

Linking Informant Discrepancies to Observed Variations in Young Children's Disruptive Behavior

Andres De Los Reyes · David B. Henry ·
Patrick H. Tolan · Lauren S. Wakschlag

© Springer Science + Business Media, LLC 2009

Abstract Prior work has not tested the basic theoretical notion that informant discrepancies in reports of children's behavior exist, in part, because different informants observe children's behavior in different settings. We examined patterns of observed preschool disruptive behavior across varying social contexts in the laboratory and whether they related to parent-teacher rating discrepancies of disruptive behavior in a sample of 327 preschoolers. Observed disruptive behavior was assessed with a lab-based developmentally sensitive diagnostic observation paradigm that assesses disruptive behavior across three interactions with the child with parent and examiner. Latent class analysis

identified four patterns of disruptive behavior: (a) low across parent and examiner contexts, (b) high with parent only, (c) high with examiner only, and (d) high with parent and examiner. Observed disruptive behavior specific to the parent and examiner contexts were uniquely related to parent-identified and teacher-identified disruptive behavior, respectively. Further, observed disruptive behavior across both parent and examiner contexts was associated with disruptive behavior as identified by both informants. Links between observed behavior and informant discrepancies were not explained by child impairment or observed problematic parenting. Findings provide the first laboratory-based support for the Attribution Bias Context Model (De Los Reyes and Kazdin Psychological Bulletin 131:483–509, 2005), which posits that informant discrepancies are indicative of cross-contextual variability in children's behavior and informants' perspectives on this behavior. These findings have important implications for clinical assessment, treatment outcomes, and developmental psychopathology research.

This project was supported by NIMH grants R01 MH68455 and MH62437 and National 0–3 to Lauren Wakschlag, and support from the Shaw and Children's Brain Research Foundations. The contributions of Patrick H. Tolan and David B. Henry were supported by CDC grant U49/CE 000732, NICHD grant R01 HD042030 and NIDA grant R01 DA020829.

A. De Los Reyes (✉)
Department of Psychology,
University of Maryland at College Park,
1147 Biology/Psychology Building, Room 3123H,
College Park, MD 20742, USA
e-mail: adelosreyes@psyc.umd.edu

D. B. Henry · P. H. Tolan · L. S. Wakschlag
Institute for Juvenile Research, Department of Psychiatry,
University of Illinois at Chicago,
1747 W. Roosevelt Rd., MC 747,
Chicago, IL 60608, USA

D. B. Henry
e-mail: dhenry@psych.uic.edu

P. H. Tolan
e-mail: ptolan@psych.uic.edu

L. S. Wakschlag
e-mail: lwakschlag@psych.uic.edu

Keywords Attribution Bias Context Model ·
Correspondence · Disagreement · Disruptive behavior ·
Informant discrepancies

Cross-informant discrepancies in reports of clinical symptoms present significant challenges to the assessment and identification of clinical phenomena. Meta-analyses of cross-informant agreement in both adult and youth assessments have revealed low-to-moderate correlations across ratings taken from multiple information sources of the same clinical symptoms (e.g., *r*s ranging from 0.20s to 0.60s; Achenbach et al. 2005; Achenbach et al. 1987). These discrepancies are among the most consistent, yet poorly understood phenomena in mental health research (e.g., Achenbach 2006).

Informant discrepancies are often dismissed by researchers (and even the informants themselves; see Bidaut-Russell et al. 1995) as measurement error or suggested to represent informant bias (e.g., maternal depression is associated with elevated ratings of symptoms; Richters 1992). However, cross-informant discrepancies may reflect true situation-related variation in behavior (e.g., disruptive behaviors exhibited at home but not school) and/or differences in informants' knowledge and experiences (e.g., teachers have different reference points than parents such as experience with many typical preschoolers versus their own child) (Achenbach 2006; De Los Reyes and Kazdin 2005). In order to empirically examine conceptualizations of discrepancies as reflective of, at least in part, true differences in children's behavior across contexts, this paper examines the correspondence of informant discrepancies in reports of young children's disruptive behavior (i.e., oppositional and conduct symptoms) to observed differences in disruptive behavior across laboratory contexts.

To some extent, informant discrepancies or lack thereof have been considered as reflecting meaningful clinical differences, particularly in the conceptualization of specific clinical conditions. For instance, cross-setting consistency in symptoms is considered an indication of severity and impairment for most disorders and for some is required for meeting diagnostic criteria (*Diagnostic and statistical manual of mental disorders* [DSM-IV]; American Psychiatric Association 2000). For example, because oppositional defiant disorder often emerges early in life and can be confined within parent-child interactions, questions have been raised as to whether this specificity to one relationship context may reflect problems in parenting rather than a clinical disorder in the child (Moffitt et al. 2007; Wakschlag and Danis in press). In the debate about the validity of preschool disruptive behavior disorders, it has also been suggested that inconsistency across situations may indicate a transient developmental perturbation rather than a clinical pattern (Campbell 2002). For instance, if inconsistency in disruptive behavior is a marker for transiency in its expression, then it follows that situation-specific disruptive behavior should dissipate over time and may not warrant intervention.

A key factor in the dearth of knowledge available on the clinical significance of informant discrepancies is the prevailing notion in clinical assessment research that discrepancies largely reflect measurement error or informant bias (e.g., Richters 1992; Youngstrom et al. 1999). Although such measurement issues are certainly a key consideration in understanding informant discrepancies, alternative explanations have been posited. Specifically, discrepancies may reflect true differences in the situations which typically serve as the bases for informants' ratings of behavior (Achenbach et al. 1987; Kraemer et al. 2003). For example, parents may base their ratings on how their

children behave at home (e.g. when asked to clean up toys or go to bed), and teachers may base their ratings on how the same children behave at school (e.g. in interactions with peers). Indeed, parents and teachers are often relied on in order to gather information about children's behavior from home and school contexts, respectively (see Kraemer et al. 2003). However, this too is multifaceted. Not only do home and school situations vary in the extent, activity purpose, and the relationship between the child and adult; the relationship between the child being rated and the adult may fundamentally differ as well. For instance, parents likely base their ratings on a very limited set of experiences between themselves and their child. Conversely, teachers may have 20 or more new students each year and thus have limited time and range of activities with any given child; even if they are in their classroom or under their care the whole school day. Thus, teacher ratings may be more likely based in a broad sense of normative expectations of children but less variation in setting, structure of interactions, and roles during those interactions. These situational variations may result in discrepancies between parent and teacher ratings (see Kerr et al. 2007).

Consistent with the idea that informant discrepancies reflect situational variations in expressions of behavior, prior work has examined whether discrepancies reflect informants' perceptual biases or behavior that is expressed differently across informants. For instance, theoretical work in behavior genetics suggests that multiple informant strategies for assessing behavior in twin studies provide the opportunity to disentangle informant bias from informant-specific observations of behavior (Bartels et al. 2007). Thus, in this literature informant discrepancies might not mean that one informant provides less reliable ratings than the other. Rather, children may behave differently in front of informants, perhaps reflective of them expressing behaviors differently across the varied contexts within which informants observe behavior (e.g., home, school; Bartels et al. 2007).

We acknowledge that some of the inconsistencies among informants' ratings may be solely due to measurement error. However, we postulate that cross-informant discrepancies represent, in part, true variation among children in how much their behavior is bound to a given social context. Thus, variations across informants' ratings of children's behavior represent not only measurement artifacts but also meaningful heterogeneity in clinical patterns. These reflect differences in the actual behavior these children display in their interactions with multiple informants. Consistent with this work, we have previously proposed an explanatory framework for interpreting the clinical meaning of informant discrepancies (De Los Reyes and Kazdin 2005). Specifically, the Attribution Bias Context Model posits that when informant discrepancies arise, these discrepancies relate meaningfully to perceptual and situational differences

among informants in how and under what circumstances they observe behavior. However, these tenets of the Attribution Bias Context Model and in particular the specific notion that discrepancies reflect meaningful variations in the behavior being rated have not been empirically tested. In light of the difficulty in disentangling the effects of rater bias, situational variation and unique informant perspectives, parsing these issues clearly requires mapping informant discrepancies onto variations in observed behavior that are rated *independently of either informant* (see Richters 1992).

Demonstrating systematic linkages between informant discrepancies and parallel contextual differences in independently observed manifestations of young children's disruptive behavior is a critical next step for elucidating the clinical meaning of informant discrepancies. However, identifying these relations on its own would not definitively establish that informant discrepancies reflect real-world cross-situation variations in behavior. Here we explicate two particularly strong competing explanations. The first is that discrepancies may be a marker for severity of dysfunction. In this case, informant discrepancies would be most likely in the case of milder forms of disruptive behaviors, because the child being rated discrepantly by informants exhibits less impairing behavior overall. Therefore, the nature and extent of this child's disruptive behavior may be more likely to be "missed" or differentially interpreted across varying situations and thus across various informants observing behavior across situations. Conversely then, informant agreement would be more likely evident across ratings of more severely impaired children. Indeed, their disruptive behaviors would likely be exhibited at higher frequencies and in more extreme forms and thus would be more likely to be seen and rated by multiple informants.

A second alternative explanation is that discrepancies between parental and non-parental informants reflect *characteristics of the parent* rather than context-specific aspects of child behavior. This is particularly salient in the case of discrepancies in ratings of child disruptive behavior, because problematic parenting has consistently been shown to contribute to disruptive behavior patterns (Tolan et al. 1997). Specifically, variations in the nature and extent of disruptive behavior across various situations might result from characteristics in people primarily interacting with the child in particular situations (e.g., parent at home exhibiting harsh discipline strategies) that are not exhibited by people primarily interacting with the child in other situations (e.g., teacher at school that might face legal sanctions for exhibiting the same kind of discipline strategies exhibited by parent at home). Thus, in the present study a key factor might be the presence of harsh parental discipline, which may induce misbehavior exhibited by a child that by virtue of parental behavior, only exhibits disruptive behavior during parent-child interactions (see Granic and Patterson 2006). Harsh parental discipline might create a circum-

stance in which a child who is not otherwise disruptive exhibits disruptive behavior with his or her parent.

Although informant discrepancies are prevalent across assessments of a range of childhood psychopathologies and developmental periods, here we provide an empirical test of our hypothesis within the context of a particular disorder (disruptive behavior) during a specific developmental period (early childhood). We do so for a number of reasons. Specifically, we believe that putting our theory to the test during a developmental period in which identification of clinical patterns is especially challenging due to high rates of normative misbehavior (Campbell 2002; Wakschlag et al. 2007a) is a particularly rigorous threshold to pass for confirmatory evidence. Furthermore, unlike the pervasive developmental disorders (e.g., Ozonoff et al. 2005), cross-situational patterns are not currently incorporated into disruptive behavior classification systems such as DSM-IV. Yet, there is evidence that varying contextual patterns of disruptive behavior may have partially distinct etiologies and differing prognoses (e.g., greater overt versus covert levels of disruptive behavior in child-onset versus adolescent-onset subtypes of conduct disorder; see Kazdin and De Los Reyes 2007). As such, elucidating the extent to which informant discrepancies reflect meaningful contextual variation has important implications for advancing clinical conceptualizations of disruptive behavior disorders in children.

Most critically, previously a barrier to empirically examining the relation between contextual variation in behavior and informant discrepancies has been the dearth of direct, clinically feasible assessment methods for measuring contextual variations in behavior independently from a particular informant. Thus, we focus on informant discrepancies in the assessment of young children's disruptive behavior because by doing so we can take advantage of recent advances in developmental measurement that allow for the examination of variations in young children's disruptive behavior across varying interactional contexts within a laboratory setting (Wakschlag et al. 2008a, b). Prior work has called attention to the promise of "performance based measurement" for informing the identification of clinically relevant behaviors in children (Frick 2000). However, these assessments typically do not incorporate measurement of contextual variations in behavior. More recently, these methods have been developed. Specifically, the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS; Wakschlag et al. 2008a) was developed to distinguish the normative misbehavior of early childhood from disruptive behavior symptoms. Of relevance to the study of informant discrepancies is that the DB-DOS assesses disruptive behavior across different interactions, which vary by interactive partner (with parental and non-parental adult [clinical examiner]). Thus, the DB-DOS provides an opportunity to assess how children's display of

disruptive behavior varies as a function of the nature of the child's relationship with the adult and by the social demands of the interaction. Specifically, DB-DOS interactions between the child and his or her parent are intended to yield a proxy measure of how a child behaves in home situations with adults with whom they have a long history of interactions. Conversely, structured interactions between the child and a clinical examiner are designed to yield a proxy measure of how a child behaves outside of the home with adults with whom the child has a much shorter history of interactions (e.g., teacher). As a result, we surmise that the nature of laboratory interactions between a clinical examiner and a child might map onto real-world interactions between that same child and his or her teacher, whereas parent-child laboratory interactions might map onto real-world parent-child interactions. In this way, the DB-DOS provides a unique opportunity to test whether real-world discrepancies in reports of child behavior by parent and teacher are paralleled in independent ratings of contextual variations in child behavior observed in the laboratory with parental and non-parental adults. Along these lines, we have previously demonstrated that variability in observed social skills across the DB-DOS contexts predicts preschoolers' real-world functioning over and above their average skill level (Dirks et al. 2009).

In summary, in this paper, we use direct observations of contextual variation in young children's disruptive behavior to "put theory to the test" in regard to the clinical relevance of informants' rating discrepancies. In particular, we address the following questions:

- (1) Do cross-situational variations in laboratory observations of young children's disruptive behavior during interactions with parental and non-parental adults meaningfully correspond to parent-teacher discrepancies in reporting of disruptive behavior symptoms in these same children? Here we hypothesized that variations on the DB-DOS in terms of observed disruptive behavior with parental and non-parental adults would relate to parent-teacher rating discrepancies. In particular, we hypothesized that disruptive behaviors observed with the clinical examiner (lab-based behavioral patterns with non-parental adult) would parallel teacher ratings of disruptive behavior (reflecting real-world patterns of disruptive behavior with non-parental adults). Thus, we hypothesized that children exhibiting high levels of observed disruptive behavior during the DB-DOS examiner contexts only would be more likely to be teacher-identified only (i.e., children identified as being high on disruptive behavior symptoms based only on teacher report). Conversely, we hypothesized that children exhibiting high levels of disruptive behavior only during the

DB-DOS parent context would be more likely to be parent-identified only (i.e., children identified as being high on disruptive behavior symptoms based only on parent report). Finally, we hypothesized that children who display disruptive behavior across *both* DB-DOS parent and examiner contexts would be more likely to have parent and teacher ratings of disruptive behavior that correspond. We assessed these parent-teacher discrepancies using various informant-specific definitions of disruptive behavior symptoms (e.g., parent identifies disruptive behavior symptoms that teacher does not and vice versa) as well as combinational rules used in prior work for identifying clinical cases (e.g., "and" rules; see Youngstrom et al. 2003).

- (2) Can these discrepancy patterns be explained solely by taking child impairment and/or quality of observed parenting into account? Here we hypothesized that patterns of correspondence would not be accounted for by children's functional impairment or by observed problematic parenting behavior.

Method

Participants

Participants were derived from the Observing Young Children and Families Study of the Chicago Preschool Project. The central aim of this study was to validate the psychometric properties of the DB-DOS (for details see Wakschlag et al. 2008a). The sample for this study was comprised of 327 low-income preschoolers (ages 3–5) and their biological mothers (details of the sample have been previously described; Wakschlag et al. 2008a). (Nine children from the larger Chicago Preschool Project sample of 336 were excluded because of incomplete DB-DOS data.) Eligibility criteria were (a) being between the ages of 3–5, (b) growing up in a low-income environment due to increased rates of early emerging disruptive behavior within this psychosocially at-risk environment, (c) participation of a biological mother, and (d) attendance in preschool or out-of-home day care at least three hours per day three times per week. This latter criterion is noteworthy for this study of preschoolers, because unlike older children, many young children are not yet in school or out-of-home care. We required this because of the importance of cross-informant information for making clinical distinctions during a developmental period in which making these distinctions is challenging.

Participants were recruited from two Midwestern universities, from a larger sample of 336 children and their mothers. In order to examine individual differences in disruptive behavior along a continuum of such behavior,

Table 1 Demographic characteristics ($n=327$)

Children in the sample were recruited for this study based on whether they were clinically referred for behavioral concerns, not referred for behavioral concerns but were considered to exhibit such concerns, or not referred for such concerns and were considered to not exhibit disruptive behavior concerns (for further information on study recruitment see Wakschlag et al. 2008a).

Characteristic	Baseline demographic statistics
Child age	3 years (114, 34.9%) 4 years (105, 32.1%) 5 years (108, 33%) $M (SD)=3.98 (0.82)$
Child gender	Male (179, 54.7%) Female (148, 45.3%)
Ethnicity (% African American)	273 (83.5%)
Referral status	Not referred, no behavioral concerns (98, 30%) Not referred, behavioral concerns (101, 30.9%) Referred, behavioral concerns (128, 39.1%)

the sample was designed to represent the full spectrum of disruptive behavior. Thus, at each of the two universities, children were recruited from two sources: (a) referred children from a preschool disruptive behavior clinic ($n=128$, 39.1%) and (b) non-referred children from Pediatric clinics of these same universities ($n=199$, 60.9%). In order to further enhance behavioral heterogeneity, approximately 50% of this non-referred group ($n=101$) was recruited based on the presence of behavioral concerns (by parents or other adults) in the absence of mental health services utilization. Complete baseline demographic and DB-DOS data were available for these 327 children (see Table 1).

For the present study, we utilized three different subsamples of the total participant sample for three sets of analyses (see Fig. 1). First, in order to construct statistical models of children's disruptive behavior on the DB-DOS, we relied on participants with full DB-DOS data ($n=327$). Second, ninety-one percent of these children ($n=298$) participated in a 1-year follow-up DB-DOS assessment, of which 292 were included in tests of the stability of statistical models constructed based on baseline DB-DOS data. These baseline and 1-year follow-up DB-DOS samples did not differ from each other in terms of disruptive behavior symptoms, impairment, or demographic characteristics (Wakschlag et al. 2008a, b). Third, tests of the relations between variations in observed disruptive behavior and parent-teacher rating discrepancies were conducted based on the 288 participants with complete information at baseline on children's disruptive behavior via parent and teacher report, as well as via the DB-DOS. The demographic characteristics of the children from the total sample ($n=327$) did not significantly differ from the characteristics of the 288 children with complete baseline parent, teacher, and DB-DOS data.

Measures

Disruptive behavior symptoms were assessed by parent and teacher reports using developmentally validated measures

of DSM-IV DBD symptoms for preschool children. Identifying clinically significant disruptive behavior symptoms was defined as meeting symptom criteria for DSM-IV disruptive behavior disorders (i.e., Oppositional Defiant Disorder [ODD; 4 or more symptoms], Conduct Disorder [CD; 3 or more symptoms], or Disruptive Behavior Disorder Not Otherwise Specified [defined *a priori* as 3 or more symptoms across ODD and CD]; for details see Wakschlag et al. 2008b). Thus, clinically significant disruptive behavior symptoms were present if a child had

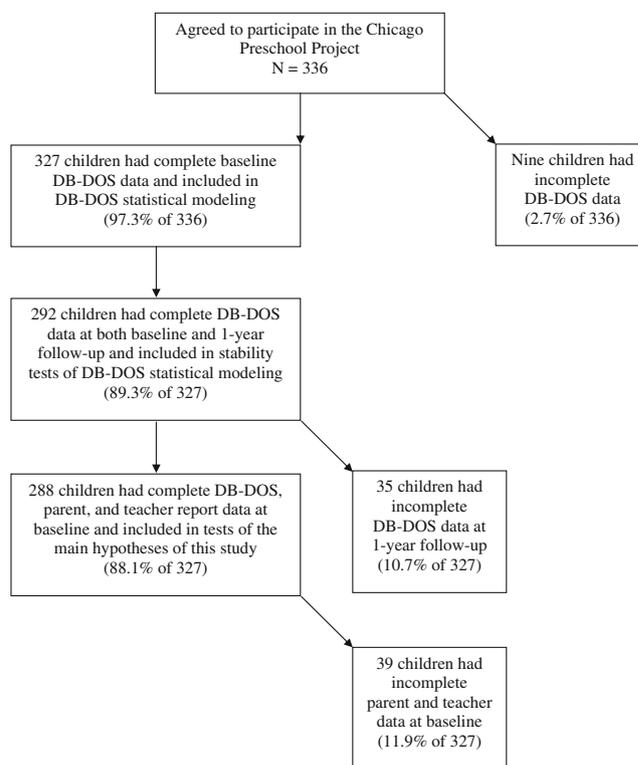


Fig. 1 Flow chart of the total sample and the different subsamples employed in the statistical modeling of the DB-DOS, longitudinal stability of DB-DOS statistical models, and the primary tests of relations between observed disruptive behavior on the DB-DOS and parent-teacher rating discrepancies of disruptive behavior

three or more DBD symptoms endorsed by an informant. The same definitions were applied for both informants.

By maternal report Mothers were administered the Kiddie Disruptive Behavior Disorders Schedule (K-DBDS), a diagnostic interview developed to assess disruptive behavior in children (Keenan et al. 2007). Here we used a developmentally enhanced version of the K-DBDS, which was designed to elucidate distinctions between symptoms and normative misbehaviors in preschool children by incorporating more detailed information about quality of symptomatic behavior (e.g., severity, difficulty modulating) into determinations of clinical significance (Keenan et al. 2001). The K-DBDS has demonstrated adequate reliability and validity (for details see Keenan et al. 2007). Mean symptoms by maternal report were 2.12 ($SD=2.77$). Eighty-eight (30.6%) children had clinically significant disruptive behavior symptoms by maternal report (see Table 2).

By teacher report Teacher ratings of disruptive behavior symptoms were taken with the Early Child Symptom Inventory (ECI; Gadow and Sprafkin 1996), a DSM-IV based checklist that assesses symptoms in young children. The ECI generates categorical and continuous symptom scores. Scores have demonstrated good test-retest reliability ($r=.56$ for ODD and $r=.41$ for CD) and differentiate clinic-referred from nonreferred children (Gadow et al. 2001). Mean symptoms by teacher report were 1.71 ($SD=3.37$). Sixty (20.8%) children had clinically significant disruptive behavior symptoms by teacher report (see Table 2).

Observed disruptive behavior Disruptive behavior was assessed in the laboratory with the DB-DOS (Wakschlag et al. 2008a). The DB-DOS is a 50-min structured laboratory observation with three interactional contexts: one Parent context and two Examiner contexts. The two Examiner contexts were administered by the same examiner but varied by level of support provided (i.e., “Examiner Engaged” context, where examiner is present and actively engaged with child and “Examiner Busy” context, where

examiner is present but “busy with other work” and not actively engaging with the child). Parent and Examiner contexts are designed to involve parallel “presses” for child misbehavior (e.g., frustration task, compliance task) (for a description see Wakschlag et al. 2007b, 2008a). Behaviors on the DB-DOS are coded separately across the three interactions using a system of 21 codes encompassing the Domains of *Problems in Behavioral Regulation* (i.e., noncompliant, provocative and inflexible behavior) and *Problems in Anger Modulation* (i.e., intense, difficult to modulate angry/irritable behavior).

Distinctions between normative misbehavior and disruptive behavior are made within the DB-DOS coding system by defining qualitative “breakpoints” that mark the shift from typical to atypical behavior. Ordinal ratings are made along a clinical continuum: *Normative Variation* (0=normative behavior, 1=normative misbehavior) and *Clinically Concerning* (2=of concern, 3=atypical). Scoring yields Domain scores for each Interaction, yielding 6 scores (e.g., *Behavioral Regulation* in the Examiner Engaged Context). The DB-DOS has demonstrated adequate item-level interrater reliability (mean weighted kappa=.68), as well as domain- and context-level internal consistency (coefficient alpha range of 0.82–0.93) and 4-week test-retest reliability (intraclass correlation coefficients range of 0.61–0.85) (Wakschlag et al. 2008a).

Drawing on previous validation work (Wakschlag et al. 2008b), for statistical modeling of observed behavior on the DB-DOS, we employed DB-DOS problem scores. Specifically, we created dichotomous problem scores differentiating high and low levels of Anger Modulation and Behavioral Regulation problems within and across the three adult-child interactions. These dichotomous problem scores were constructed such that each DB-DOS item rated as a 2 or 3 (behaviors of clinical concern) was given a score of 1 and ratings of 0/1 were given a score of 0. Dichotomized items were then summed to generate a problem count for each problem domain (Anger Modulation, Behavioral Regulation) within each DB-DOS context. These problem counts were subsequently dichotomized at

Table 2 Cross-tabulation of parent and teacher identifications of disruptive behavior in preschool children ($n=288$)

Teacher	Parent		Total
	Low disruptive behavior symptom children	High disruptive behavior symptom children	
Low disruptive behavior symptom children	164 (56.9%)	64 (22.2%)	228 (79.2%)
High disruptive behavior symptom children	36 (12.5%)	24 (8.3%)	60 (20.8%)
Total	200 (69.4%)	88 (30.6%)	288
Parent-teacher agreement	$\chi^2(1)=3.19, ns$ $\phi=0.10, ns; \kappa=0.10, ns$		

Percentages in parentheses reflect percentages of children in the discrepancy sub-sample ($n=288$).

the upper quartile to create 6 problem scores (0=no/low observed problems, 1=high observed problems). These dichotomous scores were subsequently employed to identify variations in disruptive behavior within and across DB-DOS interactional contexts. We have previously demonstrated that DB-DOS problem scores (a) differentiate preschoolers with- and without DBDs, (b) predict persistent impairment over time, and (c) add incremental utility in prediction above and beyond DSM-IV DBD symptoms (Wakschlag et al. 2008b).

Impairment To provide a global assessment of functional impairment, mothers and teachers independently completed at the baseline assessment the non-clinician version of the Children's Global Assessment Scale (C-GAS), an impairment measure for children from preschool through adolescence (Setterberg et al. 1992). The C-GAS has been used in multiple studies of preschool disruptive behavior with children as young as 2 years of age (Lavigne et al. 1998; Wakschlag and Keenan 2001). Scores range from 1–100, with scores of 60 or below indicating impairment. Based on the “60 or below” criterion and for children with complete data from parent, teacher, and the DB-DOS ($n=288$), 182 (63.2%) were not identified as impaired by either parent or teacher, 36 (12.5%) were identified by parent only, 47 (16.3%) were identified by teacher only, and 23 (7%) were identified as impaired by both parent and teacher. We employed the separate impairment identifications by parent only, teacher only, or both ($n=106$) to test whether impairment accounted for variations in behavior on the DB-DOS.

Parenting Style Mothers' behavior during the DB-DOS Parent context administered at the baseline assessment was coded using the Parenting Clinical Observation Schedule (P-COS; Hill et al. 2008). Like DB-DOS codes, P-COS codes are global, ordinal ratings assessing parenting behaviors along a continuum from competent to clinically concerning. Parenting behaviors on the P-COS are rated in three Domains: parental *Responsive Involvement*, *Constructive Discipline*, and *Problematic Discipline*. This resulted in 26% of mothers being categorized as exhibiting problematic parenting, defined as being rated as exhibiting Problematic Discipline and not demonstrating competence on either the Responsive Involvement or Constructive Discipline Domains (0=not problematic, 1=problematic). This approach enabled us to take multiple dimensions of parenting (both competence and problems) into account simultaneously to capture “style” rather than focusing on a single dimension or type of behavior. In the present sample, we have previously demonstrated the clinical significance of this score, with problematic parenting style predicting changes in child DBD status over time (Hill et al. 2008).

Procedures

Details regarding the study were first provided at initial phone contact with the mother. Informed consent was formally obtained from the mother at the laboratory visit. Institutional Review Boards at both universities affiliated with the study approved all procedures (Wakschlag et al. 2008a). Before the laboratory visit, mothers were mailed a packet of questionnaires regarding child and family functioning. One research assistant administered the K-DBDS interview to mothers while a second research assistant conducted assessments with the child; mothers provided their impairment ratings as part of the K-DBDS interview. Mother and child then participated in the DB-DOS. The mothers completed additional questionnaires while the children completed the DB-DOS with the Examiner. Administration of the DB-DOS contexts was done in a standard order (Parent, followed by Examiner Engaged, and Examiner Busy). Administering the Parent context first provided an opportunity for the child to “warm up” within the context of this mildly taxing paradigm. Administering the Examiner Engaged context before the Examiner Busy context further facilitated the child's “warming up” to the examiner. This is because this order allowed the examiner-child interactions to begin in a more ecologically valid manner than if the examiner-child tasks were in reverse (i.e., examiner-child interactions begin with examiner not actively interacting with the child). DB-DOS coders were independent of examiners and were blind to child disruptive behavior status. Questionnaires were also mailed to teachers after maternal consent was obtained.¹ Assessment procedures at baseline and 1-year follow-up

¹ Two issues arise in relation to the procedures. The first is the length of time that the teachers tended to know the children for whom they were reporting at the time of the assessment. The second is the length of time between parent and teacher disruptive behavior assessments. On average, teachers knew the children for 9.52 ($SD=10.49$) months, and the mean lag (in days) between parent and teacher assessments was 36.48 ($SD=32.04$). Approximately 80% of teachers knew the child for 12 months or less, and approximately 60% of the parent and teacher assessments had a lag of administration of 32 days or less. Given the range on these variables, we were interested in examining whether they accounted for variance in a dimensional measure of parent-teacher disagreement. To assess disagreement on a continuous scale, we used the parallel C-GAS reports of child impairment completed by parent and teacher and calculated a standardized difference score, consistent with current recommendations (De Los Reyes and Kazdin, 2004). Specifically, we converted each of the dimensional parent and teacher C-GAS scores into z scores and then subtracted the teacher z score from the parent z score. The standardized difference score of parent and teacher reports of child impairment did not significantly correlate with either the length of time that the teacher knew the child, $r=-0.05$, nor the lag between parent and teacher assessments, $r=0.01$. Thus, the findings were inconclusive as to whether these variables were related to parent-teacher disagreements on reports of child behavior.

were identical. Families were paid \$60 and \$70 respectively for their participation in the two study visits (with a \$10 bonus provided if questionnaires were completed in advance).

Results

Patterns of Parent-Teacher Agreement on Ratings of Disruptive Behavior Symptoms

To examine parent-teacher agreement, kappa and phi coefficient statistics were calculated based on whether preschoolers were classified as exhibiting high parent- and/or teacher-rated levels of disruptive behavior. Parent-rated and teacher-rated disruptive behavior symptom categories and measures of agreement are presented in Table 2. Consistent with prior work (e.g., Achenbach 2006; De Los Reyes and Kazdin 2004, 2005, 2006a; De Los Reyes and Prinstein 2004), rates of parent-teacher correspondence were modest. Of the 124 children identified by parents and/or teachers as exhibiting disruptive behavior symptoms (hereafter referred to as “informant-disruptive preschoolers”), only 8.3% ($n=24$) were identified by both parent and teacher. The remaining 100 informant-disruptive preschoolers were identified by the parent or teacher, but not both simultaneously.

To examine the relations between observed behavior on the DB-DOS and parent-teacher rating discrepancies, we grouped children based on patterns of informant correspondence and discordance on disruptive behavior symptoms. Consistent with prior work on combining information from multiple informants (see Youngstrom et al. 2003) we created four groups of informant-disruptive preschoolers. We created these groups based on the pattern of disruptive behavior endorsement across the two informants, or children who were identified by: (a) Neither Parent nor Teacher, (b) Parent Only, (c) Teacher Only, or (d) Both Parent and Teacher. Thus, informant-disruptive preschoolers were identified based on definitions “a”, “b”, “c”, and “d”, and each were constructed to be mutually exclusive of each other, based on the informant. Further, definition “d” was constructed based on combinational rules for identifying clinical cases (Youngstrom et al. 2003). Specifically, definition “d” was constructed to conform to the “and” rule of identifying clinical cases. Therefore, definition “d” differed from other definitions of positively identifying informant-disruptive preschoolers (“b” [Parent Only] and “c” [Teacher Only]) in its requirement that a positive identification of disruptive behavior could only be made if parent and teacher both identified a child as disruptive. We expected children identified as disruptive under definition “d” to only be predicted by their

performance on the DB-DOS when the child was observed exhibiting disruptive behavior pervasively or across parent- and examiner-child interactions.

Latent Class Analysis of Patterns of Observed Disruptive Behavior on the DB-DOS

To identify contextual variation in patterns of children’s observed disruptive behavior we conducted exploratory latent class analyses (LCA; McCutcheon 1987) on the six dichotomous DB-DOS measures (two problem domain measures [Anger Modulation, Behavioral Regulation] X three interactional contexts [Parent, Examiner Engaged, Examiner Busy]). Like cluster analysis, LCA attempts to identify groups of cases based on similar patterns of indicator variables. Like confirmatory factor analysis, LCA includes tests of the absolute and relative fit of models. Latent class analysis uses categorical or ordinal variables to produce classes within which there is local independence of indicators (i.e., indicator variables are statistically independent within levels of each latent class). Thus, LCA is a person-centered approach that allowed us to identify classes of children exhibiting similar patterns of disruptive behavior across the six DB-DOS context-specific ratings. Probabilities provided by a latent class solution may be used to assess the confidence with which cases are assigned, and to assign new cases based on a solution (McCutcheon 1987).

Based on our hypothesis that children’s behavior in the laboratory with parental and non-parental adults would parallel real-world interactions with parent and teacher, we expected that the LCA would identify the following latent classes of preschoolers on the DB-DOS: (a) none/low across Parent and Examiner contexts, (b) high in Parent but not Examiner context, (c) high in Examiner but not Parent context, and (d) high in Parent and Examiner contexts.

We tested one- through five-class solutions, evaluating the fit and interpretability of each. The four-class solution fit the data best, $\chi^2(36)=34.88$, ns , $L^2=39.66$, ns . Figure 2 reports the fit statistics and provides a graphical representation of the four-class solution. Table 3 reports the frequencies, percentages, and probabilities of correct latent class assignment for each of the four latent classes. We also report in Table 3 the latent class probabilities of the four classes and the conditional probabilities for each of the 6 indicators. *Chi square* analyses suggested that the class assignments from this four-class LCA solution were nearly identical to the four-class LCA solution identified when controlling for child gender, age, and referral status, $\chi^2(9)=841.06$, $p<0.001$, kappa=0.92, Cramer’s $V(3)=0.92$.

Consistent with our hypothesis, we identified the following behavioral profiles: (a) none/low in any context (Not Disruptive), (b) parent only (Disruptive with Parent), (c) examiner only (Disruptive with Examiner), and (d) with

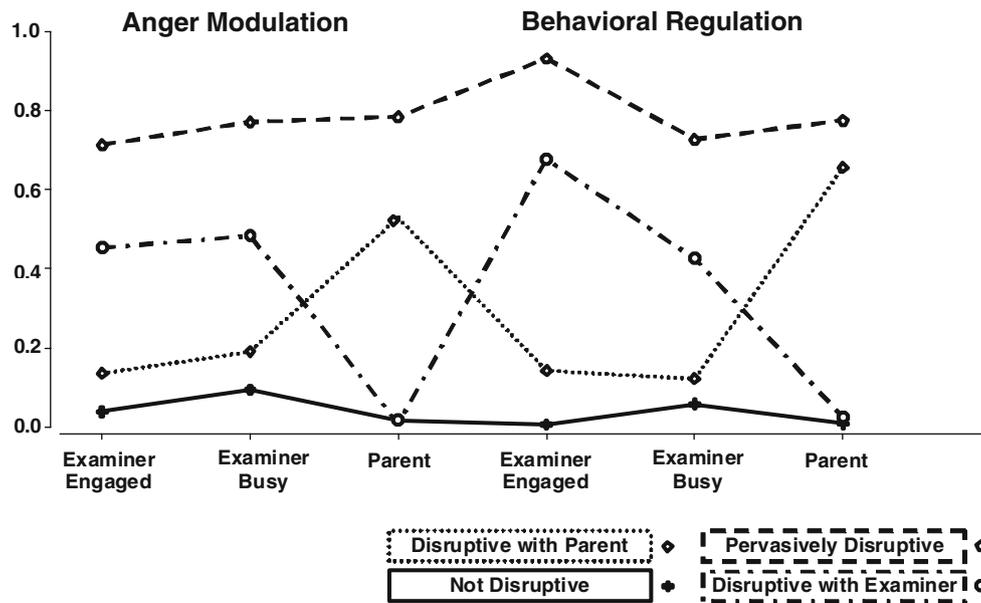


Fig. 2 Latent class solution of observed behavior on the Disruptive Behavior Diagnostic Observation Schedule ($n=327$). The x-axis denotes the three measures of observed Anger Modulation and three measures of observed Behavioral Regulation across examiner-child interactions (when the examiner is either engaged or disengaged with the child) and parent-child interactions, respectively. The y-axis denotes the probability of observing high versus low levels of observed Anger Modulation and/or Behavioral Regulation. Latent classes are denoted

via plotlines along the following classes of disruptive behavior: **a** low disruptive behavior across interactions (Not Disruptive, $n=153$), **b** high disruptive behavior with parent only and not the clinical examiner (Disruptive with Parent, $n=96$), **c** high disruptive behavior with the clinical examiner only and not the parent (Disruptive with Examiner, $n=49$), and **d** high disruptive behavior with parent and clinical examiner (Pervasively Disruptive, $n=29$). The fit statistics of this latent class solution are as follows: $\chi^2(36)=34.88, ns, L^2=39.66, ns$

parent and examiner (Pervasively Disruptive). These patterns were similar regardless of the domain of observed disruptive behavior (i.e., Anger Modulation and Behavior Regulation performed similarly in terms of patterns of variability across interactions), and across the two Examiner contexts (see Fig. 2). Stated another way, our LCA found that children that exhibited similar probabilities of disruptive behavior across both DB-DOS Examiner contexts and low probabilities within the DB-DOS Parent context were classified into a single latent class. Patterns across the two DB-DOS domains (i.e., Behavioral Regulation and Anger Modulation) were similar and indicated a general, rather than domain-specific, pattern of disruptive behavior that varied across and within adult-child interactions, regardless of domain.

As a further test of our latent class solution, we tested its longitudinal stability by using the baseline LCA classification information to obtain probabilities of class membership for each child with DB-DOS data at 1-year follow-up. Specifically, we used the patterns of disruptive behavior exhibited at the baseline assessment and their corresponding latent classification as a basis for assigning a latent classification to the patterns of disruptive behavior at 1-year follow-up. Just as in the baseline LCA, each child was assigned to the latent class having the highest probability or highest correspondence to their pattern of disruptive behavior within and across interactions. The chi square tests of the cross-tabulation of

the baseline and 1-year follow-up class memberships showed a high level of stability overall (Table 4). Specifically, with three degrees of freedom, the Cramer’s V of the relation of .30 is interpreted as a large effect size for the Cramer’s V metric (see Gravetter and Wallnau 2006). Further, although a greater proportion of children were classified in the Not Disruptive class at 1-year follow-up, children tended to be classified in the same latent classes at 1-year follow-up if baseline classification information was used.² For example, adjusted standardized residual (ASR) statistics representing movement between time periods into and out of latent classifications suggested that children with baseline membership in the Disruptive with Parent class were more likely to remain in that class at 1-year follow-up than they were to shift over to classification in the Disruptive with Examiner class (see Table 4) (for further information on interpreting ASRs see Haberman 1978). This finding was consistent with observations of children with baseline membership in the Disruptive with Examiner class. The one exception was the Pervasively Disruptive class, for which a decrease in class

² As an aside, the increase of children identified in the Not Disruptive class is consistent with the decrease at 1-year follow-up of high disruptive behavior symptoms children identified by both parents (high disruptive behavior=61 [21.2%]; low disruptive behavior=227 [78.8%]) and teachers (high disruptive behavior=45 [15.6%]; low disruptive behavior=243 [84.4%]).

Table 3 Prevalence of high disruptive behavior across interactions and identification of latent classes of disruptive behavior on the DB-DOS, baseline ($n=327$)

Measured variables				
DB-DOS Domain	<i>Examiner Engaged</i>	<i>Examiner Busy</i>	<i>Parent</i>	
High on Anger Modulation	65 (19.9%)	82 (25.1%)	84 (25.7%)	
High on Behavioral Regulation	80 (24.5%)	65 (19.9%)	97 (29.7%)	
Latent variables				
Latent class	<i>N (%)</i>	<i>Latent class probabilities</i>		<i>Mean assignment probability</i>
Not Disruptive	153 (46.8%)	0.41	0.84	
Disruptive with Parent	96 (29.4%)	0.34	0.93	
Disruptive with Examiner	49 (15%)	0.16	0.87	
Pervasively Disruptive	29 (8.8%)	0.09	0.87	
Total	327 (100%)	1.00	0.87	
Conditional probabilities for measured variables				
Measured variable	Latent classes			
	<i>Not Disruptive</i>	<i>Disruptive with Parent</i>	<i>Disruptive with Examiner</i>	<i>Pervasively Disruptive</i>
Examiner Engaged-Anger Modulation	High: 0.07	High: 0.23	High: 0.38	High: 0.32
	Low: 0.49	Low: 0.37	Low: 0.11	Low: 0.03
Examiner Engaged-Behavioral Regulation	High: 0.01	High: 0.19	High: 0.46	High: 0.34
	Low: 0.54	Low: 0.38	Low: 0.07	Low: 0.01
Examiner Busy-Anger Modulation	High: 0.15	High: 0.25	High: 0.32	High: 0.27
	Low: 0.49	Low: 0.37	Low: 0.11	Low: 0.03
Examiner Busy-Behavioral Regulation	High: 0.11	High: 0.20	High: 0.36	High: 0.32
	Low: 0.48	Low: 0.37	Low: 0.12	Low: 0.03
Parent-Anger Modulation	High: 0.03	High: 0.70	High: 0	High: 0.27
	Low: 0.54	Low: 0.21	Low: 0.22	Low: 0.03
Parent-Behavioral Regulation	High: 0.01	High: 0.75	High: 0.01	High: 0.23
	Low: 0.57	Low: 0.17	Low: 0.23	Low: 0.03

DB-DOS=Disruptive Behavior Diagnostic Observation Schedule; Not Disruptive=Low probability of disruptive behavior across parent-child and examiner-child interactions; Disruptive with Parent=High probability of disruptive behavior across parent-child and not examiner-child interactions; Disruptive with Examiner=High probability of disruptive behavior across examiner-child and not parent-child interactions; Pervasively Disruptive=High probability of disruptive behavior across both parent-child and examiner-child interactions; Conditional probabilities are to be interpreted across the row of a given indicator and within each value; Probabilities sum to 100% in each row (e.g., on the indicator Parent-Anger Modulation, High values across the row of four latent class High values total 100% probability).

membership was observed at 1-year follow-up. However, the majority of children in the baseline Pervasively Disruptive class that did shift classes at 1-year follow-up remained in an observed disruptive behavior class (i.e., either Disruptive with Parent or Disruptive with Examiner classes); these children were quite unlikely to be classified in the Not Disruptive class. Taken together, these findings suggest moderate stability of the classes across time.

Hypothesis 1: Patterns of Observed Disruptive Behavior on the DB-DOS Relate to Parent-Teacher Rating Discrepancies of Disruptive Behavior

We conducted a preliminary test of the relations between the four-class LCA solution based on the baseline DB-DOS data and the four groups of

informant-disruptive preschoolers, or definitions “a” (Neither Parent nor Teacher), “b” (Parent Only), “c” (Teacher Only), and “d” (Both Parent and Teacher) mentioned previously (Table 5). The *chi square* tests of the cross-tabulation of the DB-DOS class memberships and informant-disruptive preschooler status suggested a significant relation. Similar to the ASRs reported in Table 4, DB-DOS class memberships followed a distribution parallel to the distribution of children in the informant-disruptive groupings based on parent and teacher report. Thus, children in the Not Disruptive class were more likely to be placed in the Neither Parent nor Teacher group than in any other group, and children in the Disruptive with Parent, Disruptive with Examiner, and Pervasively Disruptive classes were more likely to be placed in the Parent Only, Teacher Only, and Both Parent

Table 4 Cross-tabulation of baseline and 1-year follow-up identification of latent classes of disruptive behavior on the DB-DOS ($n=292$)

Baseline	1-Year follow-up				Total
	Not Disruptive	Disruptive with Parent	Disruptive with Examiner	Pervasively Disruptive	
Not Disruptive	133 (45.5%) ASR=5.7	1 (0.3%) ASR=-3.6	2 (0.7%) ASR=-4.1	1 (0.3%) ASR=-0.5	137 (46.9%)
Disruptive with Parent	73 (25%) ASR=-0.4	11 (3.8%) ASR=3.0	4 (1.4%) ASR=-1.6	0 ASR=-1.1	88 (30.1%)
Disruptive with Examiner	29 (9.9%) ASR=-2.6	2 (0.7%) ASR=-0.4	10 (3.4%) ASR=3.9	0 ASR=-0.7	41 (14%)
Pervasively Disruptive	11 (3.8%) ASR=-6.2	4 (1.4%) ASR=2.0	9 (3.1%) ASR=5.0	2 (0.7%) ASR=3.5	26 (8.9%)
Total	246 (84.2%)	18 (6.2%)	25 (8.6%)	3 (1%)	292
Stability of classifications, predicted by baseline classifications ($n=292$)		$\chi^2(9)=80.03$, Cramer's $V(3)=0.30$, kappa=0.19, $p<0.001$			

DB-DOS=Disruptive Behavior Diagnostic Observation Schedule; ASR=Adjusted Standardized Residual, denoting the movement or stability of children's latent classifications between the baseline assessment and assessment at 1-year follow-up. Positive ASRs at or above 1.96 should be interpreted as indication of a significant likelihood (or for negative values at or below -1.96, an unlikelihood) that a given classification at baseline would receive the classification corresponding to the ASR at 1-year follow-up (see Haberman 1978); Not Disruptive=Low probability of disruptive behavior across parent-child and examiner-child interactions; Disruptive with Parent=High probability of disruptive behavior across parent-child interactions; Disruptive with Examiner=High probability of disruptive behavior across examiner-child interactions; Pervasively Disruptive=High probability of disruptive behavior across both parent-child and examiner-child interactions. Latent classifications at 1-year follow-up were created with DB-DOS data taken approximately 1 year later, and based on the original latent class solution arrived at based on baseline data. Although 1-year follow-up data was available for 298 children, tests of the stability of latent classifications between baseline and 1-year follow-up were based on those children that had data available at both time points ($n=292$; see Fig. 1).

Table 5 Cross-tabulation of baseline latent classes of disruptive behavior on the DB-DOS and groups of informant disruptive preschoolers based on parent and teacher report ($n=288$)

DB-DOS latent classification	Informant-disruptive preschooler status				Total
	Neither Parent nor Teacher	Parent Only	Teacher Only	Both Parent and Teacher	
Not Disruptive	91 (31.6%) ASR=3.5	24 (8.3%) ASR=-1.6	13 (4.5%) ASR=-1.3	6 (2.1%) ASR=-2.2	134 (46.5%)
Disruptive with Parent	45 (15.6%) ASR=-1.0	26 (9.0%) ASR=2.1	9 (3.1%) ASR=-0.7	6 (2.1%) ASR=-0.5	86 (29.9%)
Disruptive with Examiner	24 (8.3%) ASR=-0.5	7 (2.4%) ASR=-1.2	9 (3.1%) ASR=1.7	5 (1.7%) ASR=0.7	45 (15.6%)
Pervasively Disruptive	4 (1.4%) ASR=-4.0	7 (2.4%) ASR=1.0	5 (1.7%) ASR=1.4	7 (2.4%) ASR=4.0	23 (8%)
Total	164 (56.9%)	64 (22.2%)	36 (12.5%)	24 (8.3%)	288
Relations between DB-DOS latent classifications and informant-disruptive preschooler status		$\chi^2(9)=35.97$, Cramer's $V(3)=0.20$, kappa=0.16, $p<0.001$			

DB-DOS=Disruptive Behavior Diagnostic Observation Schedule; ASR=Adjusted Standardized Residual, denoting the relations between children's latent classifications and their informant-disruptive status. Positive ASRs at or above 1.96 should be interpreted as indication of a significant likelihood (or for negative values at or below -1.96, an unlikelihood) that a given latent classification at baseline would correspond to a given identification of disruptive behavior symptoms by parent, teacher, both informants, or neither informants (see Haberman 1978); Not Disruptive=Low probability of disruptive behavior across parent-child and examiner-child interactions; Disruptive with Parent=High probability of disruptive behavior within parent-child and not examiner-child interactions; Disruptive with Examiner=High probability of disruptive behavior within examiner-child and not parent-child interactions; Pervasively Disruptive=High probability of disruptive behavior across both parent-child and examiner-child interactions; Neither Parent nor Teacher=No informant positively identified disruptive behavior; Parent Only=Parent and not teacher identified disruptive behavior; Teacher Only=Teacher and not parent identified disruptive behavior; Both Parent and Teacher=Parent and teacher simultaneously identified disruptive behavior.

and Teacher groups, respectively. Taken together, these findings support a linkage between observed behavior and informant discrepancies.

To conduct the primary test of our hypotheses, the group of preschoolers comprising the Neither Parent nor Teacher definition of informant-disruptive preschooler status was the reference group in a multinomial logistic regression, consistent with recent work (e.g., Clifton and Pilkonis 2007; Odgers et al. 2007; Rose et al. 2007; Storr et al. 2007). In this regression, dummy codes for the baseline DB-DOS latent classifications were entered as the predictor (1=Not Disruptive, 2=Disruptive with Parent, 3=Disruptive with Examiner, 4=Pervasively Disruptive) and the four definitions of informant-disruptive preschoolers (1=Neither Parent nor Teacher, 2=Parent Only, 3=Teacher Only, 4=Both Parent and Teacher) were entered as the outcome. Results are reported in Table 6.

As predicted, observed disruptive behavior on the DB-DOS predicted classifications of informant-disruptive preschoolers. Further, context-specific disruptive behavior observed on the DB-DOS predicted informant-disruptive preschooler status, depending on the “match” between the DB-DOS interaction and the informant identifying the preschooler as disruptive. Specifically, relative to the Neither Parent nor Teacher group, Parent Only group members exhibited a high probability of being identified as disruptive on the DB-DOS during parent interactions only and not examiner interactions only. In contrast, children in the Teacher Only group exhibited a high probability of being identified as disruptive on the DB-DOS during examiner interactions only and not parent interactions only. As expected, children in the Both Parent

and Teacher group exhibited a high probability of being identified as disruptive on the DB-DOS across parent and examiner interactions, and not specifically within parent or examiner interactions. As can be seen in Table 6, disruptive behavior on the DB-DOS across parent and examiner interactions also related to identifications of children in the Parent Only and Teacher Only groups. However, relative to the Neither Parent nor Teacher group, the odds ratio for the Both Parent and Teacher group was between three and four times the odds ratios observed for the Parent Only and Teacher Only groups.

Hypothesis 2: Links between Informant Discrepancies and Observed Behavior Cannot be Explained by Children’s Impairment or Problematic Parenting

Similar to tests of the relation between parent-teacher rating discrepancies and observed disruptive behavior, children’s impairment was assessed using both parent and teacher reports. Thus, to address the child impairment portion of this hypothesis we conducted *chi square* analyses comparing the distributions of children who were identified at baseline as impaired by parent only ($n=36$), teacher only ($n=47$), or both parent and teacher ($n=23$) (based on a C-GAS rating at or below 60), and comparing distributions of impaired children within the Pervasively Disruptive class to all other DB-DOS classes. Thus, in these analyses we only included those children identified by parent and/or teacher as impaired ($n=106$), and examined the relation between the distribution of impaired children to the distribution of child DB-DOS classifications. There was a non-significant association between the distribution of

Table 6 Multinomial logistic regression analysis differentiating high- and low-disruptive behavior symptom children, based on latent class assignment ($n=288$)

Latent class assignment	Odds ratios for disruptive behavior ratings, parent and teacher report separately	Odds ratios for disruptive behavior ratings, parent and teacher report combined
Disruptive with Parent	Parent Only OR=2.19, $p<0.05$, CI=1.13, 4.24 Teacher Only OR=1.40, $p=0.47$, CI=0.55, 3.52	Both Parent and Teacher OR=2.02, $p=0.24$, CI=0.61, 6.62
Disruptive with Examiner	Parent Only OR=1.10, $p=0.83$, CI=0.42, 2.87 Teacher Only OR=2.62, $p<0.05$, CI=1.00, 6.86	Both Parent and Teacher OR=3.16, $p=0.07$, CI=0.88, 11.24
Pervasively Disruptive	Parent Only OR=6.63, $p<0.01$, CI=1.79, 24.55 Teacher Only OR=8.75, $p<0.01$, CI=2.08, 36.84	Both Parent and Teacher OR=26.54, $p<0.001$, CI=6.04, 116.64

Odds ratios reported in table reflect results of multinomial logistic regression in which disruptive behavior status was employed as the dependent variable (i.e., Neither Parent nor Teacher, Parent Only, Teacher Only, Both Parent and Teacher [AND Rule]) with Neither Parent nor Teacher as the reference category, and latent class assignment was employed as the independent variable (i.e., Not Disruptive, Disruptive with Parent, Disruptive with Examiner, Pervasively Disruptive); Not Disruptive=Low probability of disruptive behavior across parent-child and examiner-child interactions; Disruptive with Parent=High probability of disruptive behavior within parent-child and not examiner-child interactions; Disruptive with Examiner=High probability of disruptive behavior within examiner-child and not parent-child interactions; Pervasively Disruptive=High probability of disruptive behavior across both parent-child and examiner-child interactions; OR=Odds ratio; Parent Only OR=Odds ratio based on parent-identified disruptive behavior (when teacher does not identify such behavior); Teacher Only OR=Odds ratio based on teacher-identified disruptive behavior (when parent does not identify such behavior); Both Parent and Teacher OR=Odds ratio based on both parent- and teacher-identified disruptive behavior; CI=95% Confidence interval for odds ratios.

impaired children via parent and/or teacher report and the distribution of the DB-DOS classes, $\chi^2(6)=9.36$, $p=0.15$, Cramer's $V(2)=0.21$. Further, there were non-significant associations between child impairment by informant when the distribution of impaired children in the Pervasively Disruptive class was individually compared to the distributions of impaired children in the Not Disruptive class, $p=0.66$, Disruptive with Parent class, $p=0.63$, and Disruptive with Examiner class, $p=0.11$. Thus, this pattern of findings did not support impairment as explaining the relation between discrepancies and observed behavior.

Problematic parenting was assessed using behavioral observations taken of parents during the DB-DOS Parent context tasks. Thus, we examined the problematic parenting portion of this hypothesis by conducting the LCA with baseline data reported previously, but including observed problematic parenting on the P-COS as a covariate. Consistent with the original LCA, the four-class latent solution once again yielded the best model fit, $\chi^2(96)=95.68$, ns , $L^2=100.39$, ns . The structure and composition of the classes (i.e., Not Disruptive [$n=153$], Disruptive with Parent [$n=94$], Disruptive with Examiner [$n=51$], Pervasively Disruptive [$n=29$]) were virtually identical to the classes from the original LCA, $\chi^2(9)=961.63$, $p<0.001$, $kappa=0.99$, Cramer's $V(3)=0.99$.³ Thus, the structure of the latent class solution reported previously was robust to controlling for quality of observed parenting behavior.

Discussion

The current study provides an initial empirical test of a substantive conceptualization of informant discrepancies as clinically meaningful and not merely comprised of measurement error or indicative of relative clinical severity. Specifically, we tested whether laboratory observed situational variations in young children's disruptive behavior systematically correspond to patterns of independent parent-teacher discrepancies on disruptive behavior ratings in these same children. We identified systematic situational variability in children's disruptive behavior observed during a laboratory task, related to who was interacting with the child (parent vs. examiner), and with corresponding relations to who was reporting disruptive behavior symptoms (parent vs. teacher).

We found substantial variations in when observed disruptive behavior occurred, with 29.4% of preschoolers displaying disruptive behavior in interactions with parent only, 15% with examiner only, and 8.8% across both types

of interactions. These groups of children were stable over time. This suggests meaningfully different patterns of exhibition of such behavior by setting or context, suggesting consideration from our theoretical model (Attribution Bias Context Model; see De Los Reyes and Kazdin 2005) that variation in ratings may provide important distinctions between where and how such behavior occurs among a group all exhibiting substantial disruptive behavior.

Building on this foundation, we demonstrated that this variability in children's disruptive behavior observed in the laboratory mapped onto discrepancies between parent and teacher ratings of disruptive behavior in real-world settings. Specifically, parallels were found between observed behavior with parental and non-parental adults on the DB-DOS and parent-teacher discrepancies on disruptive behavior symptom ratings. Identifying these parallels in variations in children's behavior across parent, teacher, and laboratory observation ratings suggests that when parent-teacher rating discrepancies arise, that these discrepancies reflect real-world variations in children's behavior exhibited across contexts. That is, when informants differ in both their ratings of the same child's behavior and the situations within which they can be presumed to primarily observe the child's behavior (e.g., parent at home, teacher at school), these discrepancies might reflect that the child's behavior varies across situations, and that informants' ratings are reflecting these variations in behavior. That these patterns could not be accounted for by child impairment and observed problematic parenting provides further empirical support of the contextual variability we observed and the informant discrepancies that mapped onto this variability.

Based on our findings that contextual variation in laboratory observations of young children's disruptive behavior map on to parent-teacher rating discrepancies, a key question arises: *Why might children's behavior vary across interactional contexts?* As mentioned previously, this question has recently been posed in the behavioral genetics literature, which commonly encounters differences in estimates of genetic versus environmental contributions to variance in behavior, depending on the informant rating behavior (Bartels et al. 2007). Bartels and colleagues (2007) argue that informant discrepancies in tests of genetic and environmental influences on behavior represent, in part, unique expressions of the behavior, depending on the informant observing and rating the behavior. In this sense, inconsistencies do not suggest weaker effects, but rather, different effects. Similarly, children's inconsistent displays of disruptive behavior within and across contexts might suggest multiple and unique manifestations of the behavior. Stated another way, context-specific disruptive behavior may arise from at least partially distinct etiological factors, relative to disruptive behavior pervasively exhibited across contexts.

³ Only two cases differed in their classifications between the original LCA and the LCA controlling for observed parenting. These two cases were classified in the Disruptive with Examiner class in the observed parenting LCA whereas they were classified in the Disruptive with Parent class in the original LCA.

This interpretation is in line with findings from Baker and colleagues (2007) who found distinct variance estimates of shared environment or informant effects, depending on whether the informant was the caregiver or teacher. In that study, it was possible to disentangle teacher ratings with regards to shared environment and teacher-specific effects, because twins in the sample varied on whether they shared the same classroom environment, and thus the same teacher. This investigation identified greater similarity in teacher ratings of twins in the same versus different classrooms, accounting for 28.1% of the variance in shared environmental effects. However, this same twin similarity effect could not be identified when examining caregiver and child reports of antisocial behavior: Caregiver-child similarity in twin ratings was no more similar for twins in the same versus different classroom settings. The present study expands upon these findings by providing a careful and strong test of disentangling informant and context variation and in doing so suggests the validity of the contention here that informant discrepancies provide meaningful information about heterogeneity in how and when problem behaviors occur. We encourage future research to examine the relations between observed behavior and informant discrepancies using observational paradigms that systematically vary the relationship between the child being rated and the adults with whom they interact.

Limitations

There are limitations to the present study. First, our findings were based on dichotomous ratings of disruptive behavior, functional impairment, and observed parenting. Large informant discrepancies have been observed regardless of measurement scale or format (De Los Reyes and Kazdin 2005). Further, prior work suggests that the scores we employed to identify disruptive behavior, functional impairment, and problematic parenting behaviors differentiate clinically relevant from normative behaviors both concurrently and over time (Hill et al. 2008; Wakschlag et al. 2008b). As a result, there was a strong rationale present to conduct this initial study based on dichotomous scores. Nevertheless, future work ought to replicate and extend our findings based on discrepancies between continuous scores.

Second, we employed parent and teacher measures of disruptive behavior in children that varied in structure and format (parent interview, teacher questionnaire). However, this was mitigated by reliance on two measures that did not vary in their assessment of DSM-IV symptoms of childhood disruptive behavior, and the two measures were selected because of their developmental appropriateness for assessment of disruptive behavior in young children. Further, levels of parent-teacher agreement on these measures were consistent with the magnitudes of agreement

observed in prior work (Achenbach 2006; Achenbach et al. 1987). At the same time, we examined parent-teacher discrepancies between informants' measures that were not completely identical in format and structure. Thus, we were only able to examine parent-teacher discrepancies on dichotomous identifications of clinically relevant disruptive behavior, as opposed to parent-teacher discrepancies using continuous scores, or differences between "total scores" from each measure (see De Los Reyes et al. 2008; De Los Reyes and Kazdin 2004). Future work would benefit from examining parallel parent and teacher measures that could be examined using both dichotomous and continuous discrepancies measures.

Research and Clinical Implications

Consistent with the findings and the work of others (Achenbach 2006; Bartels et al. 2007; Kraemer et al. 2003; De Los Reyes and Kazdin 2005) the results from this study suggest that informant discrepancies can fruitfully be utilized as an opportunity to identify and understand variation in manifestations of developmental psychopathology. While this study cannot identify just what the implications are for clinical understanding nor is it meant to reveal the causal basis for informant discrepancies, we think the results do support more careful and substantive consideration of context specificity and informant discrepancies in relation to disruptive behavior disorders, and perhaps more generally. It may be that informant discrepancies are the result of a complex mixture of seriousness of disorder, contextual basis for symptom expression, and systematic differences among informants in their perspectives on the symptoms being rated. It may also be that informant discrepancies can provide meaningful insight into differential responses to interventions. Specifically, as multiple informants' ratings are prized in assessing outcomes of interventions (Kazdin 2003), it may be useful to reconsider what informant-specific variations in results mean. Not surprisingly, inconsistent findings commonly arise across multiple informants' outcome ratings taken within the same study, and these inconsistencies are often interpreted as either measurement error or as evidence of the ineffectiveness of the intervention being studied (De Los Reyes and Kazdin 2006b). Alternatively, our findings suggest that future work should examine whether these inconsistencies reveal variations in changes in behavior attributable to intervention effects, or point to the circumstances in which interventions yield beneficial effects (De Los Reyes and Kazdin 2008).

An example may be helpful of how inconsistent findings might inform intervention research and intervention development. Consider a study that reveals that an intervention reduces child disruptive behavior based on teacher- but not

parent-rated outcomes. Interpreting inconsistencies as yielding meaningful information might result in concluding that these findings indicate true differences in the extent to which behavior improved in various settings. For example, perhaps the intervention effectively targets contingencies influencing behavior in non-home settings (e.g., deviant peer interactions; teacher's poor classroom behavior management strategies). This might lead to future work typified by improved specification of variations in behavior, and revisions to the intervention that specifically target this behavioral variation. For instance, a controlled trial might be conducted to test whether targeting behaviors in a situation-specific manner (intervention techniques specifically targeting deviant peer interactions or teachers' classroom management skills) yields more pervasive effects than a situation-nonspecific version of the intervention (intervention as usual). Additionally, observational measures such as the DB-DOS might be employed in future intervention research as independent measures by which to examine whether informant-based discrepancies in reports of intervention outcomes are reflective of meaningful situation-specific treatment changes. The investigation of the clinical meaning of informant discrepancies within an intervention context and in particular the relations between discrepancies and changes in behavior over the course of an intervention may be a particularly fruitful avenue for future research.

Acknowledgements We gratefully acknowledge contributions of our collaborators on the DB-DOS study, Margaret Briggs-Gowan, Alice Carter, Barbara Danis, Carri Hill, Kate Keenan, Helen Egger, Dominic Cicchetti, and Bennett Leventhal and our student, Melanie Dirks.

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. doi:10.1111/j.0963-7214.2006.00414.x.
- Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological Bulletin, 131*, 361–382. doi:10.1037/0033-2909.131.3.361.
- 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. doi:10.1037/0033-2909.101.2.213.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- 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. doi:10.1037/0021-843X.116.2.219.
- 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. doi:10.1037/1082-989X.12.4.451.
- Bidaut-Russell, M., Reich, W., Cottler, L. B., Robins, L. N., Compton, W. M., & Mattison, R. E. (1995). The Diagnostic Interview Schedule for Children (PC-DISC v.3.0): Parents and adolescents suggest reasons for expecting discrepant answers. *Journal of Abnormal Child Psychology, 23*, 641–659. doi:10.1007/BF01447667.
- Campbell, S. B. (2002). *Behavior problems in preschool children: Clinical and developmental issues* (2nd ed.). New York: Guilford.
- Clifton, A., & Pilkonis, P. A. (2007). Evidence for a single latent class of Diagnostic and Statistical Manual of Mental Disorders borderline personality pathology. *Comprehensive Psychiatry, 48*, 70–78. doi:10.1016/j.comppsy.2006.07.002.
- De Los Reyes, A., Goodman, K. L., Kliewer, W., & Reid-Quinones, K. R. (2008). Whose depression relates to discrepancies? Testing relations between informant characteristics and informant discrepancies from the perspectives of both informants. *Psychological Assessment, 20*, 139–149. doi:10.1037/1040-3590.20.2.139.
- De Los Reyes, A., & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment, 16*, 330–334. doi:10.1037/1040-3590.16.3.330.
- De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483–509. doi:10.1037/0033-2909.131.4.483.
- De Los Reyes, A., & Kazdin, A. E. (2006a). Informant discrepancies in assessing child dysfunction relate to dysfunction within mother-child interactions. *Journal of Child and Family Studies, 15*, 643–661. doi:10.1007/s10826-006-9031-3.
- De Los Reyes, A., & Kazdin, A. E. (2006b). Conceptualizing changes in behavior in intervention research: The range of possible changes model. *Psychological Review, 113*, 554–583. doi:10.1037/0033-295X.113.3.554.
- 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. doi:10.1111/j.1467-8721.2008.00546.x.
- 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. doi:10.1207/s15374424jccp3302_14.
- Dirks, M. A., Henry, D. B., Hill, C., & Wakschlag, L. S. (2009). *Contextual variations in preschoolers' social skills during laboratory observations: Relationships with disruptive behavior*. Manuscript submitted for publication.
- Frick, P. J. (2000). Laboratory and performance-based measures of childhood disorders: Introduction to the special section. *Journal of Clinical Child Psychology, 29*, 475–478. doi:10.1207/S15374424JCCP2904_2.
- Gadow, K., & Sprafkin, J. (1996). *Early childhood inventories (ECI) manual*. Stony Brook: Checkmate Plus.
- Gadow, K., Sprafkin, J., & Nolan, E. (2001). DSM-IV symptoms in community and clinic preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1383–1392. doi:10.1097/00004583-200112000-00008.
- Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review, 113*, 101–131. doi:10.1037/0033-295X.113.1.101.
- Gravetter, F. J., & Wallnau, L. B. (2006). *Statistics for the behavioral sciences* (7th ed.). Belmont: Thomson/Wadsworth.
- Haberman, S. J. (1978). *Analysis of qualitative data. Volume I. Introductory topics*. New York: Academic.
- Hill, C., Maskowitz, K., Danis, B., & Wakschlag, L. S. (2008). Validation of a clinically sensitive, observational coding system

- for parenting behaviors: The Parenting Clinical Observation Schedule (P-COS). *Parenting, Science and Practice*, 8, 153–185. doi:10.1080/15295190802045469.
- Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston: Allyn & Bacon.
- Kazdin, A. E., & De Los Reyes, A. (2007). Conduct disorder. In R. J. Morris & T. R. Kratochwill (Eds.), *The practice of child therapy* (4th ed.), pp. 207–247. Philadelphia: Taylor & Francis.
- Keenan, K., Wakschlag, L. S., & Danis, B. (2001). *Kiddie-disruptive behavior disorder schedule* (version 1.1). Available at: kkeenan@yoda.bsd.uchicago.edu
- Keenan, K., Wakschlag, L. S., Danis, B., Hill, C., Humphries, M., Duax, J., et al. (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46, 457–468. doi:10.1097/CHI.0b013e31803062d3.
- Kerr, D. C. R., Lunkenheimer, E. S., & Olson, S. L. (2007). Assessment of child problem behaviors by multiple informants: A longitudinal study from preschool to school entry. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 48, 967–975. doi:10.1111/j.1469-7610.2007.01776.x.
- 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. doi:10.1176/appi.ajp.160.9.1566.
- Lavigne, J., Arend, R., Rosenbaum, D., Binns, H., Christoffel, K., & Gibbons, R. (1998). Psychiatric disorders with onset in the preschool years: I. Stability of diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, 37, 1246–1254.
- McCutcheon, A. L. (1987). *Latent class analysis*. Newbury Park: Sage.
- Moffitt, T. E., Arseneault, L., Jaffee, S. R., Kim-Cohen, J., Koenen, K. C., Odgers, C. L., et al. (2007). Research review: DSM-V conduct disorder: Research needs for an evidence base. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 49, 3–33.
- Odgers, C. L., Moretti, M. M., Burnette, M. L., Chauhan, P., Waite, D., & Reppucci, N. D. (2007). A latent variable modeling approach to identifying subtypes of serious and violent female juvenile offenders. *Aggressive Behavior*, 33, 339–352. doi:10.1002/ab.20190.
- Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 34, 523–540. doi:10.1207/s15374424jccp3403_8.
- Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin*, 112, 485–499. doi:10.1037/0033-2909.112.3.485.
- Rose, J. S., Chassin, L., Presson, C., Sherman, S. J., Stein, M. D., & Col, N. (2007). A latent class typology of young women smokers. *Addiction (Abingdon, England)*, 102, 1310–1319. doi:10.1111/j.1360-0443.2007.01889.x.
- Setterberg, S., Bird, H., Gould, M., Schaffer, D., & Fisher, P. (1992). *Parent and interviewer versions of the children's global assessment scale*. Unpublished measure. New York: Columbia University.
- Storr, C. L., Accornero, V. H., & Crum, R. M. (2007). Profiles of current disruptive behavior: Association with recent drug consumption among adolescents. *Addictive Behaviors*, 32, 248–264. doi:10.1016/j.addbeh.2006.03.045.
- Tolan, P., Gorman-Smith, D., Huesmann, L., & Zelli, A. (1997). Assessment of family relationship characteristics: A measure to explain risk for antisocial behavior and depression among urban youth. *Psychological Assessment*, 9, 212–223. doi:10.1037/1040-3590.9.3.212.
- Wakschlag, L. S., Leventhal, B. L., & Thomas, B. (2007a). Disruptive behavior disorders and ADHD in preschool children: Characterizing heterotypic continuities for a developmentally informed nosology for DSM-V. In W. Narrow, M. First, P. Sirovatka & D. Regier (Eds.), *Age and gender considerations in psychiatric diagnosis: A research agenda for DSM-V*, pp. 243–258. Arlington: American Psychiatric Association.
- Wakschlag, L. S., Briggs-Gowan, M. J., Carter, A. S., Hill, C., Danis, B., Keenan, K., et al. (2007b). A developmental framework for distinguishing disruptive behavior from normative misbehavior in preschool children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 48, 976–987. doi:10.1111/j.1469-7610.2007.01786.x.
- Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., et al. (2008a). 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 and Adolescent Psychiatry*, 47, 622–631. doi:10.1097/CHI.0b013e31816c5b5b.
- Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B. L., Keenan, K., et al. (2008b). Observational assessment of preschool disruptive behavior. Part II. Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child and Adolescent Psychiatry*, 47, 632–641. doi:10.1097/CHI.0b013e31816c5c10.
- Wakschlag, L. S., & Danis, B. (in press). Characterizing early childhood disruptive behavior: Enhancing developmental sensitivity. In C. H. Zeanah (Ed.), *Handbook of infant mental health* (3rd ed.). New York: Guilford.
- Wakschlag, L. S., & Keenan, K. (2001). Clinical significance and correlates of disruptive behavior symptoms in environmentally at-risk preschoolers. *Journal of Clinical Child Psychology*, 30, 262–275. doi:10.1207/S15374424JCCP3002_13.
- 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. doi:10.1023/A:1023244512119.
- Youngstrom, E. A., Izard, C., & Ackerman, B. (1999). Dysphoria-related bias in maternal ratings of children. *Journal of Consulting and Clinical Psychology*, 67, 905–916. doi:10.1037/0022-006X.67.6.905.