

Strategic objectives for improving understanding of informant discrepancies in developmental psychopathology research

ANDRES DE LOS REYES

University of Maryland at College Park

Abstract

Developmental psychopathology researchers and practitioners commonly conduct behavioral assessments using multiple informants' reports (e.g., parents, teachers, practitioners, children, and laboratory observers). These assessments often yield inconsistent conclusions about important questions in developmental psychopathology research, depending on the informant (e.g., psychiatric diagnoses and risk factors of disorder). Researchers have theorized why informant discrepancies exist and advanced methodological models of informant discrepancies. However, over 50 years of empirical data has uncovered little knowledge about these discrepancies beyond that they exist, complicate interpretations of research findings and assessment outcomes in practice, and correlate with some characteristics of the informants providing reports (e.g., demographics and mood levels). Further, recent studies often yield take-home messages about the importance of taking a multi-informant approach to clinical and developmental assessments. Researchers draw these conclusions from their work, despite multi-informant approaches to assessment long being a part of best practices in clinical and developmental assessments. Consequently, developmental psychopathology researchers and practitioners are in dire need of a focused set of research priorities with the key goal of rapidly advancing knowledge about informant discrepancies. In this paper, I discuss these research priorities, review work indicating the feasibility of conducting research addressing these priorities, and specify what researchers and practitioners would gain from studies advancing knowledge about informant discrepancies in developmental psychopathology research.

For decades, researchers and practitioners have both observed and struggled to interpret discrepancies among informants' reports of child and adolescent psychopathology (i.e., collectively referred to here as "children" unless otherwise specified; De Los Reyes, 2011; Renk, 2005). These discrepant reports also occur in assessments of risk factors of child psychopathology, raising significant interpretive issues (e.g., family conflict, youth victimization, parenting behaviors, and social competence; see De Los Reyes & Prinstein, 2004; Gonzales, Cauce, & Mason, 1996; Goodman, De Los Reyes, & Bradshaw, 2010; Lorenz, Melby, Conger, & Xu, 2007; Renk & Phares, 2004; Taber, 2010). In both of these cases, it is common knowledge that even when one administers the same measure about a behavior to different informants (e.g., parents, teachers, or children), these informants provide reports that yield different research conclusions. Informant discrepancies often arise in the findings of research examining such crucial issues as (a) which children are experiencing psychosocial dysfunction, (b) which children respond to treatment for such dysfunction, and (c) which children are at risk for developing such dysfunction (De Los Reyes & Kazdin, 2005). These informant discrepancies occur in multiple research and practice settings in psychology (e.g., clinical, counseling, developmental, and educa-

tional), as well as in epidemiology, criminal justice, social work, psychiatry, and primary care settings (Achenbach, 2006).

The topic of informant discrepancies is of particular relevance for understanding findings in developmental psychopathology research. Much of the evidence supporting the efficacy of treatments for children rests on multiple informants' outcome reports (for a review, see Weisz, Jensen Doss, & Hawley, 2005). Further, it is common to observe inconsistencies in the outcomes of controlled trials testing psychological interventions, depending on the informant (e.g., Casey & Berman, 1985; De Los Reyes & Kazdin, 2009; Koenig, De Los Reyes, Cicchetti, Scahill, & Klin, 2009). These discrepancies also commonly arise in the outcomes of behavior genetics studies and prospective longitudinal research generally (e.g., Achenbach, 2011; Derks, Hudziak, van Beijsterveldt, Dolan, & Boomsma, 2006; Dirks, Boyle, & Georgiades, 2011).

The common observation that different informants provide reports that yield different assessment outcomes can be traced to work conducted over half a century ago (Lapouse & Monk, 1958). Over 25 years ago, Achenbach, McConaughy, and Howell (1987) published a seminal review of 119 studies that had examined informant discrepancies; their main findings and conclusions are reported in Table 1. The Achenbach et al. (1987) review has been cited by over 2,000 peer-reviewed journal articles (Institute for Scientific Information, 2011). In addition, since this meta-analysis, many narrative and quantitative reviews have tracked and synthesized the findings of subsequent studies on informant

Address correspondence and reprint requests 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: adlr@umd.edu.

Table 1. Main findings and conclusions from Achenbach et al. (1987) and from other recent articles

Achenbach et al. (1987)		
	Main Findings	Conclusions and Recommendations
	(a) Different informants' reports of the same behavior commonly exhibit low to moderate levels of correspondence, (b) reports of two informants observing children in the same setting correspond more so than the reports of two informants observing children in different settings, (c) informants' reports of younger children correspond more so than reports of older children, and (d) informants' reports of externalizing behaviors (aggression) correspond more so than reports of internalizing behaviors (anxiety).	"Different informants are needed for different situations. . . there is no royal road or preeminent gold standard for phenomena that are inevitably affected by assessment procedures and other situational variables." (p. 227–228)
Recent Articles Examining Informant Discrepancies		
Authors	Summary of Main Findings	Conclusions and Recommendations
Collishaw, Goodman, Ford, Rabe-Hesketh, & Pickles (2009)	Informants vary in whether their reports of child psychopathology relate to associated features of child psychopathology.	"Assessments of child mental health problems should routinely use multiple informants, as assessments based on single informants may over- or underestimate the importance of child, family and community risk factors." (p. 579)
Kassam-Adams, Garcia-España, Miller, & Winston (2006)	The correspondence between parent and child reports of the child's psychopathology is low, and the level of correspondence is related to parents' own psychopathology.	"This study underscores the importance of obtaining children's self-report when assessing children who have experienced a potentially traumatic event." (p. 1492)
Mitsis, McKay, Schulz, Newcorn, & Halperin (2000)	Correspondence between parent and teacher reports of the child's psychopathology is low.	"Because of the <i>DSM-IV</i> requirement of cross-situationality of symptoms for a diagnosis of ADHD, an optimal evaluation should consist of direct input from multiple informants." (p. 312)
Randazzo, Landsverk, & Ganger (2003)	Discrepancies exist between parent reports of child psychopathology and the reports of other informants, and the parents' depressive symptoms relate to these discrepancies.	"If available, the report of the identified client/patient, a teacher, a peer, or a foster parent should be solicited. Based on these findings, it would be safest to take into account at least two sources in addition to parental reports." (p. 1350)
Rosnati, Montirosso, & Barni (2008)	Reports of child psychopathology completed by mothers and fathers of adoptive children corresponded more than reports completed by mothers and father of biological children.	"The problematic feature of behavior is situation specific, and its perception is not independent of the informant's characteristics . . . more than one informant is needed in order to obtain a more reliable and comprehensive picture of the child's adjustment . . . Each parent gives a unique, and not interchangeable, perspective on the child's problems." (p. 548)

discrepancies. Briefly, studies have addressed such topics on discrepancies as (a) their impact on interpreting research findings, (b) their associated features, and (c) the role of informant biases in behavioral reports (e.g., Achenbach, 2006; De Los Reyes & Kazdin, 2005, 2006, 2008, 2009; Duhig, Renk, Epstein, & Phares, 2000; Grills & Ollendick, 2002; Kraemer et al., 2003; Owens, Goldfine, Evangelista, Hoza, & Kaiser, 2007; Richters, 1992). Thus, the informant discrepancies research summarized by Achenbach et al. (1987) focused primarily on describing the magnitude of informant discrepancies, for which informant pairs were discrepancies at their largest (e.g., parent and child, teacher and parent) and for which behavioral domains were informant discrepancies present. In contrast, the many narrative and quantitative reviews conducted since Achenbach et al. (1987) indicate that research since 1987 has primarily focused on understanding

the associated features of informant discrepancies, potential biases in informants' reports, and how informant discrepancies impact the conclusions drawn from research studies.

In light of all the empirical work on informant discrepancies and the myriad reviews that have synthesized this empirical work, a perusal of empirical articles citing Achenbach et al. (1987) reveals an astonishing observation, which is reported in Table 1. Specifically, regardless of the focus of these articles, the ultimate conclusions drawn from this empirical work and thus recommendations to researchers and practitioners mirror those that Achenbach et al. (1987) provided over 25 years ago. In other words, recent empirical work echoes findings long known from prior meta-analytic reviews, begging the question, *Have 25 years of empirical work on informant discrepancies improved how researchers and practitioners conduct multi-informant assessments?*

To be clear, the purpose of [Table 1](#) is not to report and summarize an exhaustive list of articles examining informant discrepancies. Further, these studies did not address identical research questions and thus varied quite a bit in scope and focus. At the same time, that these articles have been published in recent years is an important observation.

It would be a mistake to conclude that no work since Achenbach et al. (1987) has improved our understanding of informant discrepancies. Prior theoretical and methodological work has provided researchers guidance on how to statistically model multi-informant reports. For example, Kraemer et al. (2003) modeled informant discrepancies as a function of three components: (a) variation in the perspectives informants have of children's behavior (e.g., self vs. other), (b) variation in children's behavior across meaningful settings (e.g., home vs. school), and (c) the extent to which the behaviors assessed are consistently expressed across informants' perspectives and settings. Thus, in the model proposed by Kraemer et al. (2003), parents' reports are modeled as arising from an observer perspective of children's behavior expressed in the home context, teachers' reports as arising from an observer perspective of behavior expressed in a nonhome context (i.e., school), and children's reports as arising from a self perspective of behavior expressed in both home and nonhome contexts. The implication of this methodological approach is that when informants provide discrepant reports, these discrepancies may point to a meaningful interaction between two factors. The first is that informants vary systematically in where they observe children's behavior. The second is that children vary systematically in where they express the behaviors being assessed.

Recent theoretical and methodological work has also focused on how informant discrepancies may be incorporated within frameworks seeking to explain how youth maladjustment develops. For example, recent theoretical work has illuminated how discrepancies between parent and youth reports of youth victimization may represent features of the relationship between parents and youths that increase risk for poor youth outcomes (Goodman et al., 2010). In addition, in the behavior genetics literature, researchers have taken structural equation modeling approaches to managing multi-informant reports that allow for testing whether nonshared variance among informants' reports reflects unique information being contributed by informants' reports rather than mere measurement error (Baker, Jacobson Raine, Lozano, & Bezdjian, 2007; Bartels, Boomsma, Hudziak, van Beijsterveldt, & van den Oord, 2007; Derks et al., 2006; Tackett, Waldman & Lahey, 2009). In this literature, informant discrepancies are modeled not as "systematic error terms" but as valuable systematic variation. Specifically, variation unique to specific informants' reports represents contextual differences in the contributions of one's environment as well as genetic predispositions to variance in expressions of specific behaviors. In other words, researchers have partitioned variance to reflect genetic influences on behavior, shared environmental influences, and nonshared environmental influences, in addition to measurement error (Bartels et al., 2007). As a consequence, informant discrepan-

cies in the behavior genetics literature yield insights as to the extent to which children's environments affect their behavior differently, depending on where the informants providing reports observe children's behavior (e.g., Derks et al., 2006).

Sound theoretical and methodological models of informant discrepancies are necessary to advance understanding of informant discrepancies. Recent work has made great strides in this regard. At the same time, such modeling may be insufficient to gain a deeper understanding of how to conduct multi-informant assessments and interpret the informant discrepancies that often arise from these assessments. Two observations of the clinical literature support this contention. First, mental health professionals identify specific informants (e.g., parents, teachers, or youths) as "optimal" informants for assessing particular kinds of children's behavior (e.g., children's internalizing and externalizing problems; Loeber, Green, & Lahey, 1990). Second, despite the availability of multiple informants' reports and the theoretical and methodological models described previously, it has become common practice for controlled trial researchers to gauge treatment response with a single or primary outcome measure, as opposed to use of multiple outcome measures (De Los Reyes, Kundey, & Wang, 2011). In both cases, these practices fly in the face of well-accepted tenets of multi-informant assessment, including the idea that informants each have unique perspectives on children's behavior and thus can provide valuable reports about such behavior (see Hunsley & Mash, 2007). As discussed below, theoretical and methodological models of informant discrepancies have been advanced despite the lack of convincing empirical data supporting the idea that informant discrepancies can be meaningfully interpreted. Therefore, before researchers and practitioners begin to collectively view informant discrepancies in a meaningful light, perhaps what is required is the coupling of models of informant discrepancies and empirical data pointing to the utility of informant discrepancies in the interpretation of multi-informant assessment outcomes.

Purpose

In sum, the main recommendations provided by current empirical research on informant discrepancies reaffirm those made long ago by earlier work. Thus, one can argue that more than 50 years of empirical research on informant discrepancies has contributed little knowledge on how to interpret informant discrepancies when they arise. As researchers and practitioners, we know that informant discrepancies ought to remind us of the importance of collecting information from more than one informant. However, has empirical work informed us on how to interpret or construct clinical and developmental assessments, beyond merely indicating that we should use more than one informant? For instance, do informant discrepancies signal that we should modify our assessment procedures to better understand their origins when they arise? When informant discrepancies occur, does their presence influence researchers' perceptions of the veracity of informants' reports or affect practitioners' decisions about patient care?

Informant discrepancies research has yet to yield firm answers to important questions about conducting and interpreting multi-informant clinical assessments. As a result, it is difficult to point to improvements in best practices in multi-informant clinical assessments that are the direct result of recent empirical work on these topics (i.e., improvements that could not have very well resulted from empirical work conducted on or before the late 1980s). Consequently, researchers need an empirical research agenda focused on improving our understanding of informant discrepancies in developmental psychopathology research. This paper outlines such an agenda. Specifically, I outline research priorities that, if addressed, may yield substantial insights into what informant discrepancies mean and how they can be used to improve our understanding of behavioral domains assessed in developmental psychopathology research.

An Agenda for Informant Discrepancies in Developmental Psychopathology Research

The scientific agenda I outline in this paper focuses on two broad research objectives that can be addressed in various assessment settings (e.g., controlled trials, longitudinal studies, and practice settings) and by examining various psychopathology domains and risk factors (e.g., parenting, adolescent anxiety, preschoolers' disruptive behavior, and family conflict). For each objective, I outline the problem and highlight recent work demonstrating the feasibility of conducting empirical work that addresses the objective. Further, I highlight the benefits to be reaped by researchers and practitioners if we conduct studies consistent with this agenda.

Strategic Objective 1: Identify When the Evidence Supports Treating Informant Discrepancies (or Consistencies Between Informants) in Reports of Specific Behaviors as Signals of Inconsistencies (or Consistencies) in Contextual Expressions of These Behaviors

Description of problem

Researchers have made some attempts to theorize why informant discrepancies exist. I outline these theoretical models below.

Informant biases. Specifically, some researchers have attributed discrepancies to informant biases. Perhaps the most widely studied form of bias has been informants' mood-congruent reporting biases. Here, the idea is that a negative mood state prompts an informant to report more negative as opposed to positive or neutral information about the behaviors being assessed (Richters, 1992). Recent work has cast much doubt on this as an explanation for discrepancies. Specifically, many times the relation between informants' mood states and discrepancies no longer exists (or the relation lessens to a great extent) when taking into account other informant characteristics (e.g., family functioning and youth clinical severity; De Los Reyes & Kazdin, 2005; De Los Reyes, Youngstrom, Pabón, et al., 2011). Further, even when re-

searchers identify these relations, they explain very little of the total variance in discrepancies between informants' reports (e.g., Youngstrom, Izard, & Ackerman, 1999).

Although informants' mood-congruent biases might not contribute to informant discrepancies to a great extent, this should not be taken to imply that informants be considered *unbiased* observers and thus reporters of behavior. For example, multiple informants may differ in their mind-sets or motivations to provide behavioral reports and thus may differ in their reports about certain behaviors due to these motivations (e.g., parents unwilling to report high levels of disruptive behavior that teachers are willing to report; see Achenbach, 2011). As another example, consider that informants such as parents often provide reports about child psychopathology (e.g., children's depressive symptoms) for which their own behavior (e.g., parenting practices) may play a role in the development and expression of such psychopathology (for reviews, see Goodman & Gotlib, 1999; Granic & Patterson, 2006). For these reasons, much research has been dedicated to examining informant biases (see De Los Reyes & Kazdin, 2005).

Measurement error. Another model of informant discrepancies that has gained some attention is measurement error. Under this view, informant discrepancies exist in part because the measures completed by informants vary in how reliably they assess constructs of interest (Fisher et al., 2006; Krosnick, 1999). By virtue of the effects of measurement error, informant discrepancies might also indicate that informants' reports vary in estimates of their validity (e.g., their ability to relate to or predict constructs of interest; for reviews see Dirks et al., 2011; Dirks, De Los Reyes, Briggs-Gowan, Cella, & Wakschlag, 2012). However, any one measure of validity cannot be assumed to be a "gold standard" metric for gauging validity (see Borsboom, Mellenbergh, & van Heerden, 2004). What this means is that any one informant's report may very well be deemed "more" or "less" valid as another informant's report, depending on the validity metric used (De Los Reyes, 2011).

With regard to measurement error, psychometric research and theory highlights three components of measurement error, and each have applicability to the study of informant discrepancies: (a) random error; (b) transient error (i.e., error observed due to a rater characteristic [e.g., mood or emotion level] that is present on one measurement occasion but not or present to a different extent at another measurement occasion); and (c) systematic error (for reviews, see Nunnally & Bernstein, 1994; Schmidt & Hunter, 1996). First, random error has important implications for informant discrepancies research, in that measurement of the differences between informants' reports can only be as reliable as the individual informants' reports (Rogosa, Brandt, & Zimowski, 1982; Rogosa & Willett, 1983). That is, too much random error variance within each of the informants' reports truncates the ability to reliably assess differences between reports. As a consequence of this, it is important that informant discrepancies be studied in reference to reports taken from well-established, validated measures (see De Los Reyes, Youngstrom, Pabón, et al., 2011).

Second, although typically the focus of changes in scores between two measurement occasions (e.g., over time), the presence of transient error can be mistaken for “real” differences between informants’ reports. That is, two informants’ reports may appear to differ because they tend to observe behavior in different settings (e.g., parents and teachers). However, these informants may also differ in other ways on a particular measurement occasion (e.g., parent may be experiencing low mood due to a “bad day” at work and the teacher may be experiencing no such effects of low mood). Third, another measurement error component, systematic error, can become problematic when two or more informants evidence a consistent pattern of differences across multiple behavioral reports. Informant discrepancies can be statistically modeled and also appear to reflect “real” differences across multiple measurements (e.g., De Los Reyes, Alfano, & Beidel, 2011; De Los Reyes, Goodman, Kliwer, & Reid-Quiñones, 2010; De Los Reyes, Youngstrom, Pabón, et al., 2011). When differences between informants’ reports result from systematic measurement error, estimating the stability of the differences is insufficient to interpret the differences as having value (e.g., reflective of variations in behaviors expressed across settings). Thus, it is important to examine whether informant discrepancies have meaning to them by examining them relative to external criteria (e.g., independent measures of the behaviors rated by informants; see De Los Reyes, Henry, Tolan, & Wakschlag, 2009).

As in all measures of all psychological constructs, measurement error undoubtedly plays some role in assessments of informant discrepancies. However, to assert that informant discrepancies are primarily accounted for by measurement error would be inconsistent with the extant data. Specifically, researchers consistently identify low to moderate rates of correspondence between informants’ reports across multiple behavioral domains (see Achenbach, 2006). These correspondence rates consistently remain in the low to moderate range, even when informants complete well-researched and validated measures and researchers hold item content constant across reports (e.g., Achenbach & Rescorla, 2001; Comer & Kendall, 2004). Low correspondence arises even when the same factor structure explains variance in the reports of multiple informants who complete the same measure (e.g., Baldwin & Dadds, 2007).

In sum, if informant discrepancies were primarily explained by measurement error, how would one explain the empirical evidence indicating that these discrepancies arise and at a consistent rate across well-researched, reliable, and valid instruments? Further, how would one explain that these discrepancies arise even when these well-researched instruments include parallel forms administered to multiple informants and the forms hold crucial measurement components constant across forms (e.g., item content, scaling, or value labels on scales)? Under a variety of circumstances, psychometric evidence indicates that measurement error alone cannot provide a complete and parsimonious account of informant discrepancies.

Variations in expressions of behavior and informants’ perspectives on behavior. The explanation of informant discrepancies that has received relatively greater acceptance is that informants systematically vary in where they observe the behaviors being assessed, and children vary in where they express the behaviors being assessed (Achenbach et al., 1987; De Los Reyes & Kazdin, 2005; De Los Reyes, Thomas, Goodman, & Kunder, 2013; Kraemer et al., 2003). The intuitive appeal of this explanation of discrepancies is that it maps onto useful interpretations of commonly observed informant discrepancies. For example, this interpretation allows a researcher or practitioner to use parent (primarily observe home behavior) and teacher (primarily observe school behavior) reports as reflections of how the child behaves in different settings (De Los Reyes, 2011).

However, attributing informant discrepancies to systematic differences in both informants’ observational settings and the settings in which children express the behaviors being assessed requires evidence supporting such an attribution. With few exceptions, this evidence does not exist because little research has demonstrated the two conditions necessary to interpret discrepancies in this way. First, the discrepancies between reports have to evidence some form of reliability or consistency. Second, when informant discrepancies exist, they have to “map onto” setting-based variations in an independent measure of the behavior being assessed (i.e., variations in a measure not completed by the informants providing reports). Further, the converse observation should also arise: when multiple informants provide consistent reports, these consistencies should correspond to independent measures that indicate that the behavior is expressed consistently across settings. In other words, before researchers and practitioners can draw any meaning from informant discrepancies, empirical work needs to demonstrate not only that they can be reliably assessed but also that it is valid to meaningfully interpret them.

Initial supportive evidence

Achenbach et al. (1987) found that pairs of informants who typically observe children’s behavior in the same setting (e.g., mother and father, teachers and clinicians) evidence larger correlations between their reports than pairs of informants who typically observe children’s behavior in different settings (e.g., parents and teachers). Further, numerous researchers have examined variations in various assessment outcomes or research findings (e.g., prevalence rates, risk factors, and treatment outcomes) as a function of informant (for reviews, see De Los Reyes & Kazdin, 2005, 2006). However, only until recently have studies of informant discrepancies addressed two objectives. The first is examining whether informant discrepancies can be reliably assessed. The second is whether informant discrepancies are valid representations of variations in children’s behavior across settings, where validity is indexed using external or independent measurements of variations in children’s behavior across settings.

First, the available evidence supporting the consistency in measurements of informant discrepancies comes from recent

work examining discrepancies between parent and child reports. For instance, a study of a multisite clinic intake sample revealed high internal consistency estimates of the differences between parent and child reports on the eight syndrome scales assessed on the Child Behavior Checklist and the Youth Self-Report (Achenbach & Rescorla, 2001; De Los Reyes, Youngstrom, Pabón, et al., 2011). Two other studies have demonstrated with longitudinal data that the discrepancies between parent and child reports of child (e.g., social anxiety) and family behavior (e.g., whether parents monitor their child's whereabouts) are stable over multiple assessment points (De Los Reyes, Alfano, & Beidel, 2010; De Los Reyes, Goodman, et al., 2010). Thus, preliminary evidence supports consistency within and between measurements of informants' discrepant reports.

Second, the limited evidence supporting informant discrepancies as valid representations of how children vary in their expressions of behaviors being assessed comes from studies of informants' reports of externalizing behaviors or peer relations (e.g., children's aggressive and oppositional behavior or social peer networks in school settings). Specifically, De Los Reyes and et al. (2009) examined whether discrepancies between parent and teacher reports of preschool children's disruptive behaviors relate to assessments of these behaviors when measured across various laboratory controlled interactions between the child and multiple adults. The laboratory measure assessed children's disruptive behavior across interactions between children and their parents and children and unfamiliar clinical examiners (Wakschlag et al., 2008). As a result, this laboratory measure yielded a representation of how children in the sample interacted with parental and non-parental adults. Thus, this index of observed behavior allowed for tests of whether (a) children varied in the laboratory in whether they were disruptive in interactions with parental and/or nonparental adults and (b) variations in disruptive behavior in the laboratory "matched" variations in children's disruptive behavior as expressed by informants' reports of behavior expressed outside of the laboratory.

De Los Reyes et al. (2009) identified substantial discrepancies in both parent and teacher reports of children evidencing high levels of disruptive behaviors, as well as substantial variation in observed behavior in the laboratory. Further, parent-teacher reporting discrepancies mapped onto laboratory variations in children's disruptive behavior. First, "parent-only" identifications of disruptive behavior mapped onto identifications of children who behaved disruptively in the laboratory with parents but not clinical examiners. Second, "teacher-only" identifications of disruptive behaviors mapped onto identifications of children who behaved disruptively in the laboratory with clinical examiners and not parents. Overall, identifications of disruptive behavior across both parent and teacher reports mapped onto identifications of children who behaved disruptively in the laboratory with both adults.

A second piece of supportive evidence comes from a study examining parent and teacher reports of relatively older children's aggressive behavior (Hartley, Zakriski, & Wright, 2011). Specifically, in a sample of 123 children (mean age

= 13.30 years), parents and teachers completed parallel survey reports of children's aggressive behavior. On a separate set of parallel measures, parents and teachers completed reports of how aggressively children react to specific social experiences encountered in the settings within which parents and teachers observe children (e.g., being teased or "bossed around" by peers, or being reprimanded or warned by adults). By focusing on the social experiences that elicit children's aggressive behavior, Hartley et al. (2011) found that parent and teacher reports exhibited progressively higher magnitudes of correspondence with increased similarity of children's social experiences across contexts, as reported by parents and teachers.

Supportive evidence of informant discrepancies as representations of setting differences in informants' observations of behaviors also comes from studies of peer and teacher reports in school settings. Specifically, discrepancies between teacher and peer reports of networks of peer relationships (i.e., social networks) can be explained, in part, by developmental variations in the density and stability of these networks and classroom characteristics (Neal, Cappella, Wagner, & Atkins, 2011). That is, the increased density and stability of children's social networks partially explained the observation that as children's grade level increased, the correspondence between teacher and peer social network reports increased. One way to interpret these findings is that the reliability of children's peer network reports may increase with grade level and that this increased reliability explains the relation. However, correspondence between teacher and peer social network reports fluctuated as a function of measures of classroom characteristics that often exhibit variability within and between grade level, such as official records of classroom size and independent observers' ratings of teachers' behavior management in the classroom.

The effects of environmental circumstances on informants' reports have recently received experimental support. Although not focusing on the discrepancies between informants' reports, De Los Reyes and Marsh (2011) found that after reading vignettes of children described as expressing behaviors typically indicative of conduct disorder, clinicians were more likely to view these behaviors as leading to a conduct disorder diagnosis if the children were living in environments that pose risk for the disorder, relative to environments posing no such risk. Findings across the studies reviewed previously (i.e., De Los Reyes, Alfano et al., 2010; De Los Reyes, Goodman et al., 2010; De Los Reyes et al., 2009; De Los Reyes & Marsh, 2011; De Los Reyes, Youngstrom, Pabón, et al., 2011; Hartley et al., 2011; Neal et al., 2011) were drawn from studies conducted in multiple research settings (longitudinal, controlled trials and laboratory observations, community mental health clinics; randomized experiments) and demonstrate that informant discrepancies are systematic, predictable, and can be linked to expressions of behaviors within particular settings. As such, these findings indicate that under some circumstances, it is feasible to examine whether informant discrepancies meaningfully reflect how children behave across and within settings, and how in-

formants use information about settings to provide behavioral reports.

Future directions

The evidence reviewed previously indicates that informant discrepancies can at times be interpreted as representing distinct differences between informants in how or under what circumstances they observe the behaviors being assessed. However, evidence supporting the validity of this interpretation is limited to (a) assessments of children's behaviors that are directly visible to outside observers and (b) examinations of discrepancies between informants who fundamentally differ in the contexts in which they observe the behaviors being assessed (e.g., parents and teachers; teachers and peers). That is, much work has documented low to moderate levels of correspondence. These levels of correspondence have been identified among various informants' reports, as well as for reports of children's behaviors that are relatively difficult to directly observe (e.g., anxiety and mood) and for interpersonal behaviors (e.g., family conflict and relationship quality; see Achenbach, 2006; Duhig et al., 2000; Grills & Ollendick, 2002; Treutler & Epkins, 2003). Unlike the research reviewed previously for parent and teacher reports of aggressive and disruptive behavior, we know little of whether researchers and practitioners may meaningfully interpret the discrepancies between reports of other informants and for other behavioral domains.

However, circumstantial evidence in both the child anxiety and the marital conflict literature indicates that it is feasible to address these questions for assessments of some of these behaviors and informants' reports of these behaviors. For example, in structured diagnostic interview assessments of child anxiety disorders, interviews based on parent and child reports correspond more so when the reports are about directly observable anxiety behaviors expressed in nonschool settings (e.g., behavioral avoidance expressed at home) relative to internal anxiety behaviors such as worry expressed in school settings (Comer & Kendall, 2004). In addition, levels of correspondence between behavioral coders' single-session laboratory observations of hostile behavior and married and cohabitating couples' survey reports of hostile behavior in the past month are lower than correspondence levels between these same behavioral codes and couples' survey reports when based on behavior during the laboratory observation (Lorenz et al., 2007).

Similar to the studies reviewed previously, future research should examine whether parent and child reports of child anxiety and multiple reports of marital and/or family conflict meaningfully relate to variations in these behaviors as assessed on independent measures. The preliminary evidence cited previously, although promising, does not rule out a number of alternative explanations for the observed relations. For example, couples' hostile behavior may fundamentally differ under laboratory versus nonlaboratory circumstances in terms of such salient characteristics as the frequency and intensity of expres-

sions of such behavior. Thus, these discrepant reports must be examined in relation to independent measures of the behaviors for which informants provide reports. For child anxiety, such a measure might involve laboratory tasks meant to elicit representations of the child's anxiety as expressed at home (e.g., parent-child interactions) and school (e.g., giving a verbal report in class or unstructured play with other children).

Regarding marital conflict, one commonly used method of assessing such conflict involves constructing a laboratory observation assessment in which family members complete a questionnaire of topics about which they typically experience conflict, and assessors then select from their questionnaire reports the topics identified as potential sources of conflict (see Donenberg & Weisz, 1997; Kiecolt-Glaser & Newton, 2001). Those sources of conflict selected by the assessor are then used as points of discussion during a laboratory task. During this task, the assessor instructs the couple to discuss the topic for a set period of time (e.g., 6 min), with the goal of "coming to a resolution" on the topic. Thus, one independent measure through which to further examine any meaning inherent in the discrepancies between couples' reports of hostile behavior may be to construct laboratory tasks in which assessors systematically or randomly vary the kinds of topics discussed during the task. This variation in topic selection might occur at the level of topic domain (e.g., finances, in-laws, or household chores), and/or whether informants report that the topic results in hostile discussions that occur in public (e.g., family gatherings) versus private (e.g., home) settings. The variation in assessors' selections of topics to discuss during the task (and whether informants were discrepant in their questionnaire reports about these topics) can then be used to examine whether informant discrepancies in reports of hostile behavior systematically relate to variations in independent raters' behavioral codings of hostile behavior during the laboratory tasks. In sum, future work should continue to empirically examine assessment occasions for which researchers and practitioners wish to deduce that informant discrepancies reflect differences in the circumstances within which informants observe the behaviors being assessed.

Strategic Objective 2: Examine the Relation Between Informant Discrepancies and How Researchers and Practitioners Make Decisions Regarding Research Design, Data Interpretation, and Patient Care

Description of problem

What do researchers and practitioners do when they encounter informant discrepancies in their work? Does the presence of informant discrepancies impact how researchers and practitioners make decisions regarding data interpretation and patient care? There is reason to believe that informant discrepancies have important implications for interpreting assessment outcomes in research and practice settings. As mentioned previously, mental health professionals hold particular views as to which specific informants are "optimal" informants for as-

sessing particular domains of children's behavior (Loeber et al., 1990). Further, discrepancies commonly arise between informants' reports of whether treatments improve behaviors targeted for treatment (De Los Reyes & Kazdin, 2006). Controlled trial researchers now identify a single measure in advance of the trial as the primary outcome measure to gauge treatment response; this measure is typically completed by a mental health professional (De Los Reyes, Kundey, et al., 2011).

Consistent with prior work, when informant discrepancies arise, mental health professionals might consistently rely more on the informants' reports they tend to view as "optimal" when making clinical decisions. That is, if mental health professionals believe that there exist "optimal" informants for assessing specific children's behaviors, why would they find benefit in making decisions that integrate data from multiple informants' reports when discrepancies arise? Under these circumstances, one might hypothesize mental health professionals would toss out reports they view as suboptimal and instead decide to place absolute reliance on one informant's report. If such a hypothesis was supported by the data, this would have a large impact on the findings drawn from assessments conducted in developmental psychopathology research and practice. Specifically, if mental health professionals systematically rely on one informant, then any information drawn from mental health professionals' reports may essentially be redundant with that one informant's report. If this practice occurs often and across assessments of multiple domains of children's behavior, it may have enormous implications for the cost-effectiveness of multi-informant assessments. Such a practice would essentially eliminate the utility of any reports that do not factor into clinical decisions. Yet, this issue has received little empirical attention.

Initial supportive evidence

There is limited correlational evidence that informant discrepancies have implications for the decisions clinicians make when confronted with informant discrepancies. For instance, in a sample of children and parents seeking treatment in a community mental health clinic, children and parents varied widely as to which problems they identified as warranting treatment (Hawley & Weisz, 2003). In this study, parents were more likely than children to identify children's behavior as warranting treatment. Conversely, children were more likely than parents to identify problems in the family as warranting treatment. Clinicians' reports of problems warranting treatment systematically corresponded more with parents than with children when the identified problem concerned the child's behavior and vice versa when the identified problem concerned the family.

The findings observed by Hawley and Weisz (2003) have since been observed in other assessment circumstances beyond community mental health clinic settings. For example, in outpatient treatment settings, clinicians rate problems in adolescents' functioning identified by parents and not adolescents as more severe than problems in such functioning iden-

tified by adolescents and not parents (Kramer et al., 2004). Further, in a controlled trial of child and adolescent social anxiety treatments, clinicians rated a lack of treatment improvement when parents reported persistent problems over the course of treatment that were not reported by either children or blinded laboratory raters (De Los Reyes, Alfano, et al., 2011). However, in this study when either children or blinded raters reported persistent problems that were not reported by parents, clinicians' ratings did not reflect a lack of treatment improvement. One recent study suggests a possible reason for these relations. In an intake clinic sample of parent and child reports of the child's behavioral and emotional functioning, clinical interviewers viewed children's reports as more or less reliable, depending on whether children's reports evidenced greater problem levels relative to parents' reports (De Los Reyes, Youngstrom, Swan, et al., 2011). That is, when children's reports evidenced greater problem levels relative to parents' reports, interviewers were relatively more likely to identify these children as reliable informants than as unreliable informants. The converse was true of instances in which children's reports evidenced less problem levels relative to parents' reports (i.e., interviewers were more likely to identify these children as unreliable than reliable informants). However, in this study, interviewers' reports of parents' reliability were unrelated to whether parents' reports of their children's problems were lower or higher relative to children's self-reports.

All of these studies provided clear psychometric evidence for the measures that the researchers had informants complete (e.g., citation support of prior work and/or reliability and validity evidence from the sample they examined). Thus, differences among informants' reports in terms of their psychometric qualities could not be used to parsimoniously explain or justify clinicians' ratings in relation to informant discrepancies. In other words, for none of these studies was it the case that clinicians were simply systematically relying on the informant providing the most reliable and valid reports relative to other informants' reports. All of the informants' reports used in these studies evidenced adequate psychometric properties as measures of the behaviors being assessed.

Future directions

Work to date examining these issues has been exclusively correlational. One cannot deduce the direction of the relation between discrepancies and clinical decisions. Therefore, the extant data cannot be used to determine whether informant discrepancies influence clinical decisions regarding patient care (e.g., case formulations or identifying treatment responders). Two lines of research may address these issues.

Experimental research on the effect of informant discrepancies on clinical decision making. First, knowledge of informant discrepancies would be greatly enhanced by experimental work in which researchers expose clinicians to multi-informant assessment outcomes (e.g., reports of hypo-

thetical children “completed” by parents, children, and teachers) that randomly vary the nature of informant discrepancies (e.g., parent provides higher ratings on a problem behavior scale than the child’s self-ratings on the same scale and vice versa). After reviewing these assessment outcomes, clinicians might be asked to provide their own impressions of the functioning of the assessed child as well as their impressions of the reliability of the informants providing reports. A key focus of such research might be to examine whether clinicians rely on one informant’s report over other informants’ reports when informant discrepancies arise (e.g., relying on parent reports over other informants’ reports). This work would be able to decipher whether discrepancies among informants’ reports causally influence clinicians to rely on a subset of the available informants.

Training clinicians to integrate reports taken in multi-informant assessments. Second, if experimental work demonstrates a causal relation between informant discrepancies and clinical decisions in the assessment of child psychopathology, future work might focus on developing techniques for reducing the tendencies of clinicians, when informant discrepancies arise, to systematically rely on one informant’s report over other informants’ reports. That is, can clinicians be trained to make clinical decisions that, on average, integrate multiple reports, or at least provide ratings that correlate with each report at similar magnitudes?

For instance, training might consist of providing clinicians psychoeducation on principles of evidence-based assessment (e.g., reliability, validity, and clinical utility of multi-informant data) and the psychometric soundness of multiple informants’ reports of children’s behavior. Further, this psychoeducation might be integrated into exercises in which clinicians articulate meaningful reasons why informant discrepancies arise as well as reasons that might hinder meaningful interpretations of these discrepancies. An example of such a set of exercises might consist of clinicians reviewing mock clinical reports that include multi-informant assessment data in which they observe discrepancies between parent and teacher reports of a child patient’s behavior. Clinicians may be instructed to articulate reasons for why parents and teachers provide discrepant reports. Clinicians could be prompted to provide meaningful reasons for the discrepancies (e.g., parents and teachers observe the child in different contexts) and reasons that point to methodological confounds in reports (e.g., parent may be experiencing intellectual functioning and/or memory deficits). In articulating these reasons for discrepant reports, clinicians may be prompted to identify collateral information they may collect to test their hypotheses as to why they observed discrepancies between parent and teacher reports (e.g., other informants’ reports and performance-based or laboratory task data). The goal of such training would be to have these principles of evidence-based decision making generalize to clinic settings in which clinicians draw interpretations from the discrepancies among multiple informants’ reports.

In any event, this form of training may increase the likelihood that, when confronted with informant discrepancies in their clinic work, clinicians make active attempts to integrate multi-informant assessments as opposed to relying on reports from a single informant. Training clinicians to reduce their systematic reliance on a single informant’s report among multiple available reports would be consistent with the reasons why these multi-informant assessments are conducted in the first place (e.g., to understand how the behaviors being assessed vary across different settings or circumstances; Achenbach, 2006). In addition, such training might radically increase the cost-effectiveness of multi-informant assessments.

What Developmental Psychopathology Research and Practice Would Gain From Following This Agenda

In outlining directions for future informant discrepancies research, it is important to discuss the tangible benefits to be had if future work in this area were to head in these directions. First, addressing the two gaps in knowledge noted previously may result in improving how researchers and practitioners solicit reports from multiple informants. Second, these improvements may come about through the development of programs for training informants on how to provide behavioral reports. Third, research consistent with the proposed agenda may lead to improvements in methods for interpreting assessments of treatment outcomes. Fourth, agenda-driven research may result in more reliable identifications of the circumstances in which informant discrepancies are best explained by measurement error.

Changing how researchers and practitioners collect informants’ behavioral reports

Perhaps the most tangible benefit of addressing the research priorities discussed previously would be a fundamental change in how researchers and practitioners conduct clinical and developmental assessments. That is, currently informants provide reports with, at best, passive instructions on how to provide these reports (e.g., assessment time scales, such as the last 2 weeks, defined scale labels and values, and instructions that make clear that there are no right or wrong answers). As mentioned previously, even when these instructions are kept constant across informants’ reports, informant discrepancies nonetheless arise. Further, without more specific instructions on how to provide behavioral reports, discrepancies between informants’ reports can arise for many reasons. These reasons include informants varying in how they interpret or define scale labels and values (e.g., *often*), and whether they base their reports on behaviors they observe in specific settings or circumstances (e.g., behaviors expressed at home or school versus peer interactions outside of home or school).

For many assessment occasions, researchers and practitioners have a good idea about why informant discrepancies arise. For instance, informant discrepancies may arise be-

cause the behaviors being assessed meaningfully vary across specific settings or contexts and informants also vary in whether they observe behavior within these specific settings. Thus, researchers and practitioners might restructure assessments to allow informants to provide reports in a way that could be linked with whether children contextually vary in their behavior. Such a linkage may lessen the ambiguity inherent in observing informant discrepancies. Further, what if this restructuring of multi-informant assessments could be accomplished across many assessment occasions in developmental psychopathology research and practice? If so, perhaps the sense of befuddlement that currently accompanies researchers' and practitioners' efforts to interpret discrepant reports may be more easily managed. How might we undertake restructuring assessment procedures? To this issue I now turn.

Developing procedures for training informants to provide behavioral reports

Informant discrepancies have often been attributed to variations in the form and function of behaviors across settings. Collecting information about the settings in which behaviors occur during behavioral assessments may significantly advance the quality of informants' reports. For example, this information may greatly inform diagnosis and treatment planning in that it would focus attention on the specific settings (e.g., parent-child interactions or peer relations) in which therapeutic techniques should be applied. If a researcher knew of the settings in which his or her treatment produced effects, then he or she might modify the treatment to maximize these effects (e.g., by specifically targeting those contexts within which the treatment was most effective).

I outlined directions for research on identifying instances in which informants, in the absence of training, provide reports that meaningfully correspond to where children express the behaviors being assessed. Specifically, links between informants' reports and setting-specific expressions of behavior may be reflected in two ways. First, consistencies between informants' reports may reflect instances in which the assessed behavior is consistently expressed across settings. Second, inconsistencies between reports may reflect instances in which there are setting-specific expressions of the assessed behavior (see De Los Reyes et al., 2009). Work demonstrating these links between informant discrepancies and setting-based variations in assessed behaviors would provide "feasibility data" on whether it would be a useful endeavor to train informants to systematically incorporate setting information into their behavioral reports. In this way, researchers may further increase the reliability of the information they receive from informants' reports on the settings in which behaviors occur.

For instance, such a procedure may take the form of an interview-led assessment program. Through this program, an assessor might guide informants to identify settings relevant to where they observe behaviors contained in the items they will complete (e.g., peer interactions, giving class presentations, or what the child does after school). Informants can

be prompted to identify settings that they perceive as "great examples" of where they typically observe behaviors about which they will provide reports. Informants can then be trained to use this information to provide behavioral reports. Such training may take the form of providing informants with operational definitions of labels on item response scales and the number of settings that would be needed to enact an item response for that particular point on the scale (e.g., the label "Some" equals two settings relevant to the item "My child worries"). Further, informants may be provided with the same instructions for using setting information to make item responses. Doing so would equate informants in the decision rules guiding their item responses. As a result, use of these decision rules consistently across informants may reduce the number of possible interpretations for why discrepancies would arise between informants' reports. In addition, if informants could be trained to use these decision rules consistently and at the same level of reliability, researchers and practitioners may attain a level of certainty that the reasons for the discrepancies arise from the fact that informants view the behaviors being assessed in different settings. The resulting information can meaningfully inform interpretations of multi-informant assessments. As mentioned previously, researchers and practitioners often conduct multi-informant assessments with an interest in treating individual informants' reports as reflections of the unique settings in which they observe the behaviors being assessed (Hunsley & Mash, 2007; Kraemer et al., 2003). Preliminary evidence indicates that one can implement a program similar to that described previously. Recent experimental work has tested a program that trains informants to systematically incorporate setting information when providing reports about risk factors for juvenile delinquency (i.e., parental knowledge of adolescents' whereabouts and activities; De Los Reyes, Ehrlich, et al., 2013). In this study and relative to reports completed in the absence of training, informants' reports completed after training became more discrepant from each other.

The findings of De Los Reyes, Ehrlich, et al. (2013) are consistent with the idea that informants' reports disagree, in part, because they observe the behaviors being assessed in different settings. As such, this preliminary work on training informants to provide setting-sensitive behavioral reports has important treatment implications. This training can be used to clarify the circumstances in which interventions exert effects. In other words, setting-sensitive assessments may increase understanding of the settings in which interventions work. An example may be a study testing the effects of an intervention that focuses on changing parental knowledge behaviors to prevent adolescent delinquency. Perhaps the outcomes of this study supported the intervention's effectiveness based on mother reports and not adolescent reports. In the absence of setting information about informants' parental knowledge reports, the causes of these discrepant outcomes might be uncertain. However, setting information linked to the parental knowledge assessments could yield valuable insight as to why the study yielded inconsistent mother- and

adolescent-reported outcomes. For example, the effects based on mother reports might have been due to mothers perceiving that the intervention changed what they know about their adolescent's whereabouts with neighborhood children on the weekend. Conversely, the lack of effects based on adolescent reports might have been due to adolescents perceiving no changes in their mother's knowledge of what they do with friends after school. The setting information provided via mother and adolescent reports could then aid intervention researchers in modifying the intervention program to increase parental knowledge in settings not currently addressed by the intervention (e.g., adolescents' whereabouts after school). Therefore, training informants to use setting information when making behavioral reports may serve to clarify research findings in controlled trials and developmental psychopathology research. In sum, I recommend that future work be dedicated to testing the feasibility and utility of informant training procedures in developmental psychopathology assessments.

Clarifying assessments of treatment response and identifying effective treatments

In light of the discrepancies among informants' reports, how can researchers and practitioners properly identify those who respond to treatment and, ultimately, which treatments effectively change the behaviors they were developed to treat? Researchers have historically engaged in practices that do not account for informant discrepancies in treatment outcome reports, such as using methods of data aggregation that essentially treat the discrepancies as measurement error (De Los Reyes & Kazdin, 2008) or using a single measure to quantify treatment outcomes (De Los Reyes, Kundey, et al., 2011). In other words, in response to informant discrepancies, controlled trial researchers have resorted to suboptimal methods for assessing treatment response. Further, methodologies for assessing treatment response vary widely across controlled trials, making it difficult to assess whether treatment effects replicate across studies testing the same treatment (De Los Reyes & Kazdin, 2006, 2009).

However, what if future work results in researchers identifying the specific assessment circumstances in which they can meaningfully interpret informant discrepancies in treatment outcome reports? Accomplishing this goal might result in researchers and practitioners being able to use informant discrepancies to discern the settings within which treatments "work" and those settings within which they do not. Recent theoretical and methodological work, along with the preliminary work on training informants, may be integrated to improve identifications of treatment response. Theoretical and methodological models already exist for statistically modeling reports about children's behavior as a function of informant and the settings in which informants observe children's behavior (for examples, see De Los Reyes & Kazdin, 2006; Kraemer et al., 2003). For example, the Kraemer et al. (2003) model partitions informants' reports into variance accounted for by informants' per-

spective, the settings in which informants observe the behavior, and the extent to which the behaviors assessed are consistently expressed across informants' perspectives and settings. Use of this model undoubtedly improves clarity in why informants' outcome reports might yield inconsistent conclusions within controlled trials. In line with this modeling approach, I previously outlined recent work experimentally testing whether informants can be trained to provide behavioral reports based on the settings in which they observe the behaviors being assessed (De Los Reyes, Ehrlich, et al., 2013). It could be that training informants to provide setting-sensitive behavioral reports may allow researchers to attain a more precise estimate of variance among informants' reports attributed to setting. This increased precision in turn may allow researchers to attain more precise estimates of whether informant discrepancies in observed outcomes within controlled trials reflect that interventions vary in the settings within which they exert their intended effects. In sum, existing theoretical and methodological models, coupled with empirical work seeking to improve how informants incorporate setting information in their behavioral outcome reports, may provide the impetus for researchers and practitioners to systematically incorporate informant discrepancies in interpretations of treatment outcome data.

Developing consensus assessment batteries for use within controlled trials

Work on understanding when informant discrepancies may meaningfully inform interpretations of treatment outcome reports may aid researchers to come to a consensus on how to comprehensively test outcomes in controlled trials. One reason for why controlled trials vary widely in assessment methods may be that researchers do not have much of an idea of what to do when different reports yield different conclusions. As a result, some researchers proceed to using multiple outcome informants and "seeing what happens," whereas others rely on a single measure to assess treatment response (e.g., De Los Reyes, Alfano, et al., 2011).

A clearer understanding of what informant discrepancies mean may allow researchers to identify standardized batteries for use in controlled trials. Whether one might be able to use informant discrepancies as an interpretive tool in controlled trials work may depend on such factors as the behavior targeted for treatment, the severity of this behavior, and the timing in the development of the treated behavior. In any event, for those circumstances in which informant discrepancies are identified as meaningful interpretive tools, the result of using informant discrepancies in controlled trials work may be an increased ability to identify patterns of consistent and inconsistent outcomes in studies testing the same or similar treatments.

Conversely, there may be instances in which intervention researchers do not require comprehensive assessment batteries. An example may be independent observers' ratings of teachers' classroom management performance for school-based prevention programs of disruptive behavior that seek to improve teachers' classroom management skills. Another

example may be standardized achievement test scores for interventions seeking to improve academic performance. In both of these circumstances, intervention outcomes and treatment responses may be sufficiently calibrated using metrics that adequately approximate highly specified intervention targets, without the use of multiple informants' reports. Here, future research on informant discrepancies may yield strategies for using multiple informants' reports not as outcome measures per se but as "measurement checks" for gauging whether outcome assessments specifically gauge treatment response for the intervention target. For example, parent and teacher reports of disruptive behavior may be used to assess whether independent observers' ratings of teachers' classroom management performance correspond to a greater extent with informants' reports taken for behavior in classroom settings (e.g., teacher report) relative to home settings (e.g., parent report). Here, the absolute level of correspondence would not be the key issue, because prior work indicates that each of these metrics will correspond only modestly (Achenbach et al., 1987). However, the differences in magnitudes of correspondence (e.g., that independent observers' reports correspond more with teacher reports than they do with parent reports) can yield supportive evidence that the key metric of treatment response for an intervention study is specifically targeting the intended construct of interest. In any event, the point to be made here is that informant discrepancies research that follows the research agenda outlined above may improve the quality of outcome assessments in controlled trials research.

Developing procedures to determine when to treat discrepancies as error

An important implication of following the informant discrepancies research agenda outlined in this paper is that science will undoubtedly uncover those instances in which informant discrepancies are best explained by measurement error. These discoveries may prove quite valuable for interpreting the outcomes of these multi-informant assessments. For instance, a variety of sophisticated statistical modeling techniques (e.g., structural equations modeling; Holmbeck, Li, Schurman, Friedman, & Coakley, 2002) and methods for integrating multi-informant reports (e.g., combinational algorithms such as AND/OR decision rules; Youngstrom, Findling & Calabrese, 2003) treat information not shared among the multiple reports (i.e., unshared variance) as unique variance. As mentioned previously, although one can identify some circumstances in which this unique variance is meaningfully modeled using such techniques as structural equations modeling (see Bartels et al., 2007), for the most part researchers treat unique variance in multi-informant modeling as mea-

surement error (see Holmbeck et al., 2002; De Los Reyes & Kazdin, 2006; De Los Reyes, Kundey, et al., 2011).

Identifying those specific circumstances in which informant discrepancies can be interpreted as measurement error will inform researchers and practitioners about the instances in which they would be justified in using methods that treat informant discrepancies as measurement error. In this way, research conducted as described above would greatly aid in clarifying problems encountered in developmental psychopathology research and practice.

Conclusions

Researchers and practitioners from various disciplines commonly encounter discrepancies among informants' behavioral reports (e.g., clinical, counseling, developmental, and educational psychology; epidemiology; criminal justice; social work; psychiatry; and primary care). The omnipresence of these discrepancies begets countless circumstances that require researchers and practitioners to make a single decision using multiple pieces of information that do not "tell the same story." These decisions vary in whether they address questions posed in practice settings (e.g., *Does this child's behavior warrant treatment? Is the treatment I am implementing working?*), clinical research settings (e.g., *Do psychological and pharmacological treatments for children's social anxiety differ in their treatment response rates?*), and nonclinical research settings (e.g., *Does maladaptive parenting pose risk for the development of children's aggressive behavior?*). In none of these settings has any research uncovered a definitive behavioral or biological index of the behaviors being assessed; this reality necessitates the use of multiple reports that almost inevitably yield inconsistent assessment outcomes.

However, empirical work highlights that researchers and practitioners continue to encounter difficulties in understanding how to use multiple informants' reports when discrepancies arise (e.g., De Los Reyes, Alfano, et al., 2011; De Los Reyes, Kundey, et al., 2011; De Los Reyes, Youngstrom, Swan, et al., 2011; Hawley & Weisz, 2003). In addition, we continue to have a poor understanding of how to structure assessment procedures so as to increase the utility of discrepant assessment outcomes (e.g., De Los Reyes, Ehrlich, et al., 2013; De Los Reyes & Kazdin, 2006). In response to these crucial gaps, this paper delineated an agenda to guide research on informant discrepancies in developmental psychopathology research. The hope is that research targeting the objectives outlined in this paper will lead to the development of assessment methods that allow researchers and practitioners to actively incorporate informant discrepancies into how they administer multi-informant assessments and interpret the outcomes of these assessments.

References

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. (2011). Definitely more than measurement error: But how should we understand and deal with informant discrepancies? *Journal of Clinical Child and Adolescent Psychology*, 40, 80–86.

- 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.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Baker, L. A., Jacobson, K. C., Raine, A., Lozano, D. I., & Bezdjian, S. (2007). Genetic and environmental bases of childhood antisocial behavior: A multi-informant twin study. *Journal of Abnormal Psychology*, *116*, 219–235.
- Baldwin, J. S., & Dadds, M. R. (2007). Reliability and validity of parent and child versions of the Multidimensional Anxiety Scale for Children in community samples. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 252–260.
- Bartels, M., Boomsma, D. I., Hudziak, J. J., van Beijsterveldt, T. C. E. M., & van den Oord, E. J. C. G. (2007). Twins and the study of rater (dis)agreement. *Psychological Methods*, *12*, 451–466.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*, 1061–1071.
- Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin*, *98*, 388–400.
- Collishaw, S., Goodman, R., Ford, T., Rabe-Hesketh, S., & Pickles, A. (2009). How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? *Journal of Child Psychology and Psychiatry*, *50*, 571–580.
- Comer, J. S., & Kendall, P. C. (2004). A symptom-level examination of parent–child agreement in the diagnosis of anxious youths. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*, 878–886.
- De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *40*, 1–9.
- De Los Reyes, A., Alfano, C. A., & Beidel, D. C. (2010). The relations among measurements of informant discrepancies within a multisite trial of treatments for childhood social phobia. *Journal of Abnormal Child Psychology*, *38*, 395–404.
- De Los Reyes, A., Alfano, C. A., & Beidel, D. C. (2011). Are clinicians' assessments of improvements in children's functioning "global"? *Journal of Clinical Child and Adolescent Psychology*, *40*, 281–294.
- De Los Reyes, A., Ehrlich, K. B., Swan, A. J., Luo, T., Van Wie, M., & Pabón, S. C. (2013). An experimental test of whether informants can report about child and family behavior based on settings of behavioral expression. *Journal of Child and Family Studies*, *22*, 177–191.
- De Los Reyes, A., Goodman, K. L., Kliewer, W., & Reid-Quinones, K. (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. (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. (2006). 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. (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.
- De Los Reyes, A., Kundey, S. M. A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review*, *31*, 829–838.
- De Los Reyes, A., & Marsh, J. K. (2011). Patients' contexts and their effects on clinicians' impressions of conduct disorder symptoms. *Journal of Clinical Child and Adolescent Psychology*, *40*, 479–485.
- De Los Reyes, A., & Prinstein, M. J. (2004). Applying depression–distortion hypotheses to the assessment of peer victimization in adolescents. *Journal of Clinical Child and Adolescent Psychology*, *33*, 325–335.
- De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, *9*, 123–149.
- De Los Reyes, A., Youngstrom, E. A., Pabón, S. C., Youngstrom, J. K., Feeny, N. C., & Findling, R. L. (2011). Internal consistency and associated characteristics of informant discrepancies in clinic referred youths age 11 to 17 years. *Journal of Clinical Child and Adolescent Psychology*, *40*, 36–53.
- De Los Reyes, A., Youngstrom, E. A., Swan, A. J., Youngstrom, J. K., Feeny, N. C., & Findling, R. L. (2011). Informant discrepancies in clinical reports of youths and interviewers' impressions of the reliability of informants. *Journal of Child and Adolescent Psychopharmacology*, *21*, 417–424.
- Derks, E. M., Hudziak, J. J., van Beijsterveldt, T. C. E. M., Dolan, C. V., & Boomsma, D. I. (2006). Genetic analyses of maternal and teacher ratings on attention problems in 7-year-old Dutch twins. *Behavior Genetics*, *36*, 833–844.
- Dirks, M. A., Boyle, M. H., & Georgiades, K. (2011). Psychological symptoms in youth and later socioeconomic functioning: Do associations vary by informant? *Journal of Clinical Child and Adolescent Psychology*, *40*, 10–22.
- Dirks, M. A., De Los Reyes, A., Briggs-Gowan, M. J., Cella, D., & Wakschlag, L. S. (2012). Embracing not erasing contextual variability in children's behavior: Theory and utility in the selection and use of methods and informants in the assessment of developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *53*, 558–574.
- Donenberg, G. R., & Weisz, J. R. (1997). Experimental task and speaker effects on parent–child interactions of aggressive and depressed/anxious children. *Journal of Abnormal Child Psychology*, *25*, 367–387.
- Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice*, *7*, 435–453.
- Fisher, S. L., Bucholz, K. K., Reich, W., Fox, L., Kuperman, S., Kramer, J., et al. (2006). Teenagers are right—Parents do not know much: An analysis of adolescent–parent agreement on reports of adolescent substance use, abuse, and dependence. *Alcoholism: Clinical and Experimental Research*, *30*, 1699–1710.
- Gonzales, N. A., Cauce, A. M., & Mason, C. A. (1996). Interobserver agreement in the assessment of parental behavior and parent–adolescent conflict: African American mothers, daughters, and independent observers. *Child Development*, *67*, 1483–1498.
- Goodman, K. L., De Los Reyes, A., & Bradshaw, C. P. (2010). Understanding and using informants' reporting discrepancies of youth victimization: A conceptual model and recommendations for research. *Clinical Child and Family Psychology Review*, *13*, 366–383.
- Goodman, S. H., & Gotlib, I. H. (1999). Risk for psychopathology in the children of depressed mothers: A developmental model for understanding mechanisms of transmission. *Psychological Review*, *106*, 458–490.
- Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review*, *113*, 101–131.
- Grills, A. E., & Ollendick, T. H. (2002). Issues in parent–child agreement: The case of structured diagnostic interviews. *Clinical Child and Family Psychology Review*, *5*, 57–83.
- Hartley, A. G., Zakriski, A. L., & Wright, J. C. (2011). Probing the depths of informant discrepancies: Contextual influences on divergence and convergence. *Journal of Clinical Child and Adolescent Psychology*, *40*, 54–66.
- 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.
- Holmbeck, G. N., Li, S. T., Schurman, J. V., Friedman, D., & Coakley, R. M. (2002). Collecting and managing multisource and multimethod data in studies of pediatric populations. *Journal of Pediatric Psychology*, *27*, 5–18.
- Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology*, *3*, 29–51.
- Institute for Scientific Information. (2011). *Researcher ID A1712-2002-S*. Retrieved July 10, 2011, from www.ISIHighlyCited.com
- Kassam-Adams, N., Garcia-España, J. F., Miller, V. A., & Winston, F. (2006). Parent–child agreement regarding children's acute stress: The role of parent acute stress reactions. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*, 1485–1493.

- Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: His and hers. *Psychological Bulletin*, *127*, 472–503.
- 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.
- 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.
- Kramer, T. L., Phillips, S. D., Hargis, M. B., Miller, T. L., Burns, B. J., & Robbins, J. M. (2004). Disagreement between parent and adolescent reports of functional impairment. *Journal of Child Psychology and Psychiatry*, *45*, 248–259.
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, *50*, 537–567.
- Lapouse, R., & Monk, M. A. (1958). An epidemiologic study of behavior characteristics of children. *American Journal of Public Health*, *48*, 1134–1144.
- Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as informants of childhood psychopathology. *Journal of Clinical Child Psychology*, *19*, 136–143.
- Lorenz, F. O., Melby, J. N., Conger, R. D., & Xu, X. (2007). The effects of context on the correspondence between observational ratings and questionnaire reports of hostile behavior: A multitrait, multimethod approach. *Journal of Family Psychology*, *21*, 498–509.
- Mitsis, E. M., McKay, K. E., Schulz, K. P., Newcorn, J. H., & Halperin, J. M. (2000). Parent–teacher concordance for DSM-IV attention-deficit/hyperactivity disorder in a clinic-referred sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 308–313.
- Neal, J. W., Cappella, E., Wagner, C., & Atkins, M. S. (2011). Seeing eye to eye: Predicting teacher–student agreement on classroom social networks. *Social Development*, *20*, 376–393.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw–Hill.
- 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.
- Randazzo, K. V., Landsverk, J., & Ganger, W. (2003). Three informants' reports of child behavior: Parents, teachers, and foster parents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*, 1343–1350.
- Renk, K. (2005). Cross-informant ratings of the behavior of children and adolescents: The “gold standard.” *Journal of Child and Family Studies*, *14*, 457–468.
- Renk, K., & Phares, V. (2004). Cross-informant ratings of social competence in children and adolescents. *Clinical Psychology Review*, *24*, 239–254.
- 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.
- Rosnati, R., Montiroso, R., & Barni, D. (2008). Behavioral and emotional problems among Italian international adoptees and non-adopted children: Father's and mother's reports. *Journal of Family Psychology*, *22*, 541–549.
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, *1*, 199–223.
- Taber, S. M. (2010). The veridicality of children's reports of parenting: A review of factors contributing to parent–child discrepancies. *Clinical Psychology Review*, *30*, 999–1010.
- Tackett, J. L., Waldman, I. D., & Lahey, B. B. (2009). Etiology and measurement of relational aggression: A multi-informant behavior genetic investigation. *Journal of Abnormal Psychology*, *118*, 722–733.
- 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.
- Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., et al. (2008). Observational assessment of preschool disruptive behavior: Part I. Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*, 622–631.
- 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.
- 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., Izard, C., & Ackerman, B. (1999). Dysphoria-related bias in maternal ratings of children. *Journal of Consulting and Clinical Psychology*, *67*, 905–916.