, and youths) from whom to collect clinical information about youths (Loeber et al. 1990). Additionally, these views of optimal informants may play a role in how practitioners use the outcomes of patient evaluations when making decisions about patients (e.g., treatment planning and assessing functional impairment). For example, in clinic settings when parent reports of youth behavior disagree with other informants' reports of youth behavior, practitioners' impressions systematically agree more with parent reports (Hawley and Weisz 2003; Kramer et al. 2004; De Los Reyes Alfano et al. 2011). In fact, even when other informants report youth behavioral and emotional concerns that are not reported by parents, practitioners rate the youth's functioning in line with parent reports (Kramer et al. 2004; De Los Reyes Alfano et al. 2011). This work provides circumstantial evidence to suggest that when discrepancies arise between a parent's report of youth behavior and another informant's report of youth behavior, practitioners might attribute the discrepancies to the parent being "right" or reliable and the other informant being "wrong" or not reliable.

Purpose

This study tests whether informant discrepancies relate to interviewers' impressions of the reliability of informants as reporters of youths' emotional and behavioral functioning. To test our hypotheses, we studied a well-characterized multisite clinic sample for whom prior work indicated that discrepancies between parent and youth reports were both stable across assessments and did not relate to factors commonly viewed as indicative of informant bias (e.g., informants' mood symptoms, and family stress and functioning) (De Los Reyes Youngstrom et al. 2011). In this sample, raters sequentially interviewed both parents and youths separately (i.e., interviewers were blind to other standardized assessments). After the interviews, interviewers separately rated the reliability of parents and youths as informants as "Good," "Fair," or "Poor."

Hypotheses

We expected to find that when parents and youths disagreed in their reports of youths' emotional and behavioral functioning, the direction of the discrepancies (i.e., which informant reported greater degrees of youth problems than the other) would relate to interviewers' ratings of the youth as a "Fair" or "Poor" informant relative to a "Good" informant. Specifically, we expected to find that when parents reported greater degrees of youth problems than youths self-reported, interviewers would be more likely to rate the youth as a "Fair" or "Poor" informant relative to a "Good" in-

formant. Further, when youths self-reported greater degrees of problems than parents reported about youths, we expected to find that interviewers would be unlikely to rate the youth as a "Fair" or "Poor" informant relative to a "Good" informant. However, as mentioned previously, prior work indicates that no matter how parents report relative to other informants (e.g., report greater or lesser degrees of emotional and behavioral problems in a youth), practitioners systematically rate the youth's functioning consistent with parent reports (Kramer et al. 2004; De Los Reyes Alfano et al. 2011). Thus, we only expected informant discrepancies to relate to interviewers' ratings of youths' reliability.

Method

Participants

Participants were 328 youths and their parents seeking outpatient services, recruited from a previous study of 420 youths between 11 and 17 years old and their parents (De Los Reyes Youngstrom et al. 2011). Youth participants met diagnostic criteria for between 0 and 8 *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) (American Psychiatric Association 1994), diagnoses (Median=3) as determined by the Kiddie Schedule for Affective Disorders and Schizophrenia-Plus (KSADS-PL-Plus) interview (Kaufman et al. 1997; Geller et al. 2001; Nottelmann 2001). Using the hierarchical approach indicated in DSM-IV and operationalized to be consistent with prior work by this group (Youngstrom et al. 2001), 37% of youths met primary diagnostic criteria (as determined through diagnostic consensus meetings) for a unipolar depressive disorder; 34% for attention-deficit/hyperactivity disorder (ADHD) or disruptive behavior disorder without the presence of a mood disorder; 5% for bipolar I disorder; 15% for bipolar II disorder, cyclothymic disorder, or bipolar not otherwise specified; and 9% met criteria for other DSM-IV diagnoses. Additionally, 52% met criteria for ADHD in addition to other diagnoses. Youths who were found to have or suspected of having pervasive developmental disorders, psychiatric disorders due to general medical conditions, or some form of mental retardation were excluded.

Measures

Parents reported on youths' emotional and behavioral functioning using the Child Behavior Checklist (CBCL) (Achenbach 1991a; Achenbach and Rescorla 2001). Youths reported on their own emotional and behavioral functioning using the Youth Self-Report (YSR), a measure derived from the CBCL (Achenbach 1991b). The YSR and CBCL assess eight syndrome scales: (1) anxious/depressed, (2) withdrawn/depressed, (3) somatic complaints, (4) thought problems, (5) social problems, (6) attention problems, (7) rule-breaking, and (8) aggressive problems. We calculated raw score totals from these scales only using the items shared across the CBCL and YSR. We used these raw score totals as opposed to the standardized *T* score totals to ensure that we assessed reporting discrepancies holding item content constant across parent and youth reports. To assess discrepancies between parent and youth reports on these scales, we converted each of the parent and youth scores on common items into *z* scores, with these *z* score conversions based on reports taken from parents and youths in the sample. We then subtracted each youth *z* score on one scale from the parent *z* score on the same scale. These eight standardized difference scores were used to examine patterns of discrepancies and in particular instances in which parents reported qualitatively greater degrees of youth emotional and behavioral problems than youths self-reported and vice

versa (De Los Reyes and Kazdin 2004; De Los Reyes et al. 2008; De Los Reyes Goodman et al. 2010; De Los Reyes Youngstrom et al. 2011). Our methods of examining qualitative differences between informants' reports are methodologically consistent with recent recommendations to calculate statistical interactions between informants' reports when assessing informant discrepancies (Laird and Weems 2011; R.D. Laird, pers. comm., April 23, 2010).

Parents and youths participated in the KSADS-PL-Plus interview mentioned previously. The KSADS-PL-Plus is a combination of the KSADS-PL and mood items from the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U KSADS). Interviewers administering the KSADS-PL-Plus included predoctoral psychology interns ($n=165$ cases; 50%), doctoral graduate students ($n=65$ cases; 20%), postgraduate doctoral-level interviewers ($n=51$ cases; 16%), and postgraduate master's-level interviewers ($n=47$ cases; 14%). After the administration of the KSADS-PL-Plus, interviewers rated the reliability of parents and youths as informants based on their observations of parents and youths and their reports during the interview. Specifically, at the end of the summary sheet of the KSADS-PL-Plus that interviewers use to catalog the lifetime diagnoses endorsed by the informant during the interview, interviewers encountered an item prompting them to rate the reliability of the information provided by the informant (i.e., an item that simply presented the prompt "Reliability of Information") as "Good," "Fair," or "Poor." Further, each of these ratings was made by a single interviewer who interviewed both parent and youth, and each rating was taken separately for each informant (i.e., rating was for "absolute reliability" of each informant and not whether one informant was "more reliable" than the other informant). Importantly, interviewers made this rating as a single subjective evaluation of each informant's reliability and they received no training on how to operationally define informant reliability. This also meant that interviewers were not provided with information on guidelines for determining informant reliability or on identifying "optimal informants" such as those delineated elsewhere (e.g., Loeber et al. 1990). In this way, these reliability judgments would generalize to judgments as observed in routine practice settings. The Clinical Global Impressions Scale Severity of Illness score (CGI) was used by the interviewer to rate the youth's clinical severity (National Institute of Mental Health 1985). Finally, youths and parents completed a contact sheet and packet with information about youth age, gender, and ethnicity/race.

Procedure

A consecutive case series of families seeking outpatient services participated in a day-long interview assessing youth and family functioning. Before the study, youths and parents assented and consented, respectively, to participate. Parents and youths were then interviewed separately by a highly trained interviewer using the KSADS-PL-Plus. Parents were interviewed first. When not being interviewed, youths and parents completed a battery of measures assessing youth and parent mood symptoms and functioning as well as family functioning. Participants were then debriefed about the study and provided feedback about assessment outcomes. When requested, a summary of the KSADS-PL-Plus was sent to a healthcare provider.

Analyses

For this study we drew from prior work in this sample on the patterns of discrepancies between parent and youth reports. Spe-

cifically, in a previous study of 420 parent-youth dyads, latent profile analyses revealed that dyads varied in the magnitude (i.e., size of discrepancies) and direction (i.e., who reported greater degrees of youth problems relative to the other informant) of reporting discrepancies (Bartholomew et al. 2002; De Los Reyes Youngstrom et al. 2011). In De Los Reyes Youngstrom et al. (2011), we identified four profiles of dyads that varied in their patterns of reporting discrepancies and these four dyads could be categorized into two higher order groups. Two classes of dyads included parents who reported either far greater or only slightly greater degrees of youth problems across the eight syndrome scales of the CBCL than youths self-reported on these same scales on the YSR. Two classes of dyads evidenced the polar opposite patterns (i.e., youth was the informant self-reporting on the YSR far greater and slightly greater degrees of problems than the parent reported on the CBCL). In this study, we tested our hypotheses using these two higher-order dyad groupings (i.e., Group 1=parents consistently reported greater degrees of youth problems than youths self-reported; Group 2=youths consistently self-reported greater degrees of youth problems than parents reported about youths).

The present study focused on a reduced sample of 328 of these 420 classified dyads that had complete data on the youth and parent reports of youths' emotional and behavioral functioning as well as clinical interviewer reports of youths' clinical severity and informant reliability. For these dyads, a trained interviewer rated the reliability of parents and youths as informants, using procedures described previously and after administration of the KSADS-PL-Plus to each informant. For this study we were interested in comparing the two groups of dyads in which parents reported greater degrees of youth emotional and behavioral problems than the youths self-reported ($n=163$) to the two groups of dyads in which youths self-reported greater degrees of problems than parents reported about youths ($n=165$). The 328 dyads did not differ from the rest of the 420 dyads on the reporting discrepancies patterns, nor did they vary as a function of youth age, gender, ethnicity, youth clinical severity, or assessment site.

We first calculated frequencies and percentages (nominal), and means and standard deviations (continuous) for all covariates used in the main tests of our hypotheses. We subsequently calculated frequencies and percentages of the patterns of reporting discrepancies and interviewer ratings of parent and youth reliability, as well as a cross-tabulation of the relation between these two variables.

For the main tests of our hypotheses, we conducted nominal logistic regressions in which interviewers' reliability reports served as the dependent variable. A dichotomized variable representing instances in which parents reported greater degrees of problem behavior than youths self-reported versus the opposite reporting pattern (youth > parent) served as the key independent variable. We also entered as covariates a number of youth demographic characteristics (youth age, gender, and ethnicity/race) and youth clinical severity (CGI Severity of Illness score). For these analyses, the two continuous variables (youth age and clinical severity) were centered before analyses. Finally, we conducted two regressions, one for interviewer reports of youth reliability and another for parent reliability.

Results

Preliminary analyses

Demographic and clinical characteristics of the analytic sample indicated that the average youth in the sample was roughly 13 years

TABLE 1. DEMOGRAPHIC AND CLINICAL CHARACTERISTICS (N=328)

Characteristic	Statistics
Youth age	M (SD)=13.5 (1.8)
Youth gender	Male (173, 52.7%) Female (155, 47.3%)
Youth ethnicity (% African-American)	224 (68.3%)
Youth Clinical Severity ^a	M (SD)=3.9 (.94)

^aYouth Clinical Severity=CGI Severity of Illness Score.
SD=standard deviation; m=mean.

of age and rated just under “Moderately Ill” on the CGI. The sample had slightly greater numbers of male youths relative to female youths and was predominantly African-American (Table 1). Additionally, as noted in Table 2, chi-square tests revealed a significant relation between patterns of reporting discrepancies and interviewer ratings of youth but not parent reliability. Specifically, dyads were nearly evenly split in terms of how many could be characterized by parents reporting greater degrees of youth problems relative to youth self-reported problems and vice versa. Yet, when parents reported greater degrees of youth problems than youths self-reported, far more interviewers rated youths as “Poor” or “Fair” informants relative to when youths self-reported greater degrees of problems than parents reported about youths. In contrast, the distributions of parents rated by interviewers as “Poor,” “Fair,” or “Good” informants were similar regardless of who reported greater degrees of youth problems.

Predicting interviewers' reliability reports by reporting discrepancies patterns

We present in Table 3 nominal regression analyses representing the main tests of our hypotheses. For interviewer ratings of youth reliability, the only control variables that significantly contributed to the statistical model were youth age and clinical severity. Specifically, using the interviewer rating of “Good” as the contrast group, youths rated by the interviewer as “Poor” or “Fair” informants were significantly more likely to be rated as presenting with a greater illness severity on the CGI. Using the interviewer rating of “Good” as the contrast group, youths higher in age were significantly unlikely to be rated by the interview as “Fair” informants.

Using the interviewer rating of “Good” as the contrast group, when parents consistently reported greater degrees of youth problems than the youth self-reported, interviewers were at significantly greater odds of rating youths as “Poor” or “Fair” informants. Conversely, when youths consistently self-reported greater degrees of problems than the parents reported about youths, interviewers were significantly unlikely to rate youths as “Poor” or “Fair” informants.

For interviewer ratings of parent reliability, the only control variable that significantly contributed to the statistical model was youth age. Using the interviewer rating of “Good” as the contrast group, parents with youths higher in age were significantly likely to be rated by the interview as “Poor” informants. In contrast to our observations for interviewer ratings of youth reliability, patterns of discrepancies between parent and youth reports did not relate to interviewer ratings of parent reliability.

Discussion

Main findings

When faced with informant discrepancies, interviewers often discount the reliability of an informant's report. However, these informant discrepancies relate to interviewer ratings of the reliability of youths but do not relate to such ratings for parents. These findings are likely to generalize because we examined a well-characterized heterogeneous multisite clinic sample in which we have previously identified no hint of evidence suggesting unreliability or bias on the part of either parent or youth reports (De Los Reyes Youngstrom et al. 2011). Then, why did we identify these relations? Our observations are consistent with prior work that indicates that no matter how parents report relative to other informants (e.g., report greater or lesser degrees of youth's emotional and behavioral problems), practitioners systematically rate the youth's functioning consistent with parent reports (Kramer et al. 2004; De Los Reyes Youngstrom et al. 2011). These effects might be due to parents often being the primary referral source for youths' evaluations. Based on this, an interviewer likely anticipates that the parent will report concerns with their youth's emotional and behavioral functioning. If the parent interview confirms this and the youth endorses few problems on the interview, the differences between parent and youth reports may be sufficient for an interviewer to then ultimately judge the reliability of the youth as an informant. Results indicate that interviewers use discrepancies between parent and youth reports as proxies for the reliability of youths' reports but not for parents' reports. If a parent reports problems that the youth does not, then an interviewer likely perceives this youth as not reliable; these relations cannot be explained by youths' demographic characteristics or interviewer impressions of youths' clinical severity (Table 3).

Limitations

To assess informant discrepancies we relied on multiple indices of discrepancies as measured by standardized difference scores (i.e., informants' reports converted to *z* scores with youth report subtracted from parent report) (De Los Reyes and Kazdin 2004; De Los Reyes et al. 2008). One concern is that these measures do not directly assess informants' perceptions of these discrepancies and such an assessment could result in different conclusions. Future research should replicate and extend our findings using other methods of measuring discrepancies. In particular, we encourage researchers to develop standardized measures of informants' views of the extent to which their reports disagree with the reports of other informants.

We did not assess whether variations in perceived reliability of informants relate to variations in interviewers' clinical decisions such as treatment planning and diagnostic formulations. We recommend that future research examines whether interviewers' perceptions of the reliability of informants affect their clinical decisions about patient care.

Interviewers worked first with the parent when the youth was younger than 11 years. For youths 11 years and older (i.e., the current sample), families were given their choice about who would do the interview first; 90% chose to have the parent complete the interview first. Further, interviewers were given no formal instruction on when to make their reliability ratings. As a consequence of these factors, interviewers did not follow a set protocol with regard to when they provided reliability ratings. That is, interviewers varied in when they completed their ratings; some interviewers rated informant reliability after interviewing both parent

TABLE 2. CROSS-TABULATION OF PATTERNS OF DISCREPANCIES BETWEEN PARENT AND YOUTH REPORTS AND INTERVIEWER RATINGS OF INFORMANT RELIABILITY (N=328)

Latent profile assignment	Interviewer reports of youth reliability			
	Poor	Fair	Good	Total
Parent reports greater than youth	28 (8.5%)	90 (27.4%)	45 (13.7%)	163 (49.7%)
Youth reports greater than parent	13 (4%)	71 (21.6%)	81 (24.7%)	165 (50.3%)
Total	41 (12.5%)	161 (49.1%)	126 (38.4%)	328
Cross-tabulation	$\chi^2 (2) = 18.00, p < 0.001$; Cramer's V = 0.23, $p < 0.001$			

Latent profile assignment	Interviewer reports of parent reliability			
	Poor	Fair	Good	Total
Parent reports greater than youth	10 (3%)	61 (18.6%)	92 (28%)	163 (49.7%)
Youth reports greater than parent	15 (4.6%)	58 (17.7%)	92 (28%)	165 (50.3%)
Total	25 (7.6%)	119 (36.3%)	184 (56.1%)	328
Cross-tabulation	$\chi^2 (2) = 1.06, ns$; Cramer's V = 0.05, ns			

ns = not statistically significant.

and youth, whereas other interviewers completed each rating after the conclusion of each interview. This variability by rater creates the potential for our findings being attributable to different raters systematically varying in when or how they rated the reliability of both informants (i.e., clustering effects). Thus, we calculated intra-class correlation coefficients (ICCs) for the interviewer reliability ratings for youths and parents. The ICC for interviewer ratings of the youth's reliability was 0.03, and the ICC for interviewer ratings of the parent's reliability was 0.07. Further, the average cluster size for each of the reliability ratings was 11. These figures translated to design effects estimates (i.e., indices of the extent of clustering effects) of 1.3 for the interviewer ratings of the youth's reliability and 1.7 for the interviewer ratings of the parent's reliability. Previous work using Monte Carlo simulations indicates

that one should be concerned with clustering effects if design effect estimates are at or above 2 (Muthen and Satorra 1995). Thus, the data indicate that clustering effects do not present confounds to our interpretations of the findings. Nevertheless, we encourage future work to (1) counterbalance the order in which interviewers administer interviews to informants and (2) instruct interviewers to provide informant reliability ratings immediately after the interview.

The focus of our investigation was on interviewer impressions of parent and youth informant reports. This is in keeping with prior work examining how informant discrepancies relate to whether practitioners agree with some informants more than others (Hawley and Weisz 2003; Kramer et al. 2004; De Los Reyes Alfano et al. 2011). However, practitioners rely on many informants' reports

TABLE 3. NOMINAL LOGISTIC REGRESSION ANALYSIS DIFFERENTIATING INTERVIEWERS' RATINGS OF YOUTH AND PARENT RELIABILITY AS INFORMANTS, BASED ON PATTERNS OF DISCREPANCIES BETWEEN PARENT AND YOUTH REPORTS (N=328)^{a,b}

Latent profile assignment	OR for interviewer ratings of informant reliability: poor versus good	OR for interviewer ratings of informant reliability: fair versus good
Interviewer reports of youth reliability		
Covariates	Gender OR=0.51, ns Ethnicity OR=1.16, ns Age OR=0.83, ns CGI OR=1.63, $p < 0.05$	Gender OR=0.61, ns Ethnicity OR=0.69, ns Age OR=0.78, $p < 0.001$ CGI OR=1.34, $p < 0.05$
Parent reports greater than youth	OR=3.27, $p < 0.01$, CI=1.49, 7.19	OR=1.79, $p < 0.05$, CI=1.08, 2.98
Youth reports greater than parent	OR=0.30, $p < 0.01$, CI=0.14, 0.67	OR=0.55, $p < 0.05$, CI=0.33, 0.92
Interviewer reports of parent reliability		
Covariates	Gender OR=0.78, ns Ethnicity OR=0.37, ns Age OR=1.31, $p < 0.05$ CGI OR=0.80, ns	Gender OR=1.26, ns Ethnicity OR=0.63, ns Age OR=1.12, ns CGI OR=1.15, ns
Parent reports greater than youth	OR=0.70, ns, CI=0.29, 1.69	OR=1.08, ns, CI=0.67, 1.76
Youth reports greater than parent	OR=1.43, ns, CI=0.59, 3.46	OR=0.92, ns, CI=0.56, 1.49

^aOR reflect regressions in which interviewers' reliability ratings were employed as dependent variables (i.e., "Good," "Fair," and "Poor," with "Good" as the reference category), and reporting pattern as the independent variable (i.e., "Youth Reports Greater than Parent" and "Parent Reports Greater than Youth").

^bIn order to attain OR estimates for both groups, regressions were conducted twice for each rating (i.e., once with "Parent Reports Greater than Youth" coded "0" and once with "Youth Reports Greater than Parent" coded "0").

OR=odds ratio; CGI=Clinician Global Impressions Severity of Illness score; CI=95% confidence interval for odds ratios; ns=not statistically significant.

other than parents and youths (e.g., teachers and other adults with whom a child may interact) when they evaluate youths in various settings (e.g., pediatricians in office settings, and counselors in school settings). Further, this current study focused on informant discrepancies in reports of youths' emotional and behavioral functioning, broadly construed. Similar findings might not emerge for informant discrepancies in other areas of clinical assessment (e.g., family functioning; see Hawley and Weisz 2003) and perhaps for different forms of youth's emotional and behavioral functioning (e.g., internalizing vs. externalizing functioning). We encourage future work to examine whether our findings generalize to how various mental health professionals perceive the reliability of other informants' reports and within assessments of other youth and family behavioral domains.

Conclusions

When faced with informant discrepancies, practitioners and researchers should refrain from immediately judging the reliability of the informants. Alternatively, prior research can be used to form simple rubrics about informant discrepancies. In turn, these rubrics can aid in hypothesizing as to the nature of any informant discrepancies that may arise. That is, as practitioners and researchers we can use the consistent nature of informant discrepancies to interpret assessment outcomes and improve patient care.

For instance, before conducting an evaluation, practitioners (and researchers designing studies) might make three predictions based on prior research and the subject of the evaluation: (1) whether informants providing reports will disagree in their reports; (2) if informants disagree, the direction of the disagreement (who will report greater degrees of problems than the other informant); and (3) what information might the disagreements yield regarding the nature and expression of the youths' emotional and behavioral

functioning (e.g., specific settings in which problems are expressed). In the presence of informant discrepancies, practitioners and researchers could then structure their evaluations so that they may test these predictions.

To that effect, in Table 4 we present data reported by Achenbach and Rescorla (2001) for select syndrome scores from parent, youth, and teacher forms assessing youths' emotional and behavioral functioning. Specifically, the table outlines data that, for instance, a practitioner evaluating a youth in an outpatient clinic setting can use to (1) predict levels of agreement between two informants' reports; (2) predict which informant will report greater degrees of youth problems than the other informant; and (3) advance hypotheses as to why the reports would disagree.

It would be helpful here to highlight an example of how to apply the data in Table 4. For instance, suppose a practitioner evaluates a youth for disruptive behavior and the referral source (the parent) complains primarily of problematic parent–youth interactions. Here, the practitioner might rely on parent and teacher reports as proxy data to understand whether the youth is disruptive in both home and school settings. Further, the referral question might suggest that parent and teacher reports will disagree, and that parent reports will suggest the presence of disruptive behavior that the teacher does not notice as prominently.

Under these circumstances, the practitioner can structure an observational assessment in the clinic, such as directing the parent and youth to discuss by themselves a topic about which they commonly disagree at home. Additionally, the practitioner might structure an independent interaction with the youth in which the practitioner and youth discuss by themselves the same topic that the youth discussed previously with his or her parent. Within this interaction, the practitioner might challenge the youth's views on the topic to see if the youth engages in similar interactions with the practitioner as he or she does with the parent. Much like the parent and teacher reports,

TABLE 4. SAMPLE RUBRIC FOR A PRACTITIONER INTERESTED IN PREDICTING THE NATURE AND EXTENT OF INFORMANT DISCREPANCIES AS OBSERVED IN REPORTS GATHERED IN OUTPATIENT REFERRAL SETTINGS^{a-c}

<i>Informant pair</i>	<i>Magnitude of agreement</i>	<i>Direction of reports</i>	<i>What might disagreements represent?</i>
Parent-Teacher	A/D: low (0.19) ATT: medium (0.44) RBB: medium (0.38)	A/D: P > T ATT: P > T RBB: P > T	Youths are expressing problem behavior in home settings or within parent–youth interactions to a greater degree than in school settings.
Parent-Youth	A/D: medium (0.45) ATT: medium (0.48) RBB: large (0.55)	A/D: P > Y ATT: P > Y RBB: P > Y	Youths may express a combination of observable forms of anxiety (avoidance of schoolmates), attention problems (difficulty completing tasks), and rule-breaking behavior (opposing teacher commands) and covert forms (e.g., worry) in home settings. Parents may base their reports on both observable behaviors and some of the covert behaviors. Conversely, youths may focus their reports only on covert behaviors expressed in home settings.
Teacher-Youth	A/D: low (0.16) ATT: medium (0.30) RBB: medium (0.32)	A/D: T > Y ATT: T > Y RBB: T > Y	Youths may express observable forms of anxiety (avoidance of schoolmates), attention problems (difficulty completing tasks), and rule-breaking behavior (opposing teacher commands) in school settings and on which teachers base their reports. Conversely, youths may focus their reports on covert behaviors expressed in home settings, which may have a lower base-rate of expression.

^aMagnitudes of agreement based on effect sizes of small ($r=0.10-0.29$), medium ($r=0.30-0.49$), and large ($r=0.50$ and above), consistent with estimates reported elsewhere (Cohen 1988).

^bPredictions of informant-based reporting agreement and directions of reports (i.e., which informants report greater degrees of youth problems relative to other informants) are based on data reported for the Empirically-Based Scales of Anxious/Depressed (A/D), Attention Problems (ATT), and Rule-Breaking Behavior (RBB) of the Child Behavior Checklist, Teacher Report Form, and Youth Self-Report for parents (P), teachers (T), and youths (Y), respectively (Achenbach and Rescorla 2001).

^cReporting agreement estimates based on Table 9-2 of Achenbach and Rescorla (2001) and reporting direction estimates based on Appendix D T scores for youths 12–18 (parent and teacher reports) and youths 11–18 (youth report).

this observational assessment can serve as proxy data for the practitioner to decide how an assessment independent of the parent and teacher reports suggests a youth behaves with both parental and nonparental (e.g., practitioner and teacher) adults.

In light of our recommendations, it is important to acknowledge two realities of clinical practice. First, comprehensive assessments already are being administered in clinical practice and administered often. However, when findings are inconsistent it is often the case that practitioners and researchers believe that they must decide which perspective is “right” and which is “wrong” (Hawley and Weisz 2003; Kramer et al. 2004; De Los Reyes Alfano et al. 2011). Instead, we argue that an integrative assessment rubric can aid practitioners in determining whether disagreements between informants’ reports reflect meaningful differences in how or whether youths’ emotional and behavioral problems vary in their expression across settings.

Second, with limited resources, practitioners may often have insufficient time to administer multi-informant assessments and interpret them following our recommendations. Under these circumstances, we recommend that practitioners engage in assessment processes that, nonetheless, serve to prevent their discounting specific informant’s reports in favor of other reports. For instance, rather than collecting multiple intake reports at once, a practitioner can collect the single report from the informant for whom the setting in which they primarily observe the youth most closely matches the referral question (e.g., teacher report if primary problem appears to be aggressive behavior at school). If the report confirms the presence of the referral problem, then the practitioner can tailor treatment efforts around this report and continue to monitor treatment progress on this one report. Once the practitioner identifies a successful treatment response, the practitioner can administer a report to one other informant (e.g., parent) to identify whether (1) treatment response generalized to another setting (e.g., home) or (2) further concerns in another setting remain that warrant treatment. The same procedure used to assess treatment response with the initial informant’s report can then be repeated with the report of the next informant and any subsequent informants that follow. This sequential use of multi-informant assessments would preserve the comprehensive nature of patient evaluations, and at the same time prevent a practitioner from having to decide at any one assessment point on which of two or more informants’ reports to rely to make decisions as to patient care. Regardless of the nature of patient evaluations, incorporating principles of what informant discrepancies might represent into best practices in clinical assessments of youths’ emotional and behavioral functioning may result in increased cost-effectiveness and clinical utility of the data gathered from such assessments.

Clinical Significance

Prior work indicates that informant discrepancies in reports of youths’ emotional and behavioral functioning reveal important information about the settings within which youths express emotional and behavioral problems. Yet, our findings suggest that discrepancies between reports provided by parents and youths relate to the extent to which interviewers perceive youths (but not parents) as reliable reporters of the youth’s emotional and behavioral functioning. Importantly, these informant discrepancies occur often in both research and practice settings, and many times practitioners and researchers can use prior work to predict the kinds of reporting discrepancies that they will observe in evaluations of these problems. Thus, practitioners and researchers can effectively

increase the cost-effectiveness and clinical utility of multi-informant clinical evaluations by anticipating informant discrepancies and predicting what they may represent before conducting these evaluations.

Acknowledgments

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.

Disclosures

Dr. De Los Reyes does not have any financial relationships with any pharmaceutical company. Dr. E. Youngstrom receives or has received travel support from Bristol-Myers Squibb and acted as a consultant for Bristol-Myers Squibb and Lundbeck. Ms. Swan does not have any financial relationships with any pharmaceutical company. Dr. J. Youngstrom does not have any financial relationships with any pharmaceutical company. Dr. Feeny does not have any financial relationships with any pharmaceutical company. 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, Forest, GlaxoSmithKline, Johnson & Johnson, KemPharm Lilly, Lundbeck, Neuropharm, Novartis, Noven, Organon, Otsuka, Pfizer, Rhodes Pharmaceuticals, Sanofi-Aventis, Schering-Plough, Seaside Therapeutics, Sepracore, Shire, Solvay, Sunovion, Supernus Pharmaceuticals, Validus, and Wyeth.

References

- Achenbach TM: Manual for the Child Behavior Checklist 4–18 and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry, 1991a.
- Achenbach TM: Manual for the Youth Self Report Form and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry, 1991b.
- Achenbach TM, McConaughy SH, Howell CT: Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychol Bull* 101:213–232, 1987.
- Achenbach TM, Rescorla LA: Manual for the ASEBA School-Age Forms and Profiles. Burlington, VT: University of Vermont, Department of Psychiatry, 2001.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. (DSM-IV). Washington, DC: American Psychiatric Association; 1994.
- Bartholomew DJ, Steele F, Moustaki I, Galbraith JI: *The Analysis and Interpretation of Multivariate Data for Social Scientists*. Philadelphia, PA: Chapman and Hall/CRC, 2002.
- Cohen J: *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Mahwah, NJ: Erlbaum, 1988.
- De Los Reyes A: Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *J Clin Child and Adol Psychol* 40:1–9, 2011.
- De Los Reyes A, Alfano CA, Beidel DC: The relations among measurements of informant discrepancies within a multisite trial of treatments for childhood social phobia. *J Abnorm Child Psychol* 38:395–404, 2010.
- De Los Reyes A, Alfano CA, Beidel DC: Are clinicians’ assessments of improvements in children’s functioning “global”? *J Clin Child Adol Psychol* 40:281–294, 2011.

- De Los Reyes A, Goodman KL, Kliewer W, Reid-Quiñones KR: Whose depression relates to discrepancies? Testing relations between informant characteristics and informant discrepancies from both informants' perspectives. *Psychol Assess* 20:139–149, 2008.
- De Los Reyes A, Goodman KL, Kliewer W, Reid-Quiñones KR: The longitudinal consistency of mother-child reporting discrepancies of parental monitoring and their ability to predict child delinquent behaviors two years later. *J Youth Adolesc* 39:1417–1430, 2010.
- De Los Reyes A, Henry DB, Tolan PT, Wakschlag LS: Linking informant discrepancies to observed variations in young children's disruptive behavior. *J Abnorm Child Psychol* 37:637–652, 2009.
- De Los Reyes A, Kazdin AE: Measuring informant discrepancies in clinical child research. *Psychol Assess* 113:554–583, 2004.
- De Los Reyes A, Kazdin AE: Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychol Bull* 131:483–509, 2005.
- De Los Reyes A, Kazdin AE: Conceptualizing changes in behavior in intervention research: The range of possible changes model. *Psychol Rev* 113:554–583, 2006.
- De Los Reyes A, Kazdin AE: When the evidence says, "Yes, no, and maybe so": Attending to and interpreting inconsistent findings among evidence-based interventions. *Curr Dir Psychol* 17:47–51, 2008.
- De Los Reyes A, Kazdin AE: Identifying evidence-based interventions for children and adolescents using the range of possible changes model: A meta-analytic illustration. *Beh Mod* 33:583–617, 2009.
- De Los Reyes A, Youngstrom EA, Pabon SC, Youngstrom JK, Feeny NC, Findling RL: Internal consistency and associated characteristics of informant discrepancies in clinic referred youth's age 11 to 17 years. *J Clin Child Adol Psychol* 40:36–53, 2011.
- Fisher SL, Bucholz KK, Reich W, Fox L, Kuperman S, Kramer J, Hesselbrock V, Dick DM, Nurnberger JI, Edenberg HJ, Bieru LJ: Teenagers are right-parents do not know much: An analysis of adolescent-parent agreement on reports of adolescent substance use, abuse, and dependence. *Alcohol Clin Exp Res* 30:1699–1710, 2006.
- Geller B, Zimmerman B, Williams M, Bolhofner K, Craney JL, DelBello MP, Soutullo C: Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. *J Am Acad Child Adol Psychiatry* 40:450–455, 2001.
- Hartley AG, Zakriski AL, Wright JC: Probing the depths of discrepancies: Contextual influences on divergence and convergence. *J Clin Child and Adol Psychol*, 40:54–66, 2011.
- Hawley KM, Weisz JR: Child, parent, and therapist (dis)agreement on target problems in outpatient therapy: The therapist's dilemma and its implications. *J Consult Clin Psychol* 71:62–70, 2003.
- Hunsley J, Mash EJ: Evidence-based assessment. *Ann Rev Clin Psychol* 3:29–51, 2007.
- Johnson SB, Wang C: Why do adolescents say they are less healthy than their parents think they are? The importance of mental health varies by social class in a nationally representative sample. *Pediatrics* 121:e307–e313, 2008.
- Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N: Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. *J Am Acad Child Adol Psychiatry* 36:980–988, 1997.
- Kramer TL, Phillips SD, Hargis MB, Miller TL, Burns BJ, Robbins JM: Disagreement between parent and adolescent reports of functional impairment. *J Child Psychol Psychiatry* 45:248–259, 2004.
- Krosnick JA: Survey research. *Ann Rev Psychol* 50:537–567, 1999.
- Laird RD, Weems CF: The equivalence of regression models using difference scores and models using separate scores for each informant: Implications for the study of informant discrepancies. *Psychol Assess* 23:388–397, 2011.
- Loeber R, Green SM, Lahey BB: Mental health professionals' perception of the utility of children, mothers, and teachers as informants of childhood psychopathology. *J Clin Child Psychol* 19:136–143, 1990.
- Muthen BO, Satorra A: Complex sample data in structural equation modeling. *Sociol Methodol* 25: 267–316, 1995.
- National Institute of Mental Health: Rating scales and assessment instruments for use in pediatric psychopharmacology research. *Psychopharmacol Bull* 21:839–843, 1985.
- Nottelmann E: National Institute of Mental Health research roundtable on prepubertal bipolar disorder. *J Am Acad Child Adol Psychiatry* 40:871–878, 2001.
- Perrin EC, Lewkowicz C, Young MH: Shared vision: Concordance among fathers, mothers, and pediatricians about unmet needs of children with chronic health conditions. *Pediatrics* 105:277–285, 2000.
- Richters JE: Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychol Bull* 112:485–499, 1992.
- Youngstrom EA, Findling RL, Danielson CK, Calabrese JR: Discriminative validity of parent report of hypomanic and depressive symptoms on the General Behavior Inventory. *Psychol Assess* 13:267–276, 2001.

Address correspondence to:
Andres De Los Reyes, Ph.D.

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